Físicas

Primer Curso

ÁLGEBRA LINEAL

Juan A. Navarro González

31 de enero de 2023

Índice

1. Preliminares

Un **conjunto** X está totalmente determinado por sus elementos, que se pueden contar, y su **cardinal** $|X|$ es el número de elementos de X (puede ser infinito).

El **conjunto vacío**, que no tiene ningún elemento, se denota $\emptyset$. Otros conjuntos son:

El conjunto de los **Números naturales**	$\mathbb{N} = \{0, 1, 2, 3, \ldots\}$
El conjunto de los **Números enteros**	$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$
El conjunto de los **Números racionales**	$\mathbb{Q} = \{a/b \colon \exists a, b \in \mathbb{Z}, b \neq 0\}$
El conjunto de los **Números reales**	$\mathbb{R} = \{\text{números decimales infinitos}\}$

Si X es un conjunto, $x \in X$ significa que x es un elemento de X.

Si X e Y son conjuntos, $Y \subseteq X$ significa que Y es un **subconjunto** de X, que todos los elementos de Y son elementos de X; i.e. $y \in Y \Rightarrow y \in X$. Si Z es otro subconjunto de X, la **unión** $Y \cup Z$ e **intersección** $Y \cap Z$ son los siguientes subconjuntos de X:

$$Y \cup Z := \{x \in X \colon x \in Y \text{ ó } x \in Z\} \quad , \quad Y \cap Z := \{x \in X \colon x \in Y \text{ y } x \in Z\}.$$

1.1. Números Complejos

Definiciones: El conjunto $\mathbb{C}$ de los **números complejos** está formado por las parejas ordenadas de números reales $x + yi$ (donde $x \in \mathbb{R}$ se llama **parte real** de z e $y \in \mathbb{R}$ se llama **parte imaginaria**) que se suman y multiplican con las siguientes reglas ($i^2 = -1$):

$$(x_1 + y_1 i) + (x_2 + y_2 i) := (x_1 + x_2) + (y_1 + y_2)i,$$
$$(x_1 + y_1 i) \cdot (x_2 + y_2 i) := (x_1 x_2 - y_1 y_2) + (x_1 y_2 + x_2 y_1)i,$$

y cada número real x se identifica con el número complejo $x + 0i$, de modo que $\mathbb{R} \subset \mathbb{C}$.

Estas dos operaciones son asociativas, $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$ y $(z_1 z_2) z_3 = z_1 (z_2 z_3)$, conmutativas, $z_1 + z_2 = z_2 + z_1$ y $z_1 z_2 = z_2 z_1$, y además $z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$.

Si $x, y \in \mathbb{R}$, el **conjugado** de $z = x + yi$ es el número complejo $\bar{z} := x - yi$, de modo que un número complejo z es real si y sólo si $z = \bar{z}$.

Si $z = x + yi$ no es nulo, tenemos que $z\bar{z} = x^2 + y^2 > 0$, así que su inverso existe y es

$$z^{-1} = \frac{x}{x^2 + y^2} - \frac{y}{x^2 + y^2} i = \frac{\bar{z}}{z\bar{z}}.$$

El **cociente** de números complejos se define como $u/z := uz^{-1} = (u\bar{z})/(z\bar{z})$.

Propiedades de la Conjugación: $\bar{\bar{z}} = z$, $\overline{z + u} = \bar{z} + \bar{u}$, $\overline{zu} = \bar{z}\bar{u}$, $\overline{z^{-1}} = \bar{z}^{-1}$, $\overline{u/z} = \bar{u}/\bar{z}$.

Demostración: Veamos las dos últimas. Como $1 = zz^{-1}$, tenemos que $1 = \overline{zz^{-1}} = \bar{z} \cdot \overline{z^{-1}}$.
Luego $\overline{z^{-1}} = \bar{z}^{-1}$, y $\overline{u/z} = \overline{uz^{-1}} = \bar{u}\bar{z}^{-1} = \bar{u}/\bar{z}$.

Definición: Si $x, y \in \mathbb{R}$, el **módulo** del número complejo $z = x + yi$ es el número real

$$|z| := +\sqrt{z \cdot \bar{z}} = +\sqrt{x^2 + y^2} \quad , \quad |z|^2 = z \cdot \bar{z}$$

y es el valor absoluto cuando z es real. Un número complejo z es nulo si y sólo si $|z| = 0$.

Propiedades del Módulo: $|z| = |\bar{z}|$, $|zu| = |z| \cdot |u|$, $|\frac{1}{z}| = \frac{1}{|z|}$, $|\frac{u}{z}| = \frac{|u|}{|z|}$, $|z + u| \leq |z| + |u|$.

Demostración: Veamos la última. Pongamos $z = x + yi$ con $x, y \in \mathbb{R}$.

$$z + \bar{z} = x + yi + x - yi = 2x \leq 2|x| = 2\sqrt{x^2} \leq 2\sqrt{x^2 + y^2} = 2|z|.$$

$$\begin{aligned} |z + u|^2 &= (z + u)(\overline{z + u}) = (z + u)(\bar{z} + \bar{u}) = |z|^2 + |u|^2 + z\bar{u} + \bar{z}u = |z|^2 + |u|^2 + z\bar{u} + \overline{z\bar{u}} \\ &\leq |z|^2 + |u|^2 + 2|z\bar{u}| = |z|^2 + |u|^2 + 2|z| \cdot |\bar{u}| = |z|^2 + |u|^2 + 2|z| \cdot |u| = (|z| + |u|)^2 . \end{aligned}$$

Definición: Para cada número real $t \in \mathbb{R}$ pondremos

$$\boxed{e^{ti} := \cos t + i \,\operatorname{sen} t}$$

donde el seno y coseno se consideran en radianes para que $\frac{\mathrm{d}}{\mathrm{d}t}(e^{it}) = ie^{it}$. Luego $e^0 = 1$, y

$$e^{2\pi i} = 1 \quad \textbf{fórmula de Euler} \text{ (1707-1783).}$$

Las fórmulas del seno y coseno de una suma expresan que

$$\boxed{e^{ti}e^{t'i} = e^{(t+t')i}}$$

$$e^{ti}e^{t'i} = (\cos t + i \,\operatorname{sen} t)(\cos t' + i \,\operatorname{sen} t') = \big((\cos t)(\cos t') - (\operatorname{sen} t)(\operatorname{sen} t')\big) +$$
$$+\, i\big((\cos t)(\operatorname{sen} t') + (\operatorname{sen} t)(\cos t')\big) = \cos(t+t') + i \,\operatorname{sen}(t+t') = e^{(t+t')i}.$$

En general $|e^{ti}| = \sqrt{\cos^2 t + \operatorname{sen}^2 t} = 1$, y todo número complejo de módulo 1 es $e^{\theta i}$ para algún número real θ, único salvo la adición de un múltiplo entero de 2π.

Definición: Si $z \in \mathbb{C}$ es de módulo $\rho \neq 0$, entonces el módulo de z/ρ es 1, así que

$$\boxed{z = \rho e^{\theta i} = \rho(\cos\theta + i \,\operatorname{sen}\theta)}$$

para algún número real $\theta = \arg z$, llamado **argumento** de z (bien definido salvo la adición de un múltiplo entero de 2π). Cuando $z = x + yi$, con $x, y \in \mathbb{R}$, por definición el argumento de z es cualquier número real θ que cumpla

$$\cos\theta = \frac{x}{\rho} = \frac{x}{\sqrt{x^2+y^2}} \quad , \quad \operatorname{sen}\theta = \frac{y}{\rho} = \frac{y}{\sqrt{x^2+y^2}} \quad , \quad \operatorname{tg}\theta = \frac{y}{x} \cdot$$

$$\arg z = \pm \arccos(x/\rho) \ \ , \text{ donde el signo es el de la parte imaginaria } y.$$

Ejemplos: Si $a \in \mathbb{R}$ es positivo, $\arg a = 0$, $\arg(ai) = \pi/2$, $\arg(-a) = \pi$, $\arg(-ai) = 3\pi/2$.

Propiedades del Argumento: $\arg(z \cdot z') = (\arg z) + (\arg z')$, $\arg z^{-1} = \arg \bar{z} = -\arg z$, $\arg(u/z) = \arg u - \arg z$.

Demostración: Se sigue de las igualdades $zz' = (\rho e^{\theta i})(\rho' e^{\theta' i}) = \rho\rho' e^{(\theta+\theta')i}$, $|zz'| = \rho\rho'$; y de que $z^{-1}z = 1$ y $z\bar{z} = |z|^2$ tienen argumento nulo, porque son números reales positivos.

Funciones Elementales: Dado un número complejo $z = x + yi$, $x, y \in \mathbb{R}$, pondremos

$$e^z := e^x e^{yi} = e^x(\cos y + i \operatorname{sen} y)$$

de modo que $e^z \neq 0$, y $e^{z'+z} = e^{z'}e^z$ para cualesquiera números complejos z', z.

Definimos el seno y coseno de un número complejo z por las igualdades

$$\cos z := \frac{e^{zi} + e^{-zi}}{2} \quad , \quad \operatorname{sen} z := \frac{e^{zi} - e^{-zi}}{2i} \cdot$$

Cuando $e^u = z$, decimos que u es un **logaritmo** (neperiano) de z, y ponemos $u = \ln z$.

El logaritmo de un número complejo no nulo $z = \rho e^{\theta i} = e^{\ln\rho} e^{\theta i} = e^{\ln\rho + \theta i}$ es

$$\ln z = \ln\rho + (\theta + 2k\pi)i \quad , \quad k \in \mathbb{Z},$$

y toma infinitos valores, que se diferencian en un múltiplo entero de $2\pi i$.

Por ejemplo, $\cos i = \frac{1}{2}(e^{-1} + e) \approx 1'543$ es un número real mayor que 1, y $\ln(-1) = \pi i$.

Raíces n-ésimas Complejas: Dado un número complejo no nulo a, pondremos

$$a^z := e^{z \ln a} = e^{z(\ln a + 2k\pi i)} \quad , \quad k \in \mathbb{Z},$$

y puede tomar infinitos valores.

Sea $z = \rho e^{\theta i}$ un número complejo no nulo de módulo ρ y argumento θ.

Sus raíces n-ésimas complejas ($n \in \mathbb{N}$, $n \geq 2$) son

$$z^{1/n} = e^{\frac{1}{n}(\ln \rho + (\theta + 2k\pi)i)} = \boxed{\sqrt[n]{\rho}\, e^{\left(\frac{\theta+2k\pi}{n}\right)i} = \left(\sqrt[n]{\rho}\, e^{\frac{\theta}{n}i}\right) e^{\frac{2k\pi}{n}i}} \quad , \quad k \in \mathbb{Z},$$

y, como k y $k+n$ dan la misma raíz n-ésima, basta tomar $k = 0, \ldots, n-1$.

