# Cálculo para la ingeniería

Salvador Vera

29 de abril de 2003

# Índice general

**3. Derivación de funciones de varias variables. 5**

3.1. Derivadas parciales. . . . . 5

3.1.1. Introducción. . . . . 5

3.1.2. Definición. . . . . 6

3.1.3. Funciones de más de dos variables . . . . . 10

3.1.4. Razón de cambio . . . . . 11

3.1.5. Interpretación geométrica de las derivadas parciales . . . . . 12

3.2. Derivadas parciales de órdenes superiores . . . . . 14

3.3. Derivadas direccionales. . . . . 18

3.3.1. Derivadas direccionales . . . . . 18

3.3.2. Relación entre la derivada direccional y las derivadas parciales . . . . 20

3.4. Diferenciabilidad . . . . . 22

3.4.1. Definición . . . . . 22

3.4.2. Diferenciabilidad y derivadas direccionales . . . . . 34

3.4.3. La derivada según una dirección curva . . . . . 35

3.5. Gradiente . . . . . 35

3.5.1. Definición . . . . . 35

3.5.2. Vector gradiente y derivada direccional . . . . . 37

3.5.3. Gradiente y curvas de nivel . . . . . 38

3.6. Plano tangente . . . . . 39
3.6.1. Vectores normales . . . . . 39
3.6.2. Plano tangente . . . . . 41
3.6.3. Recta tangente y plano normal a una curva en el espacio . . . . . 45
3.6.4. La diferencial como aproximación del incremento . . . . . 47
3.7. Funciones vectoriales y matriz Jacobiana . . . . . 50
3.7.1. Funciones vectoriales de variable vectorial . . . . . 50
3.7.2. Continuidad de las funciones vectoriales . . . . . 50
3.7.3. Derivadas parciales de funciones vectoriales . . . . . 53
3.7.4. Funciones vectoriales diferenciables . . . . . 53
3.8. Regla de la cadena . . . . . 56
3.8.1. Funciones compuestas, inversas e implícitas de una variable . . . . . 56
3.8.2. Composición de funciones vectoriales . . . . . 58
3.8.3. Regla de la cadena. Perspectiva teórica: Diferencial . . . . . 61
3.8.4. Regla de la cadena. Perspectiva práctica: Parciales . . . . . 63
3.8.5. Regla de la cadena. Perspectiva general: Matriz jacobiana . . . . . 69
3.9. Funciones implícitas . . . . . 74
3.9.1. Funciones de una variable . . . . . 74
3.9.2. Funciones de dos variables . . . . . 78
3.10. Extremos de las funciones de varias variables . . . . . 81
3.10.1. Introducción . . . . . 81
3.10.2. Definiciones . . . . . 82
3.10.3. Criterio del hessiano . . . . . 85
3.10.4. Extremos condicionados. Multiplicadores de Lagrange . . . . . 87
3.10.5. Máximos y mínimos absolutos . . . . . 89

# Capítulo 3

# Derivación de funciones de varias variables.

Salvador Vera Ballesteros
www.satd.uma.es/matap/svera

## 3.1. Derivadas parciales.

### 3.1.1. Introducción.

**Funciones de una variable.** Para una función de una variable $f : \mathcal{D} \subseteq \mathbb{R} \to \mathbb{R}$ definida en el intervalo abierto $\mathcal{D}$ de $\mathbb{R}$, se define la derivada de $f$ en $x_0 \in \mathcal{D}$, denotada por $f'(x_0)$, como el valor del siguiente límite, si existe y es finito.

$$f'(x_0) = \frac{dy}{dx}(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

Si $f'(x_0)$ existe, su valor nos da la pendiente de la recta tangente a la gráfica de la función $y = f(x)$ en el punto $\big(x_0, f(x_0)\big)$. El valor de $f'(x_0)$ también representa la *razón de cambio instantáneo* de $y$ respecto de $x$.

**Nota:** La razón de definir la derivada en un punto perteneciente a un conjunto *abierto* $\mathcal{D}$ (dominio de la función), es para poder asegurar que para $h \in \mathbb{R}$ pequeño, se tenga $(x_0 + h) \in \mathcal{D}$ y que así tenga sentido la expresión $f(x_0 + h)$ que aparece en la definición de derivada.

La importancia de estudiar el concepto de derivada radica en que a partir de la derivada de una función en un punto se puede obtener información sobre el comportamiento de la propia función, aunque esta información es sólo *local*, es decir, a partir de $f'(x_0)$ obtenemos información sobre el comportamiento de $f$, pero sólo alrededor de $x_0$. Así, por ejemplo, el simple hecho de la existencia de $f'(x_0)$ señala un comportamiento suave de la gráfica de la función en los alrededores del punto $\big(x_0, f(x_0)\big)$; el signo de $f'(x_0)$ señala el crecimiento o decrecimiento de la función alrededor del punto, etc. Esta información, aunque sólo es local, es muy importante. No obstante, a partir de la función derivada $f'(x)$ podemos obtener una información más *global* del comportamiento de la función.

**Funciones de dos variables.** El concepto de derivada se puede generalizar a funciones de dos variables $z = f(x, y)$. La idea intuitiva responde a la siguiente cuestión: ¿Cómo se va a ver afectada una función de dos variables por una variación de una de sus variables, mientras que la otra variable permanece fija?. Podemos responder a esta cuestión considerando cada vez sólo una de las variables. Esto es, hacemos la derivada de la función cada vez con respecto a una variable, manteniendo constante la otra. Este proceso se conoce como derivación parcial, y su resultado se llama derivada parcial de la función respecto a la variable independiente elegida. Para ello partimos de la idea del concepto de derivada de funciones de una variable *"el límite, cuando el incremento de la variable tiende a cero, del cociente del incremento de la función dividido entre el incremento de la variable"*. Suponemos que una de las variables es constante e incrementamos sólo la otra, es decir, hacemos la derivada suponiendo que la función depende sólo de una de las variables. Con ello se reduce la discusión al caso uni-dimensional considerando una función de varias variables como una función de una sola variable (cada variable separadamente), manteniendo fijas las demás.

### 3.1.2. Definición.

Supongamos que en cierto entorno del punto $(x_0, y_0)$ está dada la función $z = f(x, y)$. Si fijamos la variable $y$: $y = y_0$, obtenemos una función de una sola variable $x$: $z = f(x, y_0)$. La derivada habitual de esta función en el punto $x = x_0$ se llama derivada parcial de la función $f(x, y)$ en el punto $(x_0, y_0)$, respecto de $x$, y se denota por $\dfrac{\partial f(x_0, y_0)}{\partial x}$.
De esta forma,

$$\frac{\partial f(x_0, y_0)}{\partial x} \stackrel{\text{def}}{=} \left.\frac{\partial f(x, y_0)}{\partial x}\right]_{x=x_0}$$

**nota:** Señalemos que la designación de la derivada parcial respecto de la variable $x$ por $\dfrac{\partial f(x_0, y_0)}{\partial x}$ es tradicional. Aunque es más correcto escribir $\dfrac{\partial f}{\partial x}(x_0, y_0)$, ya que $\dfrac{\partial f}{\partial x}$ es un símbolo único, que designa la nueva función, cuyo valor se analiza en el punto $(x_0, y_0)$.

Y teniendo en cuenta la definición de derivada de una sola variable, resulta la siguiente,

**Definición 3.1.** *Dada una función de dos variables $f : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}$ definida en el abierto $\mathcal{D}$ de $\mathbb{R}^2$, se define la derivada parcial de $f$ con respecto a $x$ en el punto $\mathbf{p} = (x_0, y_0)$ de $\mathcal{D}$ como el valor del siguiente límite, si existe y es finito.*

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}$$

*Del mismo modo, se define la derivada parcial de $f$ con respecto a $y$ en $\mathbf{p}$, como el siguiente límite, si existe y es finito.*

$$\frac{\partial f}{\partial y}(x_0, y_0) = \lim_{k \to 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k}$$

**Cálculo de derivadas parciales aplicando la definición.**

**Ejemplo 3.1.** *Calcular las dos derivadas parciales de la función $f(x, y) = xy + x - y$ en el punto $\mathbf{p}(0, 0)$.*

*Solución.*

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{h\cdot 0+h-0-0}{h} =$$
$$= \lim_{h\to 0}\frac{h}{h} = \lim_{h\to 0} 1 = 1$$

$$\frac{\partial f}{\partial y}(0,0) = \lim_{k\to 0}\frac{f(0,0+k)-f(0,0)}{k} = \lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k} = \lim_{k\to 0}\frac{0\cdot k+0-k-0}{k} =$$
$$= \lim_{k\to 0}\frac{-k}{k} = \lim_{k\to 0} -1 = -1$$

**Ejemplo 3.2.** *Calcular las dos derivadas parciales en el punto* $\mathbf{p}(0,0)$ *de la función:*

$$f(x,y) = \begin{cases} \dfrac{xy^2}{x^2+y^2} & si \;\; (x,y)\neq(0,0) \\ 0 & si \;\; (x,y)=(0,0) \end{cases}$$

*Solución.*

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{\dfrac{h\cdot 0^2}{h^2+0^2}-0}{h} =$$
$$= \lim_{h\to 0}\frac{0}{h} = \lim_{h\to 0} 0 = 0$$

$$\frac{\partial f}{\partial y}(0,0) = \lim_{k\to 0}\frac{f(0,0+k)-f(0,0)}{k} = \lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k} = \lim_{k\to 0}\frac{\dfrac{0\cdot k^2}{0^2+k^2}-0}{k} =$$
$$= \lim_{k\to 0}\frac{0}{k} = \lim_{k\to 0} 0 = 0$$

**Ejemplo 3.3.** *Calcular las dos derivadas parciales de la función* $f(x,y) = 2x^2 - xy + y^2$ *en el punto* $\mathbf{p}(1,0)$.

*Solución.*

$$\frac{\partial f}{\partial x}(1,0) = \lim_{h\to 0}\frac{f(1+h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{2(1+h)^2-(1+h)\cdot 0+0^2-(2-0+0)}{h} =$$
$$= \lim_{h\to 0}\frac{2(1+2h+h^2)-2}{h} = \lim_{h\to 0}\frac{2+4h+2h^2-2}{h} = \lim_{h\to 0}\frac{4h+2h^2}{h} =$$
$$= \lim_{h\to 0}(4+2h) = 4$$

$$\frac{\partial f}{\partial y}(1,0) = \lim_{k\to 0}\frac{f(1,0+k)-f(1,0)}{k} = \lim_{k\to 0}\frac{f(1,k)-f(1,0)}{k} = \lim_{k\to 0}\frac{2\cdot 1^2-1\cdot k+k^2-(2)}{k} =$$
$$= \lim_{0\to 0}\frac{2-k+k^2-2}{k} = \lim_{k\to 0}\frac{-k+k^2}{k} = \lim_{k\to 0}(-1+k) = -1$$

**La función derivada parcial**

Si hallamos las derivadas parciales de una función de dos variables $z = f(x, y)$ en un punto genérico $(x, y)$ de su dominio, obtenemos, a su vez, funciones de dos variables, llamadas funciones derivadas parciales. Así:

$$\frac{\partial f}{\partial x}(x,y) = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h}, \qquad \frac{\partial f}{\partial y}(x,y) = \lim_{k\to 0} \frac{f(x,y+k) - f(x,y)}{k}$$

**Ejemplo 3.4.** *Hallar, aplicando la definición, las derivadas parciales de la función:*

$$f(x,y) = x^2y^3 + 5$$

*Solución.*

$$\begin{aligned}
\frac{\partial f}{\partial x}(x,y) &= \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h} = \lim_{h\to 0} \frac{(x+h)^2y^3 + 5 - (x^2y^3+5)}{h} = \\
&= \lim_{h\to 0} \frac{x^2y^3 + 2hxy^3 + h^2y^3 + 5 - x^2y^3 - 5}{h} = \lim_{h\to 0} \frac{2hxy^3 + h^2y^3}{h} = 2x\,y^3 \\
\frac{\partial f}{\partial y}(x,y) &= \lim_{k\to 0} \frac{f(x,y+k) - f(x,y)}{k} = \lim_{k\to 0} \frac{x^2(y+k)^3 + 5 - (x^2y^3+5)}{k} = \\
&= \lim_{k\to 0} \frac{x^2y^3 + 3x^2y^2k + 3x^2yk^2 + x^2k^3 + 5 - x^2y^3 - 5}{k} = \\
&= \lim_{k\to 0} \frac{3x^2y^2k + 3x^2yk^2 + x^2k^3}{k} = \lim_{k\to 0} \left(3x^2y^2 + 3x^2yk + x^2k^2\right) = 3x^2y^2
\end{aligned}$$

**Cálculo de derivadas parciales mediante las reglas de derivación.**

Si observamos los resultados del ejemplo anterior, podemos concluir que no es necesario aplicar la definición para calcular las derivadas parciales, sino que se pueden calcular aplicando las reglas ordinarias de derivación. Esto se deduce de la propia definición, ya que de la definición de las derivadas parciales, como derivadas ordinarias con la condición de que se han fijado todas las variables excepto una, respecto de la cual se toma la derivada, se deduce que al calcular las derivadas parciales se pueden utilizar las reglas del cálculo de las derivadas ordinarias. Es decir,

$\frac{\partial f}{\partial x}(x,y)$ Se puede calcular mediante las reglas de derivación, es decir, como una derivada ordinaria, de la función $f$ respecto de la variable $x$, suponiendo $y$ constante (es decir, como si la función $f$ dependiera sólo de $x$, porque $y$ es un número).

$\frac{\partial f}{\partial y}(x,y)$ Se puede calcular mediante las reglas de derivación, es decir, como una derivada ordinaria, de la función $f$ respecto de la variable $y$, suponiendo $x$ constante (es decir, como si la función $f$ dependiera sólo de $y$, porque $x$ es un número).

**Ejemplo 3.5.** *Calcular las dos derivadas parciales de las siguientes funciones:*

1. $f(x,y) = x^2 + 2xy^2 - y^3$

   *a*) Fijemos en la fórmula la variable $y$, es decir, supongamos que $y$ es un número, con lo cual obtenemos una función de una sola variable $x$; y calculando su derivada tendremos:

$$\frac{\partial f}{\partial x} = 2x + 2y^2$$

*b*) De la misma forma, fijemos ahora la variable $x$ es decir, supongamos que $x$ es un número, con lo cual obtenemos una función de una sola variable $y$; y calculando su derivada tendremos:

$$\frac{\partial f}{\partial y} = 4xy - 3y^2$$

2. $z = (x^2 + y^2)e^{-xy}$

$$\frac{\partial z}{\partial x} = 2xe^{-xy} + (x^2 + y^2)(-ye^{-xy}) = (2x - x^2y - y^3)e^{-xy}$$
$$\frac{\partial z}{\partial y} = 2ye^{-xy} + (x^2 + y^2)(-xe^{-xy}) = (2y - x^3 - xy^2)e^{-xy}$$

3. $z = xye^{x/y}$

$$\frac{\partial z}{\partial x} = ye^{x/y} + xy(\frac{1}{y}e^{x/y}) = (y + x)e^{x/y}$$
$$\frac{\partial z}{\partial y} = xe^{x/y} + xy(\frac{-x}{y^2}e^{x/y}) = (x - \frac{x^2}{y})e^{x/y} = \frac{x(y - x)e^{x/y}}{y}$$

**Ejemplo 3.6.** *Dada la función:* $f(x, y) = 4x^3y^2 - 4x^2 + y^6 + 1$. *Hallar las dos derivadas parciales en el punto* $(1, 1)$

*Solución.* Calculamos las derivadas parciales mediante las reglas de derivación y luego sustituimos las coordenadas del punto.

$$\frac{\partial f}{\partial x}(x, y) = 12x^2y^2 - 8x \quad \rightarrow \quad \frac{\partial f}{\partial x}(1, 1) = 12 - 8 = 4$$
$$\frac{\partial f}{\partial y}(x, y) = 8x^3y + 6y^5 \quad \rightarrow \quad \frac{\partial f}{\partial y}(1, 1) = 8 + 6 = 14$$

**Notación:** Se utilizan las siguientes notaciones:

$$z_x = z'_x = \frac{\partial z}{\partial x} = \frac{\partial f}{\partial x} = \frac{\partial f}{\partial x}(x, y) = f_x = f'_x = f_x(x, y) = D_x f(x, y) = D_1 f(x, y)$$

$$z_y = z'_y = \frac{\partial z}{\partial y} = \frac{\partial f}{\partial y} = \frac{\partial f}{\partial y}(x, y) = f_y = f'_y = f_y(x, y) = D_y f(x, y) = D_2 f(x, y)$$

Cuando se trate de un punto genérico $(x, y)$, normalmente no indicaremos las coordenadas, simplemente pondremos $f_x$. Sin embargo, cuando se trate de un punto concreto $(x_0, y_0)$, siempre las indicaremos, para saber de qué punto se trata. En este caso, en algunas ocasiones utilizaremos la siguiente notación:

$$\frac{\partial f}{\partial x}(x_0, y_0) = \left.\frac{\partial f}{\partial x}\right]_{(x_0,y_0)}$$

### 3.1.3. Funciones de más de dos variables

El concepto de derivada parcial se puede generalizar a funciones de cualquier número de variables. Bastará con derivar la función suponiendo que depende, en cada caso, de una sola variable, con las demás fijas.

**Ejemplo 3.7.** *Hallar las derivadas parciales de la función:*

$$f(x, y, z) = e^{-\pi z} \cos 4x \ \operatorname{sen} 6y$$

*Solución.* En cada caso, suponemos constante dos de las variables, con lo cual se obtiene una función de una sola variable, respecto de la cual se deriva.

$$\frac{\partial f}{\partial x} = -4e^{-\pi z} \operatorname{sen} 4x \ \operatorname{sen} 6y$$
$$\frac{\partial f}{\partial y} = 6e^{-\pi z} \cos 4x \ \cos 6y$$
$$\frac{\partial f}{\partial z} = -\pi e^{-\pi z} \cos 4x \ \operatorname{sen} 6y$$

**Notación vectorial para definir las derivadas parciales.** El concepto de derivada parcial se extiende a funciones de tres variables $f(x, y, z)$ de la siguiente forma:

$$\frac{\partial f}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y, z) - f(x, y, z)}{\Delta} = \lim_{t \to 0} \frac{f(x, y + t, z) - f(x, y, z)}{t}$$

Análogamente se calculan las demás derivadas parciales.
Para los incremento, $\Delta x$, $\Delta y$, etc., en vez de usar las letras $h$, $k$, etc., se pueden usar $h_1$, $h_2$, etc., o bien, podemos usar siempre la letra $t$.
Del mismo modo se extiende a funciones de cuatro variables $f(x, y, z, u)$

$$\frac{\partial f}{\partial y} = \lim_{t \to 0} \frac{f(x, y + t, z, u) - f(x, y, z, u)}{t}$$

Desde el punto de vista teórico, para definir las derivadas parciales en el caso general de funciones de $n$ variables $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$, conviene usar notación vectorial. Para ello consideramos los vectores $\mathbf{e}_1, \ \mathbf{e}_2, \ \cdots, \mathbf{e}_n \in \mathbb{R}^n$ de la base canónica de $\mathbb{R}^n$, cuyas componentes son:

$$\mathbf{e}_1 = (1, 0, \cdots, 0), \ \mathbf{e}_2 = (0, 1, \cdots, 0), \cdots, \mathbf{e}_n = (0, 0, \cdots, 1)$$

con lo cual, para cualquier número $t \in \mathbb{R}$, resulta:

$$t\mathbf{e}_1 = (t, 0, \cdots, 0), \ t\mathbf{e}_2 = (0, t, \cdots, 0), \cdots, t\mathbf{e}_n = (0, 0, \cdots, t)$$

de donde, para cualquier punto $\mathbf{p} \in \mathbb{R}^n$, de coordenadas $\mathbf{p} = (x_1, x_2, \cdots, x_n)$, se tiene:

$$\mathbf{p} + t\mathbf{e}_i = (x_1, x_2, \cdots, x_n) + (0, \cdots, t, \cdots, 0) = (x_1, x_2, \cdots, x_i + t, \cdots, x_n)$$

Es decir, $\mathbf{p} + t\mathbf{e}_i$ tiene las mismas coordenadas que $\mathbf{p}$, salvo la i-ésima coordenada que se ha incrementado en $t$. En consecuencia, podemos enunciar la siguiente definición:

**Definición 3.2 (Derivadas parciales).** *Sea $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto*[1] *$\mathcal{D}$ de $\mathbb{R}^n$, y sea $\mathbf{p} \in \mathcal{D}$. Se define la derivada parcial de $f$ con respecto a su i-ésima variable en el punto $\mathbf{p}$, como el siguiente límite, si existe y es finito:*

$$\frac{\partial f}{\partial x_i}(\mathbf{p}) = \lim_{t \to 0} \frac{f(\mathbf{p} + t\mathbf{e}_i) - f(\mathbf{p})}{t}$$

[1] La razón de definir las derivadas parciales en un punto perteneciente a un conjunto *abierto* $\mathcal{D}$ (dominio de la función), es para poder asegurar que para $t \in \mathbb{R}$ pequeño, se tenga $\mathbf{p} + t\mathbf{e}_i \in \mathcal{D}$ y que así tenga sentido la expresión $f(\mathbf{p} + t\mathbf{e}_i)$ que aparece en la definición de derivada parcial.

Del mismo modo puede definirse la derivada parcial en un punto genérico $\mathbf{x} \in \mathcal{D}$, como

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{e}_i) - f(\mathbf{x})}{t}$$

### 3.1.4. Razón de cambio

Las derivadas parciales también pueden interpretarse como la razón de cambio instantáneo de la función respecto de una variable, mientras la otra permanece fija. Así,

$\dfrac{\partial f}{\partial x}(x_0, y_0)$ Se puede interpretar como la *razón de cambio instantáneo* de la función $f$ cuando se conserva *fija* la variable $y$ y *varía* la $x$.

$\dfrac{\partial f}{\partial y}(x_0, y_0)$ Se puede interpretar como la *razón de cambio instantáneo* de la función $f$ cuando se conserva *fija* la variable $x$ y *varía* la $y$.

Es decir, las derivadas parciales miden la velocidad de variación *parcial* de la función con respecto a cada variable, cuando las demás se mantienen fijas.

**Ejemplo 3.8.** *Un cilindro recto tiene* 4 *cm. de radio y* 20 *cm. de altura. Hallar la razón de cambio del volumen del cilindro respecto del radio y respecto de la altura.*

*Solución.* Tenemos que $V = \pi r^2 h$, luego

$$\frac{\partial V}{\partial r} = 2\pi r h \quad \rightarrow \quad \frac{\partial V}{\partial r}(r = 4, h = 20) = 2\pi \cdot 4 \cdot 20 = 160\pi \ cm^3/cm \ \ de \ \ r$$

$$\frac{\partial V}{\partial h} = \pi r^2 \quad \rightarrow \quad \frac{\partial V}{\partial h}(r = 4, h = 20) = 16\pi \ cm^3/cm \ \ de \ \ h$$

En el primer caso, si mantenemos fija la altura e incrementamos el radio, se produce un incremento del volumen de $160\pi \ cm^3/cm \ \ de \ \ r$. Mientras que en el segundo caso, si mantenemos fijo el radio e incrementamos la altura, se produce un incremento del volumen de $16\pi \ cm^3/cm \ \ de \ \ h$

### 3.1.5. Interpretación geométrica de las derivadas parciales

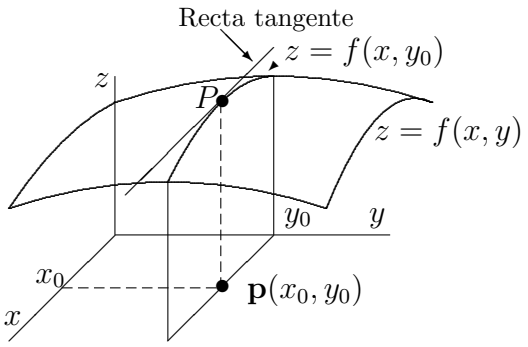



Figura 3.1: Curva $z = f(x, y_0)$.

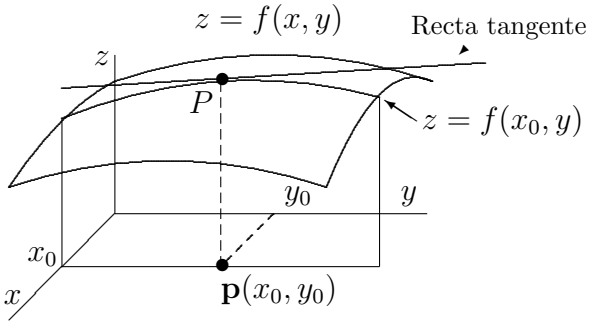



Figura 3.2: Curva $z = f(x_0, y)$.

Desde el punto de vista geométrico, la función $g(x) = f(x, y_0)$ representa la curva que se obtiene de la intersección de la superficie $z = f(x, y)$ con el plano $y = y_0$

$$\left.\begin{array}{l} z = f(x, y) \\ y = y_0 \end{array}\right\} z = f(x, y_0) = g(x)$$

La derivada parcial de la función $f$, respecto de la variable $x$, en el punto $\mathbf{p}$ representa la pendiente de la tangente a la curva $g(x) = f(x, y_0)$ en el punto $P$ correspondiente de la gráfica, es decir, la inclinación de la superficie en la dirección del eje $x$.
Análogamente, la función $g(y) = f(x_0, y)$ representa la curva que se obtiene de la intersección de la superficie $z = f(x, y)$ con el plano $x = x_0$

$$\left.\begin{array}{l} z = f(x, y) \\ x = x_0 \end{array}\right\} z = f(x_0, y) = g(y)$$

La derivada parcial de la función $f$, respecto de la variable $y$, en el punto $\mathbf{p}$ representa la pendiente de la tangente a la curva $g(y) = f(x_0, y)$ en el punto $P$ correspondiente de la gráfica, es decir, la inclinación de la superficie en la dirección del eje $y$.

**Ejemplo 3.9.** *Hallar la pendiente a la superficie* $f(x, y) = -\dfrac{x^2}{2} - y^2 + \dfrac{25}{8}$ *en el punto* $P(1/2, 1, 2)$*, en las direcciones de los ejes* $x$ *e* $y$*.*

*Solución.*

$$\begin{array}{lllll} \dfrac{\partial f}{\partial x} = -x & \rightarrow & \dfrac{\partial f}{\partial x}(1/2, 1) = -1/2 & \rightarrow & \tan\alpha = -1/2 \\ \dfrac{\partial f}{\partial y} = -2y & \rightarrow & \dfrac{\partial f}{\partial y}(1/2, 1) = -2 & \rightarrow & \tan\beta = -2 \end{array}$$

**Continuidad y derivadas parciales**

La existencia de las derivadas parciales no garantiza la continuidad de una función.

En efecto, la existencia de $f_x$ depende del comportamiento de la función sólo en la dirección del eje $x$, y la existencia de $f_y$ del comportamiento de la función sólo en la dirección del eje $y$, mientras que la continuidad depende del comportamiento de la función en todos los puntos del entorno. Esto significa que una función puede tener derivadas parciales en un punto aunque no sea continua en dicho punto.
Tenemos pues, que de la continuidad de las funciones de $n$ variables en un punto dado, no se deriva la existencia de sus derivadas parciales en ese punto. Pero es más, cuando $n \geq 2$ incluso de la existencia de todas las derivadas parciales en cierto punto, no se deduce la continuidad en ese punto. (Recordemos que para $n = 1$, es decir, para las funciones de una variable, de la existencia de la derivada en un punto se deriva también que la función es continua en ese punto).

**Ejemplo 3.10.** *Estudiar la continuidad y calcular las dos derivadas parciales en el punto* $\mathbf{p}(0,0)$ *de la función* $f : \mathbb{R}^2 \to \mathbb{R}$ *definida por:*

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & si \ \ (x,y) \neq (0,0) \\ 0 & si \ \ (x,y) = (0,0) \end{cases}$$

*Solución.*

1. Continuidad: La función no es continua en el punto $\mathbf{p}(0,0)$, ya que no existe el límite en dicho punto. En efecto, si nos acercamos al punto mediante las rectas $y = mx$ resulta:
$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2} = \lim_{\substack{(x,y)\to(0,0)\\ y=mx}} \frac{xy}{x^2+y^2} = \lim_{\substack{(x,y)\to(0,0)\\ y=mx\\ x\to 0}} \frac{xmx}{x^2+m^2x^2} = \lim_{x\to 0} \frac{m}{1+m^2} = \frac{m}{1+m^2}$$
luego el límite no existe ya que depende del valor de m. Es decir, según la recta por la que nos aproximemos al punto tendríamos un valor del límite u otro.

2. Existencia de las derivadas parciales. A pesar de que la función no es continua en el punto $\mathbf{p}(0,0)$, las derivadas parciales en dicho punto existen. En efecto:
$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0} \frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0} \frac{\dfrac{h\cdot 0}{h^2+0^2}-0}{h} = 0$$
$$\frac{\partial f}{\partial y}(0,0) = \lim_{k\to 0} \frac{f(0,0+k)-f(0,0)}{k} = \lim_{k\to 0} \frac{f(0,k)-f(0,0)}{k} = \lim_{k\to 0} \frac{\dfrac{0\cdot k^2}{0^2+k^2}-0}{k} = 0$$

**Ejemplo 3.11.** *Estudiar la continuidad y calcular las dos derivadas parciales en el punto* $\mathbf{p}(0,0)$ *de la función* $f : \mathbb{R}^2 \to \mathbb{R}$ *definida por:*

$$f(x,y) = \begin{cases} 1 & si \ xy \neq 0 \\ 0 & si \ xy = 0 \end{cases}$$

*Solución.*

1. Continuidad: La función no es continua en el punto $\mathbf{p}(0,0)$, ya que no existe el límite en dicho punto. En efecto, si nos acercamos al punto mediante la recta $y = x$ resulta:

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{\substack{(x,y)\to(0,0)\\ y=x}} \Big[1\Big] = 1 \neq f(0,0) = 0$$

2. Existencia de las derivadas parciales. A pesar de que la función no es continua en el punto $\mathbf{p}(0,0)$, las derivadas parciales en dicho punto existen. En efecto:

$$\frac{\partial f}{\partial x}(0,0) = \frac{\partial}{\partial x}\Big[f(x,0)\Big] = \frac{\partial}{\partial x}\Big[0\Big] = 0$$
$$\frac{\partial f}{\partial y}(0,0) = \frac{\partial}{\partial y}\Big[f(0,y)\Big] = \frac{\partial}{\partial y}\Big[0\Big] = 0$$

## 3.2. Derivadas parciales de órdenes superiores

### Derivadas parciales de segundo orden

Las derivadas parciales de una función de dos variables $f(x,y)$, son, a su vez, funciones de dos variables, $f_x(x,y)$, $f_y(x,y)$ y, por lo tanto, podemos obtener de ellas, nuevamente, sus derivadas parciales. Llamadas derivadas parciales de segundo orden.

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$

**Notación:** Para simplificar los paréntesis usaremos la siguiente notación:

$$(f_x)_x = f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = D_1\left(D_1 f\right) = D_{11} f$$

$$(f_x)_y = f_{xy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} = D_2\left(D_1 f\right) = D_{21} f$$

$$(f_y)_x = f_{yx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y} = D_1\left(D_2 f\right) = D_{12} f$$

$$(f_y)_y = f_{yy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y \partial y} = D_2\left(D_2 f\right) = D_{22} f$$

Para evitar confusión con el orden de derivación ($f_{xy} = D_{21} f$), utilizaremos el siguiente criterio: se empieza derivando por la variable *"más cercana"* a la función.

**Ejemplo 3.12.** *Hallar las derivadas parciales de segundo orden de la función*

$$f(x,y) = \operatorname{sen}(x^2 y)$$

*Solución.*

$f_x = 2xy\cos(x^2 y)$

$f_{xx} = \big(f_x\big)_x = 2y\cos(x^2 y) + 2xy\big(-2xy\operatorname{sen}(x^2 y)\big) = 2y\cos(x^2 y) - 4x^2 y^2 \operatorname{sen}(x^2 y)$

$f_{xy} = \big(f_x\big)_y = 2x\cos(x^2 y) + 2xy\big(-x^2\operatorname{sen}(x^2 y)\big) = 2x\cos(x^2 y) - 2x^3 y\operatorname{sen}(x^2 y)$

$f_y = x^2\cos(x^2y)$

$f_{yx} = \big(f_y\big)_x = 2x\cos(x^2y) + x^2\big(-2xy\,\mathrm{sen}(x^2y)\big) = 2x\cos(x^2y) - 2x^3y\,\mathrm{sen}(x^2y)$

$f_{yy} = \big(f_y\big)_y = x^2\big(-x^2\,\mathrm{sen}(x^2y)\big) = -x^4\,\mathrm{sen}(x^2y)$

**Derivadas parciales cruzadas.** Las derivadas parciales $f_{xy}$ y $f_{yx}$, se llaman derivadas parciales *cruzadas* y cuando son continuas coinciden.

**Teorema 3.1 (Teorema de Schwartz).** *Sea $f : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función definida en el abierto $\mathcal{D}$ de $\mathbb{R}^2$. Si las derivadas parciales $f_{xy} : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ y $f_{yx} : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ existen y son funciones continuas en $D$, entonces*

$$f_{xy} = f_{yx}$$

El teorema de Schwartz también puede enunciarse en los siguientes términos: *Si $f_x$, $f_y$, y $f_{xy}$ son continuas en un entorno del punto $(x_0, y_0)$, entonces existe $f_{yx}(x_0, y_0)$ y se verifica $f_{yx}(x_0, y_0) = f_{xy}(x_0, y_0)$.*

**Ejemplo 3.13.** *Hallar las derivadas parciales de segundo orden de la función*

$$f(x, y) = x^2ye^{x^2+y^2}$$

*Solución.*

$f_x = 2xye^{x^2+y^2} + x^2y\big(2xe^{x^2+y^2}\big) = (2x^3y + 2xy)e^{x^2+y^2}$

$f_{xx} = (6x^2y + 2y)e^{x^2+y^2} + (2x^3y + 2xy)\big(2xe^{x^2+y^2}\big) = (4x^4y + 10x^2y + 2y)e^{x^2+y^2}$

$f_{xy} = (2x^3 + 2x)e^{x^2+y^2} + (2x^3y + 2xy)\big(2ye^{x^2+y^2}\big) = (4x^3y^2 + 2x^3 + 4xy^2 + 2x)e^{x^2+y^2}$

$f_y = x^2e^{x^2+y^2} + x^2y\big(2ye^{x^2+y^2}\big) = (2x^2y^2 + x^2)e^{x^2+y^2}$

$f_{yx} = (4xy^2 + 2x)e^{x^2+y^2} + (2x^2y^2 + x^2)\big(2xe^{x^2+y^2}\big) = (4x^3y^2 + 2x^3 + 4xy^2 + 2x)e^{x^2+y^2}$

$f_{yy} = 4x^2ye^{x^2+y^2} + (2x^2y^2 + x^2)\big(2ye^{x^2+y^2}\big) = (4x^2y^3 + 6x^2y)e^{x^2+y^2}$

**Ejemplo 3.14.** *Comprobar que las derivadas parciales cruzadas de la siguiente función, en el punto $(0, 0)$ no coinciden.*

$$f(x, y) = \begin{cases} \dfrac{x^3y - xy^3}{x^2 + y^2} & si \quad (x, y) \neq (0, 0) \\ 0 & si \quad (x, y) = (0, 0) \end{cases}$$

*Solución.*

Si $(x, y) \neq (0, 0)$, las derivadas parciales son:

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\left(\frac{x^3y - xy^3}{x^2 + y^2}\right) = \frac{(3x^2y - y^3)(x^2 + y^2) - (x^3y - xy^3)2x}{(x^2 + y^2)^2} =$$

$$= \frac{3x^4y + 3x^2y^3 - x^2y^3 - y^5 - 2x^4y + 2x^2y^3}{(x^2 + y^2)^2} = \frac{x^4y + 4x^2y^3 - y^5}{(x^2 + y^2)^2}$$

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}\left(\frac{x^3y - xy^3}{x^2+y^2}\right) = \frac{(x^3-3xy^2)(x^2+y^2)-(x^3y-xy^3)2y}{(x^2+y^2)^2} =$$

$$= \frac{x^5+x^3y^2-3x^3y^2-3xy^4-2x^3y^2+2xy^4}{(x^2+y^2)^2} = \frac{x^5-4x^3y^2-xy^4}{(x^2+y^2)^2}$$

En el punto $(0,0)$, las derivadas parciales son:

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{\frac{h^3\cdot 0 - h\cdot 0^3}{h^2+0^2}-0}{h} = 0$$

$$\frac{\partial f}{\partial x}(0,0) = \lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k} = \lim_{k\to 0}\frac{\frac{0^3\cdot k - 0\cdot k^3}{0^2+k^2}-0}{k} = 0$$

De donde, aplicando directamente la definición de derivadas parciales, resulta:

$$\frac{\partial^2 f}{\partial y\partial x}(0,0) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(0,0) = \lim_{k\to 0}\frac{\frac{\partial f}{\partial x}(0,k)-\frac{\partial f}{\partial x}(0,0)}{k} = \lim_{k\to 0}\frac{\frac{0^4\cdot k+4\cdot 0^2k^3-k^5}{(0^2+k^2)^2}-0}{k} = -1$$

$$\frac{\partial^2 f}{\partial x\partial y}(0,0) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(0,0) = \lim_{h\to 0}\frac{\frac{\partial f}{\partial y}(h,0)-\frac{\partial f}{\partial y}(0,0)}{h} = \lim_{h\to 0}\frac{\frac{h^5-4h^3\cdot O-h\cdot 0^4}{(h^2+0^2)^2}-0}{h} = +1$$

**Derivadas parciales de tercer orden**

Si hacemos las derivadas parciales de tercer orden, resultan $2^3 = 8$ derivadas, pero si son continuas se reducen a $3+1=4$ derivadas distintas:

$$f\begin{cases} f_x\begin{cases} f_{xx}\begin{cases} f_{xxx} \\ f_{xxy}\end{cases} \\ f_{xy}\begin{cases} f_{xyx} \\ f_{xyy}\end{cases}\end{cases} \\ f_y\begin{cases} f_{yx}\begin{cases} f_{yxx} \\ f_{yxy}\end{cases} \\ f_{yy}\begin{cases} f_{yyx} \\ f_{yyy}\end{cases}\end{cases}\end{cases}$$

Si son continuas $\rightarrow$

$$\begin{tikzcd}[row sep=small] & & & f_{xxx} \\ & & f_{xx} \arrow[ur] \arrow[dr] & \\ & f_x \arrow[ur] \arrow[dr] & & f_{xxy} \\ f \arrow[ur] \arrow[dr] & & f_{xy} \arrow[ur] \arrow[dr] & \\ & f_y \arrow[ur] \arrow[dr] & & f_{xyy} \\ & & f_{yy} \arrow[ur] \arrow[dr] & \\ & & & f_{yyy} \end{tikzcd}$$

Es decir, si las derivadas parciales son continuas no importa el orden de derivación, sino el número de veces que se ha derivado respecto de cada variable.

**Ejemplo 3.15.** *Hallar las derivadas parciales de tercer orden de la función:*

$$f(x,y) = x^2 + 2xy^2 - y^3$$

*Solución.*

$$
\begin{tikzcd}[row sep=small]
 & & & f_{xxx} = 0 \\
 & & f_{xx} = 2 \arrow[ur] \arrow[dr] & \\
 & f_x = 2x + 2y^2 \arrow[ur] \arrow[dr] & & f_{xxy} = 0 \\
f(x,y) = x^2 + 2xy^2 - y^3 \arrow[ur] \arrow[dr] & & f_{xy} = 4y \arrow[ur] \arrow[dr] & \\
 & f_y = 4xy - 3y^2 \arrow[ur] \arrow[dr] & & f_{xyy} = 4 \\
 & & f_{yy} = 4x - 6y \arrow[ur] \arrow[dr] & \\
 & & & f_{yyy} = -6
\end{tikzcd}
$$

**Derivadas parciales de orden $n$**

Si hacemos las derivadas parciales de orden $n$, resultan $2^n$ derivadas, pero si son continuas se reducen a $n+1$ derivadas distintas. Es decir, si las derivadas parciales son continuas no importa el orden de derivación, sino el número de veces que se ha derivado respecto de cada variable. Ahora bien, aunque el resultado final de las derivadas parciales no depende del orden de derivación, el proceso de derivación puede resultar mucho más complicado en un orden que en otro.

**Ejemplo 3.16.** *Dada la función* $f(x,y) = \dfrac{xy}{y^2+z^2}$. *Hallar* $D_{2311}f$ *y* $D_{1132}f$

*Solución.*

$$D_{2311}f = D_{231}\big(D_1 f\big) = D_{231}\left(\frac{y}{y^2+z^2}\right) = 0$$

$$D_{1132}f = D_{113}\big(D_2 f\big) = D_{113}\left(\frac{x(y^2+z^2) - xy2y}{(y^2+z^2)^2}\right) = D_{113}\left(\frac{xz^2 - xy^2}{(y^2+z^2)^2}\right) =$$

$$= D_{11}\left(\frac{2xz(y^2+z^2)^2 - (xz^2 - xy^2)2(y^2+z^2)2z}{(y^2+z^2)^4}\right) = D_{11}\left(\frac{2xz(3y^2 - z^2)}{(y^2+z^2)^3}\right) =$$

$$= D_1\left(\frac{2z(3y^2 - z^2)}{(y^2+z^2)^3}\right) = 0$$

## 3.3. Derivadas direccionales.

### 3.3.1. Derivadas direccionales

Las derivadas parciales $f_x(x,y)$ y $f_y(x,y)$, representan, respectivamente, la pendiente de la superficie $z = f(x,y)$ en las direcciones del eje $OX$ y del eje $OY$. Para hallar la pendiente en cualquier otra dirección se utilizan las derivadas direccionales. Es decir, las derivadas parciales nos dan una medida de la variación de una función solamente en la dirección de cada eje coordenado. Es natural buscar un concepto más general de derivada a fin de que nuestras consideraciones no queden restringidas a las direcciones particulares de los ejes coordenados y nos permita estudiar la razón de incrementos en una dirección cualquiera. La derivada direccional responde a este propósito.

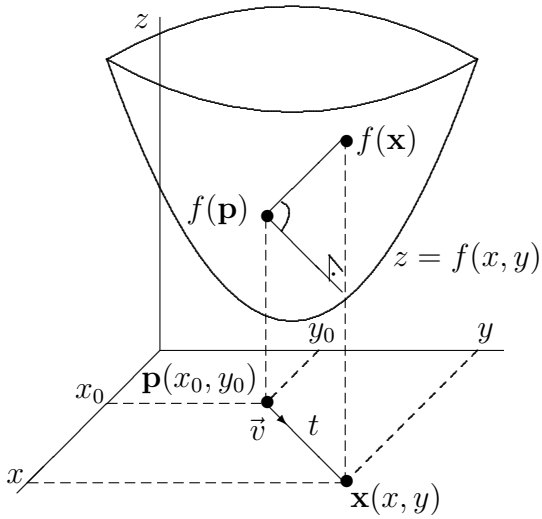



Figura 3.3: Derivada direccional.

Queremos estudiar la variación de la función $f$ en el punto $\mathbf{p}$ cuando el argumento varía en la dirección marcada por el vector $\vec{v}$. Para ello partimos de la idea del concepto de derivada de funciones de una variable *"el límite, cuando el incremento de la variable tiende a cero, del cociente del incremento de la función dividido entre el incremento de la variable"*, es decir:

$$D_{\vec{v}} f(\mathbf{p}) = \lim_{\mathbf{x}\to\mathbf{p}} \frac{f(\mathbf{x}) - f(\mathbf{p})}{t}$$

donde $\mathbf{x}$ es un punto próximo a $\mathbf{p}$ y además situado en la dirección marcada por el vector $\vec{v}$, y $t$ es la *longitud orientada* del segmento $\overline{\mathbf{px}}$, es decir, la longitud de este segmento con signo positivo, si el vector $\overrightarrow{\mathbf{px}}$ tiene la misma dirección que el vector $\vec{v}$, y con signo negativo en caso contrario.

Para la derivada direccional usaremos cualquiera de las notaciones:

$$D_{\vec{v}} f(\mathbf{p}) = \frac{\partial f}{\partial \vec{v}}(\mathbf{p})$$

**Cálculo de la derivada direccional conocido el vector director:**

1. **Vector director unitario $\vec{u}$ :** Si el vector director es unitario resulta:

$$\left.\begin{array}{r} \|\overrightarrow{\mathbf{px}}\| = |t| \\ \|\vec{u}\| = 1 \end{array}\right\} \Rightarrow \overrightarrow{\mathbf{px}} = t\,\vec{u} \Rightarrow \mathbf{x} - \mathbf{p} = t\,\vec{u} \Rightarrow \mathbf{x} = \mathbf{p} + t\,\vec{u}$$

de donde,

$$\boxed{D_{\vec{v}} f(\mathbf{p}) = \lim_{t\to 0} \frac{f(\mathbf{p} + t\,\vec{u}) - f(\mathbf{p})}{t}} \tag{3.1}$$

2. **Vector director no unitario $\vec{v}$:** Si el vector no es unitario, resulta:

$$\left.\begin{array}{r} \|\overrightarrow{\mathbf{px}}\| = |t| \\ \|\vec{v}\| \neq 1 \end{array}\right\} \Rightarrow \overrightarrow{\mathbf{px}} \neq t\,\vec{v}$$

luego no podemos hacer la sustitución anterior. Por lo tanto, si el vector no es unitario, hallamos el vector unitario de la misma dirección y sentido que el dado, y a ese nuevo vector hallado le aplicamos el resultado anterior.

$$\vec{u} = \frac{\vec{v}}{\|\vec{v}\|}$$

El concepto de derivada direccional generaliza el concepto de derivada parcial, de manera que las derivadas parciales pueden obtenerse como casos particulares de las derivadas direccionales. Así, $f_x$ es la derivada direccional en la dirección del vector $(1,0)$ y $f_y$ en la dirección del vector $(0,1)$, es decir:

$$f_x(\mathbf{p}) = D_{\vec{u}}f(\mathbf{p}) \ \text{ para } \ \vec{u} = (1,0) \qquad f_y(\mathbf{p}) = D_{\vec{u}}f(\mathbf{p}) \ \text{ para } \ \vec{u} = (0,1)$$

Se debe observar que puede existir la derivada direccional de una función, en un punto, con respecto a un vector, y sin embargo, puede suceder que no exista la derivada direccional con respecto a otro vector.

**Ejemplo 3.17.** *Hallar la derivada direccional de* $f(x,y) = x^2 + 3xy^2$ *en el punto* $\mathbf{p}(1,2)$*, en la dirección que apunta hacia el origen.*
*Solución.* Hallamos el vector director y comprobamos su módulo.

$$\vec{v} = \overrightarrow{\mathbf{po}} = \mathbf{o} - \mathbf{p} = (0,0) - (1,2) = (-1,-2) \ \rightarrow \ |\vec{v}\,| = \sqrt{1+4} = \sqrt{5} \neq 1$$

luego,

$$\vec{u} = \frac{\vec{v}}{|\vec{v}\,|} = \left(\frac{-1}{\sqrt{5}}, \frac{-2}{\sqrt{5}}\right) \ \rightarrow \ \mathbf{p} + t\,\vec{u} = (1,2) + t\left(\frac{-1}{\sqrt{5}}, \frac{-2}{\sqrt{5}}\right) = \left(1 - \frac{t}{\sqrt{5}}, 2 - \frac{2t}{\sqrt{5}}\right)$$

de donde,

$$f(\mathbf{p}) = f(1,2) = 1^2 + 3 \cdot 1 \cdot 2^2 = 1 + 12 = 13$$

$$\begin{aligned} f(\mathbf{p} + t\,\vec{u}\,) &= f\left(1 - \frac{t}{\sqrt{5}}, 2 - \frac{2t}{\sqrt{5}}\right) = \left(1 - \frac{t}{\sqrt{5}}\right)^2 + 3\left(1 - \frac{t}{\sqrt{5}}\right)\left(2 - \frac{2t}{\sqrt{5}}\right)^2 = \\ &= 1 - \frac{2t}{\sqrt{5}} + \frac{t^2}{5} + \left(3 - \frac{3t}{\sqrt{5}}\right)\left(4 - \frac{8t}{\sqrt{5}} + \frac{4t^2}{5}\right) = \\ &= 1 - \frac{2t}{\sqrt{5}} + \frac{t^2}{5} + 12 - \frac{24t}{\sqrt{5}} + \frac{12t^2}{5} - \frac{12t}{\sqrt{5}} + \frac{24t^2}{5} - \frac{12t^3}{5\sqrt{5}} = \\ &= 13 - \frac{38t}{\sqrt{5}} + \frac{37t^2}{5} - \frac{12t^3}{5\sqrt{5}} \end{aligned}$$

con lo que resulta,

$$\begin{aligned} D_{\vec{v}}f(1,2) &= \lim_{t\to 0} \frac{f(\mathbf{p} + t\,\vec{u}\,) - f(\mathbf{p})}{t} = \lim_{t\to 0} \frac{13 - \dfrac{38t}{\sqrt{5}} + \dfrac{37t^2}{5} - \dfrac{12t^3}{5\sqrt{5}} - 13}{t} = \\ &= \lim_{t\to 0}\left(\frac{-38}{\sqrt{5}} + \frac{37t}{5} - \frac{12t^2}{5\sqrt{5}}\right) = \frac{-38}{\sqrt{5}} \end{aligned}$$

El que la derivada direccional sea negativa significa que la función decrece en esa dirección.

**Cálculo de la derivada direccional conocido el ángulo director.**
Conocido el ángulo $\theta$ que forma el vector director con la dirección positiva del eje $OX$, se tienen inmediatamente las coordenadas del vector director unitario. En efecto,

$$\vec{u} = (\cos\theta, \operatorname{sen}\theta).$$

El concepto de derivada direccional se puede generalizar para cualquier número de variables. Así, para el caso general podemos enunciar la siguiente definición:

**Definición 3.3 (Derivada direccional).** *Sea $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto $\mathcal{D}$ de $\mathbb{R}^n$ y sea $\mathbf{p} \in \mathcal{D}$ un punto dado de $\mathcal{D}$. Sea $\vec{u}$ un vector unitario dado. Se define la derivada de la función $f$ en $\mathbf{p}$, en la dirección del vector $\vec{u}$, denotada $\dfrac{\partial f}{\partial \vec{u}}(\mathbf{p})$, o bien $D_{\vec{u}} f(\mathbf{p})$, como el siguiente límite, si existe y es finito.*

$$D_{\vec{u}} f(\mathbf{p}) = \lim_{t\to 0} \frac{f(\mathbf{p} + t\,\vec{u}) - f(\mathbf{p})}{t}$$

**Ejemplo 3.18.** *Hallar la derivada direccional de $f(x, y, z) = xyz$ en el punto $\mathbf{p}(1, 0, -1)$, según la dirección del vector $\vec{v} = (1, 1, 1)$.*

*Solución.* Hallamos el vector unitario $\vec{u}$ con la misma dirección y sentido que $\vec{v}$.

$$|\vec{v}\,| = \sqrt{1+1+1} = \sqrt{3} \neq 1 \quad \to \quad \vec{u} = \frac{\vec{v}}{|\vec{v}\,|} = \Big(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\Big)$$

luego,

$$\mathbf{p} + t\,\vec{u} = (1, 0, -1) + t\Big(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\Big) = \Big(1 + \frac{t}{\sqrt{3}}, \frac{t}{\sqrt{3}}, -1 + \frac{t}{\sqrt{3}}\Big)$$

de donde,

$$f(\mathbf{p}) = f(1, 1, 1) = 1 \cdot 0 \cdot (-1) = 0$$

$$\begin{aligned} f(\mathbf{p} + t\,\vec{u}) &= f\Big(1 + \frac{t}{\sqrt{3}}, \frac{t}{\sqrt{3}}, -1 + \frac{t}{\sqrt{3}}\Big) = \left(1 + \frac{t}{\sqrt{3}}\right)\left(\frac{t}{\sqrt{3}}\right)\left(-1 + \frac{t}{\sqrt{3}}\right) = \\ &= \left(\frac{t^2}{3} - 1\right)\left(\frac{t}{\sqrt{3}}\right) = \frac{t^3}{3\sqrt{3}} - \frac{t}{\sqrt{3}} \end{aligned}$$

con lo que resulta,

$$D_{\vec{v}} f(1, 1, 1) = \lim_{t\to 0} \frac{f(\mathbf{p} + t\,\vec{u}) - f(\mathbf{p})}{t} = \lim_{t\to 0} \frac{\dfrac{t^3}{3\sqrt{3}} - \dfrac{t}{\sqrt{3}} - 0}{t} = \lim_{t\to 0} \left(\frac{t^2}{3\sqrt{3}} - \frac{1}{\sqrt{3}}\right) = \frac{-1}{\sqrt{3}}$$

### 3.3.2. Relación entre la derivada direccional y las derivadas parciales

Veamos a partir de un ejemplo un resultado que justificaremos más adelante.

**Ejemplo 3.19.** *Hallar la derivada direccional de* $f(x,y) = x^2 + y^3$ *en un punto genérico* $\mathbf{p}(x,y)$*, según la dirección de un vector genérico unitario* $\vec{u} = (u_1, u_2)$*.*

*Solución.* Tenemos,

$$\mathbf{p} + t\,\vec{u} = (x,y) + t(u_1,u_2) = (x + tu_1, x + tu_2)$$

de donde,

$$\begin{aligned} f(\mathbf{p}) &= f(x,y) = x^2 + y^3 \\ f(\mathbf{p} + t\,\vec{u}) &= f(x + tu_1, x + tu_2) = (x + tu_1)^2 + (x + tu_2)^3 = \\ &= x^2 + 2xtu_1 + t^2(u_1)^2 + x^3 + 3x^2tu_2 + 3xt^2(u_2)^2 + t^3(u_2)^3 \end{aligned}$$

con lo que resulta,

$$\begin{aligned} D_{\vec{u}} f(x,y) &= \lim_{t\to 0} \frac{f(\mathbf{p} + t\,\vec{u}) - f(\mathbf{p})}{t} = \\ &= \lim_{t\to 0} \frac{x^2 + 2xtu_1 + t^2(u_1)^2 + y^3 + 3y^2tu_2 + 3yt^2(u_2)^2 + t^3(u_2)^3 - (x^2 + y^3)}{t} = \\ &= \lim_{t\to 0} \left(2xu_1 + t(u_1)^2 + 3y^2u_2 + 3yt(u_2)^2 + t^2(u_2)^3\right) = 2xu_1 + 3y^2u_2 \end{aligned}$$

El cálculo de la derivada direccional aplicando la definición resulta bastante engorroso, no obstante, el resultado de este ejemplo nos puede hacer intuir una propiedad que demostraremos más adelante (ver teorema 3.4 en la página 34) $D_{\vec{u}}\, f = f_x \cdot u_1 + f_y \cdot u_2$. Es decir, *la derivada direccional se puede obtener como la suma de los productos de las derivadas parciales por las componentes del vector unitario de dirección.* Esto nos permite obtener las derivadas direccionales de una manera mucho más fácil que aplicando directamente la definición. Sin embargo, esta fórmula no es válida para todo tipo de funciones, sino solamente para las diferenciables, de ahí que, en ocasiones, tengamos que acudir al incómodo límite de la definición.

**Ejemplo 3.20.** *Comprobar que las derivadas direccionales calculadas en los ejemplo 3.17 y 3.18 cumplen la relación anterior.*

*Solución.*

1. En el ejemplo 3.17 tenemos $f(x,y) = x^2 + 3xy^2$, $\mathbf{p}(1,2)$ y $\vec{u} = \left(\frac{-1}{\sqrt{5}}, \frac{-2}{\sqrt{5}}\right)$ luego:

$$\left.\begin{array}{l} f_x = 2x + 3y^2 \\ f_y = 6xy \end{array}\right\} \left.\begin{array}{l} f_x(1,2) = 2 + 12 = 14 \\ f_y(1,2) = 12 \end{array}\right\}$$

de donde,

$$\frac{\partial f}{\partial \vec{u}}(\mathbf{p}) = f_x(\mathbf{p}) \cdot u_1 + f_y(\mathbf{p}) \cdot u_2 = \frac{-14}{\sqrt{5}} + \frac{-24}{\sqrt{5}} = \frac{-38}{\sqrt{5}}$$

2. En el ejemplo 3.18 tenemos $f(x,y) = xyz$, $\mathbf{p}(1,0,-1)$ y $\vec{u} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ luego:

$$\left.\begin{array}{l} f_x = yz \\ f_y = xz \\ f_z = xy \end{array}\right\} \left.\begin{array}{l} f_x(1,0,-1) = 0 \\ f_y(1,0,-1) = -1 \\ f_z(1,0,-1) = 0 \end{array}\right\}$$

de donde,

$$\frac{\partial f}{\partial \vec{u}}(\mathbf{p}) = f_x(\mathbf{p}) \cdot u_1 + f_y(\mathbf{p}) \cdot u_2 + f_z(\mathbf{p}) \cdot u_3 = 0 + \frac{-1}{\sqrt{3}} + 0 = \frac{-1}{\sqrt{3}}$$

**Derivada direccional y continuidad.** La existencia de todas las derivadas direccionales de una función en un punto no garantiza la continuidad de la función en dicho punto, ya que el cálculo de las derivadas direccionales equivale a *acercarse* al punto sólo mediante rectas.

**Ejemplo 3.21.** *Estudiar la continuidad y calcular todas las derivadas direccionales en el punto* $\mathbf{p}(0,0)$ *de la función* $f : \mathbb{R}^2 \to \mathbb{R}$ *definida por:*

$$f(x,y) = \begin{cases} \dfrac{x^2 y}{x^4 + y^2} & si \;\; (x,y) \neq (0,0) \\ 0 & si \;\; (x,y) = (0,0) \end{cases}$$

*Solución.*

1. Continuidad: La función no es continua en el punto $\mathbf{p}(0,0)$, ya que no existe el límite en dicho punto. En efecto, si nos acercamos al punto mediante las parábolas $y = ax^2$ resulta:
$$\lim_{(x,y)\to(0,0)} \frac{x^2 y}{x^4 + y^2} = \lim_{\substack{(x,y)\to(0,0) \\ y=ax^2 \\ x\to 0}} \frac{x^2 a x^2}{x^4 + a^2 x^4} = \lim_{x\to 0} \frac{a}{1+a^2} = \frac{a}{1+a^2}$$
luego el límite no existe ya que depende del valor de $a$. Es decir, según la parábola por la que nos aproximemos al punto tendríamos un valor del límite u otro.

2. Existencia de las derivadas direccionales. A pesar de que la función no es continua en el punto $\mathbf{p}(0,0)$, las derivadas direccionales en dicho punto existen en todas direcciones. En efecto, sea $\vec{u} = (a,b) \in \mathbb{R}^2$ un vector unitario dado, será:
$$\begin{aligned} \frac{\partial f}{\partial \vec{u}}(0,0) &= \lim_{t\to 0} \frac{f\big((0,0) + t(a,b)\big) - f(0,0)}{t} = \lim_{t\to 0} \frac{f(ta,tb) - f(0,0)}{t} = \\ &= \lim_{t\to 0} \frac{\dfrac{(ta)^2(tb)}{(ta)^4 + (tb)^2}}{t} = \begin{cases} 0 \quad \text{si} \quad b = 0 \\ \lim\limits_{t\to 0} \dfrac{t^3 a^2 b}{t^3(t^2 a^4 + b^2)} = \dfrac{a^2}{b} \quad \text{si} \quad b \neq 0 \end{cases} \end{aligned}$$

## 3.4. Diferenciabilidad

### 3.4.1. Definición

Al generalizar un concepto de $\mathbb{R}$ a $\mathbb{R}^n$, tratamos de conservar las propiedades importantes que consideremos en el caso uni-dimensional. Por ejemplo, en $\mathbb{R}$ la existencia de la derivada en un punto $x$ implica la continuidad en el mismo. Sin embargo, para funciones de dos variable, hemos visto que la existencia de las

derivadas parciales en un punto no implica la continuidad en ese punto, ni siquiera la existencia de todas las derivadas direccionales implica la continuidad. Por esta razón las derivadas parciales, al igual que las derivadas direccionales, son una extensión en cierto modo poco satisfactoria del concepto de derivada uni-dimensional. Por tanto, parece natural el deseo de tener una noción de derivada para funciones de varias variables que implique la continuidad. Introducimos ahora una generalización más conveniente que implica la continuidad y, al propio tiempo, nos permite extender los principales teoremas de la teoría de la derivada uni-dimensional a las funciones de varias variables. El concepto que mejor sirve a tal propósito es la noción de *diferencial*.

**Generalización del concepto unidimensional.** Para funciones de una sola variable $f : \mathcal{D} \subseteq \mathbb{R} \to \mathbb{R}$ el concepto de diferenciabilidad se confunde con el concepto de derivabilidad, así, *una función $y = f(x)$ es diferenciable en un punto $x_0 \in \mathcal{D}$ sii posee derivada en ese punto.* Sin embargo, para varias variables estos dos conceptos no son equivalentes. Así, para funciones de dos variable, hemos visto que la existencia de las derivadas parciales en un punto no implica la continuidad en ese punto, ni siquiera la existencia de todas las derivadas direccionales implica la continuidad. Esto nos obliga a buscar el verdadero significado del concepto de diferenciabilidad, separándolo del concepto de derivabilidad, de manera que represente la *"suavidad"* de la función y de él se deduzca la continuidad, como ocurre en una variable.
Recordemos que una función de una variable $f : \mathcal{D} \subseteq \mathbb{R} \to \mathbb{R}$ se dice que es diferenciable en $x_0 \in \mathcal{D}$ si el siguiente límite existe y es finito:

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

En tal caso el valor del límite se llama "derivada de $f$ en $x_0$" y se denota por $f'(x_0)$. Recordemos también que dicho límite admitía una segunda expresión que era equivalente a la anterior:

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

Un primer intento para conseguir un concepto equivalente para funciones de varias variables sería copiar la definición anterior extendiéndola a la nueva situación. Esto, sin embargo, conduce a una expresión que carece de sentido. En efecto, para $n$ variables tendríamos:

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0)}{\mathbf{x} - \mathbf{x}_0}$$

donde aparece una división por un vector $\vec{v} = \mathbf{x} - \mathbf{x}_0$, operación que carece de sentido con vectores de $\mathbb{R}^n, n > 1$.
Un segundo intento sería sustituir el vector del denominador por su norma o por su longitud orientada. Pero en este caso la expresión carece de interés, ya que dicho límite sólo existe en puntos muy concretos del dominio, con lo cual se trata de una propiedad muy restrictiva de difícil verificación, y, por lo tanto, deja de representar el concepto de *función diferenciable*. En efecto, para que exista el límite,

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|}$$

su valor ha de ser 0, ya que, por ejemplo, todas las derivadas direccionales en el punto $\mathbf{x}_0$ tienen que coincidir.

Esto hace que tengamos que replantearnos el concepto de *función diferenciable* de manera que la nueva visión del concepto sea extendible a $n$ variables.

**Replanteamiento del concepto.** Partamos de la interpretación geométrica del concepto. *Una función de una variable $f(x)$ es diferenciable en un punto $x_0$ de su dominio, si su gráfica tiene recta tangente en dicho punto.* Pero, ¿qué entendemos por recta tangente a una curva en uno de sus puntos?. De todas las rectas que pasan por el punto ¿cuál es la recta tangente?. La recta tangente es una recta que *toca* a la curva en un punto, pero que, además, la curva se *aplana* en las proximidades del punto de tangencia, tratando de confundirse, "por un instante", con la propia recta. Este *aplanamiento* en los alrededores del punto de tangencia, este "*tratar de confundirse*" con la recta tangente, es lo que hace que la curva sea

*suave* y que se pueda aproximar mediante la recta tangente en los alrededores del punto de tangencia, y esto es lo que realmente caracteriza el concepto de *función diferenciable.*

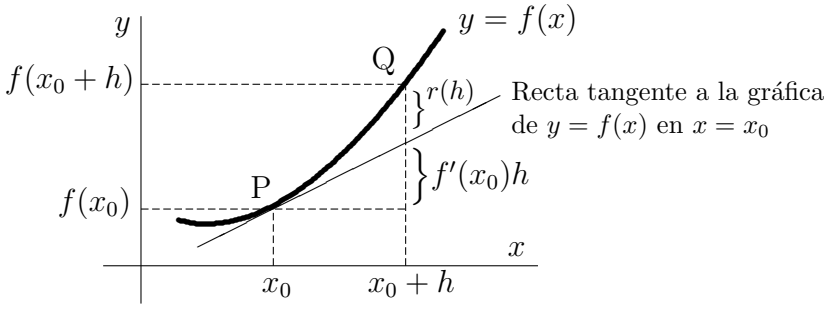



Figura 3.4: Diferencial de una función.

La recta tangente a la gráfica de la función $y = f(x)$ en el punto $P = \big(x_0, f(x_0)\big)$ viene definida por la ecuación

$$y = f(x_0) + f'(x_0)(x - x_0)$$

Y en las proximidades del punto tenemos,

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r(h)$$

donde $r(h)$ es la distancia (vertical) entre la curva y la recta tangente. Este residuo $r(h)$ es lo que nos va a permitir caracterizar el concepto de diferenciabilidad. En efecto, una primera observación nos hace ver que el residuo $r(h)$ tiende a cero a medida que $h$ tiende a cero. Sin embargo, este hecho no es importante en la definición de diferenciabilidad, pues el que $\lim_{h\to 0} r(h) = 0$ lo único que nos dice es que la función es continua en $x_0$, y seguiría valiendo cero cualquiera que fuera la recta que pase por $P$, aunque no fuera la recta tangente, e incluso, aunque la función no tuviera tangente. Lo importante, cuando se estudia la diferenciabilidad de funciones, es que el residuo $r(h)$ tiende a cero *más rápido* de lo que lo hace $h$. Esto significa que:

$$\lim_{h\to 0} \frac{r(h)}{h} = 0$$

Gráficamente, este límite viene a significar el hecho de que la curva se *"embarra"* ("*se confunde*") con la recta tangente en los alrededores del punto $P$. En otras palabras, la curva tiene que ser "suave"en $P$, para que "se pueda ver localmente como una recta"(su recta tangente).

Tratemos de expresar estos conceptos de manera algebraica, desprendiéndolos de sus significados geométricos, con objeto de poderlos generalizar a varias variables. Una recta cualquiera que pase por el punto $P = \big(x_0, f(x_0)\big)$ vendrá definida por la ecuación: $y = f(x_0) + A(x - x_0)$, o lo que es lo mismo $y = f(x_0) + Ah$ donde $A$ es la pendiente de la recta y $h = x - x_0$. Si queremos que esta recta sea la recta tangente a la curva $y = f(x)$ en el punto $P$, tendrá que cumplirse:

$$f(x_0 + h) = f(x_0) + Ah + r(h) \quad \text{donde} \quad \lim_{h\to 0} \frac{r(h)}{h} = 0$$

Por lo tanto, podemos establecer la siguiente definición de diferenciabilidad para funciones de una variable:

**Definición 3.4.** *Una función $f : D \in \mathbb{R} \to \mathbb{R}$ es diferenciable en $x_0 \in D$ si existe una constante $A$ tal que*

$$f(x_0 + h) = f(x_0) + Ah + r(h) \quad \textit{con} \quad \lim_{h\to 0} \frac{r(h)}{h} = 0$$

Esta definición es equivalente a la anterior, en efecto, despejando $A$ de la expresión anterior resulta:

$$A = \frac{f(x_0 + h) - f(x_0)}{h} - \frac{r(h)}{h}$$

de donde, al tomar límite cuando $h \to 0$ se ve la equivalencia entre la definición de diferenciabilidad y la existencia de la derivada, resultando $A = f'(x_0)$.

**Generalización para dos variables.**
*Una función de dos variable $f : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}$ es diferenciable en un punto $\mathbf{p}(x_0, y_0)$ de su dominio, si su gráfica tiene plano tangente en dicho punto.* Pero, ¿qué entendemos por plano tangente a una superficie en uno de sus puntos?. De todos los planos que pasan por el punto, ¿cuál es el plano tangente?. El plano tangente es un plano que *toca* a la superficie en un punto, pero que, además, la superficie se *aplana* en las proximidades del punto de tangencia, tratando de confundirse, "por un instante", con el propio plano. Este *aplanamiento* en los alrededores del punto de tangencia, este "*tratar de confundirse*" con el plano tangente, es lo que hace que la superficie sea *suave* y que se pueda aproximar mediante el plano tangente en los alrededores del punto de tangencia, y esto es lo que realmente caracteriza el concepto de *función diferenciable.*
Un plano cualquiera que pase por el punto $P = \big(x_0, y_0; f(x_0, y_0)\big)$ vendrá definida por la ecuación: $z = f(x_0, y_0) + A(x - x_0) + B(y - y_0)$. Si queremos que este plano sea tangente a la superficie $z = f(x, y)$ en el punto $P$, tendrá que cumplirse:

$$f\big((x_0, y_0) + (h, k)\big) = f(x_0, y_0) + Ah + Bk + r(h, k) \quad \text{donde} \quad \lim_{(h,k)\to(0,0)} \frac{r(h, k)}{\|(h, k)\|} = 0$$

Por lo tanto, podemos establecer la siguiente definición de diferenciabilidad para funciones de dos variables:

**Definición 3.5.** *Una función $f : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}$ es diferenciable en $\mathbf{p}(x_0, y_0) \in \mathcal{D}$ si existen dos constantes $A$ y $B$ tales que*

$$f\big((x_0, y_0) + (h, k)\big) = f(x_0, y_0) + Ah + Bk + r(h, k) \quad \textit{donde} \quad \lim_{(h,k)\to(0,0)} \frac{r(h, k)}{\|(h, k)\|} = 0$$

**Proposición 3.1.** *Si una función $f : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}$ es diferenciable en el punto $\mathbf{p} \in \mathcal{D}$, entonces existen sus derivadas parciales en dicho punto.*

*Demostración.* Supongamos que la función $f : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}$ es diferenciable en el punto $\mathbf{p}(x_0, y_0) \in \mathcal{D}$, entonces existen las constantes $A$ y $B$ tales que:

$$f\big((x_0, y_0) + (h, k)\big) = f(x_0, y_0) + Ah + Bk + r(h, k) \quad \text{donde} \quad \lim_{(h,k)\to(0,0)} \frac{r(h, k)}{\|(h, k)\|} = 0$$

Ahora bien, poniendo $\mathbf{h} = (h, 0)$ en la expresión anterior resulta:

$$\frac{r(h, 0)}{h} = \frac{r(h)}{h} = \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} - A$$

de donde, al tomar límite cuando $h \to 0$ resulta

$$\lim_{h\to 0} \frac{r(h)}{h} = \lim_{h\to 0} \left( \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} - A \right)$$

de donde obtenemos

$$A = \lim_{h\to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} = \frac{\partial f}{\partial x}(x_0, y_0)$$

De manera análoga, poniendo $\mathbf{h} = (0, k)$ obtenemos:

$$B = \lim_{k \to 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k} = \frac{\partial f}{\partial y}(x_0, y_0) \qquad \square$$

Se tiene entonces que una condición *necesaria* para que una función $f(x, y)$ sea diferenciable en un punto $\mathbf{p} = (x_0, y_0)$ es que existan sus derivadas parciales en ese punto. Sin embargo, esta condición no es suficiente, ya que la existencia de todas las derivadas parciales no garantiza la diferenciabilidad. No obstante, lo interesante de esta propiedad es su negación lógica, o sea: *Si una función no tiene todas sus derivadas parciales, entonces no es diferenciable.*
Como consecuencia de este resultado, podemos establecer la siguiente

**Proposición 3.2.** *La función $f : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}$ definida en un conjunto abierto $\mathcal{D}$ de $\mathbb{R}^2$, es diferenciable en el punto $\mathbf{p} = (x_0, y_0) \in \mathcal{D}$, si:*

1. *Existen las derivadas parciales de $f$ en $\mathbf{p}$*

$$A = \frac{\partial f}{\partial x}(x_0, y_0), \qquad B = \frac{\partial f}{\partial y}(x_0, y_0)$$

2. *El residuo $r(h, k)$ definido en la expresión:*

$$f\big((x_0, y_0) + (h, k)\big) = f(x_0, y_0) + Ah + Bk + r(h, k)$$

*tiene la propiedad*

$$\lim_{(h,k) \to (0,0)} \frac{r(h, k)}{\|(h, k)\|} = 0$$

Esta definición de diferenciabilidad sí que garantiza la continuidad de la función en aquellos puntos en los que es diferenciable, como sucede con las funciones de una variable.

**Nota:** Con objeto de recorda mejor la expresión, el límite que caracteriza la diferenciabilidad se puede expresar de la siguiente forma:

$$\lim_{(h,k) \to (0,0)} \frac{r(h, k)}{\|(h, k)\|} = \lim_{(h,k) \to (0,0)} \frac{f\big((x_0, y_0) + (h, k)\big) - f(x_0, y_0) - Ah - Bk}{\sqrt{h^2 + k^2}}$$

Y llamando $\Delta f = f(x_0 + h, y_0 + k) - f(x_0, y_0)$ y $\lambda(h, k) = Ah + Bk$ resulta

$$\boxed{\lim_{(h,k) \to (0,0)} \frac{r(h, k)}{\|(h, k)\|} = \lim_{(h,k) \to (0,0)} \frac{\Delta f - \lambda(h, k)}{\sqrt{h^2 + k^2}} = 0} \tag{3.2}$$

Con lo cual la proposición 3.2 se puede enunciar de la siguiente forma:

**Corolario 3.1.** *La función $f : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}$ definida en un conjunto abierto $\mathcal{D}$ de $\mathbb{R}^2$, es diferenciable en el punto $\mathbf{p} = (x_0, y_0) \in \mathcal{D}$, si:*

1. *Existen las derivadas parciales de $f$ en* $\mathbf{p}$

$$A = \frac{\partial f}{\partial x}(x_0, y_0), \qquad B = \frac{\partial f}{\partial y}(x_0, y_0)$$

2. *El siguiente límite vale cero:*

$$\lim_{(h,k)\to(0,0)} \frac{r(h,k)}{\|(h,k)\|} = \lim_{(h,k)\to(0,0)} \frac{\Delta f - \lambda(h,k)}{\sqrt{h^2+k^2}} = 0$$

*donde,* $\Delta f = f(x_0+h, y_0+k) - f(x_0, y_0)$ *y* $\lambda(h,k) = Ah + Bk$

**La diferencial**

Si la función $f : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}$ es diferenciable en el punto $\mathbf{p} = (x_0, y_0) \in \mathcal{D}$, entonces a la parte lineal en $h$ y $k$ ($\lambda(h,k) = Ah + Bk$) de la expresión del residuo

$$f\big((x_0, y_0) + (h,k)\big) = f(x_0, y_0) + Ah + Bk + r(h,k) \quad \text{con} \quad \lim_{(h,k)\to(0,0)} \frac{r(h,k)}{\|(h,k)\|} = 0$$

se le llama *diferencial* de la función $f$ en $(x_0, y_0)$ y se denota por $df(x_0, y_0)$. Así,

$$df(x_0, y_0) = Ah + Bk$$

Y dado que, como se vio en la proposición 3.1, $A$ y $B$ representan las derivadas parciales de la función $f$ en el punto $(x_0, y_0)$, resulta:

$$df(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k$$

Y teniendo en cuenta que si $f(x,y) = x$, se tiene $h = dx$ y si $f(x,y) = y$, se tiene $k = dy$, podemos escribir:

$$df(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)dx + \frac{\partial f}{\partial y}(x_0, y_0)dy$$

O de manera más general:

**Definición 3.6 (La diferencial).** *Si la función $f : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}$ es diferenciable en el conjunto abierto $\mathcal{D}$ de $\mathbb{R}^2$, entonces, para cada $\mathbf{x} = (x,y) \in \mathcal{D}$, se llama diferencial de la función $f$ en $\mathbf{x}$ y se denota por $df$, a la expresión:*

$$\boxed{df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy} \tag{3.3}$$

**Teorema 3.2.** *Si la función $f : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}$ definida en un conjunto abierto $\mathcal{D}$ de $\mathbb{R}^2$, es diferenciable en el punto $\mathbf{p} = (x_0, y_0) \in \mathcal{D}$, entonces es continua en ese punto.*

*Demostración.* Si $f$ es diferenciable en el punto $\mathbf{p} = (x_0, y_0)$ se tiene

$$f\big((x_0, y_0) + (h, k)\big) = f(x_0, y_0) + Ah + Bk + r(h, k) \quad \text{donde} \quad \lim_{(h,k)\to(0,0)} \frac{r(h,k)}{\|(h,k)\|} = 0$$

Tomando límite cuando $(h, k) \to (0, 0)$ y teniendo en cuenta que

$$\lim_{(h,k)\to(0,0)} \frac{r(h,k)}{\|(h,k)\|} = 0 \quad \Rightarrow \quad \lim_{(h,k)\to(0,0)} r(h,k) = 0$$

tenemos que

$$\lim_{(h,k)\to(0,0)} f\big((x_0, y_0) + (h, k)\big) = \lim_{(h,k)\to(0,0)} \big(f(x_0, y_0) + Ah + Bk + r(h, k)\big) = f(x_0, y_0)$$

luego la función es continua en $\mathbf{p} = (x_0, y_0)$. □

El recíproco no es cierto, ya que existen funciones continuas que no son diferenciables.

**Ejemplo 3.22.** *Estudiar la continuidad y la diferenciabilidad en el origen de la siguiente función y, en su caso, hallar su diferencial en ese punto.*

$$f(x, y) = \begin{cases} \dfrac{xy}{\sqrt{x^2 + y^2}} & si \quad (x, y) \neq (0, 0) \\ 0 & si \quad (x, y) = (0, 0) \end{cases}$$

*Solución.*

(a) Continuidad:

$$\lim_{(x,y)\to(0,0)} \frac{xy}{\sqrt{x^2 + y^2}} = \lim_{(x,y)\to(0,0)} x\frac{y}{\sqrt{x^2 + y^2}} = 0 \cdot Acot = 0 = f(0, 0)$$

luego la función es continua en $(0, 0)$

(b) Diferenciabilidad: Calculamos las derivadas parciales aplicando la definición.

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{h\to 0} \frac{f(0 + h, 0) - f(0, 0)}{h} = \lim_{h\to 0} \frac{f(h, 0) - 0}{h} = \lim_{h\to 0} \frac{\dfrac{h \cdot 0}{\sqrt{h^2 + 0^2}}}{h} = 0$$

$$\frac{\partial f}{\partial y}(0, 0) = \lim_{k\to 0} \frac{f(0, 0 + k) - f(0, 0)}{k} = \lim_{k\to 0} \frac{f(0, k) - 0}{k} = \lim_{k\to 0} \frac{\dfrac{0 \cdot k}{\sqrt{0^2 + k^2}}}{k} = 0$$

luego,

$$\lambda(h, k) = 0h + 0k = 0$$

Por otro lado,

$$\Delta f(0, 0) = f(0 + h, 0 + k) - f(0, 0) = f(h, k) - 0 = \frac{hk}{\sqrt{h^2 + k^2}}$$

luego,

$$\lim_{(h,k)\to(0,0)} \frac{r(h,k)}{\|(h,k)\|} = \lim_{(h,k)\to(0,0)} \frac{\Delta f - \lambda(h,k)}{\sqrt{h^2+k^2}} = \lim_{(h,k)\to(0,0)} \frac{\dfrac{hk}{\sqrt{h^2+k^2}} - 0}{\sqrt{h^2+k^2}} =$$

$$= \lim_{(h,k)\to(0,0)} \frac{hk}{h^2+k^2} = \lim_{\substack{k=mh\\ h\to 0}} \frac{hmh}{h^2+m^2h^2} = \frac{m}{1+m^2}$$

Luego el límite no existe, por depender de $m$, y en consecuencia la función no es diferenciable en $(0,0)$. Al no ser diferenciable resulta que $df(0,0)$ no existe, con lo cual $\lambda(h,k) = 0h + 0k = 0$ no tiene ninguna significación.

**Ejemplo 3.23.** *Estudiar la continuidad y la diferenciabilidad en el origen de la siguiente función y, en su caso, hallar su diferencial en ese punto.*

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & si \quad (x,y) \neq (0,0) \\ 0 & si \quad (x,y) = (0,0) \end{cases}$$

*Solución.*

(a) Continuidad: Nos acercamos al origen a través de la recta $y = mx$

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2} = \lim_{\substack{y=mx\\ x\to 0}} \frac{xmx}{x^2+m^2x^2} = \frac{m}{1+m^2} = f(m)$$

luego la función no es continua en $(0,0)$

(b) Diferenciabilidad: Al no ser continua la función en el punto $(0,0)$ no puede ser diferenciable en dicho punto, en consecuencia resulta que $df(0,0)$ no existe.

**Ejemplo 3.24.** *Estudiar la continuidad y la diferenciabilidad en el origen de la siguiente función y, en su caso, hallar su diferencial en ese punto.*

$$f(x,y) = \sqrt{x^2+y^2}$$

*Solución.*

(a) Continuidad: La función es continua en $(0,0)$, en efecto

$$\lim_{(x,y)\to(0,0)} \sqrt{x^2+y^2} = \sqrt{0^2+0^2} = \sqrt{0^+} = 0 = f(0,0)$$

(b) Diferenciabilidad: Las derivadas parciales en el origen no existen, en efecto

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0} \frac{\sqrt{h^2+0^2}-0}{h} = \lim_{h\to 0} \frac{\sqrt{h^2}}{h} = \lim_{h\to 0} \frac{|h|}{h}$$

y dicho límite no existe, puesto que

$$\lim_{h\to 0^+} \frac{|h|}{h} = 1 \qquad \text{y} \qquad \lim_{h\to 0^-} \frac{|h|}{h} = -1$$

Luego, la función no es diferenciable en el origen por no existir las derivadas parciales en dicho punto y ser la existencia de las derivadas parciales en un punto **p** una condición necesaria para la diferenciabilidad de la función en dicho punto. En consecuencia resulta que $df(0,0)$ no existe.