Vemos pues que *todo número complejo no nulo tiene n raíces n-ésimas complejas, que forman un polígono regular inscrito en el círculo centrado en el 0.*

En particular, las raíces n-ésimas de la unidad son $e^{\frac{2k\pi}{n}i} = (e^{\frac{2\pi}{n}i})^k$, $k = 0, \ldots, n-1$; así que las raíces n-ésimas de un número complejo no nulo se obtienen multiplicando una de ellas por las n raíces n-ésimas de la unidad, que son las sucesivas potencias de $e^{\frac{2\pi}{n}i}$.

1.2. Permutaciones

Definiciones: Sean X e Y dos conjuntos. Dar una **aplicación** $f\colon X \to Y$ es asignar a cada elemento $x \in X$ un único elemento $f(x) \in Y$, llamado **imagen** de x por la aplicación f.

Si $g\colon Y \to Z$ es otra aplicación, la **composición** de g y f es la aplicación

$$g \circ f\colon X \longrightarrow Z, \ (g \circ f)(x) := g\big(f(x)\big).$$

La **identidad** de un conjunto X es la aplicación $\mathrm{Id}_X\colon X \to X$, $\mathrm{Id}_X(x) = x$.

Sea $f\colon X \to Y$ una aplicación. Si $A \subseteq X$, ponemos

$$f(A) := \{y \in Y\colon y = f(x), \exists x \in A\} = \{f(x)\}_{x \in A} \subseteq Y$$

y si $B \subseteq Y$, ponemos $f^{-1}(B) := \{x \in X\colon f(x) \in B\} \subseteq X$.

Si $y \in Y$, puede ocurrir que $f^{-1}(y)$ no tenga ningún elemento o tenga más de uno, de modo que, en general, f^{-1} <u>no</u> es una aplicación de Y en X.

Diremos que $f\colon X \to Y$ es **inyectiva** si elementos distintos tienen imágenes distintas:

$$x, y \in X, \ f(x) = f(y) \ \Rightarrow \ x = y$$

(i.e., cuando, para cada $y \in Y$ se tiene que $f^{-1}(y)$ tiene un elemento o ninguno) y diremos que f es **epiyectiva** si todo elemento de Y es imagen de algún elemento de X:

$$y \in Y \ \Rightarrow \ y = f(x) \text{ para algún } x \in X \ ,$$

es decir, cuando $f(X) = Y$ o, lo que es igual, cuando para cada $y \in Y$ se cumple que $f^{-1}(y)$ tiene al menos un elemento. Diremos que $f\colon X \to Y$ es **biyectiva** cuando es inyectiva y epiyectiva, cuando cada elemento $y \in Y$ es imagen de un único elemento de X, de modo que $f^{-1}(y)$ tiene un único elemento, y en tal caso $f^{-1}\colon Y \to X$ <u>sí</u> es una aplicación, llamada aplicación **inversa** de f porque $f^{-1} \circ f = \mathrm{Id}_X$ y $f \circ f^{-1} = \mathrm{Id}_Y$.

Definiciones: Sea n un número natural, $n \geq 2$. Las **permutaciones** de n elementos son las aplicaciones biyectivas $\sigma\colon \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$.

El conjunto de todas las permutaciones de n elementos se denota S_n, y su cardinal es $n! = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1$. El producto de permutaciones es la composición de aplicaciones, y como son aplicaciones biyectivas, toda permutación σ tienen una permutación inversa σ^{-1}, de modo que $\sigma^{-1}(j) = i$ cuando $\sigma(i) = j$. Además, $(\sigma\tau)^{-1} = \tau^{-1}\sigma^{-1}$.

Dados $a_1, \dots, a_d \in \{1, \dots, n\}$ distintos, $(a_1 \dots a_d)$ denota la permutación $\sigma \in S_n$ tal que $\sigma(a_i) = a_{i+1}$, entendiendo que $\sigma(a_d) = a_1$, y deja fijos los restantes elementos. Diremos que $(a_1 \dots a_d)$ es un **ciclo** de longitud d, y los ciclos $(a_1 a_2)$ de longitud 2 se llaman **trasposiciones**. El inverso de un ciclo $\sigma = (a_1 \dots a_d)$ es $\sigma^{-1} = (a_d \dots a_1)$.

Diremos que dos ciclos $(a_1 \dots a_d)$ y $(b_1 \dots b_k)$ son **disjuntos** cuando $a_i \neq b_j$ para todo par de índices i, j; en cuyo caso conmutan: $(a_1 \dots a_d)(b_1 \dots b_k) = (b_1 \dots b_k)(a_1 \dots a_d)$.

Toda permutación descompone en producto de ciclos disjuntos, y también en producto de trasposiciones, porque todo ciclo es producto de trasposiciones:

$$\boxed{(a_1 a_2 a_3 \dots a_d) = (a_1 a_2)(a_2 a_3) \cdots (a_{d-1} a_d).} \qquad (1)$$

Definición: Consideremos el siguiente polinomio con coeficientes enteros:

$$\Delta(x_1, \dots, x_n) \;=\; \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

Dada una permutación $\sigma \in S_n$, los factores de $\Delta(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \prod_{i<j}(x_{\sigma(j)} - x_{\sigma(i)})$ coinciden, eventualmente salvo el signo, con los de $\Delta(x_1, \dots, x_n)$. Luego ambos polinomios coinciden o difieren en un signo, $\Delta(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \pm\Delta(x_1, \dots, x_n)$, y llamaremos **signo** de σ al número entero $\operatorname{sgn}(\sigma) = \pm 1$ tal que

$$\Delta(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \operatorname{sgn}(\sigma) \cdot \Delta(x_1, \dots, x_n). \qquad (2)$$

Teorema 1.1 *El signo de cualquier producto de permutaciones es el producto de los signos de los factores:* $\operatorname{sgn}(\tau\sigma) = (\operatorname{sgn}\tau)(\operatorname{sgn}\sigma)$.

El signo de las trasposiciones es –1.

El signo de los ciclos de longitud d es $(-1)^{d-1}$, y por eso llamaremos **pares** a las permutaciones de signo 1, e **impares** a las de signo -1.

Demostración: Sean $\sigma, \tau \in S_n$. Aplicando τ a los índices de las indeterminadas $x_1, \dots, x_n$ en la igualdad 2, obtenemos que

$$\Delta(x_{(\tau\sigma)(1)}, \dots, x_{(\tau\sigma)(n)}) = (\operatorname{sgn}\sigma) \cdot \Delta(x_{\tau(1)}, \dots, x_{\tau(n)}) = (\operatorname{sgn}\sigma)(\operatorname{sgn}\tau) \cdot \Delta(x_1, \dots, x_n).$$

Luego $\operatorname{sgn}(\tau\sigma) = (\operatorname{sgn}\sigma)(\operatorname{sgn}\tau) = (\operatorname{sgn}\tau)(\operatorname{sgn}\sigma)$.

El único factor de $\Delta(x_1, \dots, x_n)$ que cambia de signo con la trasposición (12) es el factor $(x_2 - x_1)$; luego el signo de la trasposición (12) es –1. Si (ij) es otra trasposición, tomamos una permutación τ tal que $\tau(1) = i$, $\tau(2) = j$, de modo que $(ij) = \tau \cdot (12) \cdot \tau^{-1}$, y

$$\operatorname{sgn}(ij) = \operatorname{sgn}(\tau) \cdot \operatorname{sgn}(12) \cdot \operatorname{sgn}(\tau^{-1}) = -\operatorname{sgn}(\tau) \cdot \operatorname{sgn}(\tau^{-1}) = -\operatorname{sgn}(\tau \cdot \tau^{-1}) = -1.$$

Ahora es claro que el signo de un ciclo $(a_1 \dots a_d) = (a_1 a_2)(a_2 a_3) \cdots (a_{d-1} a_d)$ es $(-1)^{d-1}$.

1.3. Matrices y Determinantes

En adelante pondremos $K = \mathbb{R}$ ó $\mathbb{C}$, llamaremos **escalares** a los elementos de K y consideraremos matrices $A = (a_{ij})$ con coeficientes en K, donde el subíndice i indica la fila y el subíndice j la columna. Dada una matriz $A = (a_{ij})$ de m filas y n columnas, su matriz **traspuesta** es $A^t := (a_{ji})$, que tiene n filas y m columnas, y su matriz **conjugada** es $\bar{A} := (\bar{a}_{ij})$. Se cumple que $(A+B)^t = A^t + B^t$, $\overline{A+B} = \bar{A} + \bar{B}$.

Si $B = (b_{jk})$ es otra matriz de n filas y r columnas, su **producto** AB es una matriz $m \times r$ cuyo coeficiente c_{ik} de la fila i y columna k es

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} = a_{i1} b_{1k} + a_{i2} b_{2k} + \ldots + a_{in} b_{nk}.$$

El producto de matrices es asociativo, distributivo y $(AB)^t = B^t A^t$, $\overline{AB} = \bar{A} \cdot \bar{B}$.

La matriz unidad I_n es la matriz $n \times n$ con todos sus coeficientes nulos, salvo los de la diagonal, que son la unidad. Si A es una matriz $m \times n$, entonces $I_m A = A$ y $AI_n = A$.

Una matriz cuadrada A de n columnas se dice que es **invertible** si existe otra matriz cuadrada B de n columnas tal que $AB = I_n = BA$, en cuyo caso tal matriz B es única y se pone $B = A^{-1}$. Si A y B son matrices invertibles $n \times n$, entonces

$$(AB)^{-1} = B^{-1}A^{-1} \quad , \quad \bar{A}^{-1} = \overline{A^{-1}} \quad , \quad (A^t)^{-1} = (A^{-1})^t \ .$$

Definición: El **determinante** de una matriz $n \times n$ cuadrada $A = (a_{ij})$ es el escalar

$$\det A = |A| := \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)}$$

y tiene las siguientes propiedades (donde $A_1, \dots, A_n$ denotan las columnas de A):

1. $|A| = |A^t|$, $|\bar{A}| = \overline{|A|}$.

2. Es lineal en cada columna (y por tanto en cada fila):
   $|A_1, \dots, A_i + B_i, \dots, A_n| = |A_1, \dots, A_i, \dots, A_n| + |A_1, \dots, B_i, \dots, A_n|$,
   $|A_1, \dots, \lambda A_i, \dots, A_n| = \lambda |A_1, \dots, A_i, \dots, A_n|$.

3. $|A_{\sigma(1)}, \dots, A_{\sigma(n)}| = (\operatorname{sgn} \sigma)|A_1, \dots, A_n|$.

4. $\begin{vmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \dots & \dots & \dots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{11} \dots a_{nn}$, $|I_n| = 1$.

5. El determinante puede calcularse desarrollando por cualquier fila o columna:
   $$|A| = a_{i1}A_{i1} + \dots + a_{in}A_{in} = a_{1j}A_{1j} + \dots + a_{nj}A_{nj},$$
   donde el **adjunto** A_{ij} es $(-1)^{i+j}$ por el determinante de la matriz que se obtiene eliminando la fila i y la columna j de la matriz A.

6. $|AB| = |A| \cdot |B|$.

7. Las matrices con determinante no nulo son invertibles.

Cuando $\sigma = (ij)$ es una trasposición, (3) afirma que el determinante es 0 cuando dos columnas (o dos filas) son iguales. Luego

$$|A_1, \dots, A_i, \dots, A_n| = |A_1, \dots, A_i + \lambda A_j, \dots, A_n| \ , \quad i \neq j,$$

así que un determinante siempre se puede calcular aplicando a las filas o columnas **transformaciones elementales** (permutarlas, multiplicar una por un escalar no nulo, o sumarle a una el producto de otra por un escalar) hasta que quede en la forma (4).

Si A es invertible, de (6) se sigue que su determinante no es nulo, y $|A^{-1}| = |A|^{-1}$.

Regla de Crámer (1704-1752)**:** *Si $A = (A_1, \dots, A_n)$ es una matriz cuadrada invertible, el sistema de ecuaciones lineales $AX = B$ tiene una única solución,*

$$x_i = \frac{|A_1, \dots, B, \dots, A_n|}{|A|} \ , \quad \text{donde la columna } B \text{ está en el lugar } i\text{-ésimo.}$$

Demostración: Si A es invertible, la única solución de $AX = B$ es $X = A^{-1}B$. Además, si $x_1, \dots, x_n$ es la solución del sistema, entonces $x_1A_1 + \dots + x_nA_n = B$ y por tanto:

$$|A_1, \dots, B, \dots, A_n| = \textstyle\sum_j x_j |A_1, \dots, A_{i-1}, A_j, A_{i+1}, \dots, A_n| = x_i |A_1, \dots, A_i, \dots, A_n|$$

porque la matriz $(A_1, \ldots, A_{i-1}, A_j, A_{i+1}, \ldots, A_n)$ tiene dos columnas iguales (las columnas i y j) cuando $i \neq j$. Luego $x_i = |A_1, \ldots, B, \ldots, A_n| / |A_1, \ldots, A_i, \ldots, A_n|$. q.e.d.

Poniendo $B = (1, 0, \ldots, 0)^t, \ldots, (0, \ldots, 0, 1)^t$ en la regla de Crámer, vemos que

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} A_{11} & \ldots & A_{1n} \\ \ldots & \ldots & \ldots \\ A_{n1} & \ldots & A_{nn} \end{pmatrix}^t .$$

Definición: El **rango** (por columnas) de una matriz A es el máximo número de columnas de A linealmente independientes, y se denota $\operatorname{rg} A$.

Teorema de Rouché-Frobënius (1832-1910, 1849-1917)**:** *Un sistema de ecuaciones lineales $AX = B$ es compatible si y sólo si* $\operatorname{rg}A = \operatorname{rg}(A|B)$ *.*

Definición: Los **menores** de orden r de una matriz A son los determinantes de las matrices formadas con los coeficientes de r filas y r columnas de A.

Teorema del Rango: *Dadas r columnas de una matriz, son linealmente independientes si y sólo si con ellas se puede formar un menor de orden r no nulo.*

Por tanto, el rango de una matriz es el mayor orden de sus menores no nulos, y el rango por filas coincide con el rango por columnas.

Sistemas de Ecuaciones Lineales: Una forma de resolver un sistema de ecuaciones lineales $AX = B$ compatible es fijar un menor no nulo de A de orden $r = \operatorname{rg}A = \operatorname{rg}(A|B)$, eliminar las ecuaciones que no entren en ese menor e igualar a parámetros las $n - r$ indeterminadas que no entren en ese menor, obteniendo así un sistema de Crámer.

Por último, si X_0 es una solución particular, $AX_0 = B$, entonces todas las soluciones se obtienen sumándole las soluciones del sistema homogéneo $AY = 0$; i.e., las soluciones son $X = X_0 + Y$, donde $AY = 0$.

En efecto, $B = A(X_0 + Y) = AX_0 + AY = B + AY$ si y sólo si $AY = 0$.

1.4. Grupos

Dados dos conjuntos X, Y, no necesariamente distintos, el **producto directo** o **cartesiano** $X \times Y$ denota el conjunto de parejas ordenadas (x, y) donde $x \in X$, $y \in Y$.

Definición: Una **operación** (interna) en un conjunto G es una aplicación $G \times G \to G$, y la imagen de un par $(a, b) \in G \times G$ se denota $a * b$, $a \cdot b$, $a + b$, ó ab sin más.

Definición: Un conjunto G con una operación $G \times G \stackrel{\cdot}{\longrightarrow} G$ es un **grupo** cuando

1. La operación es asociativa: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$, $\forall a, b, c \in G$.

2. Existe un elemento $1 \in G$, llamado **neutro** o **unidad**, tal que $a \cdot 1 = 1 \cdot a = a$, $\forall a \in G$.

3. Para cada $a \in G$ existe $a^{-1} \in G$, llamado **inverso** de a, tal que $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

y se dice que el grupo es **conmutativo** o **abeliano** cuando $a \cdot b = b \cdot a$, $\forall a, b \in G$.

Notación: En un grupo pondremos $a^0 := 1$, $a^2 := a \cdot a$, $a^3 := a \cdot a \cdot a$,..., $a^{-2} := a^{-1} \cdot a^{-1}$, $a^{-3} := a^{-1} \cdot a^{-1} \cdot a^{-1}$,..., de modo que $a^{n+m} = a^n a^m$, $\forall n, m \in \mathbb{Z}$.

Cuando el grupo es conmutativo, a menudo usaremos la notación aditiva: ponemos $a + b$ en vez de $a \cdot b$, el neutro se denota 0, el inverso $-a$ (y se llama **opuesto** de a) y pondremos $a - b := a + (-b)$, $0 \cdot a := 0$, $2a := a + a$, $3a := a + a + a$,..., $(-2)a := -a - a$,..., de modo que $(n + m)a = na + ma$, $\forall n, m \in \mathbb{Z}$.

Ejemplos:

1. Con la suma usual, $\mathbb{C}$ es un grupo conmutativos, y el neutro es el número 0.

   Con el producto usual, $\mathbb{C}_{\neq 0} := \{x \in \mathbb{C} \colon x \neq 0\}$ es un grupo conmutativo, y el neutro es el número 1.

2. El conjunto S_n de las permutaciones de n elementos, con la composición de aplicaciones, es un grupo, llamado **grupo simétrico** n-ésimo, y el neutro es la permutación identidad.

3. El conjunto $M_{m\times n}(K)$ de las matrices de m filas y n columnas con coeficientes en K ($= \mathbb{R}$ ó $\mathbb{C}$), con la suma de matrices, es un grupo conmutativo, y el neutro es la matriz que tiene todos los coeficientes nulos.

4. El conjunto $Gl(n,K) := \{A \in M_{n\times n}(K) \colon \det(A) \neq 0\}$, con el producto de matrices, es un grupo, llamado **grupo lineal** n-ésimo, y el neutro es la matriz I_n.

5. Los polinomios en una indeterminada x y coeficientes en K, con la suma de polinomios, forman un grupo conmutativo. El neutro es el polinomio de coeficientes nulos.

6. Fijado un conjunto X y un grupo abeliano $(G,+)$, el conjunto formado por las aplicaciones $f\colon X \to G$, con la suma de aplicaciones, $(f+h)(x) := f(x) + h(x)$, forman un grupo conmutativo. El neutro es la aplicación constante 0, y el opuesto de $f\colon X \to G$ es la aplicación $(-f)(x) := -\big(f(x)\big)$.

7. Dados dos grupos G_1, G_2, su producto directo o cartesiano $G_1 \times G_2$, con la operación $(a_1,a_2)\cdot(b_1,b_2) := (a_1b_1, a_2b_2)$, también es un grupo. El neutro es la pareja $(1,1)$, y $(a_1,a_2)^{-1} = (a_1^{-1}, a_2^{-1})$. Este grupo $G_1 \times G_2$ es conmutativo cuando G_1 y G_2 lo son.

   En particular K^n es grupo conmutativo con la operación
   $$(\lambda_1,\ldots,\lambda_n) + (\mu_1,\ldots,\mu_n) := (\lambda_1+\mu_1,\ldots,\lambda_n+\mu_n),$$
   el neutro es la sucesión nula $0 = (0,\ldots,0)$, y $-(\lambda_1,\ldots,\lambda_n) = (-\lambda_1,\ldots,-\lambda_n)$.

Definición: Un subconjunto H de un grupo G es un **subgrupo** si cumple las siguientes condiciones (de modo que H, con la operación de G, también es un grupo):

1. $ab \in H$, $\forall a,b \in H$
2. $1 \in H$.
3. $a^{-1} \in H$, $\forall a \in H$.

Ejemplos:

1. $\mathbb{Z}$, $\mathbb{Q}$ y $\mathbb{R}$ son subgrupos del grupo $\mathbb{C}$.

2. $\{\pm 1\}$, $\mathbb{R}_+ := \{x \in \mathbb{R} \colon x > 0\}$ y $\mathbb{R}_{\neq 0}$ son subgrupos del grupo $\mathbb{C}_{\neq 0}$.

3. Si $A \in M_{m\times n}(K)$, las soluciones del sistema de ecuaciones lineales homogéneo $AX = 0$ forman un subgrupo de K^n.

4. El grupo lineal $Gl(n,\mathbb{R})$, el **grupo ortogonal** $O(n) := \{A \in Gl(n,\mathbb{R})) \colon A^tA = I_n\}$ y el **grupo unitario** $U(n) := \{A \in Gl(n,\mathbb{C}) \colon \bar{A}^tA = I_n\}$ son subgrupos de $Gl(n,\mathbb{C})$.

5. Fijado $n \in \mathbb{N}$, los polinomios de grado $\leq n$ forman un subgrupo P_n de los polinomios.

6. Si H, H' son subgrupos de un grupo G, entonces $H \cap H'$ también es un subgrupo de G.

Definición: Una aplicación $f\colon G \to G'$ entre grupos es **morfismo de grupos** cuando

$$f(a \cdot b) = f(a) \cdot f(b),\ \forall a, b \in G.$$

En tal caso, como $f(1) = f(1 \cdot 1) = f(1) \cdot f(1)$, tendremos que $f(1)$ es el neutro de G', y como $1 = f(1) = f(a^{-1} \cdot a) = f(a^{-1}) \cdot f(a)$, tendremos que $f(a^{-1}) = f(a)^{-1}$, $\forall a \in G$.