**Ejemplo 3.25.** *Estudiar la continuidad y la diferenciabilidad de la función $f(x,y) = xy^2$ en el origen y hallar su diferencial en ese punto.*

*Solución.* Estudiemos primero la diferenciabilidad en el origen

$$f(xy) = xy^2 \left\{ \begin{array}{l} f_x(x,y) = y^2 \\ f_y(x,y) = 2xy \end{array} \right\} \left. \begin{array}{l} f_x(0,0) = 0 \\ f_y(0,0) = 0 \end{array} \right\} \lambda(h,k) = 0h + 0k = 0$$
$$\Delta f(0,0) = f(0+h, 0+k) - f(0,0) = f(h,k) - 0 = hk^2$$

luego,

$$\lim_{(h,k)\to(0,0)} \frac{r(h,k)}{\|(h,k)\|} = \lim_{(h,k)\to(0,0)} \frac{\Delta f - \lambda(h,k)}{\sqrt{h^2+k^2}} = \lim_{(h,k)\to(0,0)} \frac{hk^2 - 0}{\sqrt{h^2+k^2}} = \lim_{(h,k)\to(0,0)} k^2 \frac{h}{\sqrt{h^2+k^2}} = 0$$

Luego la función es diferenciable en $(0,0)$ y en consecuencia es continua en dicho punto. Al ser diferenciable resulta,

$$df(0,0) = \lambda(h,k) = 0h + 0k = 0$$

### Caracterización de las funciones diferenciables

Con objeto de desarrollar una intuición que permita detectar a priori la diferenciabilidad de las funciones de varias variables deben tenerse en cuenta las siguientes proposiciones

**Proposición 3.3.** *Toda función polinómica $f : \mathbb{R}^2 \to \mathbb{R}$, $\quad f(x,y) = \sum_{i,j=0}^{n} a_{ij} x^i y^j$ es diferenciable en todo punto $(x_0, y_0) \in \mathbb{R}^2$*

**Proposición 3.4.**

(a) *La suma y el producto de funciones diferenciables es otra función diferenciable.*

(b) *El cociente de dos funciones diferenciables es otra función diferenciable en todos aquellos puntos que no anulen el denominador.*

**Proposición 3.5.** *La composición de dos funciones diferenciables es otra función diferenciable.*

**Ejemplo 3.26.** *Estudiar la diferenciabilidad de las funciones*

(a) $f(x,y) = \dfrac{x^2 - xy + 1}{x^2 + y^2}$ (b) $g(x,y) = e^{x^2+y^2}$ (c) $h(x,y) = x^2 e^{x^2+y^2} + \operatorname{sen} \dfrac{1-y^2}{1+x^2}$

*Solución.*

(a) La función es diferenciable en $\mathbb{R}^2 - \{(0,0)\}$ por tratarse del cociente de dos funciones polinómicas y el único punto que anula el denominador es el $(0,0)$.

(b) La función es diferenciable en todo $\mathbb{R}^2$ por tratarse de la composición de dos funciones diferenciables.

(c) La función es diferenciable en todo $\mathbb{R}^2$ por tratarse de suma de funciones diferenciables.

## Funciones de $n$ variables.

El concepto de diferenciabilidad se puede extender a funciones de cualquier número de variables, de manera análoga a como se ha hecho al caso de dos variable.

**Definición 3.7.** *Se dice que la función $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en un conjunto abierto $\mathcal{D}$ de $\mathbb{R}^n$, es diferenciable en el punto $\mathbf{x}_0 = (x_1, \cdots, x_n) \in \mathcal{D}$, si existen las derivadas parciales de $f$ en $\mathbf{x}_0$*

$$A_i = \frac{\partial f}{\partial x_i}(\mathbf{x}_0), \quad i = 1, 2, \cdots, n$$

*y si el residuo $r(\mathbf{h})$ definido en la expresión:*

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} A_i h_i + r(\mathbf{h})$$

*(donde $\mathbf{h} = (h_1, \cdots, h_n) \in \mathbb{R}^n$ es tal que $\mathbf{x}_0 + \mathbf{h} \in \mathcal{D}$) tiene la propiedad*

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

Esta definición de diferenciabilidad también garantiza la continuidad de la función en aquellos puntos en los que es diferenciable, como sucede con las funciones de una y dos variable.

**Nota:** Con objeto de recorda mejor la expresión, el límite que caracteriza la diferenciabilidad se puede expresar de la siguiente forma:

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = \lim_{(h,k) \to (0,0)} \frac{f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) - \sum_{i=1}^{n} A_i h_i}{\sqrt{h_1^2 + \cdots + h_n^2}}$$

Y llamando $\Delta f = f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0)$ y $\lambda(\mathbf{h}) = \sum_{i=1}^{n} A_i h_i$ resulta

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = \lim_{\mathbf{h} \to \mathbf{0}} \frac{\Delta f - \lambda(\mathbf{h})}{\sqrt{h_1^2 + \cdots + h_n^2}} = 0$$

### La Diferencial

Si la función $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en el punto $\mathbf{x}_0 = (x_1, \cdots, x_n) \in \mathcal{D}$, entonces a la parte lineal en $h_i$, $i = 1, 2, \cdots, n$ ($\lambda(\mathbf{h}) = A_1 h_1 + \cdots + A_n h_n$) de la expresión del residuo

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + A_1 h_1 + \cdots + A_n h_n + r(\mathbf{h}) \quad \text{con} \quad \lim_{\mathbf{h} \to \mathbf{o}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

se le llama *diferencial* de la función $f$ en $\mathbf{x}_0$ y se denota por $df(\mathbf{x}_0)$.

$$df(\mathbf{x}_0) = \lambda(\mathbf{h}) = A_1 h_1 + \cdots + A_n h_n$$

Y dado que los coeficientes $A_i$ $i = 1, \cdots, n$ representan las derivadas parciales de la función $f$ en el punto $\mathbf{x}_0$, resulta

$$df(\mathbf{x}_0) = \frac{\partial f}{\partial x}(\mathbf{x}_0) h_1 + \cdots + \frac{\partial f}{\partial y}(\mathbf{x}_0) h_n$$

Y teniendo en cuenta que si $f(x_1, \cdots, x_n) = x_1$, se tiene $h_1 = dx_1$ y si $f(x_1, \cdots, x_n) = x_n$, se tiene $h_n = dx_n$, podemos escribir:

$$df(\mathbf{x}_0) = \frac{\partial f}{\partial x}(\mathbf{x}_0) dx_1 + \cdots + \frac{\partial f}{\partial y}(\mathbf{x}_0) dx_n$$

Y, en general, podemos enunciar la siguiente

**Definición 3.8 (La diferencial).** *Si la función $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en el conjunto abierto $\mathcal{D}$, entonces para cada $\mathbf{x} = (x_1, \cdots, x_n) \in \mathcal{D}$, se llama diferencial de la función $f$ en $\mathbf{x}$, y se denota por $df$, a la expresión*

$$df = \frac{\partial f}{\partial x} dx_1 + \cdots + \frac{\partial f}{\partial y} dx_n$$

### Condición suficiente para la diferenciabilidad

**Teorema 3.3.** *Sea $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto $\mathcal{D}$ de $\mathbb{R}^n$. Si las funciones (derivadas parciales)*

$$\frac{\partial f}{\partial x_i} : \tilde{D} \subseteq \mathbb{R}^n \to \mathbb{R}, \quad i = 1, 2, \cdots, n, \quad \tilde{D} \subseteq \mathcal{D}$$

*son continuas en el punto $\mathbf{x}_0 \in \tilde{D}$, entonces $f$ es diferenciable en $\mathbf{x}_0$*

*Demostración.* (Opcional, caso $n = 2$) Sea la función $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ tal que las funciones derivadas parciales sean continuas en un punto $\mathbf{p} = (x_0, y_0)$. Sea $\mathbf{x} = (x, y)$ un punto próximo a $\mathbf{p}$ y sea $\mathbf{q} = (x_0, y_0 + k)$. Tenemos:

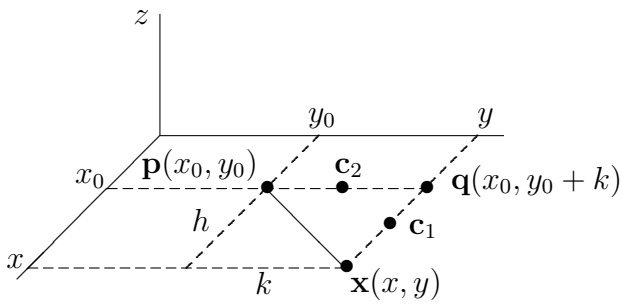



Figura 3.5: valor medio.

$$\begin{aligned}\Delta f &= f(\mathbf{x}) - f(\mathbf{p}) = \\ &= \big[f(\mathbf{x}) - f(\mathbf{q})\big] + \big[f(\mathbf{q}) - f(\mathbf{p})\big]\end{aligned}$$

Aplicando el teorema del valor medio en cada corchete, resulta

$$\Delta f = f_x(\mathbf{c_1})h + f_y(\mathbf{c_2})k$$

Y al ser las parciales continuas en el punto $\mathbf{p}$ será:

$$\begin{aligned}\Delta f &= \big(f_x(\mathbf{p}) + \epsilon_1\big)h + \big(f_y(\mathbf{p}) + \epsilon_2\big)k = \\ &= f_x(\mathbf{p})h + f_y(\mathbf{c_2})k + \epsilon_1 h + \epsilon_2 k\end{aligned}$$

Luego $f$ es diferenciable en $\mathbf{p}$ □

El recíproco de este teorema no es cierto, ya que existen funciones diferenciables en un punto $\mathbf{p}$ y sin embargo sus derivadas parciales en dicho punto no son continuas. Por lo tanto, si las derivadas parciales no son continuas, entonces no podemos asegurar nada. No obstante, este teorema permite comprobar la diferenciabilidad de una función de una manera fácil, y se puede extender a un dominio $\mathcal{D}$ de la siguiente forma: *Si las derivadas parciales son continuas en un dominio $\mathcal{D}$ entonces la función es diferenciable en ese dominio.*

Recapitulando, podemos resumir las condiciones de diferenciabilidad de la siguiente forma:

### Condiciones necesarias de diferenciabilidad:

- *Si la función es diferenciable en un punto, entonces es continua en ese punto.*

- *Si la función es diferenciable en un punto, entonces existen las derivadas parciales en ese punto.*

**Condición suficiente de diferenciabilidad:** *Si las derivadas parciales son continuas en un punto, entonces la función es diferenciable en ese punto.*

Los recíprocos de los teoremas anteriores no son ciertos. Para recordarlos podemos realizar el siguientes esquema gráfico

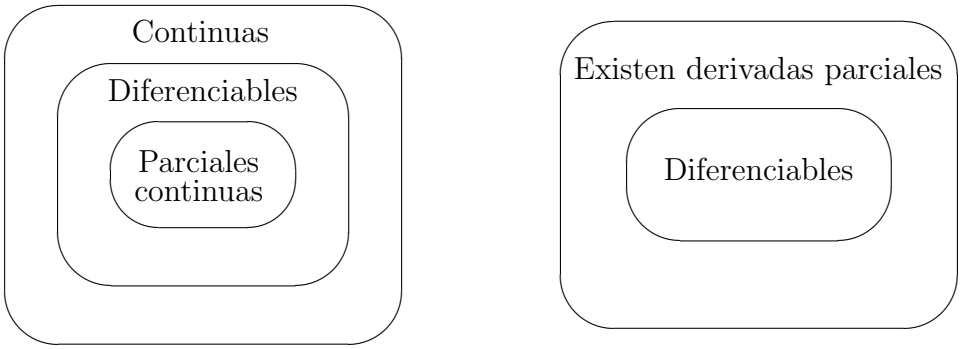



Figura 3.6: Parciales continuas $\Rightarrow$ Diferenciable $\Rightarrow$Continua  Diferenciable $\Rightarrow$Existen parciales

**Ejemplo 3.27.** *Estudiar la diferenciabilidad de las funciones*

(a) $f(x,y) = xy^2$  (b) $g(x,y) = e^{x^2+y^2}$  (c) $h(x,y,z) = \operatorname{sen}(x + y^2 - z^2)$

*Solución.*

(a) Las derivadas parciales de la función $f(x,y) = xy^2$ son:

$$f_x(x,y) = y^2 \qquad f_y(x,y) = 2xy$$

que son funciones continuas en $\mathbb{R}^2$, luego la función es diferenciable en $\mathbb{R}^2$.

(b) Las derivadas parciales de la función $g(x,y) = e^{x^2+y^2}$ son:

$$g_x(x,y) = 2xe^{x^2+y^2} \qquad g_y(x,y) = 2ye^{x^2+y^2}$$

que son funciones continuas en $\mathbb{R}^2$, luego la función es diferenciable en $\mathbb{R}^2$.

(c) Las derivadas parciales de la función $h(x,y,z) = \operatorname{sen}(x + y^2 - z^2)$ son:

$$\begin{aligned} h_x(x,y,z) &= \cos(x + y^2 - z^2) \\ h_y(x,y,z) &= 2y\cos(x + y^2 - z^2) \\ h_z(x,y,z) &= -2z\cos(x + y^2 - z^2) \end{aligned}$$

que son funciones continuas en $\mathbb{R}^3$, luego la función es diferenciable en $\mathbb{R}^3$.

### 3.4.2. Diferenciabilidad y derivadas direccionales

El cálculo de la derivada direccional aplicando el límite (3.1) que aparece en la definición resulta un poco engorroso. No obstante, para las funciones diferenciables existe la posibilidad de calcular las derivadas direccionales de una manera mucho más fácil. Como *la suma de los productos de las derivadas parciales por las componentes del vector unitario de dirección.* En efecto,

**Teorema 3.4.** *Sea $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable en el punto $\mathbf{x}_0 \in D$, y sea $\mathbf{u} = (u_1, \cdots, u_n)$ un vector unitario. Entonces*

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0) u_i$$

*Demostración.* Al ser $f$ diferenciable en $\mathbf{x}_0$ se tiene que:

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0) h_i + r(\mathbf{h}) \quad \text{donde} \quad \lim_{\mathbf{h}\to\mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

Limitemos el incremento de la función $f$ sólo a la dirección marcada por el vector $\mathbf{u}$. Para ello, escribamos el vector $\mathbf{h}$ como $\mathbf{h} = t\mathbf{u}$, $t \in \mathbb{R}$, luego será: $h_i = tu_i$, con lo cual nos queda

$$f(\mathbf{x}_0 + t\mathbf{u}) - f(\mathbf{x}_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0) t u_i + r(t\mathbf{u})$$

de donde, dividiendo po $t$, resulta

$$\frac{f(\mathbf{x}_0 + t\mathbf{u}) - f(\mathbf{x}_0)}{t} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0) u_i + \frac{r(t\mathbf{u})}{t}$$

Y teniendo en cuenta que $\|\mathbf{h}\| = \|t\mathbf{u}\| = |t|\|\mathbf{u}\| = |t|$, pues el vector $\mathbf{u}$ es unitario. Es decir, $t$ es la *longitud orientada por* $\mathbf{u}$ del vector $\mathbf{h}$, es decir, la longitud de este vector con signo positivo, si el vector $\mathbf{h}$ tiene la misma dirección que el vector $\mathbf{u}$, y con signo negativo en caso contrario. Es decir, $t = \pm\|\mathbf{h}\|$. Con lo cual decir que $\mathbf{h} \to \mathbf{0}$ equivale a decir que $t \to 0$. En consecuencia tomando límite en la expresión anterior cuando $t \to 0$, (y teniendo en cuenta que los primeros sumandos del 2º miembro no dependen de $\mathbf{h}$ ni de $t$) resulta

$$\lim_{t\to 0} \frac{f(\mathbf{x}_0 + t\mathbf{u}) - f(\mathbf{x}_0)}{t} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0) u_i \pm \lim_{\mathbf{h}\to\mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|}$$

El límite que aparece en el primer miembro es la derivada direccional de la función $f$ en el punto $\mathbf{x}_0$ según la dirección del vector unitario $\mathbf{u}$ (ver (3.1)), y el límite que aparece en el segundo miembro es cero por ser la función diferenciable. En consecuencia resulta:

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0) u_i \qquad \square$$

**Ejemplo 3.28.** *Calcular, usando las derivadas parciales, la derivada direccional de la función $f(x, y) = x^2 + y^2$ en el punto $\mathbf{p}(1, 1)$ en el sentido del vector que parte de $\mathbf{p}$ y forma un ángulo de $60^o$ con el sentido positivo del eje $OX$.*

*Solución.* Hallamos el vector unitario de dirección,

$$\vec{u} = (\cos\alpha, \operatorname{sen}\alpha) = (\frac{1}{2}, \frac{\sqrt{3}}{2})$$

Hallamos las derivadas parciales en el punto $\mathbf{p}(1, 1)$

$$\left.\begin{array}{l} f_x = 2x \\ f_y = 2y \end{array}\right\} \begin{array}{l} f_x(1,1) = 2 \\ f_y(1,1) = 2 \end{array}$$

de donde,

$$D_{\vec{u}} f(1, 1) = f_x(1, 1)u_1 + f_y(1, 1)u_2 = 2\frac{1}{2} + 2\frac{\sqrt{3}}{2} = 1 + \sqrt{3}$$

### 3.4.3. La derivada según una dirección curva

La derivada direccional también se puede aplicar para direcciones curvas. En este caso se entiende que el vector director es el vector tangente a la curva.

**Vector tangente a una curva plana**

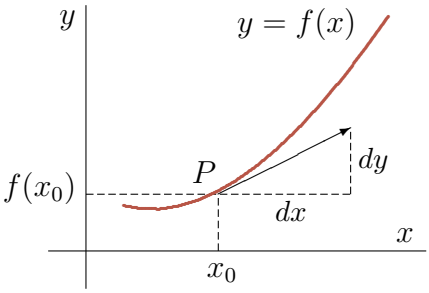



Figura 3.7: Vector tangente.

Como vector tangente podemos elegir cualquiera de los vectores siguientes, todos ellos paralelos entre sí:

$$\vec{v}_T = (dx, dy) \parallel (1, \frac{dy}{dx}) = \big(1, y'(x)\big) \parallel \big(x'(t), y'(t)\big)$$

**Ejemplo 3.29.** *Calcular la derivada del campo escalar* $z = \arctan(xy)$ *en el punto* $\mathbf{p}(1,1)$*, según la dirección de la parábola* $y = x^2$*, en el sentido del crecimiento de la abscisa.*

*Solución.* Hallamos el vector tangente unitario a la parábola $y = x^2$, en el punto $\mathbf{p}(1,1)$, con la primera componente positiva.

$$\vec{v}_T = \big(1, y'(x)\big) = (1, 2x) \quad \rightarrow \quad \vec{v}_T(1,1) = (1,2)$$

$$|\vec{v}_T(1,1)| = \sqrt{1+4} = \sqrt{5} \quad \rightarrow \quad \vec{u}_T(1,1) = (\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$$

Hallamos las derivadas parciales de la función en el punto $\mathbf{p}$

$$\left.\begin{array}{l} \dfrac{\partial z}{\partial x} = \dfrac{y}{1+x^2y^2} \\ \dfrac{\partial z}{\partial y} = \dfrac{x}{1+x^2y^2} \end{array}\right\} \left.\begin{array}{l} \dfrac{\partial z}{\partial x}(1,1) = \dfrac{1}{1+1} = \dfrac{1}{2} \\ \dfrac{\partial z}{\partial y}(1,1) = \dfrac{x}{1+1} = \dfrac{1}{2} \end{array}\right\} \quad \text{de donde,} \quad \frac{\partial z}{\partial \vec{u}_T} = \frac{1}{2}\frac{1}{\sqrt{5}} + \frac{1}{2}\frac{2}{\sqrt{5}} = \frac{3}{2\sqrt{5}}$$

## 3.5. Gradiente

### 3.5.1. Definición

Si la función es diferenciable, entonces la derivada direccional y el diferencial recuerdan un producto escalar

$$D_{\vec{u}} f = f_x u_1 + f_y u_2 = (f_x, f_y) \cdot (u_1, u_2) = \overrightarrow{\nabla f} \cdot \vec{u}$$

$$df = f_x dx + f_y dy = (f_x, f_y) \cdot (dx, dy) = \overrightarrow{\nabla f} \cdot \vec{v}_T$$

Este resultado nos hace tener en consideración el vector cuyas componentes son las derivadas parciales de una función en un punto. Así, *Dada una función diferenciable de dos variables, se llama vector gradiente de dicha función en un punto* $\mathbf{p}$*, al vector cuyas componentes son las derivadas parciales de la función en dicho punto.* Y se denota por cualquiera de los símbolos $\mathbf{grad} f(\mathbf{p})$, $\nabla f(\mathbf{p})$ o $\overrightarrow{\nabla f}(\mathbf{p})$.

$$\mathbf{grad} f(\mathbf{p}) = \nabla f(\mathbf{p}) = \overrightarrow{\nabla f}(\mathbf{p}) = \left(\frac{\partial f}{\partial x}(\mathbf{p}), \frac{\partial f}{\partial y}(\mathbf{p})\right)$$

De manera análoga se define el vector gradiente para tres o más variables

$$\mathbf{grad} f(\mathbf{p}) = \nabla f(\mathbf{p}) = \overrightarrow{\nabla f}(\mathbf{p}) = \left(\frac{\partial f}{\partial x}(\mathbf{p}), \frac{\partial f}{\partial y}(\mathbf{p}), \frac{\partial f}{\partial z}(\mathbf{p})\right)$$

Formalmente la definición es la siguiente:

**Definición 3.9.** *Sea $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable definida en el conjunto abierto $\mathcal{D}$ de $\mathbb{R}^n$. Se define el (vector)* gradiente *de la función $f$ en el punto $\mathbf{x}_0 \in \mathcal{D}$, como el vector de $\mathbb{R}^n$ dado por*

$$\mathbf{grad} f(\mathbf{x}_0) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \frac{\partial f}{\partial x_2}(\mathbf{x}_0), \cdots, \frac{\partial f}{\partial x_n}(\mathbf{x}_0)\right)$$

**Ejemplo 3.30.** *Hallar el vector gradiente de la función $f(x, y) = x^2y + xy^3$ en el punto $(-1, 2)$*

*Solución.* Hallamos las derivadas parciales y las evaluamos en el punto $(-1, 2)$

$$\left.\begin{array}{l} f_x = 2xy + y^3 \\ f_y = x^2 + 3xy^2 \end{array}\right\} \begin{array}{l} f_x(-1, 2) = -4 + 8 = 4 \\ f_y(-1, 2) = 1 - 12 = -11 \end{array}$$

de donde,

$$\overrightarrow{\nabla f}(-1, 2) = (4, -11)$$

**Ejemplo 3.31.** *Hallar el vector gradiente de la función $f(x, y, z) = \dfrac{x + y}{z}$ en un punto genérico.*

*Solución.* Hallamos las derivadas parciales

$$\left.\begin{array}{l} f_x = \dfrac{1}{z} \\ f_y = \dfrac{1}{z} \\ f_z = -\dfrac{x + y}{z^2} \end{array}\right\} \overrightarrow{\nabla f} = \left(\frac{1}{z}, \frac{1}{z}, \frac{-x - y}{z}\right)$$

**Nota.** Para funciones de una variable $y = f(x)$ teníamos que:

$$df = f'(x)dx = f'(x) \cdot h$$

Para funciones de varias variables tenemos que:

$$df = f_x dx + f_y dy = (f_x, f_y) \begin{pmatrix} dx \\ dy \end{pmatrix} = \nabla f \cdot \mathbf{h}$$

Si se comparan ambas expresiones, se observa que el gradiente, $\nabla f$, puede pensarse como la generalización del concepto de derivada para funciones de varias variables. Si bien, hay que advertir que mientras que la derivada de una función de una variable en un punto es un *número*, la derivada de una función de varias variables es un *vector*.

### 3.5.2. Vector gradiente y derivada direccional

A partir del vector gradiente, la derivada direccional de una función diferenciable $f$ en un punto $\mathbf{p}$ en la dirección del vector unitario $\vec{u}$, se puede expresar como un producto escalar

$$\frac{\partial f}{\partial \vec{u}}(\mathbf{p}) = \overrightarrow{\nabla f} \cdot \vec{u}$$

Es decir, *la derivada direccional de la función $f$ en el punto $\mathbf{p}$ en la dirección del vector unitario $\vec{u}$ es el producto escalar del vector gradiente de $f$ en el punto $\mathbf{p}$ por el vector $\vec{u}$.* Este hecho permite obtener las siguientes propiedades

**Propiedades:**

(a) *Si en un punto $\mathbf{p}$ el gradiente es cero, entonces todas las derivadas direccionales en ese punto valen cero.*

$$\nabla f(\mathbf{p}) = 0 \quad \Rightarrow \quad D_{\vec{v}} f(\mathbf{p}) = 0, \quad \text{cualquiera que sea} \quad \vec{v}$$

(b) *La derivada direccional es máxima en la dirección y sentido del gradiente, mínima en sentido contrario, y nula en la dirección perpendicular al gradiente, además, el valor máximo de esta derivada es el módulo del gradiente.*
En efecto, teniendo en cuenta que el producto escalar de dos vectores es el producto de los módulos de los vectores por el coseno del ángulo que forman, y llamando $\theta$ al ángulo que forman el gradiente $\overrightarrow{\nabla f}$ y el vector de dirección $\vec{u}$, resulta

$$\frac{\partial f}{\partial \vec{u}}(\mathbf{p}) = \overrightarrow{\nabla f} \cdot \vec{u} = \|\overrightarrow{\nabla f}\| \cdot \|\vec{u}\| \cdot cos\,\theta = \|\overrightarrow{\nabla f}\| \cdot 1 \cdot cos\,\theta = \|\overrightarrow{\nabla f}\| cos\,\theta$$

Ahora bien, $-1 \leq \cos\theta \leq 1$, luego el valor máximo del producto $\|\overrightarrow{\nabla f}\| cos\,\theta$ se obtiene cuando $\cos\theta = 1$ y el valor mínimo cuando $\cos\theta = -1$, es decir,
Cuando $\overrightarrow{\nabla f}$ y $\vec{u}$ son vectores de la misma dirección y sentido, se tiene $\theta = 0$, y por lo tanto $\cos\,\theta = 1$, de donde,

$$\frac{\partial f}{\partial \vec{u}}(\mathbf{p}) = \overrightarrow{\nabla f} \cdot \vec{u} = \|\overrightarrow{\nabla f}\| \cdot \|\vec{u}\| \cdot cos\,0 = \|\overrightarrow{\nabla f}\| \cdot 1 \cdot 1 = \|\overrightarrow{\nabla f}\|$$

Cuando $\overrightarrow{\nabla f}$ y $\vec{u}$ son vectores de la misma dirección pero de sentidos opuestos, se tiene $\theta = \pi$, y por lo tanto $\cos\,\theta = -1$, de donde,

$$\frac{\partial f}{\partial \vec{u}}(\mathbf{p}) = \overrightarrow{\nabla f} \cdot \vec{u} = \|\overrightarrow{\nabla f}\| \cdot \|\vec{u}\| \cdot cos\,\pi = \|\overrightarrow{\nabla f}\| \cdot 1 \cdot (-1) = -\|\overrightarrow{\nabla f}\|$$

Y cuando $\overrightarrow{\nabla f}$ y $\vec{u}$ son vectores perpendiculares, se tiene $\theta = \pi/2$, y por lo tanto $\cos\,\theta = 0$, de donde,

$$\frac{\partial f}{\partial \vec{u}}(\mathbf{p}) = \overrightarrow{\nabla f} \cdot \vec{u} = \|\overrightarrow{\nabla f}\| \cdot \|\vec{u}\| \cdot cos\,\pi/2 = \|\overrightarrow{\nabla f}\| \cdot 1 \cdot 0 = 0$$

Esquemáticamente:

$$\begin{aligned}
\overrightarrow{\nabla f} \nearrow\!\!\!\nearrow \vec{u} &\Rightarrow D_{\vec{u}}f(\mathbf{p}) = \|\overrightarrow{\nabla f}(\mathbf{p})\| \quad \text{Máxima} \\
\overrightarrow{\nabla f} \swarrow\!\!\!\swarrow \vec{u} &\Rightarrow D_{\vec{u}}f(\mathbf{p}) = -\|\overrightarrow{\nabla f}(\mathbf{p})\| \quad \text{mínima} \\
\overrightarrow{\nabla f} \perp \vec{u} &\Rightarrow D_{\vec{u}}f(\mathbf{p}) = 0
\end{aligned}$$

**Ejemplo 3.32.** *La temperatura en grados Celsius, sobre la superficie de una placa metálica, viene dada por $T(x,y) = 20 - 4x^2 - y^2$, midiéndose $x$ e $y$ en pulgadas. Desde el punto $(2,3)$ ¿En qué dirección crece la temperatura más rápidamente? ¿A qué razón se produce ese crecimiento?.*

*Solución*: La dirección de máximo crecimiento es la dirección del gradiente, y la razón de ese crecimiento su módulo, en consecuencia hallamos el gradiente y lo evaluamos en el punto $(2,3)$

$$\overrightarrow{\nabla f} = (T_x, T_y) = (-8x, -2y) \qquad \rightarrow \qquad \overrightarrow{\nabla f}(2,3) = (-16, 6)$$

luego la dirección de máximo crecimiento es la del vector $\vec{v} = \overrightarrow{\nabla f}(2,3) = (-16,6)$, y la razón de máximo crecimiento es el módulo de gradiente, luego,

$$D_{\vec{v}}T(2,3) = \|\overrightarrow{\nabla f}(2,3)\| = \|(-16,6)\| = \sqrt{16^2+6^2} = \sqrt{256+36} = \sqrt{292} = 17'09^{\circ}/\text{pul.}$$

### 3.5.3. Gradiente y curvas de nivel

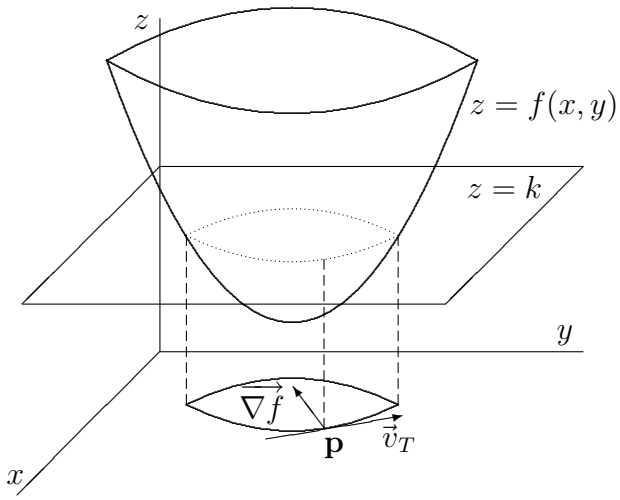



Figura 3.8: Gradiente-curva de nivel.

Las curvas de nivel de la superficie de ecuación $z = f(x,y)$ se obtienen dándole a $z$ un valor numérico concreto. En consecuencia, caminando sobre la curva de nivel, los valores de la función no cambian durante el recorrido. Es decir, al movernos por la superficie siguiendo una curva de nivel, los valores de la función se mantienen constante, y, por lo tanto, es de esperar que la derivada de la función en esa dirección sea cero. Es decir,

$$D_{\vec{u}_T}f(\mathbf{p}) = \overrightarrow{\nabla f} \cdot \vec{u}_T = 0$$

Esto significa que en cada punto de una curva de nivel el vector gradiente es perpendicular al vector tangente a la curva de nivel y, en consecuencia, *en cada punto de una superficie, el vector gradiente es perpendicular a la curva de nivel que pasa por ese punto.*

Otra manera de verlo es la siguiente: En cada punto $(x,y)$ de la curva de nivel se tiene $f(x,y) = k$, de donde, $df(x,y) = 0$, y por lo tanto, $f_x dx + f_y dy = 0$, de donde, expresándolo en forma vectorial resulta

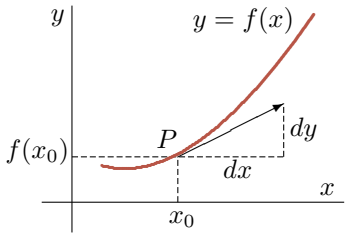
$$(f_x, f_y) \cdot (dx, dy) = 0 \quad \Rightarrow \quad \overrightarrow{\nabla f} \cdot \vec{v}_T = 0 \quad \Rightarrow \quad \overrightarrow{\nabla f} \perp \vec{v}_T$$

## 3.6. Plano tangente

### 3.6.1. Vectores normales

**Vector tangente**

(a) A una curva plana.




Como vector tangente podemos elegir cualquiera de los vectores siguientes, todos ellos paralelos entre sí:

$$\vec{v}_T = (dx, dy) \parallel (1, \frac{dy}{dx}) = \big(1, y'(x)\big) \parallel \big(x'(t), y'(t)\big)$$

(b) A una superficie. De manera análoga, a lo anterior

$$\vec{v}_T = (dx, dy, dz) \begin{cases} y = cte. \rightarrow \vec{v}_T = (dx, 0, dz) \parallel (1, 0, \dfrac{\partial z}{\partial x}) \\ x = cte. \rightarrow \vec{v}_T = (0, dy, dz) \parallel (0, 1, \dfrac{\partial z}{\partial y}) \end{cases}$$

**Vector normal a una curva plana**

Se llama vector *normal* a una curva, en un punto de la misma, al vector *perpendicular* a su recta tangente en dicho punto. Análogamente, se llama vector *normal* a una superficie, en un punto de la misma, al vector *perpendicular* a su plano tangente en dicho punto. Es evidente que si existe un vector normal entonces existen infinitos, ya que si $\vec{v}_p$ es un vector normal, entonces también lo es $\lambda\vec{v}_p$, con $\lambda \in \mathbb{R}$

**(a) Curvas dadas de forma explícitas** $y = f(x)$

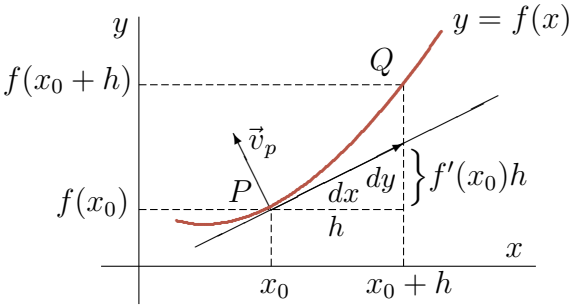



Figura 3.9: Vector normal a una curva.

Dada una función de una variable $y = f(x)$ y un punto $(x_0, y_0)$ de la misma, el vector tangente a su gráfica en dicho punto vendrá definido por $(dx, dy)$, o bien, $\big(h, f'(x_0)h\big)$, o más simplificado, $\vec{v}_T = \big(1, f'(x_0)\big)$, y en consecuencia, el vector normal a la curva en el punto $(x_0, y_0)$ vendrá definido por

$$\vec{v}_p = \big(-f'(x_0), 1\big)$$

Donde se ha tenido en cuenta que para calcular un vector perpendicular a otro conocido del plano basta con cambiar las coordenadas de orden y una de ellas de signo.

**(b) Curvas dadas en forma implícita** $F(x, y) = 0$

Dada una curva plana de ecuación $y = f(x)$, igualando a cero (o a una constante) la ecuación, podemos considerar la curva $y - f(x) = 0$ como una curva de nivel de una

función de dos variables $F(x,y) = 0$, siendo $F(x,y) = y - f(x)$, con lo cual el vector gradiente de esta función será un vector normal a la curva dada.

$$\vec{v}_p = \nabla F$$

**Ejemplo 3.33.** *Hallar el vector normal a la parábola de ecuación $y = x^2$ en el punto* $(1,1)$

*Solución.* Igualando la ecuación a cero, resulta

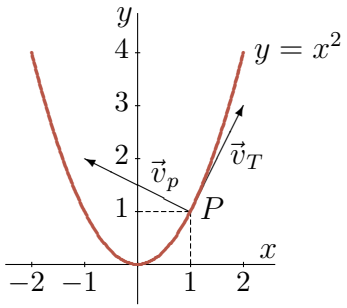



Figura 3.10: $y = x^2$

$$y = x^2 \quad \rightarrow \quad y - x^2 = 0 \quad \rightarrow \quad F(x,y) = y - x^2$$

de donde,

$$\left.\begin{array}{l} F_x(x,y) = -2x \\ F_x(x,y) = 1 \end{array}\right\} \nabla F(x,y) = (-2x, 1)$$

luego,

$$\vec{v}_p = \nabla F(1,1) = (-2,1)$$

**Vector normal a una superficie**

**(a) Superficies dadas de forma explícitas** $z = f(x,y)$

El vector normal $\vec{v}_p$ a una superficie $\mathcal{S}$ en un punto $P$ de la misma, será un vector perpendicular al plano tangente a la superficie en dicho punto. Ahora bien, el vector perpendicular al plano tangente en el punto $P$ será perpendicular a cualquier recta tangente a la superficie en dicho punto. Un vector, genérico, tangente a la superficie viene dado por las componentes $(dx, dy, dz)$. En particular, los vectores tangentes a la superficie, en el punto $P$, en las direcciones de los ejes $OX$ y $OY$ vienen dados respectivamente por:

$$\vec{v}_{Tx} = \left(1, 0, \frac{\partial f}{\partial x}(x_0, y_0)\right), \quad \vec{v}_{Ty} = \left(0, 1, \frac{\partial f}{\partial y}(x_0, y_0)\right)$$

El vector normal a la superficie ha de ser perpendicular a los dos vectores $\vec{v}_{Tx}$ y $\vec{v}_{Ty}$, luego se puede obtener a partir del producto vectorial de los mismos, así

$$\vec{v}_p = \vec{v}_{Tx} \times \vec{v}_{Ty} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & f_x(x_0,y_0) \\ 0 & 1 & f_y(x_0,y_0) \end{vmatrix} = \big(-f_x(x_0,y_0), -f_y(x_0,y_0), 1\big)$$

nota: Para que la existencia del plano tangente en el punto $P(x_0, y_0, z_0)$ esté garantizada, la función ha de ser diferenciable en el punto $\mathbf{p}(x_0, y_0)$

**(b) Superficies dadas de forma implícita** $F(x,y,z) = 0$

Dada una superficie de ecuación $z = f(x,y)$, igualando a cero (o a una constante) la ecuación, podemos considerar la superficie $z - f(x,y) = 0$ como una "superficie de nivel" de una función de tres variables $F(x,y,z) = 0$ siendo $F(x,y,z) = y - f(x,y)$, con lo cual el vector gradiente de esta función será un vector normal a la superficie dada.

$$\vec{v}_p = \nabla F$$

Es evidente que si la función viene definida de manera explícita $z = f(x, y)$, entonces ambos procedimientos coinciden. En efecto, haciendo $F(x, y, z) = z - f(x, y)$, resulta

$$\vec{v}_p = \nabla F = (F_x, F_y, F_z) = (-f_x, -f_y, 1)$$

**Ejemplo 3.34.** *Hallar un vector perpendicular al plano* $3x - y + 2z = 3$

*Solución.* Consideramos la función: $F(x, y, z) = 3x - y + 2z$, de donde

$$F_x = 3, \quad F_y = -1 \quad F_z = 2$$

luego,

$$\vec{v}_p = \nabla F = (3, -1, 2)$$

**Ejemplo 3.35.** *Hallar un vector perpendicular a la superficie* $x^2 + yz = 5$ *en el punto* $(1, 2, 2)$

*Solución.* Consideramos la función: $F(x, y, z) = x^2 + yz$, de donde

$$F_x = 2x, \quad F_y = z \quad F_z = y \quad \rightarrow \quad \nabla f = (2x, z, y)$$

luego,

$$\vec{v}_p = \nabla F(1, 2, 2) = (2, 2, 2)$$

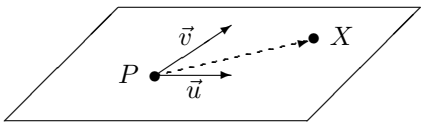
### 3.6.2. Plano tangente

**Consideraciones geométricas**

**Ecuación de un plano**

Para hallar la ecuación de una recta, de una curva, de un plano, o de cualquier lugar geométrico, el procedimiento habitual es el denominado del *"punto genérico"*, que consiste en suponer un punto genérico $X$ de la figura geométrica correspondiente y relacionarlo con los datos que disponemos. Buscaremos una relación que vincule al punto $X$ con los datos del problema y que cumplan los puntos de la figura geométrica y sólo ellos. La relación puede ser vectorial, trigonométrica o de cualquier tipo.