Teorema 1.2 *Si $f\colon G \to G'$ y $h\colon G' \to G''$ son morfismos de grupos, entonces su composición $h \circ f\colon G \to G''$ también es morfismo de grupos.*

Demostración: Si $a, b \in G$, se cumple que

$$(h \circ f)(ab) = h(f(ab)) = h(f(a) \cdot f(b)) = h(f(a)) \cdot h(f(b)) = (h \circ f)(a) \cdot (h \circ f)(b).$$

Teorema 1.3 *Sea $f\colon G \to G'$ un morfismo de grupos. Su **núcleo** $\operatorname{Ker} f = \{a \in G\colon f(a) = 1\}$ es un subgrupo de G, y su **imagen** $\operatorname{Im} f = \{a' \in G'\colon a' = f(a), \exists a \in G\} = \{f(a)\}_{a \in G}$ es un subgrupo de G'.*

Demostración: Veamos que $\operatorname{Ker} f$ es un subgrupo de G.
Si $a, b \in \operatorname{Ker} f$, entonces $ab \in \operatorname{Ker} f$ porque $f(ab) = f(a)f(b) = 1 \cdot 1 = 1$.
Si $a \in \operatorname{Ker} f$, entonces $a^{-1} \in \operatorname{Ker} f$ porque $f(a^{-1}) = f(a)^{-1} = 1^{-1} = 1$.
Por último, $1 \in \operatorname{Ker} f$ porque $f(1) = 1$.

Veamos ahora que $\operatorname{Im} f$ es un subgrupo de G'.
Si $a', b' \in \operatorname{Im} f$, por definición existen $a, b \in G$ tales que $a' = f(a)$ y $b' = f(b)$, así que $a'b' = f(a)f(b) = f(ab) \in \operatorname{Im} f$.
Si $a' \in \operatorname{Im} f$, entonces $a' = f(a)$ para algún $a \in G$, y $(a')^{-1} = f(a)^{-1} = f(a^{-1}) \in \operatorname{Im} f$.
Por último, $1 \in \operatorname{Im} f$ porque $f(1) = 1$.

Teorema 1.4 *Un morfismo de grupos $f\colon G \to G'$ es inyectivo si y sólo si* $\operatorname{Ker} f = 1$.

Demostración: Si f es inyectivo y $a \in \operatorname{Ker} f$, entonces $f(a) = 1 = f(1)$; luego $a = 1$.
Recíprocamente, supongamos que $\operatorname{Ker} f = 1$. Si $f(a) = f(b)$, donde $a, b \in G$, entonces $f(a^{-1}b) = f(a)^{-1}f(b) = 1$; luego $a^{-1}b \in \operatorname{Ker} f = 1$, así que $a^{-1}b = 1$ y por tanto $a = b$.

Ejemplos:

1. La aplicación $f\colon \mathbb{C}_{\neq 0} \to \mathbb{R}_+$, $f(z) = |z|$, es morfismo de grupos porque $|zu| = |z| \cdot |u|$; luego su núcleo $U(1) = \{z \in \mathbb{C}\colon |z| = 1\}$ es un subgrupo de $\mathbb{C}_{\neq 0}$.
2. El signo $\operatorname{sgn}\colon S_n \to \{\pm 1\}$ es morfismo de grupos porque $\operatorname{sgn}(\sigma\tau) = (\operatorname{sgn}\sigma)(\operatorname{sgn}\tau)$, así que el **grupo alternado** $A_n = \{\sigma \in S_n\colon \operatorname{sgn}\sigma = 1\}$ es un subgrupo de S_n.
3. El determinante $Gl(n, K) \to K_{\neq 0}$ es morfismo de grupos porque $|AB| = |A| \cdot |B|$, y el **grupo especial** $Sl(n, K) := \{A \in Gl(n, K)\colon |A| = 1\}$ es un subgrupo de $Gl(n, K)$.
4. Si $a \in K$, la aplicación $f\colon K \to K$, $f(x) = ax$, es morfismo de grupos: $a(x+y) = ax + ay$.
5. La aplicación $f\colon \mathbb{C} \to \mathbb{C}_{\neq 0}$, $f(z) = e^z$, es morfismo de grupos: $e^{z+z'} = e^z e^{z'}$.
6. Si G es un grupo conmutativo y $n \in \mathbb{N}$, la aplicación $f\colon G \to G$, $f(a) = a^n$, es morfismo de grupos porque $(ab)^n = a^n b^n$. En particular, cuando $G = \mathbb{C}_{\neq 0}$, vemos que las raíces n-ésimas de la unidad complejas forman un subgrupo $\{z \in \mathbb{C}\colon z^n = 1\}$ de $\mathbb{C}_{\neq 0}$.
7. Sean H y H' dos subgrupos de un grupo conmutativo G. La aplicación $s\colon H_1 \times H_2 \to G$, $s(h_1, h_2) = h_1 + h_2$; es morfismo de grupos, así que su imagen

   $$H_1 + H_2 := \{g \in G\colon g = h_1 + h_2;\ \exists h_1 \in H_1, h_2 \in H_2\} = \{h_1 + h_2\}_{h_1 \in H_1, h_2 \in H_2}$$

   también es un subgrupo de G.

2. Espacios Vectoriales

Definiciones: Un conjunto A con dos operaciones internas, que denotaremos $+$ y $\cdot$, es un **anillo** cuando

1. $(A,+)$ es un grupo conmutativo.
2. El producto es asociativo: $a\cdot(b\cdot c)=(a\cdot b)\cdot c$, $\forall a,b,c\in A$.
3. Existe un elemento $1\in A$ tal que $a\cdot 1=1\cdot a=a$, $\forall a\in G$.
4. Se cumple la propiedad distributiva: $a(b+c)=ab+ac$, $\forall a,b,c\in A$.

y se dice que un anillo es **conmutativo** cuando $a\cdot b=b\cdot a$, $\forall a,b\in A$.

Un anillo conmutativo K es un **cuerpo** cuando para cada elemento no nulo $a\in K$ existe $a^{-1}\in K$ tal que $a\cdot a^{-1}=1$. Por ejemplo, $\mathbb{Q},\mathbb{R}$, $\mathbb{C}$ y {par=0,impar=1}, con las operaciones usuales, son cuerpos; mientras que el anillo $M_{2\times 2}(\mathbb{R})$ no es conmutativo, y el anillo conmutativo $\mathbb{Z}$ no es cuerpo. En adelante pondremos $K=\mathbb{R}$ ó $\mathbb{C}$ (o en general un cuerpo), y llamaremos **escalares** a los elementos de K.

Dar una **operación externa** de K en un conjunto E es dar una aplicación $K\times E\to E$, y la imagen de un par $(\lambda,e)\in K\times E$ se suele denotar $\lambda\cdot e$, ó λe sin más.

2.1. Espacios Vectoriales y Subespacios Vectoriales

Definición: Un **K-espacio vectorial** es un grupo conmutativo E (cuya operación denotamos aditivamente $+$) con una operación externa $K\times E\to E$ que cumple

1. $\lambda(e+v)=\lambda e+\lambda v$ para todo $\lambda\in K$, $e,v\in E$.
2. $(\lambda+\mu)e=\lambda e+\mu e$ para todo $\lambda,\mu\in K$, $e\in E$.
3. $(\lambda\mu)e=\lambda(\mu e)$ para todo $\lambda,\mu\in K$, $e\in E$.
4. $1\cdot e=e$ para todo vector $e\in E$.

Los elementos de E se llaman **vectores** o **puntos**, y pondremos $v-e:=v+(-e)$, de modo que $e-e=0$, y muchas otras reglas del cálculo usuales son válidas:

$$0\cdot e=\lambda\cdot 0=0,\ \lambda(-e)=(-\lambda)e=-(\lambda e),\ \lambda(v-e)=\lambda v-\lambda e,\ (\lambda-\mu)e=\lambda e-\mu e,...$$

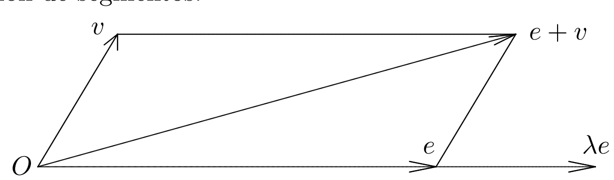
Definición: Un subgrupo V de un espacio vectorial E es un **subespacio vectorial** de E cuando $\lambda v\in V$ para todo $\lambda\in K$ y $v\in V$, de modo que V, con el producto por escalares que tenemos en E, también es un espacio vectorial.

Como $0=0\cdot e$, $-e=(-1)\cdot e$, para que un subconjunto no vacío $V\subseteq E$ sea un subespacio vectorial, es necesario y suficiente que sea estable por sumas y producto por escalares:

$$v+v',\lambda v\in V,\ \forall v,v'\in V,\ \lambda\in K.$$

Ejemplos:

1. En la Geometría Euclídea clásica, fijado un origen O, los puntos forman un espacio vectorial real cuando se suman con la regla del paralelogramo, y se multiplican por escalares según la proporción de segmentos:



Las rectas y planos que pasan por el origen O son subespacios vectoriales.

2. El grupo $K^n = K \times \overset{n}{\ldots} \times K$, con la operación externa
$$\alpha \cdot (\lambda_1, \ldots, \lambda_n) = (\alpha\lambda_1, \ldots, \alpha\lambda_n)$$
es un K-espacio vectorial. El vector 0 es $(0, \ldots, 0)$ y $-(\lambda_1, \ldots, \lambda_n) = (-\lambda_1, \ldots, -\lambda_n)$.

   Si $A \in M_{m\times n}(K)$, las soluciones del sistema de ecuaciones lineales homogéneo $AX = 0$ forman un subespacio vectorial de K^n.

3. Fijados dos números naturales positivos m y n, el producto usual de matrices por escalares $\lambda(a_{ij}) = (\lambda a_{ij})$ define una estructura de K-espacio vectorial en el grupo $M_{m\times n}(K)$. El vector 0 es la matriz con todos sus coeficientes nulos, y $-(a_{ij}) = (-a_{ij})$.

4. El espacio vectorial con un único vector (necesariamente el vector nulo) se denota 0. Todo espacio vectorial E admite los subespacios vectoriales triviales 0 y E.

5. Sean V y W dos subespacios vectoriales de un K-espacio vectorial E. Su intersección $V \cap W := \{e \in E \colon e \in V \text{ y } e \in W\}$ es un subespacio vectorial de E, y su **suma**
$$V + W := \{e \in E \colon e = v + w, \exists v \in V, w \in W\} = \{v + w\}_{v\in V, w\in W}$$
también es un subespacio vectorial de E. Si un subespacio vectorial F de E contiene a V y a W, entonces también contiene a $V + W$, porque F es un subgrupo de E.

6. Si e es un vector de un K-espacio vectorial E, entonces
$$\langle e\rangle = Ke := \{v \in E \colon v = \lambda e, \exists \lambda \in K\} = \{\lambda e\}_{\lambda\in K}$$
es un subespacio vectorial de E, y diremos que está **generado** por el vector e. Si V es un subespacio vectorial de E y $e \in V$, entonces $Ke \subseteq V$.

7. Si $e_1, \ldots, e_n$ son vectores de un espacio vectorial E, entonces
$$\langle e_1, \ldots, e_n\rangle := Ke_1 + \ldots + Ke_n = \{\lambda_1 e_1 + \ldots + \lambda_n e_n\}_{\lambda_1,\ldots,\lambda_n\in K}$$
es un subespacio vectorial de E, y diremos que está **generado** por $e_1, \ldots, e_n$. Si V es un subespacio vectorial de E y $e_1, \ldots, e_n \in V$, entonces $\langle e_1, \ldots, e_n\rangle \subseteq V$. Luego
$$\langle e_1, \ldots, e_n\rangle = \langle v_1, \ldots, v_m\rangle \Leftrightarrow e_1, \ldots, e_n \in \langle v_1, \ldots, v_m\rangle \text{ y } v_1, \ldots, v_m \in \langle e_1, \ldots, e_n\rangle.$$
En particular, para cualesquiera escalares $\lambda_1, \ldots, \lambda_n \in K$ se cumple que
$$\langle e_1, \ldots, e_n\rangle = \langle e_1, \ldots, e_n, \lambda_1 e_1 + \ldots + \lambda_n e_n\rangle.$$

8. Los polinomios en una indeterminada x y coeficientes en K, con la suma de polinomios y el producto usual por escalares, forman un K-espacio vectorial donde el 0 es el polinomio de coeficientes nulos. Fijado $n \in \mathbb{N}$, los polinomios de grado $\leq n$ forman un subespacio vectorial $P_n = K + Kx + \ldots + Kx^n = \{a_0 + a_1 x + \ldots + a_n x^n\}$.

9. Fijado un conjunto X, las funciones $\psi \colon X \to K$ forman un K-espacio vectorial, con las operaciones
$$(\psi + \phi)(x) := \psi(x) + \phi(x) \quad , \quad (\lambda\psi)(x) := \lambda\big(\psi(x)\big).$$
El vector 0 es la función que se anula en todo punto de X, y $(-\psi)(x) = -(\psi(x))$.

   Cuando $K = \mathbb{C}$ y $X = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ es un producto de intervalos en $\mathbb{R}^3$, cada función continua compleja $\psi(x, y, z) \colon X \to \mathbb{C}$ (no idénticamente nula) describe el estado de una partícula cuántica confinada en X, entendiendo que la probabilidad de encontrar a la partícula en cierta región $\Omega \subseteq X$ es
$$\frac{\displaystyle\int_\Omega |\psi(x_1, x_2, x_3)|^2 \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3}{\displaystyle\int_X |\psi(x_1, x_2, x_3)|^2 \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3}.$$

De hecho una función $\psi(x_1, x_2, x_3)$ y $\lambda\psi(x_1, x_2, x_3)$, $0 \neq \lambda \in \mathbb{C}$, describen el mismo comportamiento, así que el estado de la partícula viene dado por la recta $\mathbb{C}\psi = \langle\psi\rangle$, y puede ser representado por cualquier función $\lambda\psi(x_1, x_2, x_3)$, $0 \neq \lambda \in \mathbb{C}$.

En Mecánica Cuántica, la función (dependiente del tiempo) que describe una onda plana es $\psi(t, x_1, x_2, x_3) = e^{(k_1x_1+k_2x_2+k_3x_3-\omega t)i}$, donde $k_1, k_2, k_3, \omega \in \mathbb{R}$. En este caso, como $|\psi(t, x_1, x_2, x_3)| = 1$, la probabilidad es proporcional al volumen de Ω.

10. En general, fijado un conjunto X y un K-espacio vectorial E, las aplicaciones $f\colon X \to E$ forman un K-espacio vectorial, con las operaciones

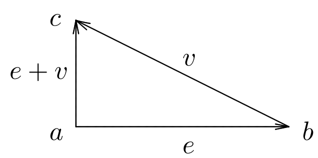
$$(f + h)(x) := f(x) + h(x) \quad , \quad (\lambda f)(x) := \lambda\big(f(x)\big),$$

y el 0 es la aplicación $f\colon X \to E$ tal que $f(x) = 0 \in E$ en todo punto $x \in X$.

11. Todo espacio vectorial complejo es también, de modo natural, un espacio vectorial real, porque todo número real es un número complejo; pero los conceptos (como el subespacio vectorial Ke generado por un vector e) son muy distintos según que los escalares considerados sean $\mathbb{C}$ o $\mathbb{R}$.

12. Si E y E' son dos K-espacios vectoriales, su **producto directo** $E \times E'$ es un K-espacio vectorial, con las operaciones $(e, e') + (v, v') := (e + v, e' + v')$, $\lambda(e, e') := (\lambda e, \lambda e')$, y el 0 es la pareja $(0, 0)$.

13. Dado un espacio vectorial E, si dibujamos una flecha con origen en un punto p y final en otro q, nos referimos al vector $\vec{pq} := q - p$. Dados tres puntos $a, b, c \in E$, si ponemos $e = b - a$, $v = c - b$, tendremos que $e + v = c - b + b - a = c - a$:



Definición: Dado un subespacio vectorial V de un espacio vectorial E y un punto $p \in E$, la **subvariedad lineal** de E que pasa por p con **dirección** V es

$$p + V := \{e \in E\colon e = p + v, \exists v \in V\} = \{p + v\}_{v \in V}.$$

Una propiedad fundamental de las subvariedades lineales es que *si una subvariedad lineal $p+V$ pasa por un punto $q \in E$, entonces $p+V = q+V$*. Por tanto, dos subvariedades lineales de igual dirección o son disjuntas o coinciden.

En efecto, si $q \in p + V$, tendremos $q = p + v$ para algún vector $v \in V$; luego

$$q + V = p + v + V = p + V,$$

pues $v + V = V$ (todo vector $v' \in V$ es $v' = v + (v' - v)$ y $v' - v \in V$).

Ejemplo: Sea $A \in M_{m\times n}(K)$ y $B \in K^m$. Si el sistema de ecuaciones lineales $AX = B$ es compatible, sus soluciones forman una subvariedad lineal de K^n, de dirección $AX = 0$, porque si X_0 es una solución particular, $AX_0 = B$, entonces todas las soluciones del sistema $AX = B$ son $X = X_0 + Y$, donde $AY = 0$.

2.2. Teoría de la Dimensión

Definiciones: Diremos que unos vectores $e_1, \ldots, e_n$ de un espacio vectorial E lo generan, o que forman un **sistema de generadores** de E cuando todo vector de E es una combinación lineal de $e_1, \ldots, e_n$ con coeficientes en K:

$$Ke_1 + \ldots + Ke_n = E.$$

Diremos que $e_1, \ldots, e_n$ son **linealmente dependientes** si $\lambda_1 e_1 + \ldots + \lambda_n e_n = 0$ para ciertos escalares $\lambda_1, \ldots, \lambda_n$ no todos nulos (si admiten una relación de dependencia lineal).

En caso contrario diremos que son **linealmente independientes**.

Es decir, $e_1, \ldots, e_n$ son linealmente independientes cuando la única combinación lineal nula es la que tiene todos los coeficientes nulos:

$$\lambda_1, \ldots, \lambda_n \in K \text{ y } \lambda_1 e_1 + \ldots + \lambda_n e_n = 0 \;\Rightarrow\; \lambda_1 = \ldots = \lambda_n = 0.$$

Definición: Diremos que una sucesión de vectores $e_1, \ldots, e_n$ de un espacio vectorial E es una **base** de E cuando tales vectores sean linealmente independientes y generen E.

En tal caso, cada vector $e \in E$ se escribe de modo <u>único</u> como combinación lineal

$$e = x_1 e_1 + \ldots + x_n e_n$$

con coeficientes en K. En efecto, si $e = x_1 e_1 + \ldots + x_n e_n = y_1 e_1 + \ldots + y_n e_n$, tendremos que $0 = x_1 e_1 + \ldots + x_n e_n - (y_1 e_1 + \ldots + y_n e_n) = (x_1 - y_1) e_1 + \ldots + (x_n - y_n) e_n$, y $x_i - y_i = 0$ para todo índice i, porque los vectores $e_1, \ldots, e_n$ son linealmente independientes.

Diremos que $(x_1, \ldots, x_n) \in K^n$ son las **coordenadas** del vector e en la base $e_1, \ldots, e_n$, y pondremos $e = (x_1, \ldots, x_n)$ cuando la base $e_1, \ldots, e_n$ se sobrentiende.

Nótese que $e_1 = (1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$. Además, las coordenadas de una suma de vectores $e + v$ se obtienen sumando las coordenadas de e y v, y que las coordenadas de λe se obtienen multiplicando por λ las coordenadas de e.

Ejemplos:

1. Los vectores $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, \ldots, 0)$, $\ldots$, $e_n = (0, \ldots, 0, 1)$ forman una base de K^n, llamada **base usual** de K^n. Las coordenadas de un vector $e = (a_1, \ldots, a_n)$ de K^n en esta base son precisamente $(a_1, \ldots, a_n)$, porque $e = a_1 e_1 + \ldots + a_n e_n$.

2. Una base de $M_{2\times 2}(K)$ está formada por las matrices

$$U_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, U_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, U_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, U_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

   y las coordenadas en esta base de una matriz $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ son $(a_{11}, a_{12}, a_{21}, a_{22})$.

   Igualmente, las matrices $m \times n$ que tienen todos sus coeficientes nulos, excepto uno que es la unidad, definen una base de $M_{m\times n}(K)$.

3. Los polinomios $1, x, \ldots, x^n$ forman una base del K-espacio vectorial $P_n = K + \ldots + Kx^n$. Las coordenadas de un polinomio $a_0 + a_1 x + \ldots + a_n x^n$ en esta base son $(a_0, \ldots, a_n)$.

4. Todo vector no nulo $e \in E$ es linealmente independiente, porque $\lambda e = 0 \Rightarrow \lambda = 0$; luego e es una base del subespacio vectorial Ke que genera.

5. Sean $\alpha_1, \ldots, \alpha_n$ números complejos distintos entre sí. Las funciones $e^{\alpha_1 x}, \ldots, e^{\alpha_n x}$ de una variable real x, son linealmente independientes, y forman por tanto una base del subespacio vectorial que generan $\mathbb{C}e^{\alpha_1 x} + \ldots + \mathbb{C}e^{\alpha_n x}$.