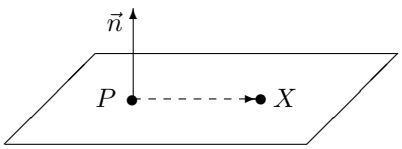
(a) Ecuación de un plano conocido un punto del mismo y dos vectores paralelos al plano, pero no paralelos entre sí.




Si unimos el punto $P$ con el punto $X$ mediante el vector $\overrightarrow{PX}$, resultará que los tres vectores $\vec{u}$, $\vec{v}$ y $\overrightarrow{PX}$ son coplanarios y por lo tanto serán linealmente dependientes y, en consecuencia, el determinante de la matriz de sus componentes será cero.

$$\left.\begin{array}{l}\overrightarrow{PX}\\ \vec{u}\\ \vec{v}\end{array}\right\}\text{Linealmente dependientes} \Longleftrightarrow \begin{vmatrix}\overrightarrow{PX}\\ \vec{u}\\ \vec{v}\end{vmatrix} = 0$$

(b) Ecuación de un plano conocido un punto del mismo y un vector normal al plano.




Si unimos el punto $P$ con el punto $X$ mediante el vector $\overrightarrow{PX}$, resultará que los vectores $\overrightarrow{PX}$ y $\vec{n}$ son perpendiculares y por lo tanto su producto escalar será cero.

$$\vec{n} \perp \overrightarrow{PX} \Longleftrightarrow \vec{n} \cdot \overrightarrow{PX} = 0$$

### Plano tangente

Desde el punto de vista geométrico se llama plano tangente a una superficie en un punto al plano que contiene a todas las rectas tangentes a la superficie en dicho punto, o mejor dicho, a las rectas tangentes de todas las curvas trazadas sobre la superficie que pasan por el punto. Si todas las tangentes no están sobre el mismo plano, entonces se dice que el plano tangente no existe.
Desde el punto de vista analítico para que exista el plano tangente a una superficie en un punto de la misma, la función que define la superficie ha de ser *diferenciable* en el punto correspondiente.

**(a) Superficies dadas de forma explícita** $z = f(x, y)$

El plano tangente ha de contener todas las rectas tangentes a la superficie en el punto correspondiente, luego, en particular, ha de contener las rectas tangentes en las direcciones de los ejes $OX$ y $OY$, por lo tanto, los vectores

$$\vec{v}_{Tx} = \left(1, 0, \frac{\partial f}{\partial x}(x_0, y_0)\right), \quad \vec{v}_{Ty} = \left(0, 1, \frac{\partial f}{\partial y}(x_0, y_0)\right)$$

serán paralelos al plano buscado.
En consecuencia el plano tangente buscado contiene al punto $\big(x_0, y_0; f(x_0, y_0)\big)$ y es paralelo a los vectores $\vec{v}_{Tx} = \big(1, 0, f_x(x_0, y_0)\big)$ y $\vec{v}_{Ty} = \big(0, 1, f_y(x_0, y_0)\big)$, por lo que su ecuación será:

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ 1 & 0 & f_x(x_0, y_0) \\ 0 & 1 & f_y(x_0, y_0) \end{vmatrix} = 0$$

de donde resulta $z - z_0 - f_x(x_0, y_0)(x - x_0) - f_y(x_0, y_0)(y - y_0) = 0$ que se puede expresar de la forma:

$$\boxed{z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)}$$

o bien, despejando $z$, y poniendo $z_0 = f(x_0, y_0)$ resulta

$$\boxed{z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)} \tag{3.4}$$

nota: Si la función no es diferenciable, entonces la ecuación anterior no representa el plano tangente, ya que en este caso el plano tangente no existe. Es decir, si la función no es diferenciable, pero tiene derivadas parciales, podemos construir la ecuación anterior, pero en este caso dicha ecuación no representa el plano tangente, sino simplemente un plano que pasa por el punto $(x_0, y_0)$.
Al mismo resultado llegamos sabiendo que el vector $\vec{v}_p = (-f_x, -f_y, 1)$ es perpendicular al plano buscado. En este caso será $\vec{v}_p \cdot \overrightarrow{\mathbf{px}} = 0$, y en consecuencia obtenemos el mismo resultado.

$$-f_x(x_0, y_0) \cdot (x - x_0) - f_y(x_0, y_0) \cdot (y - y_0) + z - z_0 = 0$$

de donde,

$$z - z_0 = f_x(x_0, y_0) \cdot (x - x_0) + f_y(x_0, y_0) \cdot (y - y_0)$$

que es el mismo resultado anterior.

**Ejemplo 3.36.** *Hallar la ecuación del plano tangente al paraboloide $z = x^2 + y^2$ en el punto $P(2, -1; 5)$*

*Solución.* Hallamos el gradiente en el punto $\mathbf{p}(2, -1)$

$$\left.\begin{array}{l} z_x = 2x \\ z_y = 2y \end{array}\right\} \left.\begin{array}{l} z_x(2,-1) = 4 \\ z_y(2,-1) = -2 \end{array}\right\} \nabla z(2,-1) = (4, -2)$$

de donde: $z - 5 = 4(x - 2) - 2(y + 1)$ o bien, simplificando, $4x - 2y - z = 5$

**(b) Superficies dadas de forma implícita** $F(x, y, z) = 0$

Dada una superficie de ecuación $z = f(x, y)$, igualando a cero (o a una constante) la ecuación, podemos considerar la curva $z - f(x, y) = 0$ como una "*superficie de nivel*" de una función de tres variables $F(x, y, z) = 0$, siendo $F(x, y, z) = y - f(x, y)$, con lo cual en cada punto $(x, y, z)$ el vector gradiente de esta función será un vector normal a la superficie dada.

$$\vec{v}_p = \nabla F$$

En consecuencia, el plano tangente a la superficie en el punto $P(x_0, y_0, z_0)$ será el plano que contiene a $P$ y tiene por vector normal a $\nabla F(x_0, y_0, z_0)$. Para hallar su ecuación, tomamos un punto genérico $X(x, y, z)$ del plano, y tendrá que ser $\nabla F(x_0, y_0, z_0) \perp \overrightarrow{PX}$, y en consecuencia su producto escalar ha de ser cero $\nabla F(x_0, y_0, z_0) \cdot \overrightarrow{PX} = 0$, luego su ecuación será:

$$\frac{\partial F}{\partial x}(x_0, y_0, z_0) \cdot (x - x_0) + \frac{\partial F}{\partial y}(x_0, y_0, z_0) \cdot (y - y_0) + \frac{\partial F}{\partial z}(x_0, y_0, z_0) \cdot (z - z_0) = 0 \quad (3.5)$$

**Recta normal**

Se llama recta normal a una superficie $\mathcal{S}$ en un punto $P(x_0, y_0, z_0)$ de la misma, a la recta que pasa por $P$ y tiene por vector director al vector normal a la superficie en dicho punto. Es decir, la recta perpendicular al plano tangente a la superficie en $P$. Teniendo en cuenta que el vector normal a la superficie en el punto $P(x_0, y_0, z_0)$ es $\nabla F(x_0, y_0, z_0)$, podemos concluir que la ecuación de la recta normal a la superficie en el punto $P$ viene definida por las ecuaciones:

$$\frac{x - x_0}{F_x(x_0, y_0, z_0)} = \frac{y - y_0}{F_y(x_0, y_0, z_0)} = \frac{z - z_0}{F_z(x_0, y_0, z_0)} \quad (3.6)$$

**Ejemplo 3.37.** *Hallar la ecuación del plano tangente y de la recta normal al paraboloide de ecuación $z^2 - 2x^2 - 2y^2 - 12 = 0$, en el punto $(1, -1, 4)$*

*Solución*: Consideramos la función $F(x, y, z) = z^2 - 2x^2 - 2y^2 - 12$ y hallamos su gradiente en el punto $(1, -1, 4)$

$$\left.\begin{array}{l} F_x = -4x \\ F_y = -4y \\ F_z = 2z \end{array}\right\} \left.\begin{array}{l} F_x(1,-1,4) = -4 \\ F_y(1,-1,4) = 4 \\ F_z(1,-1,4) = 8 \end{array}\right\} \nabla F(1,-1,4) = (-4, 4, 8) \parallel (-1, 1, 2)$$

Luego el plano tangente es: $-1(x-1)+1(y+1)+2(z-4)=0$ y simplificando resulta

$$x-y-2z+6=0$$

y la recta normal

$$\frac{x-1}{-1}=\frac{y+1}{1}=\frac{z-4}{2}$$

**Existencia del plano tangente**

Para construir la ecuación del plano tangente a una superficie en un punto lo único que necesitamos son las derivadas parciales de la función en dicho punto, sin embargo, no siempre dicha ecuación representa al plano tangente. Para que la ecuación (3.4) represente, realmente, el plano tangente a la superficie $z=f(x,y)$ en el punto $P(x_0,y_0,z_0)$ es necesario que la función $f$ sea diferenciable en el punto $\mathbf{p}(x_0,y_0)$. No debe olvidarse que las derivadas parciales pueden existir, incluso, en puntos donde la función no es ni siquiera continua, y no tiene sentido hablar de plano tangente en dichos puntos. El concepto geométrico de tangencia es el mismo que el concepto analítico de diferenciabilidad. Así, si la función $f:D\subseteq\mathbb{R}^2\to\mathbb{R}$ es diferenciable en $\mathbf{p}(x_0,y_0)$ entonces diremos que la ecuación (3.4) define al plano tangente a la superficie $z=f(x,y)$ en el punto $P\big(x_0,y_0;f(x_0,y_0)\big)$. Si la función no es diferenciable en el punto $\mathbf{p}(x_0,y_0)$, entonces el plano tangente a la superficie en el punto correspondiente no existe, con lo cual el plano obtenido con la ecuación (3.4) no representa al plano tangente en el sentido preciso que se entiende en matemáticas.

**Significado geométrico de la *tangencia***

Si la función $f$ es diferenciable en el punto $\mathbf{p}(x_0,y_0)$, será:

$$f(x_0+h,y_0+k)=f(x_0,y_0)+\frac{\partial f}{\partial x}(x_0,y_0)h+\frac{\partial f}{\partial x}(x_0,y_0)k+r(h,k)\quad\text{donde}\quad\lim_{(h,k)\to(0,0)}\frac{r(h,k)}{\|(h,k)\|}=0$$

con lo cual, poniendo $h=x-x_0$, $k=y-y_0$, la expresión anterior se convierte en:

$$f(x,y)=f(x_0,y_0)+\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0)+\frac{\partial f}{\partial x}(x_0,y_0)(y-y_0)+r(x-x_0,y-y_0)$$

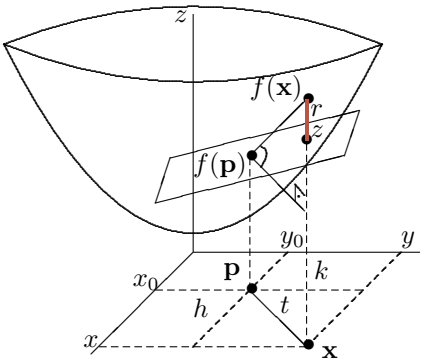


Figura 3.11: Plano tangente.

donde $$\lim_{(x,y)\to(x_0,y_0)}\frac{r(x-x_0,y-y_0)}{\|(x-x_0,y-y_0)\|}=0$$

Ahora bien, teniendo en cuenta la ecuación del plano tangente

$$z=f(x_0,y_0)+f_x(x_0,y_0)(x-x_0)+f_y(x_0,y_0)(y-y_0)$$

donde $z$ representa la altura en el punto $\mathbf{x}=(x,y)$ hasta el plano tangente.
Resulta que el residuo $r(x-x_0,y-y_0)$ se puede expresar como

$$r(x-x_0,y-y_0)=f(x,y)-z$$

Es decir, el residuo es la diferencia entre la $z$ de la función $z=f(x,y)$ en el punto $\mathbf{x}=(x,y)$ y la $z$ en el mismo punto del plano tangente a la gráfica en $\mathbf{p}=(x_0,y_0)$

El hecho de que $f$ sea diferenciable en $\mathbf{p}=(x_0,y_0)$ no sólo nos dice que $r$ es muy pequeño en torno al punto $\mathbf{p}$, sino, además, que es mucho más pequeño que la distancia de $\mathbf{p}$ a $\mathbf{x}$, es decir $r<<\|\overrightarrow{\mathbf{px}}\|$ . El hecho de que el residuo $r\to 0$ cuando $\mathbf{x}\to\mathbf{p}$ sólo nos dice que la función es continua en $\mathbf{p}$. Lo que realmente da el carácter de tangencia es el hecho de que $\dfrac{r}{\|\overrightarrow{\mathbf{px}}\|}\to 0$ cuando $\mathbf{x}\to\mathbf{p}$. Este hecho geométricamente viene a significar que la superficie se *aplana* en los alrededores del punto $\mathbf{p}$ tratando de confundirse, *por un instante* con el plano tangente.

### 3.6.3. Recta tangente y plano normal a una curva en el espacio

**(a) Curvas dadas en forma paramétrica**

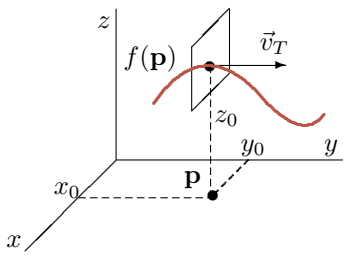



Figura 3.12: curva en el espacio

Las coordenadas de un punto genérico $X(x, y, z)$ de la curva vendrán dadas en función del parámetro $t$, resultando $X\big(x(t), y(t), z(t)\big)$. En consecuencia, el vector tangente en dicho punto genérico será $\vec{v}_T = \big(x'(t), y'(t), z'(t)\big)$. de donde resultan

(a) Recta tangente en el punto $P(x_0, y_0, z_0)$

$$\frac{x - x_0}{x'(t_0)} = \frac{y - y_0}{y'(t_0)} = \frac{z - z_0}{z'(t_0)}$$

(b) Plano normal en el punto $P(x_0, y_0, z_0)$

$$x'(t_0)(x - x_0) + y'(t_0)(y - y_0) + z'(t_0)(z - z_0) = 0$$

**Ejemplo 3.38.** *Hallar las ecuaciones de la recta tangente y del plano normal a la curva*

$$x = t \quad y = t^2 \quad z = t^3$$

*en el punto* $t = 1$

*Solución.* Hallamos las coordenadas del punto $P$ y del vector tangente $\vec{v}_T$ correspondientes al valor del parámetro $t = 1$

$$t = 1 \to P(1, 1, 1)$$

$$\left.\begin{array}{l} x'(t) = 1 \\ y'(t) = 2t \\ z'(t) = 3t^2 \end{array}\right\} \left.\begin{array}{l} x'(1) = 1 \\ y'(1) = 2 \\ z'(1) = 3 \end{array}\right\} \vec{v}_T = (1, 2, 3)$$

de donde,

(a) Recta tangente en el punto $P(1, 1, 1)$

$$\frac{x - 1}{1} = \frac{y - 1}{2} = \frac{z - 1}{3}$$

(b) Plano normal en el punto $P(1, 1, 1)$

$$1(x - 1) + 2(y - 1) + 3(z - 1) = 0$$

**Ejemplo 3.39.** *Hallar las ecuaciones de la recta tangente y del plano normal a la curva*

$$x = t - 2 \quad y = 3t^2 + 1 \quad z = 2t^3$$

*en el punto donde corta al plano* $yz$

*Solución.* En el punto donde la curva corta al plano $yz$ será $x = 0$, luego $t - 2 = 0$, de donde $t = 2$. En consecuencia.

$$t = 2 \to P(0, 13, 16)$$

$$\left.\begin{array}{l} x'(t) = 1 \\ y'(t) = 6t \\ z'(t) = 6t^2 \end{array}\right\} \left.\begin{array}{l} x'(2) = 1 \\ y'(2) = 12 \\ z'(2) = 24 \end{array}\right\} \vec{v}_T = (1, 12, 24)$$

de donde,

(a) Recta tangente en el punto $P(0, 13, 16)$

$$\frac{x}{1} = \frac{y-13}{12} = \frac{z-16}{24}$$

(b) Plano normal en el punto $P(0, 13, 16)$

$$x + 12(y-13) + 24(z-16) = 0 \quad \rightarrow \quad x + 12y + 24z - 540 = 0$$

**Curvas dadas como intersección de dos superficies**

Sean las superficies definidas por las ecuaciones $F(x, y, z) = 0$ y $G(x, y, z) = 0$.

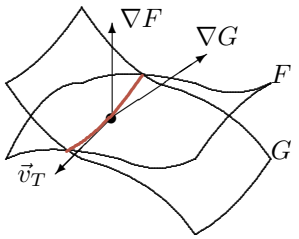



Figura 3.13: curva en el espacio

Dichas superficies se cortarán en la curva definida por el sistema de ecuaciones:

$$\left.\begin{array}{r} F(x,y,z) = 0 \\ G(x,y,z) = 0 \end{array}\right\}$$

Los vectores gradientes $\nabla F$ y $\nabla G$ en cualquier punto de la curva de corte serán normales a sus respectivas superficies y por tanto serán normales a la curva de corte, en consecuencia el vector perpendicular a ambos definido por su producto vectorial $\vec{v}_T = \nabla F \times \nabla G$ será tangente a la curva de corte, de donde, si sus componentes son $\vec{v}_T = (v_1, v_2, v_3)$ resultan

(a) Recta tangente en el punto $P(x_0, y_0, z_0)$ $\quad \dfrac{x-x_0}{v_1} = \dfrac{y-y_0}{v_2} = \dfrac{z-z_0}{v_3}$

(b) Plano normal en el punto $P(x_0, y_0, z_0)$ $\quad v_1(x-x_0) + v_2(y-y_0) + v_3(z-z_0) = 0$

**Ejemplo 3.40.** *Hallar la ecuación de la recta tangente y del plano normal a la curva definida por la intersección del elipsoide $x^2 + 4y^2 + 2z^2 = 27$ y el hiperboloide $x^2 + y^2 - 2z^2 = 11$ en el punto $P(3, -2, 1)$*

*Solución.* Hallamos los vectores gradientes en el punto $P(3, -2, 1)$

$$F(x,y,z) = x^2 + 4y^2 + 2z^2 \left\{\begin{array}{l} F_x = 2x \\ F_y = 8y \\ F_z = 4z \end{array}\right\} \left.\begin{array}{l} F_x = 6 \\ F_y = -16 \\ F_z = 4 \end{array}\right\} \nabla F(3,-2,1) = (6, -16, 4)$$

$$G(x,y,z) = x^2 + y^2 - 2z^2 \left\{\begin{array}{l} F_x = 2x \\ F_y = 2y \\ F_z = -4z \end{array}\right\} \left.\begin{array}{l} F_x = 6 \\ F_y = -4 \\ F_z = -4 \end{array}\right\} \nabla G(3,-2,1) = (6, -4, -4)$$

luego, el vector tangente será:

$$\vec{v}_T = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 6 & -16 & 4 \\ 6 & -4 & -4 \end{vmatrix} = (80, 48, 72) = 8(10, 6, 9)$$

y en consecuencia

(a) Recta tangente en el punto $P(3, -2, 1)$

$$\frac{x-3}{10} = \frac{y+2}{6} = \frac{z-1}{9}$$

(b) Plano normal en el punto $P(3, -2, 1)$

$$10(x-3) + 6(y+2) + 9(z-1) = 0$$

### 3.6.4. La diferencial como aproximación del incremento

Sea $f : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable en el conjunto abierto $\mathcal{D}$ de $\mathbb{R}^2$. Entonces, para cada $\mathbf{x} \in \mathcal{D}$ tendremos

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + \frac{\partial f}{\partial x}h_1 + \frac{\partial f}{\partial x}h_2 + r(\mathbf{h}) \quad \text{donde} \quad \lim_{\mathbf{h}\to\mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

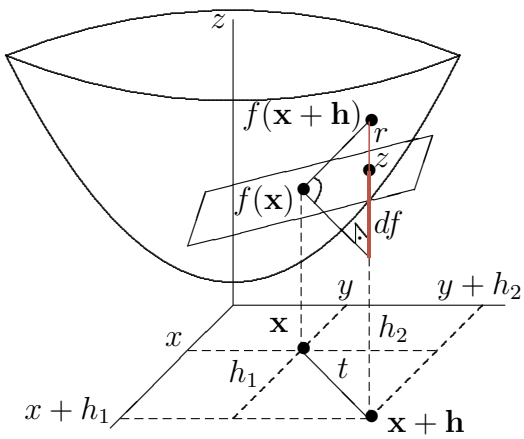


Figura 3.14: $\Delta f \approx df$

Vimos en la definición (3.3) que a la parte lineal en $h_1$ y $h_2$ de esta expresión, $\lambda(h_1, h_2) = f_x h_1 + f_y h_2$, se le llama diferencial de la función $f$ en el punto $\mathbf{x} = (x, y)$ y se denota por $df(x, y)$, o simplemente por $df$. Es decir,

$$df = \frac{\partial f}{\partial x}h_1 + \frac{\partial f}{\partial y}h_2$$

Y teniendo en cuenta que si $f(x, y) = x$, se tiene $h_1 = dx$ y si $f(x, y) = y$, se tiene $h_2 = dy$, podemos escribir:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

Con lo cual resulta

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + df(\mathbf{x}) + r(\mathbf{h})$$

Ahora bien, dado que $\lim_{\mathbf{h}\to 0} r(\mathbf{h}) = 0$ se tiene que, para $\|h\|$ pequeño, será $r(\mathbf{h}) \approx 0$ y en consecuencia

$$f(\mathbf{x}+\mathbf{h}) \approx f(\mathbf{x}) + df(\mathbf{x})$$

Es decir,

$$f(\mathbf{x}+\mathbf{h}) \approx z$$

Es decir, la $z$ de la función $z = f(\mathbf{x}+\mathbf{h})$ en el punto $\mathbf{x}+\mathbf{h} = (x+h_1, y+h_2)$, coincide, de manera aproximada, con la $z$, en el mismo punto, del plano tangente a la gráfica en un punto cercano $\mathbf{x} = (x, y)$
Lo anterior se puede expresar de la forma

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) \approx df(\mathbf{x})$$

Es decir,

$$\Delta f \approx df$$

Geométricamente se pueden dar las siguientes interpretaciones:

(a) El valor aproximado de una función $f(\mathbf{x})$, en un punto $\mathbf{x}$, se puede obtener sustituyendo el valor de $\mathbf{x}$ en la ecuación del plano tangente, en un punto cercano $\mathbf{x}_0$. Es decir, para hallar el valor aproximado de una función en un punto, calculamos la ecuación del plano tangente, en un punto cercano, y sustituimos las coordenadas del punto sobre la ecuación de dicho plano.

(b) Al pasar del punto $\mathbf{x}$ al punto cercano $\mathbf{x}+\mathbf{h}$, el incremento que sufre la función

$$\Delta f = f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x})$$

coincide, de manera aproximada, con el diferencial $df$. Es decir, *la diferencial de una función es una buena aproximación del incremento*

**Cálculo de valores aproximados**

Supongamos que queremos calcular, aunque sea de manera aproximada, el valor de una función en un punto $(x, y)$, pero no sabemos calcular el valor de la función en dicho punto, $f(x, y)$, o dicho cálculo resulta extremadamente complicado, y sin embargo, supongamos que sabemos calcular el valor de la función y el de sus derivadas parciales en un punto cercano $(x_0, y_0)$. Pues bien, podemos utilizar el valor de la función y el de su derivadas parciales en este punto $(x_0, y_0)$ para calcular el valor aproximado de la función en el punto desconocido $(x, y)$.

Para hacer cálculos aproximados de operaciones podemos utilizar cualquiera de las dos opciones:

(a) Para hallar el valor aproximado de una función en un punto, calculamos la ecuación del plano tangente, en un punto cercano, y sustituimos las coordenadas del punto sobre la ecuación de dicho plano.

$$f(x, y) \approx f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_x(x_0, y_0)(y - y_0)$$

(b) Aproximar el incremento mediante la diferencial

$$f(x, y) \approx f(x_0, y_0) + df(x_0, y_0)$$

Hay que advertir que estas aproximaciones sólo serán validas para valores muy cercanos al punto conocido y para funciones diferenciables.

**Ejemplo 3.41.** *Comparar el incremento con el diferencial del área de un rectángulo.*

*Solución.* Incrementemos los lados de un rectángulo $x$ e $y$, en $dx$ y $dy$, respectivamente. El área del rectángulo viene definida por $a = x \cdot y$, de donde,

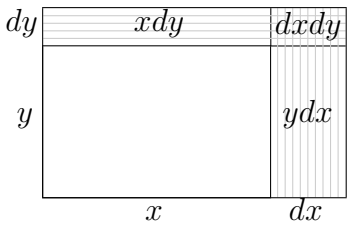



Figura 3.15: $\Delta a \approx da$

El incremento del área será: $\Delta a = ydx + xdy + dxdy$

Y el diferencial

$$\left.\begin{array}{l} \dfrac{\partial a}{\partial x} = y \\ \dfrac{\partial a}{\partial y} = x \end{array}\right\} da = ydx + xdy$$

En consecuencia, $\Delta a = da + dxdy$

**Ejemplo 3.42.** *Dada la superficie $z = xe^{xy-2}$ se pide:*

*(a) Hallar la ecuación del plano tangente a la superficie en el punto $P(2, 1, 2)$.*

(b) *Usar el plano tangente para obtener una aproximación del valor de la función en el punto* $\mathbf{q}(1'9, 1'02)$.

*Solución.* (a) El plano tangente viene definido por la ecuación:

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

para lo cual

$$\left.\begin{array}{l} f_x = e^{xy-2} + xye^{xy-2} \\ f_y = x^2 e^{xy-2} \end{array}\right\} \begin{array}{l} f_x(2,1) = e^0 + 2 \cdot 1 \cdot e^0 = 1 + 2 = 3 \\ f_y(2,1) = 4e^0 = 4 \end{array}$$

de donde, la ecuación del plano tangente es

$$z - 2 = 3(x - 2) + 4(y - 1)$$

que se simplifica en $z = 3x + 4y - 8$
(b) El valor aproximado de la función se calcula sustituyendo los valores de $x$ e $y$ en la ecuación del plano tangente, así,

$$z(1'9, 1,02) = 1'9e^{1'9 \cdot 1'02-2} \approx 3 \cdot 1'9 + 4 \cdot 1'02 - 8 = 5'7 + 4'08 - 8 = 9'78 - 8 = 1'78$$

(los cálculos suelen resultan más fáciles si se sustituye directamente sobre la ecuación del plano tangente, antes de eliminar los paréntesis, despejando previamente $z$, en efecto $z = 3 + 3(-0'1) + 4(0'02) = 3 - 0'3 + 0'08 = 1'78$)

**Ejemplo 3.43.** *Estimar* $\sqrt{2'01^2 + 1'98^2 + 1'05^2}$

*Solución.* Necesitamos una función y un punto próximo sobre el que sepamos hacer los cálculos con facilidad. En este caso, la función es $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ y el punto $\mathbf{p}(2, 2, 1)$. en consecuencia,

$$f(2, 2, 1) = \sqrt{2^2 + 2^2 + 1^1} = \sqrt{9} = 3$$

$$\left.\begin{array}{l} \dfrac{\partial f}{\partial x} = \dfrac{x}{\sqrt{x^2 + y^2 + z^2}} \\ \dfrac{\partial f}{\partial y} = \dfrac{y}{\sqrt{x^2 + y^2 + z^2}} \\ \dfrac{\partial f}{\partial z} = \dfrac{z}{\sqrt{x^2 + y^2 + z^2}} \end{array}\right\} \begin{array}{l} \dfrac{\partial f}{\partial x}(2,2,1) = \dfrac{2}{\sqrt{9}} = \dfrac{2}{3} \\ \dfrac{\partial f}{\partial y}(2,2,1) = \dfrac{2}{\sqrt{9}} = \dfrac{2}{3} \\ \dfrac{\partial f}{\partial z}(2,2,1) = \dfrac{1}{\sqrt{9}} = \dfrac{1}{3} \end{array}$$

de donde, sustituyendo en la ecuación del plano tangente:

$$z = f(x_0, y_0, z_0) + f_x(x_0, y_0, z_0)(x - x_0) + f_y(x_0, y_0, z_0)(y - y_0) + f_z(x_0, y_0, z_0)(z - z_0)$$

resulta,

$$\sqrt{2'01^2 + 1'98^2 + 1'05^2} \approx 3 + \frac{2}{3}(0'01) + \frac{2}{3}(-0'02) + \frac{1}{3}(0'05) = 3'01$$

## 3.7. Funciones vectoriales y matriz Jacobiana

### 3.7.1. Funciones vectoriales de variable vectorial

**Definición 3.10.** *Se llama función vectorial de variable vectorial a las funciones del tipo:*

$$\begin{gathered}\mathbf{f} : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^m \\ (x_1, x_2, \cdots, x_n) \mapsto (y_1, y_2, \cdots, y_m)\end{gathered}$$

Es decir,

$$\mathbf{f}(x_1, x_2, \cdots, x_n) = (y_1, y_2, \cdots, y_m) \text{ donde } \left\{ \begin{array}{l} y_1 = f_1(x_1, x_2, \cdots, x_n) \\ y_2 = f_2(x_1, x_2, \cdots, x_n) \\ \cdots \\ y_m = f_m(x_1, x_2, \cdots, x_n) \end{array} \right.$$

Tenemos pues que una función $\mathbf{f} : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es una regla que asocia a cada n-ada ordenada de números reales $(x_1, x_2, \cdots, x_n)$ de $\mathcal{D}$, una m-ada ordenada de números reales bien determinada $(y_1, y_2, \cdots, y_m)$ de $\mathbb{R}^m$. O bien, en términos vectoriales, una regla que asocia a cada vector $\mathbf{x}$ de $\mathcal{D}$, un vector bien determinado $\mathbf{y}$ de $\mathbb{R}^m$.
En otras palabras, una función vectorial es un vector cuyas componentes son funciones escalares.

$$\mathbf{f} = (f_1, f_2, \cdots, f_m)$$

Es decir,

$$\begin{gathered}\mathbf{f} : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^m \\ (x_1, x_2, \cdots, x_n) \mapsto \big(f_1(x_1, x_2, \cdots, x_n), f_2(x_1, x_2, \cdots, x_n), \cdots, f_m(x_1, x_2, \cdots, x_n)\big)\end{gathered}$$

Por ejemplo, las curvas paramétricas pueden considerarse como funciones vectoriales de $\mathbb{R}$ en $\mathbb{R}^2$. Así, si las ecuaciones paramétricas de una curva vienen definidas por:

$$\left. \begin{array}{l} x(t) = 2t + 3 \\ y(t) = t - 5 \end{array} \right\}$$

las podemos interpretar como una función vectorial de la siguiente forma:

$$\left. \begin{array}{c} \mathbb{R} \xrightarrow{\mathbf{f}} \mathbb{R}^2 \\ t \mapsto (2t + 3, t - 5) \end{array} \right\} \mathbf{f}(t) = (2t + 3, t - 5) = \big(x(t), y(t)\big)$$

Por ejemplo, la función vectorial $\mathbf{f}(x, y) = (x + 2, x - y, y + 4)$ viene definida por:

$$\left. \begin{array}{c} \mathbb{R}^2 \xrightarrow{\mathbf{f}} \mathbb{R}^3 \\ (x, y) \mapsto (x + 2, x - y, y + 4) \end{array} \right\} \mathbf{f}(x, y) = (x + 2, x - y, y + 4) = \big(f_1(x, y), f_2(x, y), f_3(x, y)\big)$$

**Campos vectoriales.** Se llaman campos vectoriales a las funciones vectoriales de $\mathbb{R}^n$ en sí mismo. Es decir, se llama campo vectorial (en $\mathbb{R}^n$) a las funciones del tipo $\mathbf{f} : \mathcal{D} \subseteqq \mathbb{R}^n \to \mathbb{R}^n$. Una práctica para visualizar los campos vectoriales consiste en jugar co la naturaleza "dual"de los elementos de $\mathbb{R}^n$; los elementos del dominio $\mathbf{x} \in \mathcal{D}$ se ven como *puntos* y los de la imagen $\mathbf{f}(\mathbf{x})$ como *vectores*. El vector $\mathbf{f}(\mathbf{x})$ se representa mediante una flecha que se puede colocar en cualquier punto del espacio, sin embargo, para tener una visualización del campo, la flecha se colocará de manera que su punto inicial sea el punto $\mathbf{x}$.

### 3.7.2. Continuidad de las funciones vectoriales

Los conceptos de límite, continuidad y diferenciabilidad pueden extenderse a funciones vectoriales. Estos conceptos siempre tienen dos visiones; una general que es más bien teórica y consiste en la traslación de los conceptos unidimensionales a niveles vectoriales, y otra por componentes que es más bien práctica y que consiste en trasladar los conceptos a las funciones coordenadas.

## Límite de una función vectorial

**Definición 3.11.** *Sean* $\mathbf{f} : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ *y* $\mathbf{x}_0 \in \mathcal{D}$ *(siendo* $\mathcal{D}$ *un abierto de* $\mathbb{R}^n$*); Se dice que* $\mathbf{f}$ *tiene por límite* $\ell$*, cuando* $\mathbf{x}$ *tiende a* $\mathbf{x}_0$*, si* $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} \|\mathbf{f}(\mathbf{x}) - \ell\| = 0$

Se escribe entonces $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \ell$.

Merece destacar que el límite vectorial se ha definido en términos numéricos. En efecto, para cada $\mathbf{x}$ se tiene que $\|\mathbf{f}(\mathbf{x}) - \ell\|$ es un número, más preciso, la aplicación $\mathbf{x} \to \|\mathbf{f}(\mathbf{x}) - \ell\|$ es una función numérica de $\mathcal{D}$ en $\mathbb{R}^+$, y es ya familiar el límite para estas funciones.

**Bola**. Cabe interpretar la definición de límite por medio de la noción de *bola*.

**Definición 3.12.** *En* $\mathbb{R}^n$*, se llama bola abierta, de centro* $\mathbf{x}_0 \in \mathbb{R}^n$ *y de radio* $r \in \mathbb{R}^*_+$*, al conjunto de puntos* $\mathbf{x}$ *de* $\mathbb{R}^n$ *tales que* $\|\mathbf{x} - \mathbf{x}_0\| < r$.

Se designa por $B(\mathbf{x}_0, r)$, o simplemente por $B$, cuando no haya temor de confusión:

$$B(\mathbf{x}_0, r) = \{\mathbf{x} \in \mathbb{R}^n / \ \|\mathbf{x} - \mathbf{x}_0\| < r\}$$

Así, por ejemplo, en la recta real $\mathbb{R}$, la bola se llama *entorno* y está formado por los puntos del intervalo $(\mathbf{x}_0 - r, \mathbf{x}_0 + r)$; en el plano euclídeo $\mathbb{R}^2$, la bola se denomina *disco* y está formada por los puntos interiores al círculo de centro $\mathbf{x}_0$ y radio $r$; en el espacio de tres dimensiones $\mathbb{R}^3$ la bola $B(\mathbf{x}_0, r)$ está formada por los puntos interiores a la esfera de centro $\mathbf{x}_0$ y radio $r$.

La noción de límite de una función puede interpretarse entonces de la siguiente manera:

**Definición 3.13.** *Se dice que* $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \ell$ *si, para toda bola* $B'$ *de centro* $\ell$ *y radio* $\epsilon > 0$*, existe una bola* $B$ *de centro* $\mathbf{x}_0$ *y radio* $r$*, incluido en* $\mathcal{D}$*, tal que* $F(B) \subsetneqq B'$.

**Coordenadas y límite**. Teniendo en cuenta que una función vectorial $\mathbf{f} : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es un vector cuyas coordenadas son funciones escalares $\mathbf{f} = (f_1, f_2, \cdots, f_m)$.
Es decir,

$$\mathbf{f}(x_1, x_2, \cdots, x_n) = \big(f_1(x_1, x_2, \cdots, x_n), f_2(x_1, x_2, \cdots, x_n), \cdots, f_m(x_1, x_2, \cdots, x_n)\big)$$

Se tiene que:

1. Si $\mathbf{f}$, con $\mathbf{f} = (f_1, f_2, \cdots, f_m)$, tiene por límite $\ell$, con $\ell = (\ell_1, \ell_2, \cdots, \ell_m)$, entonces cada una de las funciones coordenadas de $\mathbf{f}$ tienen por límite, respectivamente, $\ell_1, \ell_2, \cdots, \ell_m$.
2. Si las coordenadas de $\mathbf{f}$ tienen por límite, respectivamente, $\ell_1, \ell_2, \cdots, \ell_m$, entonces $\mathbf{f}$ tiene por límite $\ell = (\ell_1, \ell_2, \cdots, \ell_m)$.

En consecuencia, se puede enunciar el siguiente teorema:

**Teorema 3.5.** *Para que* $\mathbf{f}$ *tienda a un límite* $\ell$*, es necesario y suficiente que las coordenadas de* $\mathbf{f}$ *tengan por límites respectivamente las coordenadas de* $\ell$.

## Continuidad

**Definición 3.14.** *Sean* $\mathbf{f} : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ *y* $\mathbf{x}_0 \in \mathcal{D}$ *(siendo* $\mathcal{D}$ *un abierto de* $\mathbb{R}^n$*); Se dice que* $\mathbf{f}$ *es continua en* $\mathbf{x}_0$*, si* $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0)$.

En otros términos, $\mathbf{f}$ *es continua en* $\mathbf{x}_0$ *si* $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\| = 0$, o bien, $\mathbf{f}$ es continua en $\mathbf{x}_0$ si, para todo $\epsilon > 0$, existe $\delta > 0$ tal que:

$$\|\mathbf{x} - \mathbf{x}_0\| < \delta \quad \Longrightarrow \quad \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\| < \epsilon.$$

Sirviéndonos de la noción geométrica de *bola* podemos dar la siguiente definición: $\mathbf{f}$ *es continua en* $\mathbf{x}_0$ *si, para toda bola* $\mathcal{B}$ *de centro* $\mathbf{f}(\mathbf{x}_0)$ *existe una bola* $\mathcal{B}'$ *de centro en* $\mathbf{x}_0$*, incluida en* $\mathcal{D}$*, tal que* $\mathbf{f}(\mathcal{B}') \subseteqq \mathcal{B}$.
Se dice que $\mathbf{f}$ es continua sobre un dominio $\mathcal{D}$, si $\mathbf{f}$ es continua en todo punto $\mathbf{x}_0$ de $\mathcal{D}$.

**Coordenadas de una función continua**. El concepto de continuidad se puede establecer en términos de las funciones coordenadas. Así,

**Teorema 3.6.** *Para que* $\mathbf{f}$ *sea continua, en un dominio* $\mathcal{D}$*, es necesario y suficiente que todas las coordenadas de* $\mathbf{f}$ *lo sean.*

## Composición de funciones continuas

Aunque la composición de funciones se tratará más detalladamente en el epígrafe siguiente dedicado a la Regla de la cadena, la vemos aquí con objeto de demostrar que *la composición de dos funciones continuas es a su vez otra función continua.*

**Definición 3.15 (Composición de funciones).** *Dadas la función* $\mathbf{g} : D_g \subseteq \mathbb{R}^n \to \mathbb{R}^p$ *definida en el conjunto* $D_g$ *de* $\mathbb{R}^n$ *y la función* $\mathbf{f} : D_f \subseteq \mathbb{R}^m \to \mathbb{R}^n$ *definida en el conjunto* $D_f$ *de* $\mathbb{R}^m$*, cuyo rango está contenido en* $D_g$ *(es decir,* $\mathbf{f}(D_f) \subseteq D_g$*), entonces se puede formar la composición de ambas funciones* $\mathbf{g} \circ \mathbf{f} : D_f \subseteq \mathbb{R}^m \to \mathbb{R}^p$*, definida como:*

$$(\mathbf{g} \circ \mathbf{f})(\mathbf{u}) = \mathbf{g}\big(\mathbf{f}(\mathbf{u})\big), \quad \mathbf{u} \in D_f$$

Esquemáticamente sería.

$$\left.\begin{array}{l} \mathbf{f} : D_f \subseteq \mathbb{R}^m \to \mathbb{R}^n \\ \mathbf{g} : D_g \subseteq \mathbb{R}^n \to \mathbb{R}^p \end{array}\right\} \quad \left.\begin{array}{c} D_f \xrightarrow{\mathbf{f}} D_g \xrightarrow{\mathbf{g}} \mathbb{R}^p \\ \mathbf{g} \circ \mathbf{f} : D_f \subseteq \mathbb{R}^m \to \mathbb{R}^p \end{array}\right\} (\mathbf{g} \circ \mathbf{f})(\mathbf{u}) = \mathbf{g}\big(\mathbf{f}(\mathbf{u})\big)$$

y mediante diagramas, teniendo en consideración los dominios,

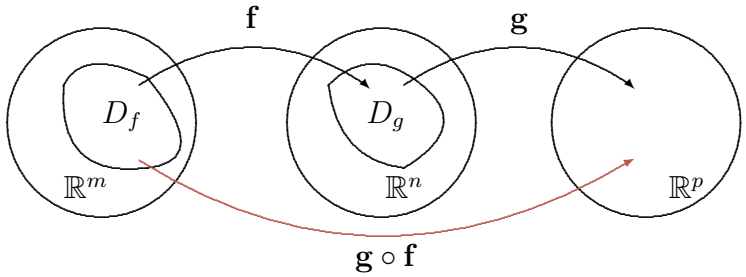



Figura 3.16: Composición de funciones

Igual que en una variable, en general, la composición de funciones no cumple la propiedad conmutativa, y sí la asociativa.