Para demostrarlo procedemos por inducción sobre n, y es claro cuando $n = 1$ porque $e^{\alpha x} \neq 0$. Cuando $n > 1$, si $\phi(x) = \lambda_1 e^{\alpha_1 x} + \ldots + \lambda_n e^{\alpha_n x} = 0$, derivando

$$0 = \phi'(x) = \alpha_1 \lambda_1 e^{\alpha_1 x} + \ldots + \alpha_n \lambda_n e^{\alpha_n x}.$$

$$0 = \alpha_n \phi(x) - \phi'(x) = (\alpha_n - \alpha_1)\lambda_1 e^{\alpha_1 x} + \ldots + (\alpha_n - \alpha_{n-1})\lambda_{n-1} e^{\alpha_{n-1} x}.$$

Por hipótesis de inducción $(\alpha_n - \alpha_1)\lambda_1 = \ldots = (\alpha_n - \alpha_{n-1})\lambda_{n-1} = 0$, y vemos que $\lambda_1 = \ldots = \lambda_{n-1} = 0$ porque $\alpha_i \neq \alpha_j$ cuando $i \neq j$. Luego $0 = \phi(x) = \lambda_n e^{\alpha_n x}$, y concluimos que también $\lambda_n = 0$.

El siguiente resultado nos da el significado geométrico de la independencia lineal:

Teorema 2.1 *Unos vectores no nulos $e_1, \ldots, e_n \in E$ son linealmente dependientes si y sólo si alguno de ellos es combinación lineal de los anteriores.*

Demostración: Si $e_1, \ldots, e_n$ admiten una relación de dependencia lineal $\lambda_1 e_1 + \ldots + \lambda_n e_n = 0$, tomamos el mayor índice i tal que $\lambda_i \neq 0$. Tenemos que $i > 1$ porque $e_1 \neq 0$, y

$$e_i = -\tfrac{\lambda_1}{\lambda_i} e_1 - \tfrac{\lambda_2}{\lambda_i} e_2 - \ldots - \tfrac{\lambda_{i-1}}{\lambda_i} e_{i-1}.$$

Recíprocamente, si existe algún índice i tal que $e_i = \mu_1 e_1 + \ldots + \mu_{i-1} e_{i-1}$, tendremos que $\mu_1 e_1 + \ldots + \mu_{i-1} e_{i-1} - e_i = 0$, y los vectores $e_1, \ldots, e_n$ son linealmente dependientes.

Lema Fundamental: *Sea E un espacio vectorial generado por n vectores $e_1, \ldots, e_n$. Si r vectores $v_1, \ldots, v_r \in E$ son linealmente independientes, entonces $r \leq n$.*

Demostración: Procedemos por reducción al absurdo. Supongamos que $r > n$.

Como v_1 es combinación lineal de $e_1, \ldots, e_n$, los vectores $v_1, e_1, \ldots, e_n$ son linealmente dependientes y generan E. Por el teorema anterior, alguno es combinación lineal de los anteriores y, después de eliminarlo, los vectores siguen generando E. Reordenando $e_1, \ldots, e_n$ podemos suponer que es e_1, de modo que $v_1, e_2, \ldots, e_n$ generan E.

Ahora $v_1, v_2, e_2, \ldots, e_n$ son linealmente dependientes y generan E, así que algún vector es combinación lineal de los anteriores (no puede ser v_2 porque los vectores v_1, v_2 son independientes) y, después de eliminarlo, los vectores siguen generando E. Reordenando $e_2, \ldots, e_n$ podemos suponer que es e_2, de modo que $v_1, v_2, e_3, \ldots, e_n$ generan E.

Repitiendo el razonamiento obtendremos que los vectores $v_1, \ldots, v_n$ generan E, de modo que $v_{n+1} = \lambda_1 v_1 + \ldots + \lambda_n v_n$. Absurdo, $v_1, \ldots, v_n, v_{n+1}$ son linealmente independientes.

Ejemplo: Si cierto menor de orden r de una matriz A no es nulo, y el rango de A es mayor que r, por el Lema fundamental alguna columna no está en el subespacio vectorial que generan las columnas correspondientes al menor, y por tanto esas $r + 1$ columnas son linealmente independientes (y lo mismo es cierto para las filas): el menor puede ampliarse hasta obtener un menor no nulo de orden $r + 1$.

Teorema 2.2 *Todas las bases de un espacio vectorial tienen igual número de vectores.*

Demostración: Sean $e_1, \ldots, e_n$ y $v_1, \ldots, v_r$ dos bases de un espacio vectorial E.

Como $v_1, \ldots, v_r \in E = Ke_1 + \ldots + Ke_n$ son linealmente independientes, por el lema fundamental $r \leq n$. Como $e_1, \ldots, e_n \in E = Kv_1 + \ldots + Kv_r$ son linealmente independientes, también tenemos que $n \leq r$; luego $n = r$.

Definición: El espacio vectorial $E = 0$ tiene dimensión 0, y la **dimensión** de un espacio vectorial $E \neq 0$ es el número de vectores de cualquier base de E, y se denota $\dim_K E$. Cuando ninguna familia finita de vectores de un espacio vectorial $E \neq 0$ sea una base de E, diremos que E tiene dimensión infinita.

La dimensión de una subvariedad lineal $X = p + V$ es la de su dirección V, y diremos que dos subvariedades lineales $p + V$, $q + W$ de E de la misma dimensión son **paralelas** cuando tienen la misma dirección, $V = W$.

Las subvariedades lineales de dimensión 1 y 2 se llaman **rectas** y **planos** respectivamente.

Ejemplos:

1. $\dim_K K^n = n$, $\dim_K M_{m\times n}(K) = mn$, y $\dim P_n = n + 1$.

2. $\dim_{\mathbb{C}}(\mathbb{C}e^{\alpha_1 x} + \ldots + \mathbb{C}e^{\alpha_n x}) = n$, cuando $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ son distintos entre sí.

3. De acuerdo con el lema fundamental, en un espacio vectorial de dimensión n no puede haber más de n vectores linealmente independientes.

4. Si un vector e no es nulo, $\dim_K(Ke) = 1$. Si además $v \notin Ke$, entonces e y v son linealmente independientes por 2.1; luego forman una base de $Ke + Kv$, y $\dim_K(Ke + Kv) = 2$.

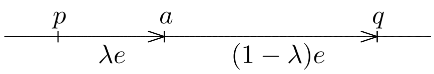
5. *Por dos puntos distintos p y $q = p + e$ pasa una única recta,*

$$p + Ke = \{p + \lambda e = (1 - \lambda)p + \lambda q\}_{\lambda\in K},$$

formada por todas las combinaciones lineales $\alpha p + \beta q$, con $\alpha + \beta = 1$.

En efecto, si una recta $a + V$ pasa por p, entonces $a + V = p + V$. Si además pasa por $q = p + e$, entonces $e \in V$ y $Ke = V$, pues en caso contrario existiría $v \in V$, $v \notin Ke$, de modo que $e, v \in V$ son linealmente independientes, en contra de que $\dim V = 1$.

En el caso real, $K = \mathbb{R}$, se define además el **segmento** de extremos p y q como el conjunto formado por los puntos $a = p + \lambda e = (1 - \lambda)p + \lambda q$, con $0 \leq \lambda \leq 1$, y diremos que tal punto a divide al segmento en la proporción $\lambda : (1 - \lambda)$ porque $a - p = \frac{\lambda}{1-\lambda}(q - a)$:



Diremos que $m = p + \frac{1}{2}e = \frac{p+q}{2}$ es el **punto medio** entre p y q, porque $m - p = q - m$

Teorema 2.3 *Todo sistema de generadores $e_1, \ldots, e_n$ de un espacio vectorial $E \neq 0$ contiene una base de E. Por tanto $n \geq \dim E$, y si además $n = \dim E$, entonces los vectores $e_1, \ldots, e_n$ ya forman una base de E.*

Demostración: Sea d el máximo número de vectores linealmente independientes que podamos extraer entre $e_1, \ldots, e_n$. Reordenando los vectores si fuera necesario, podemos suponer que $e_1, \ldots, e_d$ son linealmente independientes.

Ahora $e_{d+1}, \ldots, e_n \in \langle e_1, \ldots, e_d\rangle$, porque si $e_j \notin \langle e_1, \ldots, e_d\rangle$, entonces $e_1, \ldots, e_d, e_j$ son linealmente independientes por 2.1, lo que contradice la elección del número d.

Luego $E = \langle e_1, \ldots, e_d, \ldots e_n\rangle = \langle e_1, \ldots, e_d\rangle$, de modo que $E = \langle e_1, \ldots, e_d\rangle$, y concluimos que $e_1, \ldots, e_d$ es una base de E.

Por último, si $n = \dim E$, entonces los vectores $e_1, \ldots, e_n$ ya forman una base de E; porque una base de E no puede tener menos de n vectores según 2.2.

Corolario 2.4 *Sea $e_1, \ldots, e_m$ una base de un espacio vectorial E. Si $A \in M_{m\times n}(K)$ es la matriz que tiene por columnas las coordenadas de $v_1, \ldots, v_n \in E$ en tal base, se cumple que*

$$\boxed{\dim(Kv_1 + \ldots + Kv_m) = \operatorname{rg} A.}$$

Demostración: Pongamos $r = \operatorname{rg} A$ y $d = \dim(Kv_1 + \ldots + Kv_n)$.

Como $v_1, \ldots, v_n$ contiene una base $v_{i_1}, \ldots, v_{i_d}$ de $Kv_1 + \ldots + Kv_n$, las columnas $i_1, \ldots, i_d$ de la matriz A son linealmente independientes (unos vectores son linealmente independientes si y sólo si sus coordenadas son linealmente independientes en K^m) y por tanto $d \leq r$.

Pero también $r \leq d$, porque, por el lema fundamental, en $Kv_1 + \ldots + Kv_n$ no puede haber $d + 1$ vectores linealmente independientes.

Teorema 2.5 *Sea $E \neq 0$ un espacio vectorial de dimensión finita. Toda familia de vectores $e_1, \ldots, e_n \in E$ linealmente independiente se puede ampliar hasta obtener una base de E. Si además $n = \dim E$, entonces $e_1, \ldots, e_n$ ya forman una base de E.*

Demostración: Añadimos vectores de E hasta obtener una familia linealmente independiente $e_1, \ldots, e_n, e'_1, \ldots, e'_s$ que no pueda ampliarse de modo que lo siga siendo (el proceso termina porque, en virtud del lema fundamental, siempre $n + s \leq \dim E$).

Ahora $e_1, \ldots, e_n, e'_1, \ldots, e'_s$ ya es una base de E por 2.1.

Por último, si $n = \dim E = n + s$, entonces $s = 0$ y $e_1, \ldots, e_n$ ya es base de E.

Ecuaciones Paramétricas e Implícitas: Fijada una base de un espacio vectorial de dimensión finita E, dar ecuaciones **paramétricas** de un subespacio vectorial V de E es dar las coordenadas de un sistema de generadores de V (mejor si forman una base de V), y dar ecuaciones **implícitas** de V es dar un sistema homogéneo de ecuaciones lineales $AX = 0$ (mejor si son linealmente independientes) cuyas soluciones $X^t = (x_1, \ldots, x_n)$ sean las coordenadas de los vectores de V.

Dar ecuaciones **paramétricas** de una subvariedad lineal $X = p + V$ de E es dar las coordenadas de un punto p de X y de un sistema de generadores de la dirección V de X, y dar ecuaciones **implícitas** de X es dar un sistema de ecuaciones lineales $AX = B$ cuyas soluciones sean las coordenadas de los puntos de X,

Pasar de ecuaciones implícitas a paramétricas es sencillamente resolver el sistema, y recíprocamente, dadas las ecuaciones paramétricas de una subvariedad lineal X de dimensión d en un espacio vectorial de dimensión n, un método para hallar ecuaciones implícitas de X es despejar los d parámetros en función de las coordenadas del punto de X (usando d ecuaciones) para después sustituir esos valores en las restantes $n - d$ ecuaciones.

Un método alternativo para hallar ecuaciones implícitas de un subespacio vectorial V de base $v_1, \ldots, v_r$ es usar que, de acuerdo con 2.1, un vector $e \in E$ está en V si y sólo si los vectores $v_1, \ldots, v_r, e$ son linealmente dependientes. Por tanto, dadas ecuaciones paramétricas de V (donde suponemos que la matriz $A = (a_{ij})$ ya es de rango $r = \dim V$)

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \lambda_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \ldots + \lambda_r \begin{pmatrix} a_{1r} \\ \vdots \\ a_{nr} \end{pmatrix}$$

el vector de coordenadas $(x_1, \ldots, x_n)$ está en V precisamente cuando

$$r = \operatorname{rg} \begin{pmatrix} a_{11} & \ldots & a_{1r} & x_1 \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nr} & x_n \end{pmatrix}$$

es decir, cuando se anulen todos los menores de orden $r + 1$ (y basta considerar los que amplíen a un menor no nulo de orden r dado) lo que nos da ecuaciones implícitas de V.

Teorema 2.6 *Sea E un espacio vectorial de dimensión finita. Para todo subespacio vectorial V de E se cumple que*

$$\dim V \leq \dim E\ ,$$

y sólo se da la igualdad cuando $V = E$.

Demostración: La dimensión de V también es finita, porque si $v_1, \ldots, v_r$ es una familia linealmente independiente en V que ya no pueda ampliarse con un vector de V de modo que lo siga siendo (existe porque, si $n = \dim E$, por el lema fundamental en E no puede haber más de n vectores linealmente independientes) entonces $v_1, \ldots, v_r$ es una base de V por 2.1, de modo que $r = \dim V$.

Como $v_1, \ldots, v_r \in E$ son linealmente independientes, según 2.5 tenemos que $r \leq \dim E$, y que si $r = \dim E$, entonces $v_1, \ldots, v_r$ también es una base de E, de modo que

$$E = Kv_1 + \ldots + Kv_r = V.$$

Ejemplo: Sea E un espacio vectorial de dimensión m, y sea $A \in M_{m \times n}(K)$ la matriz que tiene por columnas las coordenadas de unos vectores $v_1, \ldots, v_n \in E$ en cierta base de E.

Por el teorema, $v_1, \ldots, v_n$ generan E cuando $m = \dim \langle v_1, \ldots, v_n \rangle = \operatorname{rg} A$.

Por otra parte, son linealmente independientes cuando $n = \dim \langle v_1, \ldots, v_n \rangle = \operatorname{rg} A$.

Luego forman una base de E cuando $\operatorname{rg} A = m = n$; es decir, cuando $|A| \neq 0$.

Teorema de Rouché-Frobënius (1832-1910, 1849-1917)**:** *Un sistema de ecuaciones lineales $AX = B$ es compatible si y sólo si* $\operatorname{rg} A = \operatorname{rg}(A|B)$ *.*

Demostración: Sean $A_1, \ldots, A_n$ las columnas de A, de modo que el sistema $AX = B$ puede escribirse $x_1 A_1 + \ldots + x_n A_n = B$, y la condición de que sea compatible significa que en K^m tenemos que $B \in \langle A_1, \ldots, A_n \rangle$; es decir, que $\langle A_1, \ldots, A_n \rangle = \langle A_1, \ldots, A_n, B \rangle$.

Ahora bien, el teorema 2.6 afirma que

$$\langle A_1, \ldots, A_n \rangle = \langle A_1, \ldots, A_n, B \rangle \;\Leftrightarrow\; \dim\langle A_1, \ldots, A_n \rangle = \dim\langle A_1, \ldots, A_n, B \rangle$$

y, de acuerdo con 2.4, esta última condición significa que $\operatorname{rg} A = \operatorname{rg}(A|B)$.

2.3. Suma Directa

Teorema 2.7 *Sean V y W dos subespacios vectoriales de dimensión finita de un espacio vectorial E. Se cumple que*

$$\boxed{\dim (V + W) = \dim V + \dim W - \dim (V \cap W).}$$

Demostración: Tomemos una base $u_1, \ldots, u_d$ de $V \cap W$, que ampliamos hasta obtener una base $u_1, \ldots, u_d, v_1, \ldots, v_m$ de V y una base $u_1, \ldots, u_d, w_1, \ldots, w_n$ de W.

Tenemos que $d = \dim (V \cap W)$, $d + m = \dim V$ y $d + n = \dim W$; así que basta ver que los vectores $u_1, \ldots, u_d, v_1, \ldots, v_m, w_1, \ldots, w_n$ forman una base de $V + W$.

Los vectores $u_1, \ldots, u_d, v_1, \ldots, v_m, w_1, \ldots, w_n$ son linealmente independientes:

Si una combinación lineal es nula, $\sum_i \lambda_i u_i + \sum_j \alpha_j v_j + \sum_k \beta_k w_k = 0$, entonces

$$\sum_j \alpha_j v_j = -\sum_i \lambda_i u_i - \sum_k \beta_k w_k \in V \cap W = \langle u_1, \ldots, u_d \rangle,$$
$$\sum_k \beta_k w_k = -\sum_i \lambda_i u_i - \sum_j \alpha_j v_j \in V \cap W = \langle u_1, \ldots, u_d \rangle.$$

Como $\langle u_1, \ldots, u_d \rangle \cap \langle v_1, \ldots, v_m \rangle = 0 = \langle u_1, \ldots, u_d \rangle \cap \langle w_1, \ldots, w_n \rangle$, concluimos que $\sum_j \alpha_j v_j = 0$ y $\sum_k \beta_k w_k = 0$, y por tanto también $\sum_i \lambda_i u_i = 0$.

Luego todos los coeficientes $\lambda_i, \alpha_j, \beta_k$ son nulos, pues las familias $u_1, \ldots, u_d$ y $v_1, \ldots, v_m$ y $w_1, \ldots, w_n$ son linealmente independientes.

Los vectores $u_1, \ldots, u_d, v_1, \ldots, v_m, w_1, \ldots, w_n$ generan $V + W$:

Como para todo vector $e \in E$ se cumple que $Ke + Ke = Ke$, tenemos que

$$\begin{aligned} V + W &= Ku_1 + \ldots + Ku_d + Kv_1 + \ldots + Kv_m + Ku_1 + \ldots + Ku_d + Kw_1 + \ldots + Kw_n \\ &= Ku_1 + \ldots + Ku_d + Kv_1 + \ldots + Kv_m + Kw_1 + \ldots + Kw_n. \end{aligned}$$

Corolario 2.8 *Si E y E' son dos K-espacios vectoriales de dimensión finita,*

$$\dim (E \times E') = \dim E + \dim E'.$$

Demostración: Pongamos $V = E \times 0$, $W = 0 \times E'$. Tenemos que $\dim E = \dim V$, $\dim E' = \dim W$, y terminamos porque $E \times E' = V + W$ y $0 = V \cap W$.

Definición: Sea E un espacio vectorial. Diremos que la suma $V_1 + \ldots + V_r$ de unos subespacios vectoriales $V_1, \ldots, V_r$ de E es **directa** si cada vector $e \in V_1 + \ldots + V_r$ descompone de modo único en la forma $e = v_1 + \ldots + v_r$, donde $v_i \in V_i$.

En tal caso, el subespacio vectorial $V_1 + \ldots + V_r$ se denota $V_1 \oplus \ldots \oplus V_r$.

Diremos que dos subespacios vectoriales V y W de E son **suplementarios** (o que W es un suplementario de V en E) cuando $E = V \oplus W$; i.e., cuando cada vector de E descompone, y de modo único, en suma de un vector de V y otro de W.

Ejemplos:

1. Si $e_1, \ldots, e_n$ es una base de un espacio vectorial E, entonces cada vector $e \in E$ descompone de modo único como combinación lineal $e = \lambda_1 e_1 + \ldots + \lambda_n e_n$; luego
$$E = Ke_1 \oplus \ldots \oplus Ke_n,$$
y un suplementario de $Ke_1 \oplus \ldots \oplus Ke_r$ en E es $Ke_{r+1} \oplus \ldots \oplus Ke_n$.

2. Para hallar un suplementario de un subespacio vectorial V de E basta ampliar una base $v_1, \ldots, v_r$ de V hasta obtener una base $v_1, \ldots, v_r, w_1, \ldots, w_s$ de E, pues un suplementario de $V = Kv_1 \oplus \ldots \oplus Kv_r$ en E es $Kw_1 \oplus \ldots \oplus Kw_s$.

3. Sean $p + V$ y $q + W$ dos subvariedades lineales de un espacio vectorial E. Dar un punto de corte es dar vectores $v \in V$, $w \in W$ tales que $p + v = q + w$; es decir, $q - p = v - w$. Por tanto, la condición necesaria y suficiente para que se corten es que $q - p \in V + W$, y el punto de corte es único cuando la suma $V \oplus W$ es directa.

Teorema 2.9 *La condición necesaria y suficiente para que la suma de dos subespacios vectoriales V y W de un espacio vectorial E sea directa es que $V \cap W = 0$.*

Demostración: Si $e \in V \cap W$, tenemos que $0 = 0 + 0 = e + (-e)$, donde $0, e \in V$ y $0, -e \in W$. Si la suma de V y W es directa, la unicidad de la descomposición del vector 0 en suma de un vector de V y otro de W implica que $e = 0$. Luego $V \cap W = 0$.