**Teorema 3.7 (Composición de funciones continuas).** *Dadas la función* $g : D_g \subseteq \mathbb{R}^n \to \mathbb{R}^p$ *definida en el conjunto* $D_g$ *de* $\mathbb{R}^n$ *y la función* $\mathbf{f} : D_f \subseteq \mathbb{R}^m \to \mathbb{R}^n$ *definida en el conjunto* $D_f$ *de* $\mathbb{R}^m$*, cuyo rango está contenido en* $D_g$ *(es decir,* $\mathbf{f}(D_f) \subseteq D_g$*), Si la función* $\mathbf{f}$ *es continua en el punto* $\mathbf{x}_0$ *y la función* $\mathbf{g}$ *es continua en el punto* $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$ *entonces la función compuesta* $\mathbf{g} \circ \mathbf{f}$ *también es continua en* $\mathbf{x}_0$.

*Demostración.* Sea $\mathbf{x}_0 \in \mathcal{D}_f$ e $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0) \in \mathcal{D}_g$. Puesto que $\mathbf{g}$ es continua en $\mathbf{y}_0$, para toda bola $\mathcal{B}$ de centro $\mathbf{g}(\mathbf{y}_0)$ existe una bola $\mathcal{B}'$, de centro $\mathbf{y}_0$, tal que $\mathbf{g}(\mathcal{B}') \subseteqq \mathcal{B}$.
Puesto que $\mathbf{f}$ es continua en $\mathbf{x}_0$, para toda bola $\mathcal{B}'$ de centro $\mathbf{f}(\mathbf{x}_0) = \mathbf{y}_0$, existe una bola $\mathcal{B}''$ de centro $\mathbf{x}_0$ tal que $\mathbf{f}(\mathcal{B}'') \subseteqq \mathcal{B}'$.
Aplicando $\mathbf{g}$ a los dos miembros de esta relación; se obtiene:

$$\mathbf{g} \circ \mathbf{f}(\mathcal{B}'') \subseteqq \mathbf{g}(\mathcal{B}') \subseteqq \mathbf{B}$$

y, por consiguiente, $\mathbf{g} \circ \mathbf{f}$ es continua en el punto $\mathbf{x}_0$. $\square$

### 3.7.3. Derivadas parciales de funciones vectoriales

Aunque a nivel teórico las derivadas parciales de funciones vectoriales pueden definirse en términos vectoriales, en la práctica, se establecen en términos de las funciones componentes. Teniendo en cuenta que una función vectorial $\mathbf{f} : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es un vector cuyas coordenadas son funciones escalares $\mathbf{f} = (f_1, f_2, \cdots, f_m)$. Es decir,

$$\mathbf{f}(x_1, x_2, \cdots, x_n) = \big(f_1(x_1, x_2, \cdots, x_n), f_2(x_1, x_2, \cdots, x_n), \cdots, f_m(x_1, x_2, \cdots, x_n)\big)$$

Se tiene que:

$$\frac{\partial \mathbf{f}}{\partial x_1}(x_1, x_2, \cdots, x_n) = \big(\frac{\partial f_1}{\partial x_1}(x_1, x_2, \cdots, x_n), \frac{\partial f_2}{\partial x_1}(x_1, x_2, \cdots, x_n), \cdots, \frac{\partial f_m}{\partial x_1}(x_1, x_2, \cdots, x_n)\big)$$
$$\cdots$$
$$\frac{\partial \mathbf{f}}{\partial x_n}(x_1, x_2, \cdots, x_n) = \big(\frac{\partial f_1}{\partial x_n}(x_1, x_2, \cdots, x_n), \frac{\partial f_2}{\partial x_n}(x_1, x_2, \cdots, x_n), \cdots, \frac{\partial f_m}{\partial x_n}(x_1, x_2, \cdots, x_n)\big)$$

**Ejemplo 3.44.** *Hallar las derivadas parciales de la función* $\mathbf{f}(x, y, z) = (xe^y \cos z, xe^y \operatorname{sen} z, \tan z)$

*Solución.* Derivando componente a componente, se tiene,

$$\begin{aligned}
\mathbf{f}_x &= (e^y \cos z, e^y \operatorname{sen} z, 0)\\
\mathbf{f}_y &= (xe^y \cos z, xe^y \operatorname{sen} z, 0)\\
\mathbf{f}_z &= (-xe^y \operatorname{sen} z, xe^y \cos z, 1 + \tan^2 z)
\end{aligned}$$

### 3.7.4. Funciones vectoriales diferenciables

En la definición 3.4 de la página 24 se estableció, para funciones numéricas de una sola variable, que: *Una función $f : D \in \mathbb{R} \to \mathbb{R}$ es diferenciable en $x_0 \in D$ si existe una constante $A$ tal que*

$$f(x_0 + h) = f(x_0) + Ah + r(h) \quad \textit{con} \quad \lim_{h \to 0} \frac{r(h)}{h} = 0$$

En tal caso, a la expresión lineal $\lambda(h) = Ah$ se le llama diferencial de $f$, y se tiene:

$$df(x_0) = \lambda(h) = Ah = f'(x_0)\, dx$$

En la definición 3.7 de la página 31 se estableció, para funciones numéricas de $n$ variables, que: *Una función $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en el punto $\mathbf{x}_0 = (x_1, \cdots, x_n) \in \mathcal{D}$, si existen $n$ constantes $A_i, \quad i = 1, 2, \cdots, n$ tales que*

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} A_i h_i + r(\mathbf{h}) \quad \textit{con} \quad \lim_{\mathbf{h} \to \mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

En tal caso, a la expresión lineal $\lambda(h_1, \cdots, h_n) = \sum_{i=1}^{n} A_i h_i$ se le llama diferencial de $f$, y se tiene:

$$df(x_0) = \lambda(\mathbf{h}) = \sum_{i=1}^{n} A_i h_i = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0) dx_i = \frac{\partial f}{\partial x_1}(\mathbf{x}_0) dx_1 + \frac{\partial f}{\partial x_2}(\mathbf{x}_0) dx_2 + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{x}_0) dx_n$$

En cada uno de los casos, la diferencial de una función en un punto aparece como una *aplicación lineal* que se aproxima a la función $f$ en ese punto. Se trata de generalizar esta importante noción a las funciones de varias variables, ya tengan imágenes reales o vectoriales.
La generalización se va a hacer por la vía de definir la diferencial como una aplicación lineal. Recuérdese que una aplicación $\lambda(h, k)$ es lineal si cumple la propiedad $\lambda(Ah, Bk) = A\lambda(h, k) + B\lambda(h, k)$, siendo $A$ y $B$ números. Recuérdese también la relación de las aplicaciones lineales con las matrices. Así, se establece la siguiente definición:

**Definición 3.16 (Funciones diferenciables).** *Una función* $\mathbf{f} : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ *se dice diferenciable en el punto* $\mathbf{x}_0 = (x_1, \cdots, x_n) \in \mathcal{D}$*, si existen una aplicación lineal* $\lambda(\mathbf{h})$ *tal que*

$$\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) = \mathbf{f}(\mathbf{x}_0) + \lambda(\mathbf{h}) + r(\mathbf{h}) \quad \textit{con} \quad \lim_{\mathbf{h}\to\mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = \mathbf{0}$$

En tal caso, a la expresión lineal $\lambda(\mathbf{h})$ se le llama diferencial de $\mathbf{f}$, y se designa por $\mathbf{df}(\mathbf{x}_0)$
**nota:** Es evidente que esta definición generaliza las dos anteriores, ya que para $f : D \subseteqq \mathbb{R} \to \mathbb{R}$ tenemos la primera definición, y para $f : \mathcal{D} \subseteqq \mathbb{R}^n \to \mathbb{R}$ la segunda.
Con objeto de eludir el *cociente* en el límite que aparece en la definición, se puede modificar la definición de función diferenciable en los siguientes términos. *Una función* $\mathbf{f} : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ *se dice diferenciable en el punto* $\mathbf{x}_0 = (x_1, \cdots, x_n) \in \mathcal{D}$*, si existen una aplicación lineal* $\lambda(\mathbf{h})$ *tal que*

$$\boxed{\begin{array}{ll} \mathbf{f}(\mathbf{x}_0 + \mathbf{h}) = \mathbf{f}(\mathbf{x}_0) + \lambda(\mathbf{h}) + \|\mathbf{h}\|\epsilon(\mathbf{h}) & (3.7) \\ \textit{con:} & \\ \lim_{\mathbf{h}\to\mathbf{0}} \epsilon(\mathbf{h}) = \mathbf{0} & (3.8) \end{array}}$$

**nota:** Muchos autores prefieren dar la definición de función diferenciable unificando los dos apartados 3.7 y 3.8 en un solo límite. Así, se tiene:

**Definición 3.17.** *Sea* $\mathbf{f} : \mathcal{D} \subseteqq \mathbb{R}^n \to \mathbb{R}^m$ *y sea* $\mathbf{x}_0$ *un punto interior de* $\mathcal{D}$*. Se dice que* $\mathbf{f}$ *es diferenciable en* $\mathbf{x}_0$ *si existe una aplicación lineal* $\lambda : \mathbb{R}^n \to \mathbb{R}^m$*, dependiente, en general, de* $\mathbf{x}_0$ *tal que*

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{x}_0) - \lambda(\mathbf{h})}{\|\mathbf{h}\|} = \mathbf{0}$$

**Proposición 3.6.** *Si la función* $\mathbf{f} : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ *es diferenciable en el punto* $\mathbf{x}_0 = (x_1, \cdots, x_n) \in \mathcal{D}$ *entonces existen sus derivadas parciales en dicho punto.*

*Demostración.* Supongamos, para simplificar la escritura, que $n = 3$. Supongamos, por tanto, que la función $\mathbf{f} : \mathcal{D} \subseteq \mathbb{R}^3 \to \mathbb{R}^m$ es diferenciable en el punto $\mathbf{x}_0 = (x_0, y_0, z_0) \in \mathcal{D}$, entonces existen una aplicación lineal $\lambda(\mathbf{h})$ tal que, para todo $\mathbf{h}$ tal que $\mathbf{x}_0 + \mathbf{h} \in \mathcal{D}$, se tiene:

$$\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{x}_0) = \lambda(\mathbf{h}) + \|\mathbf{h}\|\epsilon(\mathbf{h}) \quad \text{con} \quad \lim_{\mathbf{h}\to\mathbf{0}} \epsilon(\mathbf{h}) = \mathbf{0}$$

Designemos por $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ la base canónica de $\mathbb{R}^3$. Elijamos en primer lugar $\mathbf{h}$ del modo siguiente:

$$\mathbf{h} = (t, 0, 0) = t\mathbf{e}_1 \quad (t \in \mathbb{R}).$$

Puesto que la diferencial es lineal, se tiene $\lambda(\mathbf{h}) = \lambda(t\mathbf{e}_1) = t\lambda(\mathbf{e}_1)$, en consecuencia la relación (3.7) se escribe:

$$\mathbf{f}(x_0 + t, y_0, z_0) - \mathbf{f}(x_0, y_0, z_0) = t\lambda(\mathbf{e}_1) + |t|\,\|\mathbf{e}_1\|\epsilon(t\mathbf{e}_1)$$

y, como $\lim_{t\to 0} \epsilon(t\mathbf{e}_1) = \mathbf{0}$, se obtiene:

$$\lim_{t\to 0} \frac{\mathbf{f}(x_0 + t, y_0, z_0) - \mathbf{f}(x_0, y_0, z_0)}{t} = \lambda(\mathbf{e}_1).$$

Resulta así que la función $\mathbf{f}$ admite, en $(x_0, y_0, z_0)$, derivada parcial respecto de $x$, que vale:

$$\lambda(\mathbf{e}_1) = \frac{\partial \mathbf{f}}{\partial x}(x_0, y_0, z_0).$$

Análogamente, tomando $\mathbf{h}(0, t, 0)$ y después $\mathbf{h}(0, 0, t)$ se probaría que $\mathbf{f}$ admite, en el punto $(x_0, y_0, z_0)$, derivadas parciales con relación a $y$ y $z$, respectivamente iguales a:

$$\lambda(\mathbf{e}_2) = \frac{\partial \mathbf{f}}{\partial y}(x_0, y_0, z_0) \quad \text{y} \quad \lambda(\mathbf{e}_3) = \frac{\partial \mathbf{f}}{\partial z}(x_0, y_0, z_0). \qquad \square$$

**La diferencial.** Tomando $\mathbf{h} = (dx, dy, dz)$, y teniendo en cuenta que la diferencial es lineal, resulta:

$$\lambda(\mathbf{h}) = \lambda(\mathbf{e}_1)dx + \lambda(\mathbf{e}_2)dy + \lambda(\mathbf{e}_3)dz.$$

Es decir,

$$\mathbf{df}(\mathbf{x}_0) = \lambda(\mathbf{h}) = \frac{\partial \mathbf{f}}{\partial x}dx + \frac{\partial \mathbf{f}}{\partial y}dy + \frac{\partial \mathbf{f}}{\partial z}dz$$

Y para el caso de $n$ variables, resulta:

$$\mathbf{df}(\mathbf{x}_0) = \lambda(\mathbf{h}) = \frac{\partial \mathbf{f}}{\partial x_1}dx_1 + \frac{\partial \mathbf{f}}{\partial x_2}dx_2 + \cdots + \frac{\partial \mathbf{f}}{\partial x_n}dx_n$$

**Coordenadas y diferenciabilidad de las funciones vectoriales.** Los conceptos de continuidad y diferenciabilidad de funciones vectoriales se establecen en términos de las funciones componentes. Es decir, se dice que la función $f = (f_1, f_2, \cdots, f_m)$ es continua (respectivamente, diferenciable) en el punto $\mathbf{x}_0 \in D$, si y sólo si todas y cada una de las funciones componentes $f_i : D \subseteq \mathbb{R}^n \to \mathbb{R}, \quad i = 1, 2, \cdots, m$, lo son.
Ahora bien, teniendo en cuenta que las coordenadas de la función$\mathbf{f}$ son $\mathbf{f} = (f_1, f_2, \cdots, f_m)$, será:

$$\frac{\partial \mathbf{f}}{\partial x_i} = \left(\frac{\partial f_1}{\partial x_i}, \frac{\partial f_2}{\partial x_i}, \cdots, \frac{\partial f_m}{\partial x_i}\right) \quad i = 1, 2, \cdots, n$$

de donde,

$$\begin{aligned}\mathbf{df}(\mathbf{x}_0) = \frac{\partial \mathbf{f}}{\partial x_1}dx_1 + \cdots + \frac{\partial \mathbf{f}}{\partial x_n}dx_n &= \left(\frac{\partial f_1}{\partial x_1}, \frac{\partial f_2}{\partial x_1}, \cdots, \frac{\partial f_m}{\partial x_1}\right)dx_1 + \cdots + \left(\frac{\partial f_1}{\partial x_n}, \frac{\partial f_2}{\partial x_n}, \cdots, \frac{\partial f_m}{\partial x_n}\right)dx_n = \\ &= \left(\frac{\partial f_1}{\partial x_1}dx_1 + \cdots + \frac{\partial f_1}{\partial x_n}dx_n, \cdots, \frac{\partial f_m}{\partial x_1}dx_1 + \cdots + \frac{\partial f_m}{\partial x_n}dx_n\right) = (df_1, df_2, \cdots, df_n)\end{aligned}$$

Luego el diferencial de una función vectorial se puede calcular por componentes:

$$\mathbf{df}(\mathbf{x}_0) = (df_1, df_2, \cdots, df_n) \tag{3.9}$$

**Ejemplo 3.45.** *Estudiar la continuidad y diferenciabilidad de la función $f : \mathbb{R}^2 \to \mathbb{R}^2$, definida por*

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

*Solución.* (a) La función es continua en todo $\mathbb{R}^2$, pues las funciones componentes

$$f_1,\ f_2 : \mathbb{R}^2 \to \mathbb{R}, \quad f_1(x, y) = x + y, \quad f_2(x, y) = |y| \quad \text{lo son.}$$

(b) La función no es diferenciable en el origen, pues la función $f_2(x, y) = |y|$ no lo es.

## El jacobiano

Si la función $\mathbf{f} : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en el punto $\mathbf{x}_0 = (x_1, \cdots, x_n) \in \mathcal{D}$ entonces, su diferencial será:

$$\mathbf{df}(\mathbf{x}_0) = \lambda(\mathbf{h}) = \frac{\partial \mathbf{f}}{\partial x_1}dx_1 + \frac{\partial \mathbf{f}}{\partial x_2}dx_2 + \cdots + \frac{\partial \mathbf{f}}{\partial x_n}dx_n$$

Que se puede expresar de manera matricial, de la forma:

$$\mathbf{df}(\mathbf{x}_0) = \left(\frac{\partial \mathbf{f}}{\partial x_1}, \frac{\partial \mathbf{f}}{\partial x_2}, \cdots, \frac{\partial \mathbf{f}}{\partial x_n}\right)\begin{pmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_n \end{pmatrix} = J\mathbf{f}(\mathbf{x}_0)\mathbf{dx}$$

A la matriz $J\mathbf{f}(\mathbf{x}_0)$ se le llama matriz Jacobiana de la función $\mathbf{f}$ en el punto $\mathbf{x}_0$.
Ahora bien, teniendo en cuenta que la función $\mathbf{f} = (f_1, f_2, \cdots, f_m)$, será:

$$\frac{\partial \mathbf{f}}{\partial x_i} = \left(\frac{\partial f_1}{\partial x_i}, \frac{\partial f_2}{\partial x_i}, \cdots, \frac{\partial f_m}{\partial x_i}\right) \quad i = 1, 2, \cdots, n$$

Con lo cual resulta que la matriz jacobiana es,

$$J\mathbf{f} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ & & \cdots & \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \nabla f_1 \\ \nabla f_2 \\ \vdots \\ \nabla f_m \end{pmatrix}$$

**Observaciones:**

(a) Para $m = 1$ se tiene $Jf(\mathbf{x}_0) = \nabla f(\mathbf{x}_0)$.
(b) Para $n = m$, el determinante de la matriz $Jf(\mathbf{x}_0)$ se llama *jacobiano*
(c) Si comparamos la expresión de la diferencial de una función de una variable, de una función de varias variables y de una función vectorial tenemos:
   1.- Para funciones de una variable tenemos: $df = f'(x) \cdot h$
   2.- Para funciones de varias variables tenemos: $df = \nabla f \cdot \mathbf{h}$
   3.- Para funciones vectoriales: $d\mathbf{f} = J\mathbf{f} \cdot \mathbf{h}$

   donde se ha usado, respectivamente, el producto numérico, el producto escalar y el producto matricial.
(d) Si se comparan las tres expresiones, se observa que el gradiente, $\nabla f$, puede pensarse como la generalización del concepto de derivada para funciones de varias variables, y la matriz jacobiana, $J\mathbf{f}$, para funciones vectoriales. Si bien, hay que advertir que mientras que la derivada de una función de una variable en un punto es un *número*, la derivada de una función de varias variables es un *vector*, y la derivada de una función vectorial una matriz.

**Ejemplo 3.46.** *Hallar el diferencial de la función vectorial* $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ *definida por:*

$$f(x, y) = (x^2 + 3y^2, 5x^3 + 2y^6)$$

*Solución.* El Jacobiano de la función viene definido por:

$$J\mathbf{f} = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x & 6y \\ 15x^2 & 12y^5 \end{pmatrix}$$

de donde,

$$\mathbf{df} = \begin{pmatrix} 2x & 6y \\ 15x^2 & 12y^5 \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} = \left(2xdx + 6ydx, 15x^2dx + 12y^5dy\right)$$

Al mismo resultado hubiéramos llegado operando por componentes, en efecto,

$$\mathbf{df} = (df_1, df_2) = \left(2xdx + 6ydx, 15x^2dx + 12y^5dy\right)$$

# 3.8. Regla de la cadena

## 3.8.1. Funciones compuestas, inversas e implícitas de una variable

**Composición de funciones**. Componer dos funciones consiste en aplicar la segunda función al resultado de la primera. Analíticamente también significa *sustituir una función en la otra*. Es decir, si tenemos la

función $y = f(x)$ que establece la dependencia entre $y$ y $x$, y la función $x = g(t)$, que establece la dependencia entre $x$ y $t$, podemos sustituir esta última en la primera y obtener $y = f\big(g(t)\big)$. A la función así obtenida (que envía $t$ a $y$) se le llama composición de $f$ con $g$ y se denota por $f \circ g$. Obsérvese que el orden en que se escribe la composición $f \circ g$ es el inverso al orden en que actúan las funciones (primero $g$, después $f$). Conviene tener siempre presente esta doble visualización de la composición de funciones: Como aplicación sucesiva de dos funciones, y como sustitución de la variable por una función de otra variable. En esquema sería lo siguiente:
(a) Como aplicación sucesiva de funciones:

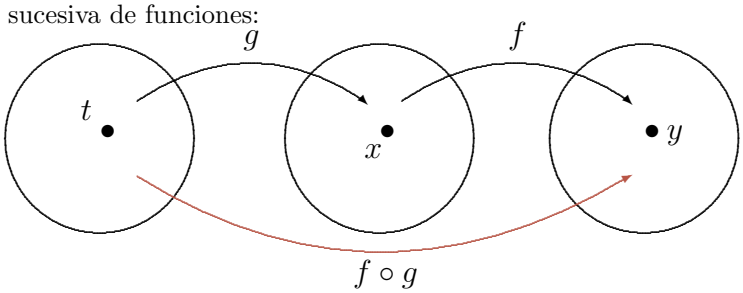



Figura 3.17: Composición de funciones

(b) Como sustitución de la variable:

$$\left.\begin{array}{l} y = f(x) \\ x = g(t) \end{array}\right\} y = f\big(g(t)\big)$$

Es evidente que para poder componer dos funciones el *rango* de la primera tiene que estar contenido en el *dominio* de la segunda $g(\mathcal{D}_g) \subseteq \mathcal{D}_f$, en caso contrario, después de aplicar $g$ no podríamos aplicar $f$. Hay que advertir que, en general, la composición de funciones no es conmutativa, es decir, en la generalidad de los casos será $f \circ g \neq g \circ f$, incluso, puede suceder que esté definida la composición en un orden y no en el otro. Sin embargo, sí se cumple la propiedad asociativa $f \circ (g \circ h) = (f \circ g) \circ h$. Desde el punto de vista formal la composición puede enunciarse de la siguiente forma: *Dadas las funciones $g : I \subseteq \mathbb{R} \to \mathbb{R}$, $f : J \subseteq \mathbb{R} \to \mathbb{R}$ (tales que $g(I) \subseteq J$), se llama composición de $f$ con $g$ y se denota $f \circ g : I \subseteq \mathbb{R} \to \mathbb{R}$, a la función $(f \circ g)(x) = f\big(g(x)\big)$.*

$$\left.\begin{array}{l} g : I \subseteq \mathbb{R} \to \mathbb{R} \\ f : J \subseteq \mathbb{R} \to \mathbb{R} \end{array}\right\} \quad \left.\begin{array}{c} I \xrightarrow{g} J \xrightarrow{f} \mathbb{R} \\ f \circ g : I \subseteq \mathbb{R} \to \mathbb{R} \end{array}\right\} (f \circ g)(x) = f\big(g(x)\big)$$

**Regla de la cadena**. La regla de la cadena nos dice que "*la composición de dos funciones diferenciables es diferenciable y su derivada es el producto de las derivadas de cada una de las funciones que se están componiendo*". Es decir, la derivada de la composición de dos funciones es el producto de las derivadas de dichas funciones, cada una de ellas aplicada en el punto correspondiente. Formalmente la regla de la cadena se puede enunciar de la siguiente forma: *Si $g$ es diferenciable en un punto $x_0 \in I$ y $f$ es diferenciable en $g(x_0) \in J$, entonces la composición $f \circ g$ es diferenciable en $x_0$, y su derivada es*

$$(f \circ g)'(x) = f'\big(g(x)\big)g'(x)$$

que se puede expresar con la notación de Leibnitz, haciendo $u = g(x)$ e $y = f(u)$

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$$

**Función inversa**. La función recíproca o inversa asigna a las imágenes los correspondientes originales (consiste en invertir las flechas). Sin embargo, la recíproca de una función no siempre es otra función, basta que dos elementos distintos tengan la misma imagen para que la recíproca no sea función, ya que asignaría a un elemento dos imágenes distintas, lo que no está permitido para las funciones, aunque sí para las correspondencias en general. Por lo tanto, para que la recíproca de una función sea otra función, dicha función ha de ser inyectiva.

$$I \underset{f^{-1}}{\overset{f}{\rightleftarrows}} J \qquad y = f(x) \quad \leftrightarrow \quad x = f^{-1}(y)$$

Desde el punto de vista analítico para obtener la inversa bastará con despejar la variable independiente y expresarla en función de la variable dependiente.
Formalmente podemos decir que *si la función $f : I \subseteq \mathbb{R} \to \mathbb{R}$ es inyectiva, con rango $J \subseteq \mathbb{R}$, entonces existe la función $f^{-1} : J \subseteq \mathbb{R} \to \mathbb{R}$, llamada "inversa"de f, tal que*

$$f^{-1}\big(f(x)\big) = x, \quad x \in I \qquad \text{y} \qquad f\big(f^{-1}(y)\big) = y, \quad y \in J$$

**Derivada de la función inversa** La derivada de la función inversa es la inversa de la derivada de la función. Formalmente se puede enunciar diciendo que *si la función $f$ es diferenciable y tal que $f'(x_0) \neq 0$, $x_0 \in I$, entonces existe la función inversa $f^{-1}$, al menos en un entorno de $f(x_0)$, dicha función también es diferenciable, y su derivada es:*

$$\big(f^{-1}\big)'\big(f(x)\big) = \frac{1}{f'(x)}$$

o con la notación de Leibnitz

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}} \quad \text{o bien,} \quad y'_x = \frac{1}{x'_y}$$

**Funciones implícitas** Si consideramos una expresión del tipo $F(x, y) = 0$ nos preguntamos ¿en que condiciones es posible despejar $y$ en términos de $x$ y establecer una función del tipo $y = f(x)$?. Evidentemente no siempre es posible obtener dicha función. Por ejemplo de $x^2 + y^2 = 1$ no se obtiene una función $y = f(x)$. Cuando de $F(x, y) = 0$ se puede despejar la variable $y$, y escribir $y = f(x)$, se dice que esta última función está dada implícitamente en $F(x, y) = 0$.

### 3.8.2. Composición de funciones vectoriales

Para funciones de varias variables la situación es, en principio, algo más complicada. Téngase en cuenta que dos campos escalares $f,\ g : D \subseteq \mathbb{R}^n \to \mathbb{R}$ no se pueden componer, pues la operación de composición entre funciones solamente se puede realizar cuando el rango de una de ellas está contenido en el dominio de la otra, y los campos escalares tienen rango en $\mathbb{R}$ y dominio en $\mathbb{R}^n$. Por lo tanto, los campos escalares solamente se podrán componer con funciones de una sola variable, por la derecha; o con funciones vectoriales, por la izquierda. Es decir,

$$\mathbb{R}^n \xrightarrow{f} \mathbb{R} \xrightarrow{g} \mathbb{R} \qquad \mathbb{R}^m \xrightarrow{\mathbf{g}} \mathbb{R}^n \xrightarrow{f} \mathbb{R}$$

Pensando en la sustitución de las variables, para componer la función $z = f(x, y)$ tendremos que sustituir las dos variables $x$ e $y$, respectivamente, por dos funciones $g_1$ y $g_2$ que las conecten con otras variables, donde $g_1$ y $g_2$ pueden ser funciones de una o varias variables (ambas de las mismas). Así, si consideramos las funciones

$$x = g_1(u, v), \quad y = g_2(u, v)$$

podemos sustituir en la función $z = f(x, y)$ y obtener la función compuesta:

$$z = f\big(g_1(u, v), g_2(u, v)\big)$$

que en esquema sería:

$$z = f(x, y) \left\{ \begin{array}{c} x = g_1(u, v) \\ y = g_2(u, v) \end{array} \right\} z = f\big(g_1(u, v), g_2(u, v)\big)$$

Ahora bien, la pareja de funciones $x = g_1(u, v)$, $y = g_2(u, v)$ puede considerarse como las componentes de *una sola* función *vectorial* $\mathbf{g} : \mathcal{D} \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, de tal manera que a cada punto $(u, v) \in \mathcal{D}$ la función $\mathbf{g}$ le asocia el punto $\mathbf{g}(u, v) \in \mathbb{R}^2$, cuyas coordenadas son $(x, y) = \big(g_1(u, v), g_2(u, v)\big)$. O sea, $\mathbf{g}(u, v) = \big(g_1(u, v), g_2(u, v)\big)$. Y esto permite interpretar la sustitución de las variables como la aplicación sucesiva

de dos funciones.
En esquema sería:

$$\begin{array}{ccccc} \mathbb{R}^2 & \xrightarrow{\mathbf{g}} & \mathbb{R}^2 & \xrightarrow{f} & \mathbb{R} \\ (u,v) & \mapsto & (x,y) & \mapsto & z \end{array}$$

de donde,

$$f \circ \mathbf{g}(u,v) = f\big(\mathbf{g}(u,v)\big) = f\big(g_1(u,v), g_2(u,v)\big)$$

**Ejemplo 3.47.** *Hallar la función compuesta de la función $f(x,y) = xy^2 + xy$ con las funciones*

$$x = g_1(u,v) = u + v \quad ey = g_2(u,v) = uv$$

*Solución.* Si queremos interpretar la composición como la aplicación sucesiva de dos funciones, consideramos la función vectorial

$$\mathbf{g}(u,v) = \big(g_1(u,v), g_2(u,v)\big) = (u+v, uv)$$

luego, en esquema, resulta

$$\left.\begin{array}{l} f(x,y) = xy^2 + xy \\ \mathbf{g}(u,v) = (u+v, uv) \end{array}\right\} \begin{array}{l} \mathbb{R}^2 \xrightarrow{f} \mathbb{R} \\ \mathbb{R}^2 \xrightarrow{\mathbf{g}} \mathbb{R}^2 \end{array} \Bigg\} \begin{array}{ccccc} \mathbb{R}^2 & \xrightarrow{\mathbf{g}} & \mathbb{R}^2 & \xrightarrow{f} & \mathbb{R} \\ (u,v) & \mapsto & (x,y) & \mapsto & z \end{array}$$

de donde, la composición buscada, será:

$$h(u,v) = (f \circ \mathbf{g})(u,v) = f\big(\mathbf{g}(u,v)\big) = f\big(g_1(u,v), g_2(u,v)\big) = f(u+v, uv) =$$
$$= (u+v)(uv)^2 + (u+v)uv = u^3v^2 + u^2v^3 + u^2v + uv^2$$

Nótese que la función resultante de la composición es una función distinta de $f$, $\mathbf{g}$, $g_1$ y $g_2$, es decir, es una nueva función $h$, tal que $f(x,y) = h(u,v) \neq f(u,v)$
En la práctica se pueden simplificar los cálculos, sustituyendo directamente:

$$f(x,y) = f(u+v, uv) = (u+v)(uv)^2 + (u+v)uv = u^3v^2 + u^2v^3 + u^2v + uv^2 = h(u,v)$$

**Ejemplo 3.48.** *Hallar la función compuesta de la función $f(x,y,z) = xy^2z^3 - xyz$ con las funciones $x = g_1(t) = e^t$, $y = g_2(t) = \operatorname{sen} t$ y $z = g_3(t) = t^2$*

*Solución.* Consideramos la función vectorial

$$\mathbf{g}(t) = \big(g_1(t), g_2(t), g_3(t)\big) = (e^t, \operatorname{sen} t, t^2)$$

luego, en esquema, resulta

$$\left.\begin{array}{l} f(x,y,z) = xy^2z^3 - xyz \\ \mathbf{g}(t) = (e^t, \operatorname{sen} t, t^2) \end{array}\right\} \begin{array}{l} \mathbb{R}^3 \xrightarrow{f} \mathbb{R} \\ \mathbb{R} \xrightarrow{\mathbf{g}} \mathbb{R}^3 \end{array} \Bigg\} \begin{array}{ccccc} \mathbb{R} & \xrightarrow{\mathbf{g}} & \mathbb{R}^3 & \xrightarrow{f} & \mathbb{R} \\ t & \mapsto & (x,y,z) & \mapsto & w \end{array}$$

de donde, la composición buscada, será:

$$h(t) = (f \circ \mathbf{g})(t) = f\big(\mathbf{g}(t)\big) = f\big(g_1(t), g_2(t), g_3(t)\big) = f(e^t, \operatorname{sen} t, t^2) =$$
$$= e^t \operatorname{sen}^2 t\,(t^2)^3 - e^t \operatorname{sen} t\, t^2 = t^6 e^t \operatorname{sen}^2 t - t^2 e^t \operatorname{sen} t$$

En la práctica se pueden simplificar los cálculos, sustituyendo directamente:

$$f(x,y,z) = f(e^t, \operatorname{sen} t, t^2) = e^t \operatorname{sen}^2 t\,(t^2)^3 - e^t \operatorname{sen} t\, t^2 = t^6 e^t \operatorname{sen}^2 t - t^2 e^t \operatorname{sen} t = h(t)$$

**Ejemplo 3.49.** *Hallar la función compuesta de la función $f(x,y) = xy^2 - xy$ con las funciones $g(t) = \sqrt{t}$*

*Solución.* Este caso es distinto de los dos anteriores, ya que aquí, para poder componer, la función $f$ tiene que actuar primero, y después la $g$. En efecto, en esquema, resulta

$$\left.\begin{array}{l} f(x,y) = xy^2 - xy \\ g(t) = \sqrt{t} \end{array}\right\} \left.\begin{array}{l} \mathbb{R}^2 \xrightarrow{f} \mathbb{R} \\ \mathbb{R} \xrightarrow{g} \mathbb{R} \end{array}\right\} \begin{array}{l} \mathbb{R}^2 \xrightarrow{f} \mathbb{R} \xrightarrow{g} \mathbb{R} \\ (x,y) \mapsto t \mapsto z \end{array}$$

de donde, la composición buscada, será:

$$h(t) = (g \circ f)(x,y) = g\big(f(x,y)\big) = \sqrt{f(x,y)} = \sqrt{xy^2 - xy}$$

Nótese que la función resultante de la composición solamente estará definida para aquellos puntos $(x,y)$ en los cuales se cumpla que $xy^2 - xy \geq 0$, es decir,

$$D_h = \{(x,y) \in \mathbb{R}^2 / \ xy^2 - xy \geq 0\}$$

En la práctica, también se pueden simplificar los cálculos sustituyendo directamente, pero en este caso, sustituímos en la función $g$:

$$g(t) = g\big(f(x,y)\big) = \sqrt{f(x,y)} = \sqrt{xy^2 - xy} = h(x,y)$$

**Caso general.** En general, para componer la función $z = f(x_1, x_2, \cdots, x_n)$ tendremos que sustituir las $n$ variables $x_1, x_2, \cdots, x_n$, respectivamente, por las funciones $g_1, g_2, \cdots g_n$ que las conecten con otras variables $u_1, u_2, \cdots, u_m$, donde $g_1, g_2, \cdots g_n$ pueden ser funciones de una o varias variables (todas de las mismas). Así, si consideramos las $n$ funciones

$$\begin{array}{l} x_1 = g_1(u_1, u_2, \cdots, u_m) \\ x_2 = g_2(u_1, u_2, \cdots, u_m) \\ \quad \vdots \\ x_n = g_n(u_1, u_2, \cdots, u_m) \end{array}$$

podemos sustituirlas en la función $z = f(x_1, x_2, \cdots, x_n)$ y obtener la función compuesta:

$$z = f\big(g_1(u_1, u_2, \cdots, u_m), g_2(u_1, u_2, \cdots, u_m), \cdots, g_n(u_1, u_2, \cdots, u_m)\big)$$

Ahora bien, las $n$ funciones

$$x_1 = g_1(u_1, u_2, \cdots, u_m), \ x_2 = g_2(u_1, u_2, \cdots, u_m), \ \cdots, \ x_n = g_n(u_1, u_2, \cdots, u_m)$$

pueden considerarse como las componentes de una sola función *vectorial* $\mathbf{g} : \mathcal{D} \subseteq \mathbb{R}^m \to \mathbb{R}^n$, de tal manera que a cada punto $(u_1, u_2, \cdots, u_m) \in \mathcal{D} \subseteq \mathbb{R}^m$ la función $\mathbf{g}$ le asocia el punto $\mathbf{g}(u_1, u_2, \cdots, u_m) = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$, cuyas coordenadas son

$$(x_1, x_2, \cdots, x_n) = \big(g_1(u_1, u_2, \cdots, u_m), g_2(u_1, u_2, \cdots, u_m), \cdots, g_n(u_1, u_2, \cdots, u_m)\big)$$

O sea,

$$g(u_1, u_2, \cdots, u_m) = \big(g_1(u_1, u_2, \cdots, u_m), g_2(u_1, u_2, \cdots, u_m), \cdots, g_n(u_1, u_2, \cdots, u_m)\big)$$

que en esquema sería:

$$\begin{array}{ccccc} \mathbb{R}^m & \xrightarrow{\mathbf{g}} & \mathbb{R}^n & \xrightarrow{f} & \mathbb{R} \\ (u_1, u_2, \cdots, u_m) & \mapsto & (x_1, x_2, \cdots, x_n) & \mapsto & z \end{array}$$

De esta manera se ve el proceso de composición de funciones de varias variables como un proceso entre dos funciones, que se puede enunciar formalmente de la siguiente forma:

**Definición 3.18 (Composición de funciones).** *Dadas la función $f : D_f \subseteq \mathbb{R}^n \to \mathbb{R}$ definida en el conjunto $D_f$ de $\mathbb{R}^n$ y la función $\mathbf{g} : D_g \subseteq \mathbb{R}^m \to \mathbb{R}^n$ definida en el conjunto $D_g$ de $\mathbb{R}^m$, cuyo rango está contenido en $D_f$ (es decir, $\mathbf{g}(D_g) \subseteq D_f$), entonces se puede formar la composición de ambas funciones $f \circ \mathbf{g} : D_g \subseteq \mathbb{R}^m \to \mathbb{R}$, definida como:*

$$(f \circ \mathbf{g})(\mathbf{u}) = f\big(\mathbf{g}(\mathbf{u})\big), \quad \mathbf{u} \in D_g$$

Esquemáticamente sería.

$$\left.\begin{array}{l} \mathbf{g} : D_g \subseteq \mathbb{R}^m \to \mathbb{R}^n \\ f : D_f \subseteq \mathbb{R}^n \to \mathbb{R} \end{array}\right\} \quad \left.\begin{array}{c} D_g \xrightarrow{\mathbf{g}} D_f \xrightarrow{f} \mathbb{R} \\ f \circ \mathbf{g} : D_g \subseteq \mathbb{R}^m \to \mathbb{R} \end{array}\right\} (f \circ \mathbf{g})(\mathbf{u}) = f\big(\mathbf{g}(\mathbf{u})\big)$$

y mediante diagramas,

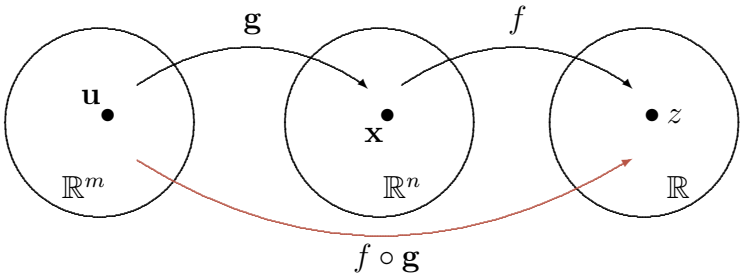



Figura 3.18: Composición de funciones

O bien, teniendo en consideración los dominios,

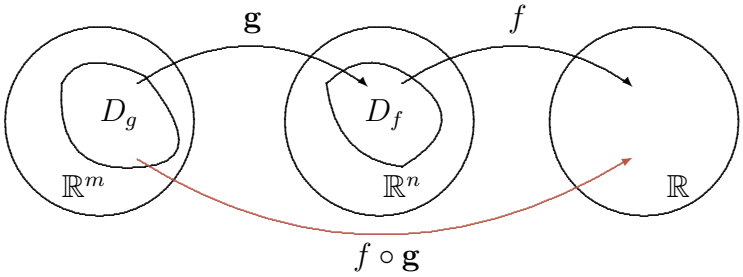



Figura 3.19: Composición de funciones

Igual que en una variable, en general, la composición de funciones no cumple la propiedad conmutativa, y sí la asociativa.

### 3.8.3. Regla de la cadena. Perspectiva teórica: Diferencial

**Teorema 3.8 (Regla de la cadena).** *Sea $\mathbf{g} : D_g \subseteq \mathbb{R}^m \to \mathbb{R}^n$ una función definida en el conjunto abierto $D_g$ de $\mathbb{R}^m$, diferenciable en $\mathbf{x}_0 \in D_g$. Sea $\mathbf{f} : D_f \subseteq \mathbb{R}^n \to \mathbb{R}^p$ una función definida en el conjunto abierto $D_f$ de $\mathbb{R}^n$, tal que $\mathbf{g}(D_g) \subseteq D_f$, diferenciable en $\mathbf{y}_0 = \mathbf{g}(\mathbf{x}_0) \in D_f$. Entonces la composición $\mathbf{f} \circ \mathbf{g} : D_g \subseteq \mathbb{R}^m \to \mathbb{R}^p$ es diferenciable en $\mathbf{x}_0$ y además, en dicho punto,*

$$\boxed{\mathbf{d}(\mathbf{f} \circ \mathbf{g}) = \mathbf{df} \circ \mathbf{dg}}$$

*Demostración.* Esquemáticamente, la situación de partida es,

$$\begin{array}{ccccc} \mathbb{R}^m & \xrightarrow{\mathbf{g}} & \mathbb{R}^n & \xrightarrow{\mathbf{f}} & \mathbb{R}^p \\ \mathbf{x}_0 & \longmapsto & \mathbf{y}_0 & \longmapsto & \mathbf{z}_0 \end{array}$$

(Si $p = 1$, la función $f$ sería numérica, que es el objeto de estudio de este curso, sin embargo estudiamos el caso general ya que el valor de $p$ no afecta a la demostración del teorema).
Por ser $\mathbf{g}$ diferenciable en $\mathbf{x}_0$, se tiene:

$$\mathbf{g}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{g}(\mathbf{x}_0) = \mathbf{dg}(\mathbf{h}) + \|\mathbf{h}\|\epsilon_1(\mathbf{h}) \quad \text{con} \quad \lim_{\mathbf{h}\to\mathbf{0}} \epsilon_1(\mathbf{h}) = \mathbf{0} \tag{3.10}$$

Consideremos ahora la imagen de $\mathbf{x}_0$ por $\mathbf{g}$: $\mathbf{y}_0 = \mathbf{g}(\mathbf{x}_0) \in \mathbf{g}(D_g) \subseteq \mathbb{R}^n$. Puesto que $\mathbf{f}$ es diferenciable en $\mathbf{y}_0$, se tiene:

$$\mathbf{f}(\mathbf{y}_0 + \mathbf{k}) - \mathbf{f}(\mathbf{y}_0) = \mathbf{df}(\mathbf{k}) + \|\mathbf{k}\|\epsilon_2(\mathbf{k}) \quad \text{con} \quad \lim_{\mathbf{k}\to\mathbf{0}} \epsilon_2(\mathbf{h}) = \mathbf{0} \tag{3.11}$$

Elijamos

$$\mathbf{k} = \mathbf{g}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{g}(\mathbf{x}_0) \tag{3.12}$$

entonces $\mathbf{k}$ es una función de $\mathbf{h}$ que tiende a $\mathbf{0}$ cuando $\mathbf{h}$ tiende hacia $\mathbf{0}$. Despejando $\mathbf{g}(\mathbf{x}_0 + \mathbf{h})$ en (3.12), se tiene: $\mathbf{g}(\mathbf{x}_0 + \mathbf{h}) = \mathbf{g}(\mathbf{x}_0) + \mathbf{k} = \mathbf{y}_0 + k$, y sustituyendo en (3.11),

$$\mathbf{y}_0 + k = \mathbf{g}(\mathbf{x}_0 + \mathbf{h}) \qquad \mathbf{y}_0 = \mathbf{g}(\mathbf{x}_0) \qquad \mathbf{k} = \mathbf{g}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{g}(\mathbf{x}_0)$$

resulta:

$$\mathbf{f}\big(\mathbf{g}(\mathbf{x}_0 + \mathbf{h})\big) - \mathbf{f}\big(\mathbf{g}(\mathbf{x}_0)\big) = \mathbf{df}\big(\mathbf{g}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{g}(\mathbf{x}_0)\big) + \|\mathbf{k}\|\epsilon_2(\mathbf{k})$$

ahora bien, teniendo en cuenta que por (3.10) es $\mathbf{g}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{g}(\mathbf{x}_0) = \mathbf{dg}(\mathbf{h}) + \|\mathbf{h}\|\epsilon_1(\mathbf{h})$, resulta

$$\mathbf{f}\big(\mathbf{g}(\mathbf{x}_0 + \mathbf{h})\big) - \mathbf{f}\big(\mathbf{g}(\mathbf{x}_0)\big) = \mathbf{df}\big(\mathbf{dg}(\mathbf{h}) + \|\mathbf{h}\|\epsilon_1(\mathbf{h})\big) + \|\mathbf{k}\|\epsilon_2(\mathbf{k})$$

y al ser $\mathbf{df}$ una aplicación lineal, será $\mathbf{df}\big(\mathbf{dg}(\mathbf{h}) + \|\mathbf{h}\|\epsilon_1(\mathbf{h})\big) = \mathbf{df}\big(\mathbf{dg}(\mathbf{h})\big) + \|\mathbf{h}\|\mathbf{df}\big(\epsilon_1(\mathbf{h})\big)$, con lo que resulta:

$$\mathbf{f}\big(\mathbf{g}(\mathbf{x}_0 + \mathbf{h})\big) - \mathbf{f}\big(\mathbf{g}(\mathbf{x}_0)\big) = \mathbf{df}\big(\mathbf{dg}(\mathbf{h})\big) + \|\mathbf{h}\|\mathbf{df}\big(\epsilon_1(\mathbf{h})\big) + \|\mathbf{k}\|\epsilon_2(\mathbf{k}) \tag{3.13}$$

de donde, para $\|\mathbf{h}\| \neq \mathbf{0}$, se tiene

$$\mathbf{f} \circ \mathbf{g}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f} \circ \mathbf{g}(\mathbf{x}_0) = \mathbf{df} \circ \mathbf{dg}(\mathbf{h}) + \|\mathbf{h}\|\Big(\mathbf{df}\big(\epsilon_1(\mathbf{h})\big) + \frac{\|\mathbf{k}\|}{\|\mathbf{h}\|}\epsilon_2(\mathbf{k})\Big)$$

luego (3.13) se puede expresar de la forma:

$$\mathbf{f} \circ \mathbf{g}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f} \circ \mathbf{g}(\mathbf{x}_0) = \mathbf{df} \circ \mathbf{dg}(\mathbf{h}) + \|\mathbf{h}\|\epsilon(\mathbf{h}) \tag{3.14}$$

donde

$$\epsilon(\mathbf{h}) = \begin{cases} \mathbf{df}\big(\epsilon_1(\mathbf{h})\big) + \dfrac{\|\mathbf{k}\|}{\|\mathbf{h}\|}\epsilon_2(\mathbf{k}) & \text{si } \mathbf{h} \neq \mathbf{0} \\ \mathbf{0} & \text{si } \mathbf{h} = \mathbf{0} \end{cases}$$

(el valor $\epsilon(\mathbf{0}) = \mathbf{0}$, se deduce al sustituir $\mathbf{h}$ por $\mathbf{0}$ en (3.13)).
Nos queda por demostrar que $\epsilon(\mathbf{h})$ tiende hacia $\mathbf{0}$ cuando $\mathbf{h}$ tiende hacia $\mathbf{0}$. Es decir, tenemos que demostrar que el siguiente límite es nulo:

$$\lim_{\mathbf{h}\to\mathbf{0}} \epsilon(\mathbf{h}) = \lim_{\mathbf{h}\to\mathbf{0}} \left(\mathbf{df}\big(\epsilon_1(\mathbf{h})\big) + \frac{\|\mathbf{k}\|}{\|\mathbf{h}\|}\epsilon_2(\mathbf{k})\right) = \lim_{\mathbf{h}\to\mathbf{0}} \mathbf{df}\big(\epsilon_1(\mathbf{h})\big) + \lim_{\mathbf{h}\to\mathbf{0}} \frac{\|\mathbf{k}\|}{\|\mathbf{h}\|}\epsilon_2(\mathbf{k})$$

Primer sumando; puesto que $\mathbf{df}$ es lineal, es continua, y se puede decir que:

$$\lim_{\mathbf{h}\to\mathbf{0}} \epsilon_1(\mathbf{h}) = \mathbf{0} \quad \Longrightarrow \quad \lim_{\mathbf{h}\to\mathbf{0}} \mathbf{df}\big(\epsilon_1(\mathbf{h})\big) = \mathbf{0}$$

Segundo sumando; puesto que $\mathbf{dg}$ es continua, existe una constante $a > 0$ tal que, para todo $\mathbf{h} \in \mathbb{R}^m$, se verifica que $\|\mathbf{dg}(\mathbf{h})\| \leq a\|\mathbf{h}\|$. Por consiguiente, según (3.12) y (3.10), se tiene:

$$\|\mathbf{k}\| = \|\mathbf{g}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{g}(\mathbf{x}_0)\| = \|\Big(\mathbf{dg}(\mathbf{h}) + \|\mathbf{h}\|\epsilon_1(\mathbf{h})\Big)\| \leq \|\mathbf{dg}(\mathbf{h})\| + \|\mathbf{h}\|\|\epsilon_1(\mathbf{h})\|$$

con lo cual,

$$\|\mathbf{k}\| \leq a\|\mathbf{h}\| + \|\mathbf{h}\|\|\epsilon_1(\mathbf{h})\| = \big(a + \|\epsilon_1(\mathbf{h})\|\big)\|\mathbf{h}\|$$

y, por tanto:

$$\lim_{\mathbf{k}\to\mathbf{0}} \epsilon_2(\mathbf{k}) = \mathbf{0} \quad \Longrightarrow \quad \lim_{\mathbf{h}\to\mathbf{0}} \frac{\|\mathbf{k}\|}{\|\mathbf{h}\|}\epsilon_2(\mathbf{k}) = \mathbf{0}$$

Y, en consecuencia, se concluye que $\lim\limits_{\mathbf{h}\to\mathbf{0}} \epsilon(\mathbf{h}) = \mathbf{0}$.
La relación (3.14) prueba, entonces que $\mathbf{f} \circ \mathbf{g}$ es diferenciable en $\mathbf{x}_0$, y también que su diferencial en ese punto es la compuesta $\mathbf{dg}\circ\mathbf{df}$ de las diferenciales de $\mathbf{g}$ y $\mathbf{f}$.Con lo cual el teorema 3.8 queda demostrado. □

### 3.8.4. Regla de la cadena. Perspectiva práctica: Parciales

En la Regla de la Cadena hay que diferenciar el aspecto teórico del aspecto práctico. El aspecto teórico presenta una gran elegancia y simplicidad, sin embargo no nos deja ver la riqueza de sus múltiples aplicaciones prácticas. Así, el teorema 3.8 puede enunciarse en términos de derivadas parciales de la siguiente forma:

**Teorema 3.9 (Regla de la cadena).** *Sea* $\mathbf{g} : D_g \subseteq \mathbb{R}^m \to \mathbb{R}^n$ *una función definida en el conjunto abierto* $D_g$ *de* $\mathbb{R}^m$*, diferenciable en* $\mathbf{x}_0 \in D_g$*. Sea* $f : D_f \subseteq \mathbb{R}^n \to \mathbb{R}$ *una función definida en el conjunto abierto* $D_f$ *de* $\mathbb{R}^n$*, tal que* $\mathbf{g}(D_g) \subseteq D_f$*, diferenciable en* $\mathbf{y}_0 = \mathbf{g}(\mathbf{x}_0) \in D_f$*. Entonces la composición* $f \circ \mathbf{g} : D_g \subseteq \mathbb{R}^m \to \mathbb{R}$ *es diferenciable en* $\mathbf{x}_0$ *y además, en dicho punto, sus derivadas parciales son:*

$$\boxed{\frac{\partial}{\partial x_j}(f \circ \mathbf{g})(\mathbf{x}_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial y_i}\big(\mathbf{g}(\mathbf{x}_0)\big)\frac{\partial g_i}{\partial x_j}(\mathbf{x}_0). \quad j = 1, 2, \ldots, m} \tag{3.15}$$

La fórmula (3.15) se comprende mejor si la desarrollamos en sus componente. En efecto, si la función $\mathbf{g}$ depende de $m$ variables, $x_1, \cdots, x_m$, $(y_1, \cdots, y_n) = \mathbf{g}(x_1, \cdots, x_m)$, tendríamos la situación siguiente:

$$\begin{array}{c} \mathbb{R}^m \xrightarrow{\mathbf{g}} \mathbb{R}^n \xrightarrow{f} \mathbb{R} \\ (x_1, \cdots, x_m) \mapsto (y_1, \cdots, y_n) \mapsto z \end{array}$$

donde $\mathbf{g}$ es una función de $n$ funciones coordenadas, cada una de ellas dependiendo de las $m$ variable $x_1, \cdots, x_m$. Al hacer la composición con $f$, se obtiene la función $f \circ \mathbf{g}$ que depende, también, de las $m$ variables $x_1, \cdots, x_m$. Para estas funciones, las derivadas parciales de las fórmulas (3.15) se puede expresar de la forma:

$$\left.\begin{array}{c} \dfrac{\partial z}{\partial x_1} = \dfrac{\partial f}{\partial y_1}\dfrac{\partial g_1}{\partial x_1} + \cdots + \dfrac{\partial f}{\partial y_n}\dfrac{\partial g_n}{\partial x_1} \\ \vdots \\ \dfrac{\partial z}{\partial x_m} = \dfrac{\partial f}{\partial y_1}\dfrac{\partial g_1}{\partial x_m} + \cdots + \dfrac{\partial f}{\partial y_n}\dfrac{\partial g_n}{\partial x_m} \end{array}\right\} \tag{3.16}$$

Es decir, de manera esquemática, pensando en términos de sustitución de las variables, y

usando la misma letras para denotar a las funciones así como a sus imágenes, tenemos:

$$z = f(y_1, \cdots, y_n) \left\{ \begin{array}{c} y_1 = y_1(x_1, \cdots, x_m) \\ \vdots \\ y_n = y_n(x_1, \cdots, x_m) \end{array} \right\} \Rightarrow \left\{ \begin{array}{c} \dfrac{\partial z}{\partial x_1} = \dfrac{\partial f}{\partial y_1}\dfrac{\partial y_1}{\partial x_1} + \cdots + \dfrac{\partial f}{\partial y_n}\dfrac{\partial y_n}{\partial x_1} \\ \vdots \\ \dfrac{\partial z}{\partial x_m} = \dfrac{\partial f}{\partial y_1}\dfrac{\partial y_1}{\partial x_m} + \cdots + \dfrac{\partial f}{\partial y_n}\dfrac{\partial y_n}{\partial x_m} \end{array} \right.$$

Y forzando un poco más la notación, resulta,

$$z = z(y_1, \cdots, y_n) \left\{ \begin{array}{c} y_1 = y_1(x_1, \cdots, x_m) \\ \vdots \\ y_n = y_n(x_1, \cdots, x_m) \end{array} \right\} \Rightarrow \left\{ \begin{array}{c} \dfrac{\partial z}{\partial x_1} = \dfrac{\partial z}{\partial y_1}\dfrac{\partial y_1}{\partial x_1} + \cdots + \dfrac{\partial z}{\partial y_n}\dfrac{\partial y_n}{\partial x_1} \\ \vdots \\ \dfrac{\partial z}{\partial x_m} = \dfrac{\partial z}{\partial y_1}\dfrac{\partial y_1}{\partial x_m} + \cdots + \dfrac{\partial z}{\partial y_n}\dfrac{\partial y_n}{\partial x_m} \end{array} \right.$$

Luego, según la regla de la cadena se tiene:

$$\boxed{\frac{\partial z}{\partial x_j} = \sum_{i=1}^{n} \frac{\partial z}{\partial y_i}\frac{\partial y_i}{\partial x_j}. \quad j = 1, 2, \ldots, m} \tag{3.17}$$

Es evidente la ventajosa simplicidad de la fórmula de las derivadas parciales de la función compuesta con esta presentación. Sin embargo, debemos hacer notar que: 1º no se hacen explícitos los puntos donde están calculadas las derivadas, y 2º se usan letras iguales para denotar cosas distintas, a saber: las funciones y sus imágenes.
Aquí se ha seguido el orden *alfabético* de las variables $x \to y \to z$. Sin embargo, en la práctica, cuando el número de variables es reducido, no se suele seguir este orden, ya que se acostumbra a utilizar letras distintas sin subíndices. Para varias variables, más acorde con la notación práctica hubiera sido seguir con el orden $t \to x \to z$, con lo que hubiera resultado:

$$z = f(x_1, \cdots, x_n) \left\{ \begin{array}{c} x_1 = x_1(t_1, \cdots, t_m) \\ \vdots \\ x_n = x_n(t_1, \cdots, t_m) \end{array} \right\} \Rightarrow \left\{ \begin{array}{c} \dfrac{\partial z}{\partial t_1} = \dfrac{\partial f}{\partial x_1}\dfrac{\partial x_1}{\partial t_1} + \cdots + \dfrac{\partial f}{\partial x_n}\dfrac{\partial x_n}{\partial t_1} \\ \vdots \\ \dfrac{\partial z}{\partial t_m} = \dfrac{\partial f}{\partial x_1}\dfrac{\partial x_1}{\partial t_m} + \cdots + \dfrac{\partial f}{\partial x_n}\dfrac{\partial x_n}{\partial t_m} \end{array} \right.$$

Veamos la materialización de esta fórmula a los distintos casos concretos.

### Regla de la cadena con una variable independiente

Si la función **g** depende de una sola variable, tendríamos la situación siguiente:

$$\begin{array}{c} \mathbb{R} \xrightarrow{\mathbf{g}} \mathbb{R}^n \xrightarrow{f} \mathbb{R} \\ t \mapsto (x_1, \cdots, x_n) \mapsto z \end{array}$$

donde $\mathbf{g}$ es una función de $n$ funciones coordenadas, cada una de ellas dependiendo de una sola variable $t$. Al hacer la composición con $f$, se obtiene la función $f \circ \mathbf{g}$ que depende, también, sólo de $t$. Para estas funciones, las derivadas parciales de las fórmulas (3.15) del teorema anterior son *derivadas totales*, quedando como:

$$\frac{d}{dt}(f \circ \mathbf{g})(t) = (f \circ \mathbf{g})'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\big(\mathbf{g}(t)\big)\frac{d\mathbf{g}_i}{dt}(t)$$

donde $g_i, i = 1, 2, \cdots, n$, son las funciones coordenadas de $\mathbf{g}$.
Que se puede expresar de la forma:

$$\frac{dz}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \cdots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt}$$

Es decir, de manera esquemática, pensando en términos de sustitución de las variables, tenemos:

$$z = f(x_1, \cdots, x_n)\left\{\begin{array}{c} x_1 = x_1(t) \\ \vdots \\ x_n = x_n(t) \end{array}\right\} \Rightarrow \frac{dz}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \cdots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt}$$

**Planteamiento práctico.** Con objeto de recordar con facilidad las fórmulas correspondientes a la regla de la cadena planteamos un segundo razonamiento (más práctico) centrado en dos variables.
Supongamos una función de dos variables $z = f(x, y)$, y supongamos que cada una de las variables $x$ e $y$ depende de una tercera variable $t$, mediante las funciones $x = x(t)$, $y = y(t)$. Podemos sustituir los valores de $x$ e $y$ en la función $z = f(x, y)$, con lo cual $z$ pasaría a depender directamente de $t$ y podríamos calcular la derivada de $z$ respecto de $t$.

$$z = f(x, y)\left\{\begin{array}{c} x = x(t) \\ y = y(t) \end{array}\right\} \Rightarrow z = f\big(x(t), y(t)\big) = h(t) \Rightarrow \frac{dz}{dt} = h'(t)$$

Esto que puede hacerse por sustitución de las variables (sustituyendo primero y derivando después), también puede hacerse por la regla de la cadena. Para obtener la fórmula correspondiente, partimos de $dz$, y a partir de ahí obtenemos $dz/dt$, por simple división. En efecto,

$$dz = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \quad \Rightarrow \quad \frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

Otra manera de obtener la fórmula es mediante un diagrama en árbol, donde la derivada de la función $z$ respecto de la variable $t$ se obtiene "sumando" todos los caminos que van de $t$ a $z$.

$$\begin{tikzcd} & x \arrow[dr, "\frac{\partial z}{\partial x}"] & \\ t \arrow[ur, "\frac{dx}{dt}"] \arrow[dr, "\frac{dy}{dt}"'] & & z \\ & y \arrow[ur, "\frac{\partial z}{\partial y}"'] & \end{tikzcd} \quad \Rightarrow \quad \frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}$$

**Ejemplo 3.50.** *Dada la función $z = x^2y - y^2$, donde $x = \operatorname{sen} t$, $y = e^t$. Hallar $dz/dt$ cuando $t = 0$,*

*(a) Por sustitución previa*

*(b) Mediante la Regla de la Cadena.*

*Solución.* (a) Mediante la sustitución previa es un proceso elemental, basta con sustituir $x$ e $y$ por sus valores y derivar una vez hecha la sustitución.

$$z = x^2y - y^2 \left\{ \begin{array}{l} x = \operatorname{sen} t \\ y = e^t \end{array} \right\} \to z = e^t \operatorname{sen}^2 t - e^{2t}$$

de donde,

$$\frac{dz}{dt} = e^t \operatorname{sen}^2 t + 2e^t \operatorname{sen} t \cos t - 2e^{2t}$$

y, en consecuencia, resulta

$$\left.\frac{dz}{dt}\right]_{t=0} = 1 \cdot 0 + 2 \cdot 1 \cdot 0 \cdot 1 - 2 \cdot 1 = -2$$

(b) Aplicando la Regla de la Cadena, resulta

$$\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = (2xy)\cos t + (x^2 - 2y)e^t$$

Que podemos expresarlo en términos de $t$, o dejarlo como está. En consecuencia, dado que

$$t = 0 \Rightarrow \left\{ \begin{array}{l} x = 0 \\ y = 1 \end{array} \right.$$

resulta:

$$\left.\frac{dz}{dt}\right]_{t=0} = 0 + (0 - 2) \cdot 1 = -2$$

**Ejemplo 3.51.** *La altura de un cono recto circular mide* 15 *cm. y aumenta a razón de* $0'2$ *cm. cada minuto. El radio de la base mide* 10 *cm. y aumenta a razón de* $0'3$ *cm. cada minuto. Hallar la variación de volumen que experimenta en la unidad de tiempo.*

*Solución.* El volumen del cono viene definido por: $V = \dfrac{1}{3}\pi r^2 h$, de donde,

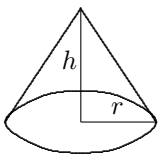



Figura 3.20: cono.

$$\frac{dV}{dt} = \frac{\partial V}{\partial r}\frac{dr}{dt} + \frac{\partial V}{\partial h}\frac{dh}{dt} = \frac{2}{3}\pi r h \frac{dr}{dt} + \frac{1}{3}\pi r^2 \frac{dh}{dt}$$

y, teniendo en cuenta que,

$$\text{para} \left\{ \begin{array}{l} h = 15 \\ r = 10 \end{array} \right\} \text{se tiene} \left\{ \begin{array}{l} h'(t) = 0'2 \\ r'(t) = 0'3 \end{array} \right.$$

resulta,

$$\left.\frac{dV}{dt}\right]_{\substack{h=15 \\ r=10}} = \frac{2}{3}\pi \cdot 10 \cdot 15 \cdot 0'3 + \frac{1}{3}\pi \cdot 100 \cdot 0'2 = \frac{-110\pi}{3}\ cm^3/min$$

**Regla de la cadena para dos variables independientes**

Si la función $\mathbf{g}$ depende de dos variables, $t_1$ y $t_2$, $(x_1, \cdots, x_n) = \mathbf{g}(t_1, t_2)$, tendríamos la situación siguiente:

$$\begin{array}{c} \mathbb{R}^2 \xrightarrow{\mathbf{g}} \mathbb{R}^n \xrightarrow{f} \mathbb{R} \\ (t_1, t_2) \mapsto (x_1, \cdots, x_n) \mapsto z \end{array}$$

donde $\mathbf{g}$ es una función de $n$ funciones coordenadas, cada una de ellas dependiendo de las dos variable $t_1$ y $t_2$. Al hacer la composición con $f$, se obtiene la función $f \circ \mathbf{g}$ que depende, también, de las variables $t_1$ y $t_2$. Para estas funciones, las derivadas parciales de las fórmulas (3.15) del teorema anterior son las siguientes:

$$\frac{\partial}{\partial t_j}(f \circ \mathbf{g})(\mathbf{x}_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\big(\mathbf{g}(\mathbf{x}_0)\big)\frac{\partial g_i}{\partial t_j}(\mathbf{x}_0). \quad j = 1, 2$$

donde $g_i, i = 1, 2, \cdots, n$, son las funciones coordenadas de $\mathbf{g}$.
Que, al igual que la fórmula (3.16), se puede expresar de la forma:

$$\left.\begin{array}{l} \dfrac{\partial z}{\partial t_1} = \dfrac{\partial f}{\partial x_1}\dfrac{\partial x_1}{\partial t_1} + \cdots + \dfrac{\partial f}{\partial x_n}\dfrac{\partial x_n}{\partial t_1} \\ \dfrac{\partial z}{\partial t_2} = \dfrac{\partial f}{\partial x_1}\dfrac{\partial x_1}{\partial t_2} + \cdots + \dfrac{\partial f}{\partial x_n}\dfrac{\partial x_n}{\partial t_2} \end{array}\right\}$$

Es decir, de manera esquemática, pensando en términos de sustitución de las variables, tenemos:

$$z = f(x_1, \cdots, x_n) \left\{ \begin{array}{c} x_1 = x_1(t_1, t_2) \\ \vdots \\ x_n = x_n(t_1, t_2) \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \dfrac{\partial z}{\partial t_1} = \dfrac{\partial f}{\partial x_1}\dfrac{\partial x_1}{\partial t_1} + \cdots + \dfrac{\partial f}{\partial x_n}\dfrac{\partial x_n}{\partial t_1} \\ \dfrac{\partial z}{\partial t_2} = \dfrac{\partial f}{\partial x_1}\dfrac{\partial x_1}{\partial t_2} + \cdots + \dfrac{\partial f}{\partial x_n}\dfrac{\partial x_n}{\partial t_2} \end{array}\right.$$

**Planteamiento práctico.** Con objeto de recordar con facilidad las fórmulas correspondientes a la regla de la cadena planteamos un segundo razonamiento (más práctico) centrado en dos variables.
Supongamos una función de dos variables $z = f(x, y)$, y supongamos que cada una de las variables $x$ e $y$ dependen de otras dos variable $s$ y $t$, mediante las funciones $x = x(s, t)$, $y = y(s, t)$. Podemos sustituir los valores de $x$ e $y$ en la función $z = f(x, y)$, con lo cual $z$ pasaría a depender directamente de $s$ y $t$ y podríamos calcular las derivadas parciales de $z$ respecto de $s$ y respecto de $t$.

$$z = f(x, y) \left\{ \begin{array}{c} x = x(s, t) \\ y = y(s, t) \end{array} \right\} \Rightarrow z = f\big(x(s, t), y(s, t)\big) = h(s, t) \Rightarrow \left\{ \begin{array}{l} \dfrac{\partial z}{\partial s} = ? \\ \dfrac{\partial z}{\partial t} = ? \end{array}\right.$$

Esto que puede hacerse por sustitución de las variables (sustituyendo primero y derivando después), también puede hacerse por la regla de la cadena. Para obtener la fórmula correspondiente, partimos de $dz$, y a partir de ahí obtenemos $\partial z/\partial s$ y $\partial z/\partial t$ por simple

división. En cada caso, $d$ se transforma en $\partial$, ya que al hacer la división suponemos que la "otra" variable permanece constante. En efecto,

$$dz = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \quad \Rightarrow \quad \begin{cases} \dfrac{\partial z}{\partial s} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial s} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial s} & t = \text{cte.} \\ \dfrac{\partial z}{\partial t} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial t} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial t} & s = \text{cte.} \end{cases}$$

Otra manera de obtener la fórmula es mediante un diagrama en árbol, donde la derivada de la función $z$ respecto de cada variable se obtiene "sumando" todos los caminos que van desde la variable correspondiente a $z$, suponiendo el resto de las variables fijas.

$$t = \text{cte.} \quad \rightarrow \quad \begin{tikzcd} & x \arrow[dr, "\frac{\partial z}{\partial x}"] & \\ s \arrow[ur, "\frac{\partial x}{\partial s}"] \arrow[dr, "\frac{\partial y}{\partial s}"'] & & z \\ & y \arrow[ur, "\frac{\partial z}{\partial y}"'] & \end{tikzcd} \quad \Rightarrow \quad \frac{\partial z}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}$$

$$s = \text{cte.} \quad \rightarrow \quad \begin{tikzcd} & x \arrow[dr, "\frac{\partial z}{\partial x}"] & \\ t \arrow[ur, "\frac{\partial x}{\partial t}"] \arrow[dr, "\frac{\partial y}{\partial t}"'] & & z \\ & y \arrow[ur, "\frac{\partial z}{\partial y}"'] & \end{tikzcd} \quad \Rightarrow \quad \frac{\partial z}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}$$

**Ejemplo 3.52.** *Dada la función $z = 2xy$, donde $x = s^2 + t^2$ e $y = s/t$.*
*Hallar $\dfrac{\partial z}{\partial s}$ y $\dfrac{\partial z}{\partial t}$*

*(a) Por sustitución de las variables.*
*(b) Por la regla de la cadena.*

*Solución.* (a) Mediante la sustitución previa es un proceso elemental, basta con sustituir $x$ e $y$ por sus valores y derivar una vez hecha la sustitución.

$$z = 2xy \left\{ \begin{array}{l} x = s^2 + t^2 \\ y = s/t \end{array} \right\} \rightarrow z = 2(s^2 + t^2)\frac{s}{t} = 2(\frac{s^3}{t} + st)$$

de donde,

$$\frac{\partial z}{\partial s} = 2(\frac{3s^2}{t} + t) = \frac{6s^2 + 2t^2}{t}$$
$$\frac{\partial z}{\partial t} = 2(\frac{-s^3}{t^2} + s) = \frac{-2s^3 + 2st^2}{t^2}$$

(b) Aplicando la Regla de la Cadena, resulta

$$\frac{\partial z}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} = (2y)(2s) + (2x)(\frac{1}{t}) = 4sy + \frac{2x}{t} = 4s\frac{s}{t} + \frac{2}{t}(s^2 + t^2) =$$
$$= \frac{4s^2}{t} + \frac{2s^2 + 2t^2}{t} = \frac{6s^2 + 2t^2}{t}$$
$$\frac{\partial z}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} = (2y)(2t) + (2x)(\frac{-s}{t^2}) = 4ty - \frac{2s}{t^2}x = 4t\frac{s}{t} - \frac{2s}{t^2}(s^2 + t^2) =$$
$$= \frac{4t^2s - 2s^3 - 2st^2}{t^2} = \frac{2st^2 - 2s^3}{t^2}$$

### 3.8.5. Regla de la cadena. Perspectiva general: Matriz jacobiana

La introducción del Jacobiano nos va a permitir utilizar el lenguaje algebraico de las matrices y las transformaciones lineales y nos va a permitir ver la regla de la cadena como un resultado de asombrosa sencillez y elegancia, como lo es en el caso de funciones de una variable.
Recordemos, primeramente, lo que establece la regla de la cadena para funciones de una sola variable: *la composición de dos funciones diferenciable es diferenciable y su derivada es el producto de las derivadas de cada una de las funciones que se están componiendo.* Queremos establecer un resultado similar para funciones vectoriales de cualquier número de variables.
Hasta ahora hemos definido los conceptos de "derivadas parciales", "derivadas direccionales", "gradiente" y "diferencial", para funciones de varias variables, pero no hemos definido lo que se entiende por "derivada"de este tipo de funciones.

**Definición 3.19 (Derivada).** *Sea la función $f : \mathcal{D} \subseteqq \mathbb{R}^n \to \mathbb{R}^m$ definida en el conjunto abierto $\mathcal{D}$ de $\mathbb{R}^n$ y sea $\mathbf{x}_0 \in \mathcal{D}$. Se dice que esta función es diferenciable en $\mathbf{x}_0$ si existe una transformación lineal $\mathbf{f}'(\mathbf{x}_0) : \mathbb{R}^n \to \mathbb{R}^m$, llamada derivada de $\mathbf{f}$ en $\mathbf{x}_0$, tal que*

$$\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) = \mathbf{f}(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)\mathbf{h} + \mathbf{r}(\mathbf{h}) \quad \textit{donde} \quad \lim_{\mathbf{h}\to\mathbf{0}} \frac{\mathbf{r}(\mathbf{h})}{\|\mathbf{h}\|},$$

*(para $\mathbf{h}$ tal que $\mathbf{x}_0 + \mathbf{h} \in \mathcal{D}$)*

NOTA: Siendo $\mathbf{r}$ una función que toma valores en $\mathbb{R}^m$, la interpretación del límite anterior sería de que todas las funciones componentes tenderán a cero cuando, al dividirlas por $\|\mathbf{h}\|$, $\mathbf{h}$ tiende a cero.
Hemos definido la derivada de una función vectorial como una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$. Toda transformación lineal tiene asociada una matriz de orden $m \times n$ que la representa.
**Matriz de una transformación lineal.** La matriz asociada a la transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ se construye de la siguiente forma: Sean $\beta_1 = \{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ y $\beta_2 = \{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \cdots, \bar{\mathbf{e}}_n\}$ las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente. Entonces el vector $T(\mathbf{e}_j)$ se puede escribir como combinación lineal de los vectores de $\beta_2$, es decir,

$$T(\mathbf{e}_j) = a_{1j}\bar{\mathbf{e}}_1 + a_{2j}\bar{\mathbf{e}}_2 + \cdots + a_{mj}\bar{\mathbf{e}}_m$$

A la matriz $A = (a_{ij}), i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n$, cuya j-ésima columna está constituida por las coordenadas de la imagen del j-ésimo vector $\mathbf{e}_j$ de la base canónica de $\mathbb{R}^n$ respecto de la base canónica de $\mathbb{R}^m$, se le llama matriz de la transformación $T$ respecto de las bases $\beta_1$ y $\beta_2$. Es decir, la transformación $T : \mathbb{R}^n \to \mathbb{R}^m$ queda unívocamente determinada cuando se da la imagen de una base de $\mathbb{R}^n$. Sean $\beta_1 = \{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ y $\beta_2 = \{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \cdots, \bar{\mathbf{e}}_n\}$ las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente. Las imágenes de los vectores de la base $\beta_1$, por estar en $\mathbb{R}^m$, se pueden expresar mediante combinación lineal de los vectores de $\beta_2$. Sean las imágenes de los vectores de la base $\beta_1$ mediante la transformación $T$ las siguientes:

$$\left.\begin{array}{r} T(\mathbf{e}_1) = a_{11}\bar{\mathbf{e}}_1 + a_{21}\bar{\mathbf{e}}_2 + \cdots + a_{m1}\bar{\mathbf{e}}_m \\ \cdots \\ T(\mathbf{e}_n) = a_{1n}\bar{\mathbf{e}}_1 + a_{2n}\bar{\mathbf{e}}_2 + \cdots + a_{mn}\bar{\mathbf{e}}_m \end{array}\right\}$$

Si $\mathbf{x}$ es un vector arbitrario de $\mathbb{R}^n$, e $\mathbf{y}$ su imagen en $\mathbb{R}^m$, será:

$$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n \quad \to \quad \mathbf{y} = y_1\bar{\mathbf{e}}_1 + \cdots + y_m\bar{\mathbf{e}}_m$$

se verificará:

$$\begin{aligned} y_1\bar{\mathbf{e}}_1 + \cdots + y_m\bar{\mathbf{e}}_m &= T(\mathbf{x}) = T(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) = x_1T(\mathbf{e}_1) + \cdots + x_nT(\mathbf{e}_n) = \\ &= x_1(a_{11}\bar{\mathbf{e}}_1 + \cdots + a_{m1}\bar{\mathbf{e}}_m) + \cdots + x_n(a_{1n}\bar{\mathbf{e}}_1 + \cdots + a_{mn}\bar{\mathbf{e}}_m) = \\ &= (a_{11}x_1 + \cdots + a_{1n}x_n)\bar{\mathbf{e}}_1 + \cdots + (a_{m1}x_1 + \cdots + a_{mn}x_n)\bar{\mathbf{e}}_m \end{aligned}$$

de donde

$$\left.\begin{array}{r} y_1 = a_{11}x_1 + \cdots + a_{1n}x_n \\ \cdots \\ y_m = a_{m1}x_1 + \cdots + a_{mn}x_n \end{array}\right\}$$

Son las ecuaciones de la transformación $T$ respecto de las bases $\beta_1$ y $\beta_2$. Las ecuaciones se pueden escribir en forma matricial del siguiente modo:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Es decir, $\mathbf{y} = A\mathbf{x}$, donde las columnas de la matriz $A$ son las coordenadas de las imágenes de los vectores de la base canónica $\beta_1$ respecto de $\beta_2$.

**Matriz jacobiana.** Sea $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función diferenciable en $\mathbf{x}_0 \in \mathcal{D}$, entonces, según la definición 3.19, será:

$$\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) = \mathbf{f}(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)\mathbf{h} + \mathbf{r}(\mathbf{h}) \quad \text{donde} \quad \lim_{\mathbf{h}\to\mathbf{0}} \frac{\mathbf{r}(\mathbf{h})}{\|\mathbf{h}\|} = \mathbf{0} \quad (\text{para } \mathbf{h} \text{ tal que } \mathbf{x}_0 + \mathbf{h} \in \mathcal{D}).$$

Considerando el vector $\mathbf{h} = t\mathbf{e}_j$, $\mathbf{e}_j$ de la base canónica de $\mathbb{R}^n$ y $t \in \mathbb{R}$, con $t$ suficientemente pequeño como para asegurar que $\mathbf{x}_0 + \mathbf{h} \in \mathcal{D}$, resulta:

$$\mathbf{f}(\mathbf{x}_0 + t\mathbf{e}_j) = \mathbf{f}(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)(t\mathbf{e}_j) + \mathbf{r}(\mathbf{h})$$

Y teniendo en cuenta que $\mathbf{f}'(\mathbf{x}_0)$ es una transformación lineal, será: $\mathbf{f}'(\mathbf{x}_0)(t\mathbf{e}_j) = t\mathbf{f}'(\mathbf{x}_0)(\mathbf{e}_j)$, luego,

$$\mathbf{f}(\mathbf{x}_0 + t\mathbf{e}_j) = \mathbf{f}(\mathbf{x}_0) + t\mathbf{f}'(\mathbf{x}_0)(\mathbf{e}_j) + \mathbf{r}(\mathbf{h})$$

de donde,

$$\mathbf{f}'(\mathbf{x}_0)(\mathbf{e}_j) = \frac{\mathbf{f}(\mathbf{x}_0 + t\mathbf{e}_j) - \mathbf{f}(\mathbf{x}_0)}{t} - \frac{\mathbf{r}(\mathbf{h})}{t}$$

y teniendo en cuenta que $\|\mathbf{h}\| = \pm t$, podemos escribir

$$\mathbf{f}'(\mathbf{x}_0)(\mathbf{e}_j) = \frac{\mathbf{f}(\mathbf{x}_0 + t\mathbf{e}_j) - \mathbf{f}(\mathbf{x}_0)}{t} \mp \frac{\mathbf{r}(\mathbf{h})}{\|\mathbf{h}\|}$$

Al tomar límite cuando $t \to 0$ (i.e. cuando $\|\mathbf{h}\| \to \mathbf{0}$) el segundo sumando del lado derecho tiende a cero, quedando sólo:

$$\mathbf{f}'(\mathbf{x}_0)(\mathbf{e}_j) = \lim_{t\to 0} \frac{\mathbf{f}(\mathbf{x}_0 + t\mathbf{e}_j) - \mathbf{f}(\mathbf{x}_0)}{t}$$

Al descomponer $\mathbf{f}$ como $(f_1, f_2, \ldots, f_n)$, y tomando límites en cada una de las coordenadas, vemos que éstos son las derivadas parciales respecto de la variable $x_j$ de cada una de las componentes. es decir,

$$\mathbf{f}'(\mathbf{x}_0)(\mathbf{e}_j) = \left(\frac{\partial f_1}{\partial x_j}(\mathbf{x}_0), \frac{\partial f_2}{\partial x_j}(\mathbf{x}_0), \ldots, \frac{\partial f_m}{\partial x_j}(\mathbf{x}_0)\right) = \sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)\bar{\mathbf{e}}_i$$

Entonces la j-ésima columna de la matriz que representa a la transformación lineal (derivada) $\mathbf{f}'(\mathbf{x}_0) : \mathbb{R}^n \to \mathbb{R}^m$ está formada por las derivadas parciales de las funciones componentes de $\mathbf{f}$ respecto de su j-ésima variable. Es decir,

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \ldots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}_0) & \ldots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}_0) \\ \vdots & \vdots & \ldots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_m}{\partial x_2}(\mathbf{x}_0) & \ldots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}$$

A esta matriz de orden $m \times n$, en cuya i-ésima linea y j-ésima columna aparece la derivada parcial $\frac{\partial f_i}{x_j}(\mathbf{x}_0)$ de la i-ésima función componente de $\mathbf{f}$ respecto de su j-ésima variable, evaluada en $\mathbf{x}_0$, se le llama *matriz*

*jacobiana* de la función $\mathbf{f}$ en $\mathbf{x}_0$ y se denota $J\mathbf{f}(\mathbf{x}_0)$. Esta es entonces la *derivada* de la función diferenciable $\mathbf{f}$ en $\mathbf{x}_0$ (identificando, como siempre, a la transformación lineal $\mathbf{f}'(\mathbf{x}_0) : \mathbb{R}^n \to \mathbb{R}^m$ con la matriz que la representa).

**Gradiente y matriz jacobiana.** Consideremos el caso $m = 1$. En este caso, la función $f : \mathcal{D} \subseteqq \mathbb{R}^n \to \mathbb{R}$ definida en el conjunto abierto $\mathcal{D}$ de $\mathbb{R}^n$ será diferenciable en $\mathbf{x}_0 \in \mathcal{D}$ si se da la transformación lineal $\mathbf{f}'(\mathbf{x}_0) : \mathbb{R}^n \to \mathbb{R}$ cuya representación matricial es la matriz $1 \times n$

$$Jf(\mathbf{x}_0) = \left( \frac{\partial f}{\partial x_1}(\mathbf{x}_0) \ \frac{\partial f}{\partial x_2}(\mathbf{x}_0) \ \cdots \ \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \right)$$

tal que

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)\mathbf{h} + r(\mathbf{h}) \quad \text{con} \quad \lim_{\mathbf{h} \to \mathbf{0}} \frac{r(\mathbf{h})}{\|\mathbf{h}\|} = 0.$$

Ahora bien, Si $\mathbf{h} = (h_1, h_2, \ldots, h_n) \in \mathbb{R}^n$ es tal que $\mathbf{x}_0 + \mathbf{h} \in \mathcal{D}$, se tiene

$$\mathbf{f}'(\mathbf{x}_0)\mathbf{h} = Jf(\mathbf{x}_0)\mathbf{h} = \left( \frac{\partial f}{\partial x_1}(\mathbf{x}_0) \ \frac{\partial f}{\partial x_2}(\mathbf{x}_0) \ \cdots \ \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \right) \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} =$$

$$= \frac{\partial f}{\partial x_1}(\mathbf{x}_0)h_1 + \frac{\partial f}{\partial x_2}(\mathbf{x}_0)h_2 + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{x}_0)h_n$$

Comparando este resultado con el obtenido en la definición 3.7 (página 31), vemos que se trata exactamente de la misma definición, sólo que en aquélla se hablaba de la existencia de las derivadas parciales y ahora hablamos de la transformación lineal $\mathbf{f}'(\mathbf{x}_0)$ (la derivada de $f$ en $\mathbf{x}_0$) o de un modo más preciso, de la matriz que la representa. Es decir, el residuo $r(\mathbf{h})$ queda determinado de la misma manera en ambas definiciones y se le pide que cumpla la misma propiedad.
Puede demostrarse que la función $\mathbf{f} : \mathcal{D} \subseteqq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $\mathbf{x}_0 \in \mathcal{D}$ (según la definición 3.19) si y sólo si sus funciones coordenadas $f_i : \mathcal{D} \subseteqq \mathbb{R}^n \to \mathbb{R}^m$ lo son (con la definición de diferenciabilidad 3.7). Más aún, teniendo en cuenta que la derivada de la i-ésima función coordenada $f_i$ en $\mathbf{x}_0$ es la matriz de orden $1 \times n$

$$Jf_i(\mathbf{x}_0) = \left( \frac{\partial f_i}{\partial x_1}(\mathbf{x}_0) \ \frac{\partial f_i}{\partial x_2}(\mathbf{x}_0) \ \cdots \ \frac{\partial f_i}{\partial x_n}(\mathbf{x}_0) \right)$$

tenemos que la derivada de la función $\mathbf{f} : \mathcal{D} \subseteqq \mathbb{R}^n \to \mathbb{R}^m$ es la matriz que en su i-ésima fila tiene la derivada de su i-ésima función componente. Esquemáticamente:

$$Jf(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \ldots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}_0) & \ldots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}_0) \\ \vdots & \vdots & \ldots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_m}{\partial x_2}(\mathbf{x}_0) & \ldots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix} = \begin{pmatrix} \leftarrow & Jf_1(\mathbf{x}_0) & \rightarrow \\ \leftarrow & Jf_2(\mathbf{x}_0) & \rightarrow \\ & \vdots & \\ \leftarrow & Jf_m(\mathbf{x}_0) & \rightarrow \end{pmatrix}$$

Obsérvese también que la matriz $1 \times n$, $Jf_i(\mathbf{x}_0)$, derivada de la función $f_i : \mathcal{D} \subseteqq \mathbb{R}^n \to \mathbb{R}$, se identifica de manera natural con el vector gradiente de $f_i$ en $x_0$

$$Jf_i(\mathbf{x}_0) = \text{grad } f_i(\mathbf{x}_o) = \left( \frac{\partial f_i}{\partial x_1}(\mathbf{x}_0), \frac{\partial f_i}{\partial x_2}(\mathbf{x}_0), \cdots, \frac{\partial f_i}{\partial x_n}(\mathbf{x}_0) \right)$$

Así, el gradiente de una función de $n$ variables (definido en 3.9) es como la *derivada* de la función (en el sentido de la definición 3.19).

**Ejemplo 3.53.** *Derivar la función* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *dada por* $f(x, y) = (x^2 + 3y^2, x^3 + 2y^5)$

*Solución.* Las funciones coordenadas vienen definidas por:

$$f_1(x,y) = x^2 + 3y^2, \quad f_2(x,y) = x^3 + 2y^5$$

luego,

$$Jf = \begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x & 6y \\ 3x^2 & 10y^4 \end{pmatrix}$$

**Ejemplo 3.54.** *Hallar la derivada en el punto* $(0,0)$ *de la función* $f : \mathbb{R}^2 \to \mathbb{R}^3$ *definida por* $f(x,y) = \big(x+y, e^{x+y}, \operatorname{sen}(x+y)\big)$.

*Solución.* Las funciones coordenadas vienen definidas por:

$$f_1(x,y) = x+y, \quad f_2(x,y) = e^{x+y}, \quad f_3(x,y) = \operatorname{sen}(x+y)$$

luego,

$$Jf = \begin{pmatrix} \dfrac{\partial f_1}{\partial x}(0,0) & \dfrac{\partial f_1}{\partial y}(0,0) \\ \dfrac{\partial f_2}{\partial x}(0,0) & \dfrac{\partial f_2}{\partial y}(0,0) \\ \dfrac{\partial f_3}{\partial x}(0,0) & \dfrac{\partial f_3}{\partial y}(0,0) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ e^{x+y} & e^{x+y} \\ \cos(x+y) & \cos(x+y) \end{pmatrix}_{\substack{x=0 \\ y=0}} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$$

**Regla de la cadena. Caso general.** Esquemáticamente la situación sería:

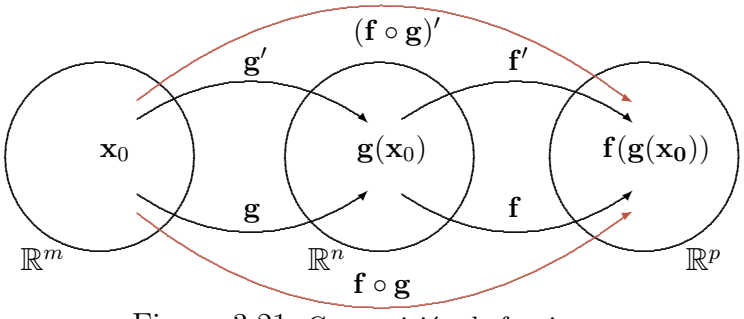



Figura 3.21: Composición de funciones

Supongamos que la función $\mathbf{g} : \mathbb{R}^m \to \mathbb{R}^n$ es diferenciable en $\mathbf{x}_0$ y la función $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$ es diferenciable en $\mathbf{g}(\mathbf{x}_0)$. Lo que dice la regla de la cadena es que la función compuesta $\mathbf{f}\circ\mathbf{g} : \mathbb{R}^m \to \mathbb{R}^p$ será diferenciable en $\mathbf{x}_0$ y que

$$(\mathbf{f}\circ\mathbf{g})'(\mathbf{x}_0) = \mathbf{f}'(\mathbf{g}(\mathbf{x}_0))\mathbf{g}'(\mathbf{x}_0)$$

entendiéndose el lado derecho de esta expresión como una composición de transformaciones lineales, o bien, en términos matriciales

$$J(f\circ g)(\mathbf{x}_0) = Jf(g(\mathbf{x}_0))Jg(\mathbf{x}_0)$$

entendiéndose el lado derecho de esta última expresión como una multiplicación de matrices. Es decir, *la (matriz que representa a la) derivada de la composición es igual al producto de las (matrices que representan a las) derivadas de las funciones componentes.*

**Ejemplo 3.55.** *Dadas las funciones* $g : \mathbb{R}^2 \to \mathbb{R}^3$ *y* $f : \mathbb{R}^3 \to \mathbb{R}^2$, *definidas por:*

$$\mathbf{g}(x,y) = (3x, xy, y^2)$$
$$\mathbf{f}(x,y,z) = (x^2 + 3y^2, 2xyz)$$

*Hallar la derivada de la composición* $\mathbf{f}\circ\mathbf{g} : \mathbb{R}^2 \to \mathbb{R}^2$

*(a) Mediante la regla de la cadena.*
*(b) Realizando previamente la composición.*

*Solución.* (a) La matriz jacobiana de la composición vendrá definida por el producto de las matrices jacobianas correspondientes, es decir,

$$
\begin{aligned}
J(\mathbf{f}\circ\mathbf{g}) &= J\mathbf{f}\big(\mathbf{g}(x,y)\big)\, J\mathbf{g}(x,y) = \\
&= \begin{pmatrix} \frac{\partial f_1}{\partial x}\big(\mathbf{g}(x,y)\big) & \frac{\partial f_1}{\partial y}\big(\mathbf{g}(x,y)\big) & \frac{\partial f_1}{\partial z}\big(\mathbf{g}(x,y)\big) \\ \frac{\partial f_2}{\partial x}\big(\mathbf{g}(x,y)\big) & \frac{\partial f_2}{\partial y}\big(\mathbf{g}(x,y)\big) & \frac{\partial f_2}{\partial z}\big(\mathbf{g}(x,y)\big) \end{pmatrix} \begin{pmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial x} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial x} \\ \frac{\partial g_3}{\partial x} & \frac{\partial g_3}{\partial x} \end{pmatrix} = \\
&= \begin{pmatrix} 2x & 6y & 0 \\ 2yz & 2xz & 2xy \end{pmatrix}_{\mathbf{g}(x,y)} \begin{pmatrix} 3 & 0 \\ y & x \\ 0 & 2y \end{pmatrix} = \\
&= \begin{pmatrix} 2(3x) & 6(xy) & 0 \\ 2(xy)(y^2) & 2(3x)(y^2) & 2(3x)(xy) \end{pmatrix} \begin{pmatrix} 3 & 0 \\ y & x \\ 0 & 2y \end{pmatrix} = \\
&= \begin{pmatrix} 6x & 6xy & 0 \\ 2xy^3 & 6xy^2 & 6x^2y \end{pmatrix} \begin{pmatrix} 3 & 0 \\ y & x \\ 0 & 2y \end{pmatrix} = \begin{pmatrix} 18x+6xy^2 & 6x^2y \\ 6xy^3+6xy^3 & 6x^2y^2+12x^2y^2 \end{pmatrix} = \\
&= \begin{pmatrix} 18x+6xy^2 & 6x^2y \\ 12xy^3 & 18x^2y^2 \end{pmatrix}
\end{aligned}
$$

(b) Aplicando previamente la composición, resulta:

$$
\begin{aligned}
(\mathbf{f}\circ\mathbf{g})(x,y) &= \mathbf{f}\big(\mathbf{g}(x,y)\big) = \mathbf{f}(3x, xy, y^2) = \big((3x)^2 + 3(xy)^2, 2(3x)(xy)(y^2)\big) = \\
&= (9x^2 + 3x^2y^2,\ 6x^2y^3)
\end{aligned}
$$

de donde, la derivada de la composición es la matriz:

$$
J(\mathbf{f}\circ\mathbf{g}) = \begin{pmatrix} 18x+6xy^2 & 6x^2y \\ 12xy^3 & 18x^2y^2 \end{pmatrix}
$$

que coincide con la obtenida mediante la regla de la cadena.

**Ejemplo 3.56.** *Dadas las funciones $g:\mathbb{R}^3\to\mathbb{R}^3$ y $f:\mathbb{R}^3\to\mathbb{R}^2$, definidas por:*

$$
\begin{aligned}
\mathbf{g}(x,y,z) &= (xy, yz, x+y+z) \\
\mathbf{f}(x,y,z) &= (x^2+y^3, y+z^2-1)
\end{aligned}
$$

*Hallar la derivada de la composición $\mathbf{f}\circ\mathbf{g}:\mathbb{R}^2\to\mathbb{R}^2$ en el punto $(1,1,1)$*

*(a) Mediante la regla de la cadena.*
*(b) Realizando previamente la composición.*

*Solución.* (a) La matriz jacobiana de la composición vendrá definida por el producto de las matrices jacobianas correspondientes. En este caso, a ser $\mathbf{g}(1,1,1) = (1,1,3)$, será,

$$
J(\mathbf{f}\circ\mathbf{g})(1,1,1) = J\mathbf{f}\big(\mathbf{g}(1,1,1)\big)\, J\mathbf{g}(1,1,1) = J\mathbf{f}(1,1,3)\, J\mathbf{g}(1,1,1)
$$

luego

$$
\begin{aligned}
J\mathbf{f}(1,1,3) &= \begin{pmatrix} 2x & 3y^2 & 0 \\ 0 & 1 & 2z \end{pmatrix}_{(1,1,3)} = \begin{pmatrix} 2 & 3 & 0 \\ 0 & 1 & 6 \end{pmatrix} \\
J\mathbf{g}(1,1,1) &= \begin{pmatrix} y & x & 0 \\ 0 & z & y \\ 1 & 1 & 1 \end{pmatrix}_{(1,1,1)} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}
\end{aligned}
$$

de donde, resulta:

$$J(\mathbf{f}\circ\mathbf{g})(1,1,1)=\begin{pmatrix}2&3&0\\0&1&6\end{pmatrix}\begin{pmatrix}1&1&0\\0&1&1\\1&1&1\end{pmatrix}=\begin{pmatrix}2&5&3\\6&7&7\end{pmatrix}$$

(b) Aplicando previamente la composición, resulta:

$$(\mathbf{f}\circ\mathbf{g})(x,y,x)=\mathbf{f}\big(\mathbf{g}(x,y,z)\big)=\mathbf{f}(xy,yz,x+y+z)=\big(x^2y^2+y^3z^3,\,yz+(x+y+z)^2-1\big)$$

de donde, la derivada de la composición es la matriz:

$$J(\mathbf{f}\circ\mathbf{g})=\begin{pmatrix}2xy^2 & 2x^2y^2+3y^2z^3 & 3y^3z^2\\ 2(x+y+z) & z+2(x+y+z) & y+2(x+y+z)\end{pmatrix}$$

de donde resulta:

$$(\mathbf{f}\circ\mathbf{g})(1,1,1)=\begin{pmatrix}2xy^2 & 2x^2y^2+3y^2z^3 & 3y^3z^2\\ 2(x+y+z) & z+2(x+y+z) & y+2(x+y+z)\end{pmatrix}_{(1,1,1)}=\begin{pmatrix}2&5&3\\6&7&7\end{pmatrix}$$

que coincide con la obtenida mediante la regla de la cadena.

El teorema que establece rigurosamente la regla de la cadena en el caso general puede enunciarse de la siguiente forma:

**Teorema 3.10.** *Sea* $\mathbf{f}:\mathcal{D}_f\subseteqq\mathbb{R}^n\to\mathbb{R}^p$ *una función definida en el abierto* $\mathcal{D}_f$ *de* $\mathbb{R}^n$ *y* $\mathbf{g}:\mathcal{D}_g\subseteqq\mathbb{R}^m\to\mathbb{R}^n$ *una función definida en el abierto* $\mathcal{D}_g$ *de* $\mathbb{R}^m$ *tal que* $\mathbf{g}(\mathcal{D}_g)\subseteqq\mathcal{D}_f$. *Si* $\mathbf{g}$ *es diferenciable en* $\mathbf{x}_0\in\mathcal{D}_g$ *y* $\mathbf{f}$ *es diferenciable en* $\mathbf{g}(\mathbf{x}_0)\in\mathcal{D}_f$ *entonces la función* $\mathbf{f}\circ\mathbf{g}:\mathcal{D}_g\subseteqq\mathbb{R}^m\to\mathbb{R}^p$ *es diferenciable en* $\mathbf{x}_0$ *y su derivada viene dada por la matriz:*

$$J(\mathbf{f}\circ\mathbf{g})(\mathbf{x}_0)=J\mathbf{f}\big(\mathbf{g}(\mathbf{x}_0)\big)\,J\mathbf{g}(\mathbf{x}_0)$$

La demostración puede verse en [**?**, Pita]

## 3.9. Funciones implícitas

### 3.9.1. Funciones de una variable

Dada la ecuación $F(x,y)=0$ nos planteamos la cuestión de si es siempre posible despejar $y$ en términos de $x$ y dejar establecida la función $y=f(x)$. La respuesta a esta cuestión es negativa. En efecto: De la expresión $x^2+y^2-1=0$ no podemos establecer *una* función $y=f(x)$ despejando $y$ en términos de $x$, ya que $x^2+y^2=1$ define dos funciones, a saber, $y=f_1(x)=\sqrt{1-x^2}$ e $y=f_2(x)=-\sqrt{1-x^2}$.
La cuestión anterior puede plantearse en términos de *curvas de nivel*. Así, podemos preguntarnos lo siguiente, dada la función $z=F(x,y)$, ¿es su nivel cero, $F(x,y)=0$, una curva que se puede ver como la gráfica de una función $y=f(x)$?.
Podemos hacer un planteamiento *local* de la cuestión anterior: Dada la función $z=F(x,y)$, sea $(x_0,y_0)$ un punto para el cual $F(x_0,y_0)=0$. De $F(x,y)=0$, ¿se puede obtener (despejando $y$ en términos de $x$) una función $y=f(x)$, *definida en una vecindad de* $x_0$, tal que $y_0=f(x_0)$. Cuando tal entorno y tal función $y=f(x)$ existen, decimos que la función

$y = f(x)$ está definida *implícitamente* por la expresión $F(x, y) = 0$, o bien que es una función implícita dada en $F(x, y) = 0$.
Gráficamente, la situación puede visualizase, en el ejemplo de la circunferencia, de la siguiente forma:

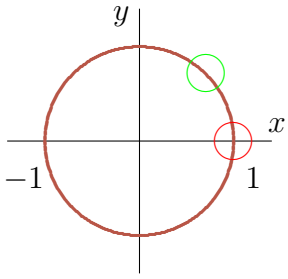



Figura 3.22: $x^2 + y^2 = 1$

La ecuación $x^2+y^2 = 1$, en general no define una sola función $y = f(x)$, sin embargo, si nos limitamos a un entorno del punto $(1/\sqrt{2}, 1/\sqrt{2})$, sí que queda definida una sola función (reduciéndonos a ese entorno), mientras que en un entorno del punto $(1, 0)$ no queda definida dicha función (ya que cada punto tendría dos imágenes, al haber dos puntos en la misma vertical).

**Planteamiento práctico.** Supongamos la ecuación $F(x, y) = 0$, si supiéramos despejar $y$ en términos de $x$, tendríamos $y = f(x)$, y podríamos calcular $\dfrac{dy}{dx} =?$.

$$F(x, y) = 0 \quad \rightarrow \quad y = f(x) \quad \rightarrow \quad \frac{dy}{dx} =?$$

El problema se presenta cuando no sabemos o no queremos despejar $y$ en términos de $x$ ¿cómo podemos calcular dicha derivada?.
Partamos de que $F(x, y) = 0$, será $dF(x, y) = 0$, de donde

$$dF = F_x dx + F_y dy = 0 \quad \rightarrow \quad \frac{dF}{dx} = F_x \frac{dx}{dx} + F_y \frac{dy}{dx} = 0 \quad \rightarrow \quad F_x + F_y \frac{dy}{dx} = 0$$

de donde,

$$\frac{dy}{dx} = \frac{-F_x(x, y)}{F_y(x, y)}$$

**Teorema 3.11 (De la función implícita).** *Supongamos la función $z = F(x, y)$. Sea $(x_0, y_0) \in \mathbb{R}^2$ un punto tal que $F(x_0, y_0) = 0$. Supongamos que la función $F$ tiene derivadas parciales continuas en alguna bola $\mathcal{B}$ con centro en $(x_0, y_0)$ y que $F_y(x_0, y_0) \neq 0$. Entonces $F(x, y) = 0$ se puede resolver para $y$ en términos de $x$ y definir así una función $y = f(x)$ con dominio en una vecindad $V$ de $x_0$, tal que $y_0 = f(x_0)$, la cual tiene derivada continua en $V$ que puede calcularse como*

$$y' = f'(x) = \frac{-F_x(x, y)}{F_y(x, y)}, \qquad x \in V$$

Esquemáticamente sería:

$$\left.\begin{array}{l} F(x_0, y_0) = 0 \\ F_x \text{ y } F_y \text{ continuas} \\ F_y(x_0, y_0) \neq 0 \end{array}\right\} \Rightarrow y = f(x), \text{ tal que } y_0 = f(x_0), \text{ y además } y' = \frac{-F_x(x, y)}{F_y(x, y)}$$

Debe tenerse en cuenta que el teorema de la función implícita es un teorema de existencia. Dice que existe una función $y = f(x)$ definida implícitamente por $F(x, y) = 0$, sin

embargo, no dice cómo se puede determinar dicha función. Es más, podría ocurrir que la estructura algebraica de la ecuación $F(x, y) = 0$ no permitiera el despeje de $y$ en términos de $x$. De ahí la importancia del teorema que afirma que, si se cumplen las hipótesis, aunque no podamos despejar $y$ en términos de $x$, tal función existe, y además el teorema nos dice cómo calcular su derivada.
También hay que advertir que se trata de un teorema *local*. Nos asegura la existencia de la función $y = f(x)$, o bien la posibilidad de despejar $y$ en términos de $x$ a partir de $F(x, y) = 0$, pero *solamente en las cercanías del punto* $(x_0, y_0)$.
También debe tenerse en cuenta que para el caso de una variable existe la posibilidad de calcular la derivada de la función implícita aplicando directamente las reglas de derivación sobre la ecuación $F(x, y) = 0$. Simplemente habrá que tener en cuenta que cada vez que derivemos la variable $y$ habrá que multiplicar por su derivada $y'$.

**Ejemplo 3.57.** *Calcular $y'$, siendo* $y^3 + y^2 - 5y - x^2 + 4 = 0$

*(a) Aplicando las reglas de derivación.*
*(b) aplicando el teorema de la función implícita.*

*Solución.* (a) Mediante las reglas de derivación tenemos:

$$3y^2y' + 2yy' - 5y' - 2x = 0 \quad \rightarrow \quad y' = \frac{2x}{3y^2 + 2y - 5}$$

(b) Mediante el teorema de la función implícita, hacemos $F(x, y) = y^3 + y^2 - 5y - x^2 + 4$, con lo cual,

$$\left.\begin{array}{l} F_x = -2x \\ F_y = 3y^2 + 2y - 5 \end{array}\right\} \rightarrow y' = \frac{-F_x}{F_y} = \frac{2x}{3y^2 + 2y - 5}$$

**Ejemplo 3.58.** *Razonar si la ecuación* $e^y = x^3 + 3y$ *determina una función* $y = f(x)$ *que cumple* $f(1) = 0$ *y que es derivable en* $x = 1$. *Calcular* $f'(1)$ *y la recta tangente a* $f(x)$ *en* $(1, 0)$.

*Solución.* El teorema de la función implícita afirma que una ecuación $F(x, y) = 0$ define una función $y = f(x)$ derivable en un punto $x_0$, si se cumplen las tres condiciones siguientes: El punto $P(x_0, y_0)$, con $y_0 = f(x_0)$, cumple la ecuación $F(P) = 0$, las derivadas parciales de $F$ en $P$ son continuas, y $F_y(P) \neq 0$.
El punto $P(1, 0)$ cumple la ecuación dada, en efecto:

$$e^0 = 1^3 + 3 \cdot 0 = 1 + 0 = 1$$

Las derivadas parciales de $F(x, y) = e^y - x^3 - 3y$ son continuas en todo $\mathbb{R}^2$ y por tanto en $P$, por tratarse de una función polinómica y una exponencial, en efecto:

$$\begin{array}{l} F_x(x, y) = -3x^2 \\ F_y(x, y) = e^y - 3 \end{array}$$

Y además: $F_y(1, 0) = e^0 - 3 = 1 - 3 = -2 \neq 0$. Luego podemos afirmar la existencia de la función $y = f(x)$ que cumple $f(1) = 0$ y que es derivable en $x = 1$.
En virtud del teorema de la función implícita tenemos:

$$f'(x) = \frac{-F_x(x, y)}{F_y(x, y)} = \frac{3x^2}{e^y - 3} \rightarrow f'(1) = \frac{3}{-2} = -\frac{3}{2}$$

y teniendo en cuenta que la ecuación de la recta tangente a la curva $y = f(x)$ en el punto $P(1,0)$ viene definida por:

$$y = f(1) + f'(1) \cdot (x-1)$$

resulta: $\quad y = 0 - \dfrac{3}{2}(x-1) \quad \rightarrow \quad 3x + 2y = 3$

**Ejemplo 3.59.** *Dada la ecuación* $x^2 + y^2 = 2$

*(a) Justificar que la ecuación* $x^2 + y^2 = 2$ *define, en un entono del punto* $(1,1)$*, una función* $y = y(x)$
*(b) Calcular* $y'(1)$*,* $y''(1)$

*Solución.* (a) Para que quede garantizada la existencia de la función $y = y(x)$, deberán cumplirse las tres condiciones siguientes:
El punto $(1,1)$ cumple la ecuación. En efecto $1^2 + 1^2 = 1 + 1 = 2$
Las derivadas parciales de la función $F(x,y) = x^2 + y^2 - 2$ son continuas en un entorno del punto $(1,1)$. En efecto,

$$F_x = 2x \quad F_y = 2y \quad \text{que son continuas en } \mathbb{R}^2$$

$F_y(1,1) \neq 0$. En efecto, $F_y(1,1) = 2 \neq 0$
Luego existe la función $y = y(x)$. Que en este caso es $y = +\sqrt{2 - x^2}$.

(b) Para calcular las derivadas podemos aplicar las reglas de derivación. En efecto,

$$x^2 + y^2 = 2 \quad \rightarrow \quad 2x + 2yy' = 0 \quad \rightarrow \quad y' = \frac{-x}{y} \quad \rightarrow \quad y'(1) = \frac{-1}{1} = -1$$

y derivando nuevamente,

$$2 + 2y'y' + 2yy'' = 0 \quad \rightarrow \quad 1 + (y')^2 + yy'' = 0 \quad \rightarrow \quad y'' = \frac{-1 - (y')^2}{y} = \frac{-2}{1} = -2$$

**nota:** Hay que aclarar que los papeles de las letras $x$ e $y$ son perfectamente intercambiables. Así, si la función $z = F(x,y)$ es tal que en el punto $(x_0, y_0)$ vale cero, $F(x_0, y_0) = 0$, que en una bola con centro en $(x_0, y_0)$ tiene derivadas parciales continuas y que su derivada *respecto de* $x$ es distinta de cero en $(x_0, y_0)$, es decir, $F_x(x_0, y_0) \neq 0$. Entonces, el teorema de la función implícita garantiza la existencia de una función $x = g(y)$ tal que $x_0 = g(y_0)$, para $y$ en una vecindad de $y_0$, y además su derivada, en esa vecindad, es:

$$\frac{dx}{dy} = g'(y) = \frac{-F_y(x,y)}{F_x(x,y)}$$

Es evidente que si en el punto $(x_0, y_0)$ se tiene

$$F_x(x_0, y_0) \neq 0 \quad \text{y} \quad F_y(x_0, y_0) \neq 0$$

entonces el teorema de la función implícita nos dice que en los alrededores de $(x_0, y_0)$, la gráfica de la curva $F(x_0, y_0) = 0$ se puede ver como la gráfica de una función $y = f(x)$ o bien como la gráfica de una función $x = f(y)$.

**Puntos regulares.** Si $z = F(x, y)$ es una función tal que $F(x_0, y_0) = 0$ y en una bola $\mathcal{B}$ con centro en $(x_0, y_0)$ las parciales $F_x$ y $F_y$ son continuas y si $F_x(x_0, y_0) \neq 0$ ó $F_y(x_0, y_0) \neq 0$ (o equivalentemente si $\big(F_x(x_0, y_0)\big)^2 + \big(F_y(x_0, y_0)\big)^2 \neq 0$) entonces $(x_0, y_0)$ se dice que es un punto *regular* de la curva $F(x, y) = 0$.
Es decir, los puntos regulares son aquellos puntos $(x_0, y_0)$ tales que en una bola con centro en ellos, a partir de $F(x, y) = 0$, se puede despejar $x$ o $y$ en términos de $y$ o $x$, respectivamente, y establecer una función con derivada continua $x = g(y)$ ó $y = g(x)$.
Si el punto $(x_0, y_0)$ no es regular de $F(x, y) = 0$, se dice que es un punto *singular*.

### 3.9.2. Funciones de dos variables

Dada una ecuación $F(x, y, z) = 0$ nos planteamos ahora la cuestión de ¿cuándo de dicha expresión se puede despejar $z$ en términos de $x$ e $y$ y formar así una función $z = f(x, y)$? Es decir, ¿cuándo el nivel cero, $F(x, y, z) = 0$, de la función $w = F(x, y, z)$ se puede ver como la gráfica de una función $z = f(x, y)$? **Planteamiento práctico.** Supongamos la ecuación $F(x, y, y) = 0$, si supiéramos despejar $z$ en términos de $x$ e $y$, tendríamos $z = f(x, y)$, y podríamos calcular $\dfrac{\partial z}{\partial x} =?$ y $\dfrac{\partial z}{\partial y} =?$.

$$F(x, y, z) = 0 \quad \rightarrow \quad z = f(x, y) \quad \rightarrow \quad \begin{cases} \dfrac{\partial z}{\partial x} =? \\ \dfrac{\partial z}{\partial y} =? \end{cases}$$

El problema se presenta cuando no sabemos o no queremos despejar $z$ en términos de $x$ e $y$ ¿cómo podemos calcular dichas derivadas?.
Partamos de que $F(x, y, z) = 0$, será $dF(x, y, z) = 0$, de donde

$$dF = F_x dx + F_y dy + F_z dz = 0 \quad \rightarrow \quad \begin{cases} \dfrac{dF}{dx} = F_x \dfrac{dx}{dx} + F_y \dfrac{dy}{dx} + F_z \dfrac{dz}{dx} = 0 \\ \dfrac{dF}{dy} = F_x \dfrac{dx}{dy} + F_y \dfrac{dy}{dy} + F_z \dfrac{dz}{dy} = 0 \end{cases}$$

Ahora bien,
(a) Para $y = cte.$, en el primer caso resulta:

$$F_x + 0 + F_z \frac{\partial z}{\partial x} = 0 \rightarrow \frac{\partial z}{\partial x} = \frac{-F_x(x, y, z)}{F_z(x, y, z)}$$

(b) Para $x = cte.$, en el segundo caso resulta:

$$0 + F_y + F_z \frac{\partial z}{\partial y} = 0 \rightarrow \frac{\partial z}{\partial y} = \frac{-F_y(x, y, z)}{F_z(x, y, z)}$$

**Teorema 3.12 (De la función implícita).** *Supongamos la función $w = F(x, y, z)$. Sea $\mathbf{p} = (x_0, y_0, z_0) \in \mathbb{R}^3$ un punto tal que $F(x_0, y_0, z_0) = 0$. Supongamos que la función $F$*

*tiene derivadas parciales continuas en alguna bola $\mathcal{B}$ con centro en $\mathbf{p}$ y que $F_z(\mathbf{p}) \neq 0$. Entonces $F(x,y,z) = 0$ se puede resolver para $z$ en términos de $x$ e $y$ y definir así una función $z = f(x,y)$ con dominio en una vecindad $V$ de $(x_0, y_0)$, tal que $z_0 = f(x_0, y_0)$, la cual tiene derivadas continuas en $V$ que pueden calcularse como*

$$\frac{\partial z}{\partial x} = \frac{-F_x(x,y,z)}{F_z(x,y,z)}, \quad \frac{\partial z}{\partial y} = \frac{-F_y(x,y,z)}{F_z(x,y,z)} \qquad (x,y) \in V$$

Esquemáticamente sería:

$$\left.\begin{array}{l} F(x_0,y_0,z_0) = 0 \\ F_x,\ F_y,\ F_z \text{ continuas} \\ F_z(x_0,y_0,z_0) \neq 0 \end{array}\right\} \Rightarrow z = f(x,y), \text{ tal que } z_0 = f(x_0,y_0), \text{ y } \left\{\begin{array}{l} \dfrac{\partial z}{\partial x} = \dfrac{-F_x(x,y,z)}{F_z(x,y,z)} \\ \dfrac{\partial z}{\partial y} = \dfrac{-F_y(x,y,z)}{F_z(x,y,z)} \end{array}\right.$$

Las derivadas parciales también pueden calcularse a partir de las reglas ordinarias de derivación, suponiendo una de las variables constante, con la única observación de multiplicar por la derivada parcial correspondiente, cada vez que derivemos $z$.

**Ejemplo 3.60.** *Calcular* $\dfrac{\partial z}{\partial x}$ *y* $\dfrac{\partial z}{\partial x}$ *siendo* $3x^2z - x^2y^2 + 2z^3 + 3yz - 5 = 0$

*(a) Mediante las reglas ordinarias de derivación.*
*(b) Mediante las fórmulas de la función implícita.*

*Solución.* (a) Mediante las reglas ordinarias de derivación tenemos:

$$\begin{array}{rcl} 6xz + 3x^2z_x - 2xy^2 + 6z^2z_x + 3yz_x = 0 & \rightarrow & z_x = \dfrac{-6xz + 2xy^2}{3x^2 + 6z^2 + 3y} \\ 3x^2z_y + 2x^2y + 6z^2z_y + 3z + 3yz_y = 0 & \rightarrow & z_y = \dfrac{2x^2y - 3z}{3x^2 + 6z^2 + 3y} \end{array}$$

(b) Mediante el teorema de la función implícita, hacemos:

$$F(x,y,z) = 3x^2z - x^2y^2 + 2z^3 + 3yz - 5,$$

con lo cual,

$$\left.\begin{array}{l} F_x(x,y,z) = 6xz - 2xy^2 \\ F_y(x,y,z) = -2x^2y + 3z \\ F_z(x,y,z) = 3x^2 + 6z^2 + 3y \end{array}\right\} \begin{array}{l} \dfrac{\partial z}{\partial x} = \dfrac{-F_x}{F_z} = \dfrac{-6xz + 2xy^2}{3x^2 + 6z^2 + 3y} \\ \dfrac{\partial z}{\partial y} = \dfrac{-F_y}{F_z} = \dfrac{2x^2y - 3z}{3x^2 + 6z^2 + 3y} \end{array}$$

**Ejemplo 3.61.** *Razonar si la ecuación* $x^3 + 2y^3 + z^3 - 3xyz - 2y + 1 = 0$; *determina una función* $z = f(x,y)$ *en un entorno del punto* $(0,1)$. *Calcular* $\dfrac{\partial f}{\partial x}(0,1)$, $\dfrac{\partial f}{\partial y}(0,1)$ *y la ecuación del plano tangente a la superficie en el punto* $(0,1,f(0,1))$.

*Solución.* El teorema de la función implícita afirma que una ecuación $F(x,y,z)=0$ define una función $z=f(x,y)$ diferenciable en un punto $\mathbf{p}=(x,y)$, si se cumplen las tres condiciones siguientes: el punto $P=(x,y,z)$ cumple la ecuación $F(P)=0$, las derivadas parciales de $F$ en $P$ son continuas, y $F_z(P)\neq 0$.
En nuestro caso, el punto $P$ que cumple la ecuación dada viene definido por las coordenadas:

$$x=0,\ \ y=1 \quad\rightarrow\quad 2+z^3-2+1=0 \quad\rightarrow\quad z=-1$$

Las derivadas parciales de $F(x,y,z)=x^3+2y^3+z^3-3xyz-2y+1$ son continuas en todo $\mathbb{R}^3$ y por tanto en $P$, en efecto:

$$\begin{aligned} F_x(x,y,z)&=3x^2-3yz\\ F_y(x,y,z)&=6y^2-3xz-2\\ F_z(x,y,z)&=3z^2-3xy \end{aligned}$$

Y además: $F_z(0,1,-1)=3\neq 0$ Luego podemos afirmar la existencia de la función $z=f(x,y)$ diferenciable en el punto $\mathbf{p}=(0,1)$.
En virtud del teorema de la función implícita tenemos:

$$\frac{\partial f}{\partial x}(x,y)=\frac{-F_x(x,y,z)}{F_z(x,y,z)}=\frac{-3x^2+3yz}{3z^2-3xy} \quad\rightarrow \frac{\partial f}{\partial x}(0,1)=\frac{-F_x(0,1,-1)}{F_z(0,1,-1)}=\frac{-3}{3}=-1$$
$$\frac{\partial f}{\partial y}(x,y)=\frac{-F_y(x,y,z)}{F_z(x,y,z)}=\frac{-6y^2+3xz+2}{3z^2-3xy} \rightarrow \frac{\partial f}{\partial y}(0,1)=\frac{-F_y(0,1,-1)}{F_z(0,1,-1)}=\frac{-4}{3}$$

y teniendo en cuenta que la ecuación del plano tangente a la superficie $z=f(x,y)$ en el punto $P(0,1,-1)$ viene definida por:

$$z=f(0,1)+\frac{\partial f}{\partial x}(0,1)\cdot(x-0)+\frac{\partial f}{\partial y}(0,1)\cdot(y-1)$$

resulta: $\quad z=-1-x-\dfrac{4}{3}(y-1)$, o bien, simplificando $3x+4y+3z=1$

**Ejemplo 3.62.** *Sea $F$ una función diferenciable y con derivadas parciales continuas tal que $F(9,2,-4)=0$ y $\nabla F(9,2,-4)=(-1,0,5)$. ¿La ecuación $F(x,y,z)=0$ define una función diferenciable $z=f(x,y)$ tal que $f(9,2)=-4$? En caso afirmativo, determina $\dfrac{\partial f}{\partial x}(9,2)$ y $\dfrac{\partial f}{\partial y}(9,2)$ y, con su ayuda, calcula la ecuación del plano tangente a la superficie $z=f(x,y)$ en el punto $(9,2,-4)$.*

*Solución.* El teorema de la función implícita afirma que una ecuación $F(x,y,z)=0$ define una función $z=f(x,y)$ diferenciable en un punto $\mathbf{p}=(x,y)$, si se cumplen las tres condiciones siguientes: el punto $P=(x,y,z)$ cumple la ecuación $F(P)=0$, las derivadas parciales de $F$ en $P$ son continuas, y $F_z(P)\neq 0$.
En nuestro caso, el punto $P$ que cumple la ecuación dada viene definido por $P(9,2,-4)$
Por hipótesis, las derivadas parciales de $F$ son continuas en todo $\mathbb{R}^3$ y por tanto en $P$. Sus valores vienen dados por las coordenadas del gradiente:

$$\nabla F=(F_x,F_y,F_z) \quad\rightarrow\quad F_x(9,2,-4)=-1,\quad F_y(9,2,-4)=0,\quad F_z(9,2,-4)=5$$

Y además: $F_z(9,2,-4) = 5 \neq 0$ Luego podemos afirmar la existencia de la función $z = f(x,y)$ diferenciable en el punto $\mathbf{p} = (9,2)$.
En virtud del teorema de la función implícita tenemos:

$$\frac{\partial f}{\partial x}(x,y) = \frac{-F_x(x,y,z)}{F_z(x,y,z)} \quad \rightarrow \frac{\partial f}{\partial x}(9,2) = \frac{-F_x(9,2,-4)}{F_z(9,2,-4)} = \frac{1}{5}$$
$$\frac{\partial f}{\partial y}(x,y) = \frac{-F_y(x,y,z)}{F_z(x,y,z)} \quad \rightarrow \frac{\partial f}{\partial y}(9,2) = \frac{-F_y(9,2,-4)}{F_z(9,2,-4)} = \frac{0}{5} = 0$$

y teniendo en cuenta que la ecuación del plano tangente a la superficie $z = f(x,y)$ en el punto $P(9,2,-4)$ viene definida por:

$$z = f(9,2) + \frac{\partial f}{\partial x}(9,2) \cdot (x-9) + \frac{\partial f}{\partial y}(9,2) \cdot (y-2)$$

resulta: $\quad z = -4 + \frac{1}{5}(x-9) + 0(y-2)$, o bien, simplificando $x - 5z = 29$

# 3.10. Extremos de las funciones de varias variables

## 3.10.1. Introducción

### Funciones de una variable

La función $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$, definida en el intervalo abierto $I$ de $\mathbb{R}$, se dice que tiene un máximo (mínimo) local o relativo en un punto $x_0 \in I$ si en un entorno $\mathcal{V}_{x_0}$ de $x_0$ se tiene $f(x_0) \geq f(x)$ ($f(x_0) \leq f(x)$, respectivamente) para todo $x$ en $\mathcal{V}_{x_0}$. En otras palabras, $f$ tiene un máximo (mínimo) local en $x_0$ si $f(x_0)$ es el valor más grande (más pequeño) de la función en torno a $x_0$.

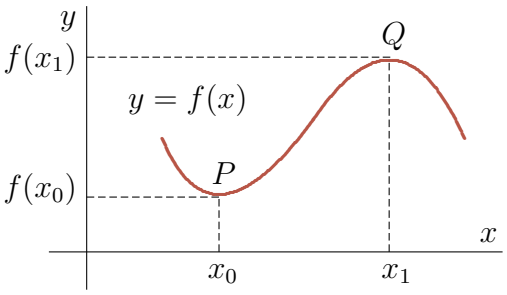



Figura 3.23: $y = f(x)$ tiene un mínimo local en $x = x_0$ y un máximo local en $x = x_1$.

Una condición necesaria para que la función $f$ tenga un extremo (máximo o mínimo) local en $x_0$ es que, si $f'(x_0)$ existe, entonces $f'(x_0) = 0$. Geométricamente esta condición significa que si la gráfica de la función es suave en $x_0$, su recta tangente en dicho punto debe ser horizontal. Es decir, en un extremo local la gráfica de la función o no tiene recta tangente, o si la tiene es una recta horizontal.

### 3.10.2. Definiciones

**Máximos y mínimos absolutos.**
Los valores $f(x_0, y_0)$ y $f(x_1, y_1)$ tal que $f(x_0, y_0) \leq f(x, y) \leq f(x_1, y_1)$ para todo $(x, y)$ en $\mathcal{D}$ se conocen como mínimo absoluto y máximo absoluto de $f$ en la región $\mathcal{D}$.

**Teorema de existencia del máximo y del mínimo absoluto.** Toda función continua, definida en una región *cerrada* y *acotada*, alcanza, en dicha región, un valor máximo absoluto y un mínimo absoluto.

**Máximo y mínimo relativo.**

**Definición 3.20 (Extremo local).** *Sea $f : \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ una función definida en el conjunto abierto $\mathcal{D}$ de $\mathbb{R}^n$. Se dice que $f$ tiene un máximo (mínimo) local o relativo en el punto $\mathbf{x}_0 \in \mathcal{D}$ si $f(\mathbf{x}_0) \geq f(\mathbf{x})$ ($f(\mathbf{x}_0) \leq f(\mathbf{x})$ respectivamente) para toda $\mathbf{x}$ en una bola $\mathcal{B}$ de centro en $\mathbf{x}_0$.*

$$f(x_0, y_0) \text{ es mínimo relativo} \quad \Leftrightarrow \quad f(x, y) \geq f(x_0, y_0) \quad \forall (x, y) \in \mathcal{B}_{\mathbf{x}_0}$$
$$f(x_0, y_0) \text{ es máximo relativo} \quad \Leftrightarrow \quad f(x, y) \leq f(x_0, y_0) \quad \forall (x, y) \in \mathcal{B}_{\mathbf{x}_0}$$

Es decir, al igual que en el caso de funciones de una variable, la función $f$ tendrá un máximo (mínimo) local en $\mathbf{x}_0 \in \mathcal{D}$ si $f(\mathbf{x}_0)$ es el valor más grande (más pequeño, respectivamente) de todos los valores de $f(\mathbf{x})$ para $\mathbf{x}$ en una bola de centro en $\mathbf{x}_0$.

**Puntos críticos.** Se llaman puntos críticos de una función a aquellos puntos en los que el gradiente vale cero o no está definido, es decir,

1. Puntos en los que todas las derivadas parciales valen cero (simultáneamente).

$$f_x(x_0, y_0) = 0 \quad \text{y} \quad f_y(x_0, y_0) = 0$$

2. Puntos en los que alguna de las derivadas parciales no está definida.

$$f_x(x_0, y_0) \quad \text{o} \quad f_y(x_0, y_0) \quad \text{no existe}$$

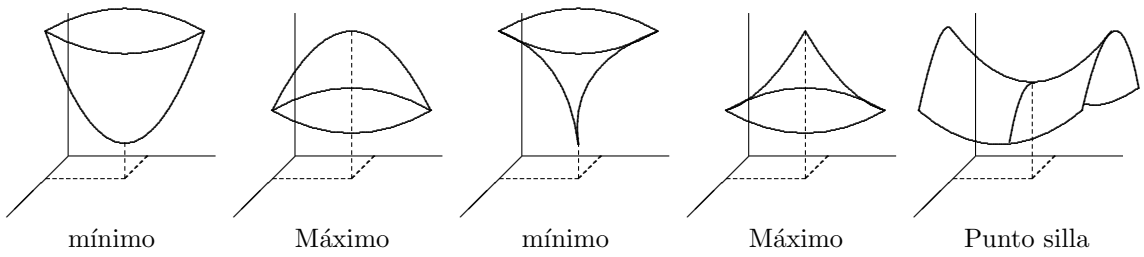



Figura 3.24: Puntos críticos.

**Teorema de los máximos y mínimos relativos.** Los extremos de una función continua se producen solamente en los puntos críticos. Es decir, en un máximo y en un mínimo relativo, las derivadas parciales o no existen o valen cero.
En efecto, $g(x) = f(x, x_0)$ tiene un extremo relativo en $x_0$ entonces $g'(x_0) = 0$ o no existe. Pero $g'(x_0) = f_x(x_0, y_0)$

De la misma forma, $h(y) = f(x_0, y)$ tiene un extremo relativo en $y_0$ entonces $h'(y_0) = 0$ o no existe. Pero $h'(y_0) = f_y(x_0, y_0)$

**Estudio de la naturaleza de los extremos relativos, a partir de la propia función.**

**Ejemplo 3.63.** *Determinar los extremos relativos de la función:*

$$f(x, y) = 2x^2 + y^2 + 8x - 6y + 20$$

*Solución.* Hallamos las derivadas parciales y determinamos los puntos críticos: Puntos en los que alguna de las derivadas parciales no está definida, y puntos en los que las dos derivadas parciales son nulas. En el caso de funciones polinómicas las parciales están siempre definidas, por tanto:

$$\left.\begin{array}{l} f_x(x,y) = 4x + 8 \\ f_y(x,y) = 2y - 6 \end{array}\right\} \left.\begin{array}{l} 4x + 8 = 0 \\ 2y - 6 = 0 \end{array}\right\} \left.\begin{array}{l} x = -2 \\ y = 3 \end{array}\right\} \mathbf{p}(-2, 3)$$

Para estudiar la naturaleza del punto crítico $\mathbf{p}(-2, 3)$ comparamos el valor de la función en el punto crítico con el valor de la función en los alrededores del punto crítico.

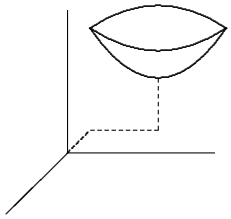


Figura 3.25: mínimo.

$$\begin{aligned} f(-2,3) &= 3 \\ f(x,y) &= 2x^2 + y^2 + 8x - 6y + 20 = 2(x^2 + 4x) + y^2 - 6y + 20 = \\ &= 2(x^2 + 4x + 4 - 4) + y^2 - 6y + 9 - 9 + 20 = \\ &= 2(x+2)^2 - 8 + (y-3)^2 - 9 + 20 = \\ &= 2(x+2)^2 + (y-3)^2 + 3 \geq 3 \end{aligned}$$

Resulta que para cualquier punto $(x, y)$ se tiene que $f(x, y) \geq f(-2, 3)$, luego $f(-2, 3)$ es un mínimo (absoluto).

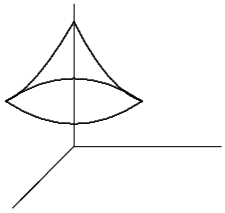
**Ejemplo 3.64.** *Determinar los extremos relativos de la función:*

$$f(x, y) = 1 - \left(x^2 + y^2\right)^{1/3}$$

*Solución.* Hallamos las derivadas parciales y determinamos los puntos críticos: Puntos en los que alguna de las derivadas parciales no está definida, y puntos en los que las dos derivadas parciales son nulas.

$$\left.\begin{array}{l} f_x(x,y) = -\dfrac{1}{3}2x\left(x^2+y^2\right)^{-2/3} = \dfrac{-2x}{3\sqrt[3]{(x^2+y^2)^2}} \\ f_y(x,y) = -\dfrac{1}{3}2y\left(x^2+y^2\right)^{-2/3} = \dfrac{-2y}{3\sqrt[3]{(x^2+y^2)^2}} \end{array}\right\}$$

Donde el único punto crítico es el punto $(0, 0)$, en donde no están definidas ninguna de las dos derivadas parciales. Para estudiar la naturaleza del punto crítico $\mathbf{p}(0, 0)$ comparamos el valor de la función en el punto crítico con el valor de la función en los alrededores del punto crítico.



$$f(0,0) = 1$$
$$f(x,y) = 1 - \sqrt[3]{x^2+y^2} \leq 1$$

Resulta que para cualquier punto $(x,y)$ del plano se tiene que $f(x,y) \leq f(0,0)$, luego $f(0,0)$ es un máximo (absoluto).

Figura 3.26: máximo.

### Criterio de los cortes con los planos verticales

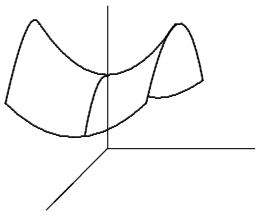
Para comprobar que en un punto crítico no existe ni máximo ni mínimo cortamos la superficie mediante planos verticales que pasen por el punto crítico, si las curvas resultantes tienen puntos por encima y por debajo del punto crítico, entonces se trata de un punto silla. El método es sólo refutativo y no permite afirmar la existencia de máximo o mínimo, sino sólo de punto silla.

**Ejemplo 3.65.** *Determinar los extremos relativos de la función:* $f(x,y) = 1 - x^2 + y^2$

*Solución.* Hallamos las derivadas parciales y determinamos los puntos críticos: Puntos en los que alguna de las derivadas parciales no está definida, y puntos en los que las dos derivadas parciales son nulas. En el caso de funciones polinómicas las parciales están siempre definidas, por tanto:

$$\left.\begin{array}{l} f_x(x,y) = -2x \\ f_y(x,y) = 2y \end{array}\right\} \left.\begin{array}{l} x = 0 \\ y = 0 \end{array}\right\} \mathbf{p}(0,0)$$

Donde el único punto crítico es el punto $(0,0)$. Para estudiar la naturaleza del punto crítico $\mathbf{p}(0,0)$ comparamos el valor de la función en el punto crítico con el valor de la función en los alrededores del punto crítico.



$$f(0,0) = 1$$
$$f(x,y) = 1 - x^2 + y^2 = \left\{\begin{array}{l} f(x,0) = 1 - x^2 < 1 \\ f(0,y) = 1 + y^2 > 1 \end{array}\right.$$

Figura 3.27: Punto silla.

Para estudiar la naturaleza del punto crítico $\mathbf{p}(0,0)$ se ha cortado la superficie mediante los dos planos verticales $y = 0$ y $x = 0$ con lo que se han obtenido las curvas $f(x,0) = 1 - x^2$ y $f(0,y) = 1 + y^2$. Para la primera, el punto $x = 0$ es un máximo; y para la segunda, el punto $y = 0$ es un mínimo. Luego el punto crítico $\mathbf{p}(0,0)$ es un punto silla.

### 3.10.3. Criterio del hessiano

**Caso particular de funciones de dos variables.**

Para estudiar la naturaleza de los puntos críticos, formamos la matriz hessiana en dichos puntos:

$$Hf(x,y) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

Y comparamos los signos de los dos determinantes principales:

$$D_1 = f_{xx}, \qquad D_2 = |Hf(x,y)| = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} = f_{xx}f_{yy} - \left(f_{xy}\right)^2$$

Resultando, para el caso de dos variables:

| H | $f_{xx}$ |
|---|---|
| + | + mínimo<br>− Máximo |
| − | Silla |
| 0 | Duda |

**Nota:** El criterio del hessiano puede fallar a la hora de estudiar la naturaleza de los extremos relativos de dos formas: Bien porque alguna de las derivadas parciales no esté definida, entonces no se puede aplicar el criterio; o bien, porque el hessiano sea cero, en cuyo caso el criterio no da información. En este caso habrá que aplicar otras técnicas, como puede ser el cortar la superficie por planos verticales.

**Ejemplo 3.66.** *Encontrar los extremos relativos de la función:*

$$f(x,y) = -x^3 + 4xy - 2y^2 + 1$$

*Solución.* Hallamos las derivadas parciales y determinamos los puntos críticos: Puntos en los que alguna de las derivadas parciales no está definida, y puntos en los que las dos derivadas parciales son nulas. En el caso de funciones polinómicas las parciales están siempre definidas, por tanto:

$$\left.\begin{array}{l} f_x(x,y) = -3x^2 + 4y \\ f_y(x,y) = 4x - 4y \end{array}\right\} \left.\begin{array}{l} -3x^2 + 4y = 0 \\ 4x - 4y = 0 \end{array}\right\} \left.\begin{array}{l} -3x^2 + 4y = 0 \\ x = y \end{array}\right\} \left.\begin{array}{l} x(-3x+4) = 0 \\ x = y \end{array}\right\} \begin{array}{l} x = 0 \\ x = \frac{4}{3} \end{array}$$

Luego los puntos críticos son $\mathbf{p}(0,0)$ y $\mathbf{q}(4/3,4/3)$.
Para estudiar la naturaleza de los puntos críticos, hallamos la matriz hessiana en cada uno de ellos:

$$Hf(x,y) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} -6x & 4 \\ 4 & -4 \end{pmatrix}$$

Con lo cual resulta:

$|Hf(0,0)| = \begin{vmatrix} 0 & 4 \\ 4 & -4 \end{vmatrix} = -16 < 0 \Rightarrow \mathbf{p}(0,0)$ es un punto silla

$|Hf(4/3,4/3)| = \begin{vmatrix} -8 & 4 \\ 4 & -4 \end{vmatrix} = 32 - 16 = 13 > 0$ y $f_{xx} = -8 < 0 \Rightarrow \mathbf{q}(4/3,4/3)$ es un máximo relativo.

**Ejemplo 3.67.** *Hallar y clasificar todos los puntos críticos de la función*

$$f(x,y) = x^4 + y^4 + 6x^2y^2 + 8x^3$$

*Solución.* Los puntos críticos son aquellos que anulan simultáneamente las dos derivadas parciales, o bien donde alguna de las derivadas parciales no existe. Al ser la función dada una función polinómica es diferenciable en todo $\mathbb{R}^2$, luego los puntos críticos vendrán dado por las soluciones del siguiente sistema de ecuaciones

$$\left.\begin{array}{l} F_x \equiv 4x^3 + 12xy^2 + 24x^2 = 0 \\ F_y \equiv 4y^3 + 12x^2y = 0 \end{array}\right\} \quad \left.\begin{array}{l} x(4x^2 + 12y^2 + 24x) = 0 \\ 4y(y^2 + 3x^2) = 0 \end{array}\right\}$$

Para resolver el sistema igualamos a cero cada factor de la primera ecuación con cada factor de la segunda ecuación:

$$\left.\begin{array}{l} 1^{\text{o}} \\ 1^{\text{o}} \end{array}\right\} \left.\begin{array}{l} x = 0 \\ y = 0 \end{array}\right\} P_1(0,0)$$

$$\left.\begin{array}{l} 1^{\text{o}} \\ 2^{\text{o}} \end{array}\right\} \left.\begin{array}{l} x = 0 \\ y^2 + 3x^2 = 0 \end{array}\right\} \left.\begin{array}{l} x = 0 \\ y^2 = 0 \end{array}\right\} \left.\begin{array}{l} x = 0 \\ y = 0 \end{array}\right\} P_1(0,0)$$

$$\left.\begin{array}{l} 2^{\text{o}} \\ 1^{\text{o}} \end{array}\right\} \left.\begin{array}{l} 4x^2 + 12y^2 + 24x = 0 \\ y = 0 \end{array}\right\} \left.\begin{array}{l} 4x^2 + 24x = 0 \\ y = 0 \end{array}\right\} \left.\begin{array}{l} 4x(x+6) = 0 \\ y = 0 \end{array}\right\} \begin{array}{l} P_1(0,0) \\ P_2(-6,0) \end{array}$$

$$\left.\begin{array}{l} 2^{\text{o}} \\ 2^{\text{o}} \end{array}\right\} \left.\begin{array}{l} 4x^2 + 12y^2 + 24x = 0 \\ y^2 + 3x^2 = 0 \end{array}\right\} \left.\begin{array}{l} 4x^2 + 12y^2 + 24x = 0 \\ y^2 = -3x^2 \end{array}\right\} \begin{array}{l} P_1(0,0) \\ \text{no admite otras soluciones} \end{array}$$

Por tanto, las soluciones son $P_1(0,0)$, $P_2(-6,0)$

Para determinar si se trata de máximo o mínimo, hallamos la matriz hessiana.

$$Hf(x,y) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 12x^2 + 12y^2 + 48x & 24xy \\ 24xy & 12y^2 + 12x^2 \end{pmatrix}$$

Con lo cual resulta que:

- $|Hf(-6,0)| = \begin{vmatrix} 144 & 0 \\ 0 & 432 \end{vmatrix} = 62208 > 0$ y $f_{xx} = +144 > 0 \Rightarrow P_2(-6,0)$ es un mínimo relativo.
- $|Hf(0,0)| = \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix} = 0 \quad \rightarrow$ el hessiano no nos determina la naturaleza del punto crítico. Para estudiar la naturaleza del punto crítico $P_1(0,0)$ cortamos la superficie mediante el plano vertical $y = x$, con lo que obtenemos la curva $z = x^4 + x^4 + 6x^4 + 8x^3 = 8x^4 + 8x^3$ y estudiamos la naturaleza del origen en esta curva. Tenemos:

  $z'_x = 32x^3 + 24x^2 \quad \rightarrow \quad z'_x(0) = 0 \quad \rightarrow x_0 = 0$ es un punto crítico

  $z''_x = 96x^2 + 48x \quad \rightarrow \quad z''_x(0) = 0$

  $z'''_x = 198x + 48 \quad \rightarrow \quad z'''_x(0) = 48 \neq 0 \quad \rightarrow x_0 = 0$ es un punto de inflexión.

  Luego el punto $P_1(0,0)$ es un punto silla, ya que en cualquier entorno suyo hay tanto puntos por encima de $f(0,0)$ como puntos por debajo.

### 3.10.4. Extremos condicionados. Multiplicadores de Lagrange

**Planteamiento del problema**. En los problemas clásicos de máximos y mínimos se trata de hacer máxima o mínima una función $f(x, y)$ sujeta a una restricción $g(x, y) = 0$. Teóricamente el problema se puede resolver despejando la variable $y$ en la ecuación $g(x, y) = 0$, y sustituyendo el valor obtenido, $y = \phi(x)$, en la en la función $f$, con lo cual el problema se reduce al cálculo de máximos y mínimos de una función de una sola variable.

$$f(x, y) = f\big(x, \phi(x)\big) = h(x)$$

El problema se presenta cuando no es posible o no es práctico despejar la variable $y$ en la ecuación $g(x, y) = 0$.

**Observación.** El extremo de la función $f(x, y)$ condicionado por la ecuación $g(x, y) = 0$, no es extremo de la función $f(x, y)$, considerada aisladamente, sino de la intersección de ambas funciones.

**Ejemplo 3.68.** *Hallar el mínimo de la función $f(x, y) = x^2 + y^2$ condicionado por la restricción $x + y - 1 = 0$.*

*Solución.* El problema puede reducirse a una sola variable, en efecto:

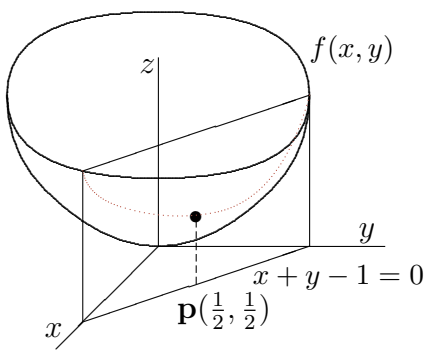



Figura 3.28:

Despejando $y$ en la restricción resulta: $y = -x + 1$ y sustituyendo el valor obtenido en $f(x, y)$, resulta una función de una sola variable.

$$f(x, y) = x^2 + (-x+1)^2 = x^2 + x^2 - 2x + 1 = 2x^2 - 2x + 1$$

Es decir,

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

de donde,

$$h'(x) = 4x - 2 = 0 \;\rightarrow\; x = 1/2 \;\rightarrow\; y = 1/2$$

Luego el mínimo de la función es

$$f(1/2, 1/2) = 1/4 + 1/4 = 1/2$$

Para comprobar que se trata de un mínimo acudimos a la derivada segunda $h''(x) = 4$, luego $h''(1/2) = +4$, luego se trata de un mínimo.

**Nota:** Este ejemplo se ha resuelto reduciéndolo a una sola variable. Con los multiplicadores de Lagrange se trata de resolver el problema sin reducirlo a una sola variable, es decir, sin despejar la variable $y$ de la ecuación $g(x, y) = 0$.

**Método de los multiplicadores de Lagrange.**

*Los extremos de la función $f(x, y)$, condicionados por la restricción $g(x, y) = 0$, se producen en los puntos críticos de la función:*

$$L(x, y, \lambda) = f(x, y) + \lambda\, g(x, y)$$

Dichos puntos críticos vendrán determinados por las soluciones del sistema:

$$\left.\begin{array}{l} L_x = 0 \\ L_y = 0 \\ L_\lambda = 0 \end{array}\right\} \left.\begin{array}{l} f_x + \lambda\, g_x = 0 \\ f_y + \lambda\, g_y = 0 \\ g(x,y) = 0 \end{array}\right\}$$

El procedimiento más cómodo para resolver el sistema consiste en eliminar $\lambda$ entre las dos primeras ecuaciones y sustituir el resultado en la tercera. En el proceso de resolución del sistema hay que procurar evitar perder soluciones en las simplificaciones. Por ejemplo, de la ecuación $\lambda\, x = x$ se obtienen dos soluciones $x = 0$ y $\lambda = 1$, mientras que si "tachamos" la $x$ perdemos la solución $x = 0$.
Para determinar su naturaleza estudiamos el signo del determinante:

$$\Delta = \begin{vmatrix} L_{\lambda\lambda} & L_{\lambda x} & L_{\lambda y} \\ L_{x\lambda} & L_{xx} & L_{xy} \\ L_{y\lambda} & L_{yx} & L_{yy} \end{vmatrix} = \begin{vmatrix} 0 & g_x & g_y \\ g_x & L_{xx} & L_{xy} \\ g_y & L_{yx} & L_{yy} \end{vmatrix} \begin{array}{l} = - \rightarrow \text{mínimo} \\ = + \rightarrow \text{Máximo} \\ = 0 \rightarrow \text{duda} \end{array}$$

Resolvamos el ejemplo 3.68 mediante el método de los multiplicadores de Lagrange.

**Ejemplo 3.69.** *Hallar el mínimo de la función $f(x,y) = x^2 + y^2$ condicionado por la restricción $x + y - 1 = 0$.*

*Solución.* Formamos la función de Lagrange:

$$L(x,y,\lambda) = f(x,y) + \lambda\, g(x,y)$$

Que, en nuestro caso, es:

$$L(x,y,\lambda) = x^2 + y^2 + \lambda(x + y - 1)$$

Los puntos críticos de esta función vendrán determinados por las soluciones del sistema:

$$\left.\begin{array}{l} L_x \equiv 2x + \lambda = 0 \\ L_y \equiv 2y + \lambda = 0 \\ L_\lambda \equiv x + y - 1 = 0 \end{array}\right\} \begin{array}{l} \left.\begin{array}{l} \lambda = -2x \\ \lambda = -2y \end{array}\right\} - 2x = -2y \Rightarrow x = y \\ x + y - 1 = 0 \rightarrow x + x = 1 \rightarrow 2x = 1 \rightarrow x = \dfrac{1}{2} \rightarrow y = \dfrac{1}{2} \end{array}$$

Luego el único punto crítico es el punto $\mathbf{p}(\frac{1}{2}, \frac{1}{2})$. Para determinar su naturaleza estudiamos el signo del determinante:

$$\Delta = \begin{vmatrix} 0 & g_x & g_y \\ g_x & L_{xx} & L_{xy} \\ g_y & L_{yx} & L_{yy} \end{vmatrix} \begin{array}{l} = - \rightarrow \text{mínimo} \\ = + \rightarrow \text{Máximo} \\ = 0 \rightarrow \text{duda} \end{array}$$

De donde,

$$\Delta L(\frac{1}{2}, \frac{1}{2}; -1) = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{vmatrix} = -2 - 2 = -4 \Rightarrow \text{mínimo}$$

**Ejemplo 3.70.** *Inscribir un rectángulo de área máxima, con los lados paralelos a los ejes de coordenadas, en la región del primer cuadrante limitada por la parábola $y = 3 - x^2$ y los ejes de coordenadas.*

*Solución.* La función a maximizar es el área del rectángulo, $a = x \cdot y$. La restricción viene determinada por el hecho de que el punto $(x, y)$ debe pertenecer a la parábola $y = 3 - x^2$, luego tenemos:

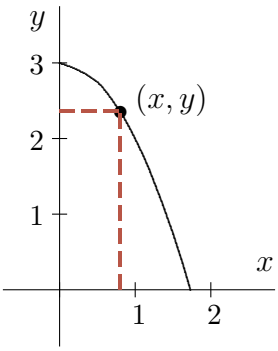


Figura 3.29: $y = 3 - x^2$

$$\left.\begin{array}{l} a = x \cdot y \\ x^2 + y - 3 = 0 \end{array}\right\} L(x, y, \lambda) = xy + \lambda(x^2 + y - 3)$$

de donde, los puntos críticos vendrán determinados por las soluciones del sistema:

$$\left.\begin{array}{l} L_x \equiv y + 2x\lambda = 0 \\ L_y \equiv x + \lambda = 0 \\ L_\lambda \equiv x^2 + y - 3 = 0 \end{array}\right\} \quad \begin{array}{c} \left.\begin{array}{c} \lambda = \frac{-y}{2x} \\ \lambda = -x \end{array}\right\} \frac{-y}{2x} = -x \Rightarrow y = 2x^2 \\ x^2 + 2x^2 - 3 = 0 \to 3x^2 = 3 \to x^2 = 1 \end{array}$$

Y al ser $x > 0$ por ser $(x, y)$ un punto del primer cuadrante, resulta:

$$x = 1 \quad \to \quad y = 2 \quad \to \quad \lambda = -1$$

Luego el único punto crítico es el punto $P(1, 2)$. Para determinar su naturaleza estudiamos el signo del determinante:

$$\Delta = \begin{vmatrix} 0 & g_x & g_y \\ g_x & L_{xx} & L_{xy} \\ g_y & L_{yx} & L_{yy} \end{vmatrix} = \begin{vmatrix} 0 & 2x & 1 \\ 2x & 2\lambda & 1 \\ 1 & 2x & 0 \end{vmatrix} \begin{array}{l} = - \to \text{mínimo} \\ = + \to \text{Máximo} \\ = 0 \to \text{duda} \end{array}$$

De donde,

$$\Delta L(1, 2; -1) = \begin{vmatrix} 0 & 2 & 1 \\ 2 & -2 & 1 \\ 1 & 1 & 0 \end{vmatrix} = 2 + 2 + 2 = 6 > 0 \Rightarrow \text{Máximo}$$

### 3.10.5. Máximos y mínimos absolutos

*Toda función continua definida en un recinto cerrado y acotado alcanza un valor máximo y un valor mínimo sobre dicho recinto.* Para hallar los máximos y mínimos absolutos de una función continua en un recinto cerrado y acotado realizaremos el siguiente proceso:

1. Hallamos los puntos críticos en el *interior* del recinto. Para ello hallamos los puntos críticos de la función, ignorando el contorno del recinto, y una vez hallados los puntos críticos seleccionamos los situados en el interior del recinto.

2. Hallamos los puntos críticos en el *contorno del recinto*. Para ello estudiamos los extremos de la función *condicionados* por el contorno; bien aplicando los multiplicadores de Lagrange, o bien por sustitución de la variable.

3. Comparamos los valores de la función en los puntos críticos hallados. El mayor corresponde al máximo y el menor al mínimo.

**Ejemplo 3.71.** *Determinar los extremos absolutos de la función* $f(x,y) = x^2 + y^2$ *en el recinto* $x^2 - 2x + y^2 - 3 \leq 0$.

*Solución.* Localizamos el recinto expresándolo en forma canónica $(x-1)^2 + y^2 \leq 4$, luego se trata del círculo $\mathcal{C}$ con centro en el punto $C(1,0)$ y radio 2.
Se trata de una función continua definida en un recinto cerrado y acotado, luego alcanzará un máximo y un mínimo en dicho recinto. Para encontrarlos seguimos los siguientes pasos:
(a) En primer lugar determinamos los puntos críticos del *interior* del recinto, igualando a cero las derivadas parciales.

$$\left.\begin{array}{l} f_x \equiv 2x = 0 \\ f_y \equiv 2y = 0 \end{array}\right\} \left.\begin{array}{l} x = 0 \\ y = 0 \end{array}\right\} P_1(0,0) \in \mathcal{C}$$

(b) En segundo lugar determinamos los puntos críticos en la frontera del recinto, para ello construimos la función lagrangiana:

$$L(x,y,\lambda) = x^2 + y^2 + \lambda(x^2 - 2x + y^2 - 3)$$

y buscamos sus puntos críticos:

$$\left.\begin{array}{l} L_x \equiv 2x + 2\lambda x - 2\lambda = 0 \\ L_y \equiv 2y + 2\lambda y = 0 \\ L_\lambda \equiv x^2 - 2x + y^2 - 3 = 0 \end{array}\right\} \left.\begin{array}{l} x + \lambda x - \lambda = 0 \\ y + \lambda y = 0 \\ x^2 - 2x + y^2 - 3 = 0 \end{array}\right\} \rightarrow \left\{\begin{array}{l} \lambda = -1 \rightarrow 1 = 0 \\ y = 0 \rightarrow \left\{\begin{array}{l} x = 3 \\ x = -1 \end{array}\right. \end{array}\right.$$

Los puntos críticos del contorno son: $P_2(3,0)$ y $P_3(-1,0)$.
(c) Comparamos los valores de la función en cada uno de los puntos críticos, el mayor será el máximo y el menor el mínimo.

$$\left.\begin{array}{l} f(0,0) = 0 \\ f(3,0) = 9 \\ f(-1,0) = 1 \end{array}\right\} \begin{array}{l} \text{mínimo} \\ \text{Máximo} \\ \\ \end{array}$$

Para determinar los máximos y mínimos absolutos de una función continua en un intervalo cerrado y acotado no es necesario determinar la naturaleza de cada uno de los puntos críticos.

**Ejemplo 3.72.** *Determinar los extremos absolutos de la función* $f(x,y) = x^2 + 3y^2$ *en el recinto* $x^2 - 2x + y^2 - 3 \leq 0$.

*Solución.* Localizamos el recinto expresándolo en forma canónica $(x-1)^2+y^2 \leq 4$, luego se trata del círculo $\mathcal{C}$ con centro en el punto $C(1,0)$ y radio 2.
Se trata de una función continua definida en un recinto cerrado y acotado, luego alcanzará un máximo y un mínimo en dicho recinto. Para encontrarlos seguimos los siguientes pasos:
(a) En primer lugar determinamos los puntos críticos del *interior* del recinto, igualando a cero las derivadas parciales.

$$\left.\begin{array}{l} f_x \equiv 2x = 0 \\ f_y \equiv 6y = 0 \end{array}\right\} \left.\begin{array}{l} x = 0 \\ y = 0 \end{array}\right\} P_1(0,0) \in \mathcal{C}$$

(b) En segundo lugar determinamos los puntos críticos en la frontera del recinto, para ello construimos la función lagrangiana:

$$L(x,y,\lambda) = x^2 + 3y^2 + \lambda(x^2 - 2x + y^2 - 3)$$

y buscamos sus puntos críticos:

$$\left.\begin{array}{l} L_x \equiv 2x + 2\lambda x - 2\lambda = 0 \\ L_y \equiv 6y + 2\lambda y = 0 \\ L_\lambda \equiv x^2 - 2x + y^2 - 3 = 0 \end{array}\right\} \left.\begin{array}{l} x + \lambda x - \lambda = 0 \\ 3y + \lambda y = 0 \\ x^2 - 2x + y^2 - 3 = 0 \end{array}\right\} \rightarrow \left\{\begin{array}{l} \lambda = -3 \rightarrow x = 3/2 \\ y = 0 \rightarrow \left\{\begin{array}{l} x = 3 \\ x = -1 \end{array}\right. \end{array}\right.$$

Para $x = \frac{3}{2}$, resulta, de la tercera ecuación, $\frac{9}{4} - 3 + y^2 - 3 = 0$, de donde, $y^2 = 6 - \frac{9}{4} = \frac{15}{4}$, y en consecuencia $y = \pm\sqrt{\frac{15}{4}}$
Luego los puntos críticos del contorno son: $P_2(3,0)$, $P_3(-1,0)$, $P_4(\frac{3}{2}, \sqrt{\frac{15}{4}})$, $P_5(\frac{3}{2}, -\sqrt{\frac{15}{4}})$.
(c) Comparamos los valores de la función en cada uno de los puntos críticos, el mayor será el máximo y el menor el mínimo.

$$\begin{array}{l} f(0,0) = 0 \rightarrow \text{mínimo} \\ f(3,0) = 9 \\ f(-1,0) = 1 \\ f\big(\frac{3}{2}, \sqrt{\frac{15}{4}}\big) = \frac{9}{4} + \frac{45}{4} = \frac{54}{4} \rightarrow \text{ Máximo} \\ f\big(\frac{3}{2}, -\sqrt{\frac{15}{4}}\big) = \frac{9}{4} + \frac{45}{4} = \frac{54}{4} \rightarrow \text{Máximo} \end{array}$$

Para determinar los máximos y mínimos absolutos de una función continua en un intervalo cerrado y acotado no es necesario determinar la naturaleza de cada uno de los puntos críticos.

**Ejemplo 3.73.** *determinar los extremos absolutos de la función* $f(x,y) = xy(1-x^2-y^2)$ *en el conjunto* $A = \{(x,y) \in \mathbb{R}^2$ *tales que* $0 \leq x \leq 1, 0 \leq y \leq 1\}$

*Solución.* La función $f(x,y) = xy(1-x^2-y^2) = xy - x^3y - xy^3$ es continua en todo $\mathbb{R}^2$. El recinto dado es el cuadrado unidad situado en el primer cuadrante. Por tanto, se trata de una función continua definida en un recinto cerrado y acotado, luego alcanzará un máximo y un mínimo absoluto. Para encomtrarlos, seguimos los siguientes pasos:

(a) En primer lugar determinamos los puntos críticos del interior del recinto, igualando a cero las derivadas parciales

$$\left.\begin{array}{l} f_x(x,y) \equiv y - 3x^2y - y^3 = 0 \\ f_y(x,y) \equiv x - x^3 - 3xy^2 = 0 \end{array}\right\} \left.\begin{array}{l} y(1-3x^2-y^2) = 0 \\ x(1-x^2-3y^2) = 0 \end{array}\right\}$$

Para resolver el sistema igualamos a cero cada factor de la primera ecuación con cada factor de la segunda ecuación:

$$\left.\begin{array}{l} 1^{\text{o}} \\ 1^{\text{o}} \end{array}\right\} \left.\begin{array}{l} y = 0 \\ x = 0 \end{array}\right\} P_1(0,0)$$

$$\left.\begin{array}{l} 1^{\text{o}} \\ 2^{\text{o}} \end{array}\right\} \left.\begin{array}{l} y = 0 \\ 1 - x^2 - 3y^2 = 0 \end{array}\right\} \left.\begin{array}{l} y = 0 \\ 1 - x^2 = 0 \end{array}\right\} \left.\begin{array}{l} y = 0 \\ x = \pm 1 \end{array}\right\} \begin{array}{l} P_2(1,0) \\ Q_1(-1,0) \text{ no pertenece al recinto} \end{array}$$

$$\left.\begin{array}{l} 2^{\text{o}} \\ 1^{\text{o}} \end{array}\right\} \left.\begin{array}{l} 1 - 3x^2 - y^2 = 0 \\ x = 0 \end{array}\right\} \left.\begin{array}{l} 1 - y^2 = 0 \\ x = 0 \end{array}\right\} \left.\begin{array}{l} y = \pm 1 \\ x = 0 \end{array}\right\} \begin{array}{l} P_3(0,1) \\ Q_2(0,-1) \text{ no pertenece al recinto} \end{array}$$

$$\left.\begin{array}{l} 2^{\text{o}} \\ 2^{\text{o}} \end{array}\right\} \left.\begin{array}{l} 1 - 3x^2 - y^2 = 0 \\ 1 - x^2 - 3y^2 = 0 \end{array}\right\} \left.\begin{array}{l} 3x^2 + y^2 = 1 \\ x^2 + 3y^2 = 1 \end{array}\right\} \left.\begin{array}{l} y^2 = 1 - 3x^2 \\ x^2 + 3 - 9x^2 = 1 \end{array}\right\} \left.\begin{array}{l} y^2 = 1 - 3x^2 \\ -8x^2 = -2 \end{array}\right\}$$

$$\left.\begin{array}{l} y^2 = 1 - 3x^2 \\ x^2 = 1/4 \end{array}\right\} \left.\begin{array}{l} y^2 = 1 - 3x^2 \\ x = \pm 1/2 \end{array}\right\} \left.\begin{array}{l} y^2 = 1 - 3/4 = 1/4 \\ x = \pm 1/2 \end{array}\right\} \left.\begin{array}{l} y = \pm 1/2 \\ x = \pm 1/2 \end{array}\right\} \begin{array}{c} P_4(1/2,1/2) \\ \text{el resto no} \end{array}$$

(b) En segundo lugar determinamos los puntos críticos de la frontera del recinto. Estos puntos críticos serán los cuatro extremos del recinto $P_1(0,0)$, $P_2(1,0)$, $P_3(0,1)$, $P_5(1,1)$, y los puntos críticos situados en los cuatro segmentos de la frontera del recinto, que los hallamos por sustitución en $z = f(x,y) = xy(1-x^2-y^2) = xy - x^3y - xy^3$, estudiando los puntos críticos de las funciones resultantes, que son de una sola variable.

$$\begin{array}{l} x = 0 \to z = 0 \to z' = 0 \text{ para cualquier valor de } y \to P_6(0,y) \\ x = 1 \to z = y - y - y^3 = -y^3 \to z' = -3y^2 = 0 \to y = 0 \to P_2(1,0) \\ y = 0 \to z = 0 \to z' = 0 \text{ para cualquier valor de } x \to P_7(x,0) \\ y = 1 \to z = x - x^3 - x = -x^3 \to z' = -3x^2 = 0 \to x = 0 \to P_3(0,1) \end{array}$$

(c) Por último comparamos los valores de la función en cada uno de los puntos críticos, el mayor será el máximo absoluto y el menor el mínimo.

$$\begin{array}{l} f(0,0) = f(1,0) = f(0,1) = 0 \\ f(0,y) = f(x,0) = 0 \\ f(1,1) = 1 - 1 - 1 = -1 \ \text{ mínimo} \\ f(1/2,1/2) = 1/4 - 1/16 - 1/16 = 2/16 = 1/8 \ \text{ Máximo} \end{array}$$

Para determinar los máximos y mínimos absolutos de una función continua en un recinto cerrado y acotado no es necesario determinar la naturaleza de cada uno de los puntos críticos.