Recíprocamente, si $V \cap W = 0$ y un vector $e \in V + W$ admite dos descomposiciones
$$e = v + w = v' + w' \ , \quad \text{donde } v, v' \in V,\ w, w' \in W$$
entonces $v' - v = w - w' \in W$. Como $v' - v \in V$, se sigue que $v' - v \in V \cap W = 0$.

Luego $0 = v' - v = w - w'$, y concluimos que $v = v'$ y $w = w'$.

Es decir, tal descomposición es única, así que la suma de V y W es directa.

Corolario 2.10 *Si la suma de dos subespacios vectoriales V y W de E es directa, entonces*
$$\boxed{\dim (V \oplus W) = \dim V + \dim W.}$$

Demostración: De acuerdo con 2.9 tenemos que $V \cap W = 0$.

Corolario 2.11 *Sea E un espacio vectorial de dimensión n y sea V un subespacio vectorial de E de dimensión d. Un suplementario de V en E es cualquier subespacio vectorial W de E de dimensión $n - d$ que cumpla que $V \cap W = 0$.*

Demostración: Como $V \cap W = 0$, tenemos que $\dim (V \oplus W) = d + (n - d) = \dim E$; luego $V \oplus W = E$ por 2.6.

3. Aplicaciones Lineales

Definición: Una aplicación $f\colon E\to E'$ entre dos K-espacios vectoriales es **K-lineal** si es morfismo de grupos, $f(e+v)=f(e)+f(v)$, $\forall e,v\in E$, y

$$f(\lambda\cdot e)=\lambda\cdot f(e)\ \text{ para todo } \lambda\in K,\ e\in E,$$

y, cuando $K=\mathbb{C}$, es **antilineal** si es morfismo de grupos y $f(\lambda\cdot e)=\bar\lambda\cdot f(e)$, $\forall\lambda\in K$, $e\in E$.

Toda aplicación lineal f cumple que $f(\lambda_1e_1+\ldots+\lambda_ne_n)=\lambda_1f(e_1)+\ldots+\lambda_nf(e_n)$, y toda aplicación antilineal f cumple que $f(\lambda_1e_1+\ldots+\lambda_ne_n)=\bar\lambda_1f(e_1)+\ldots+\bar\lambda_nf(e_n)$.

Teorema 3.1 *Sean E, E' dos K-espacios vectoriales, y $e_1,\ldots,e_n$ una base de E. Dados n vectores $e'_1,\ldots,e'_n\in E'$, se cumple que la aplicación*

$$f\colon E\longrightarrow E'\ ,\quad f(x_1,\ldots,x_n):=f(x_1e_1+\ldots+x_ne_n)=x_1e'_1+\ldots+x_ne'_n,$$

es K-lineal, y es la única aplicación K-lineal $f\colon E\to E'$ tal que $f(e_1)=e'_1,\ldots,f(e_n)=e'_n$.

Demostración: Veamos que la aplicación $f\colon E\to E'$, $f(\sum_i x_ie_i)=\sum_i x_ie'_i$, es lineal:

$$\begin{aligned}f\big(\textstyle\sum_i x_ie_i+\sum_i y_ie_i\big)&=f\big(\textstyle\sum_i(x_i+y_i)e_i\big)=\sum_i(x_i+y_i)e'_i=\\&=\textstyle\sum_i x_ie'_i+\sum_i y_ie'_i=f\big(\sum_i x_ie_i\big)+f\big(\sum_i y_ie_i\big).\\ f\big(\lambda\textstyle\sum_i x_ie_i\big)&=f\big(\textstyle\sum_i\lambda x_ie_i\big)=\sum_i\lambda_ix_ie'_i=\lambda\sum_i x_ie'_i=\lambda f\big(\sum_i x_ie_i\big).\end{aligned}$$

Ahora, si $h\colon E\to E'$ es otra aplicación K-lineal que cumple $e'_1=h(e_1),\ldots,e'_n=h(e_n)$, entonces

$$h(x_1e_1+\ldots+x_ne_n)=x_1h(e_1)+\ldots+x_nh(e_n)=x_1e'_1+\ldots+x_ne'_n=f(x_1e_1+\ldots+x_ne_n).$$

Ejemplos:

1. Dados $e_1,\ldots,e_n\in E$, la aplicación $f\colon K^n\to E$, $f(x_1,\ldots,x_n)=x_1e_1+\ldots+x_ne_n$, es K-lineal según 3.1. Su imagen es $\operatorname{Im}f=Ke_1+\ldots+Ke_n$; así que f es epiyectiva cuando $e_1,\ldots,e_n$ generan E. La condición de que $e_1,\ldots,e_n$ sean linealmente independientes significa que $\operatorname{Ker}f=0$; así que f es inyectiva cuando $e_1,\ldots,e_n$ son linealmente independientes. Por tanto, f es biyectiva cuando $e_1,\ldots,e_n$ forman una base de E.

2. Según 3.1, toda aplicación lineal $f\colon\mathbb{R}^n\to\mathbb{R}$ es $f(x_1,\ldots,x_n)=a_1x_1+\ldots+a_nx_n$ para ciertos números reales $a_1,\ldots,a_n$; y las únicas aplicaciones $f\colon\mathbb{R}\to\mathbb{R}$ que son lineales son las funciones $f(x)=ax$, donde $a\in\mathbb{R}$.

3. Dada una matriz $A\in M_{m\times n}(K)$, la aplicación $f\colon K^n\to K^m$, $f(X)=AX$, (es decir, $f(x_1,\ldots,x_n)=x_1A_1+\ldots+x_nA_n$, donde $A_1,\ldots,A_n$ son las columnas de A) es lineal según 3.1. Su núcleo $\operatorname{Ker}f$ está formado por todas las soluciones de la ecuación homogénea $AX=0$, y la condición $B\in\operatorname{Im}f$ significa que el sistema $AX=B$ es compatible.

4. Sea E el espacio vectorial complejo formado por las funciones complejas de variable real $\psi(x)=u(x)+iv(x)$ infinitamente derivables. La derivada define una aplicación $\mathbb{C}$-lineal $\frac{\mathrm{d}}{\mathrm{d}x}\colon E\longrightarrow E$, $\frac{\mathrm{d}}{\mathrm{d}x}\psi:=u'+iv'$. Además, fijada una función $\phi\in E$, la aplicación $T\colon E\longrightarrow E$, $T(\psi)=\phi\psi$, también es $\mathbb{C}$-lineal.

5. La **traza** $\operatorname{tr}\colon M_{n\times n}(K)\to K$, $\operatorname{tr}(a_{ij})=a_{11}+\ldots+a_{nn}$, es una aplicación lineal.

6. La aplicación $0\colon E\to E'$, que transforma cualquier vector $e\in E$ en el vector nulo $0\in E'$, es lineal.

7. Sea $f\colon E\to E'$ una aplicación K-lineal, y sea V un subespacio vectorial de E. La aplicación $f|_V\colon V\to E'$, llamada **restricción** de f a V, también es K-lineal.

Teorema 3.2 *Sean $f, h\colon E \to E'$ dos aplicaciones K-lineales y $\alpha \in K$. También son K-lineales las aplicaciones*

$$f + h\colon E \longrightarrow E' \ , \quad (f+h)(e) = f(e) + h(e),$$
$$\alpha f\colon E \longrightarrow E' \ , \quad (\alpha f)(e) = \alpha \cdot f(e).$$

Por tanto, el conjunto de todas las aplicaciones K-lineales de E en E' es un subespacio vectorial del espacio vectorial formado por todas las aplicaciones de E en E'.

Demostración: Si $e, v \in E$ y $\lambda \in K$, se cumple que

$$(f+h)(e+v) = f(e+v) + h(e+v) = f(e) + f(v) + h(e) + h(v) = (f+h)(e) + (f+h)(v),$$
$$(f+h)(\lambda e) = f(\lambda e) + h(\lambda e) = \lambda f(e) + \lambda h(e) = \lambda(f(e) + h(e)) = \lambda(f+h)(e),$$
$$(\alpha f)(e+v) = \alpha f(e+v) = \alpha(f(e) + f(v)) = \alpha f(e) + \alpha f(v) = (\alpha f)(e) + (\alpha f)(v),$$
$$(\alpha f)(\lambda e) = \alpha f(\lambda e) = \alpha\lambda f(e) = \lambda(\alpha f(e)) = \lambda(\alpha f)(e).$$

Teorema 3.3 *Si dos aplicaciones $f\colon E \to E'$ y $h\colon E' \to E''$ son K-lineales, también lo es su composición $h \circ f\colon E \to E''$, $(h \circ f)(e) = h\big(f(e)\big)$.*

Demostración: Como f y h son morfismos de grupos, también lo es $h \circ f$ por 1.2.

Además, para todo $\lambda \in K$, $e \in E$ tenemos que

$$(hf)(\lambda e) = h\big(f(\lambda e)\big) = h\big(\lambda \cdot f(e)\big) = \lambda \cdot h(f(e)) = \lambda \cdot (hf)(e).$$

3.1. Matriz de una Aplicación Lineal

Sea $f\colon E \to E'$ una aplicación lineal entre dos espacios vectoriales de dimensión finita, y fijemos una base $e_1, \ldots, e_n$ de E, y una base $e'_1, \ldots, e'_m$ de E'.

Según 3.1, la aplicación lineal $f\colon E \to E'$ está totalmente determinada por los vectores $f(e_1), \ldots, f(e_n) \in E'$, y para ciertos escalares $a_{ij} \in K$ tendremos que

$$f(e_j) = a_{1j}e'_1 + \ldots + a_{mj}e'_m \ , \quad 1 \leq j \leq n.$$

Definición: Llamaremos **matriz de una aplicación lineal** $f\colon E \to E'$ en las bases $e_1, \ldots, e_n$ de E y $e'_1, \ldots, e'_m$ de E' a la matriz $A = (a_{ij}) \in M_{m\times n}(K)$ cuyas columnas están formadas por las coordenadas de los vectores $f(e_1), \ldots, f(e_n)$ en la base $e'_1, \ldots, e'_m$.

Si $(x_1, \ldots, x_n)$ son las coordenadas de un vector $e \in E$ en la base $e_1, \ldots, e_n$, tendremos que las coordenadas $(x'_1, \ldots, x'_m)$ del vector

$$f(e) = f(x_1e_1 + \ldots + x_ne_n) = x_1f(e_1) + \ldots + x_nf(e_n) \in E'$$

en la base $e'_1, \ldots, e'_m$ fijada en E' vienen dadas por las siguientes ecuaciones

$$\left\{\begin{array}{l} x'_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ x'_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ x'_m = a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{array}\right. \quad , \quad \begin{pmatrix} x'_1 \\ \vdots \\ x'_m \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Es decir, si X denota las coordenadas del vector e en la base $e_1, \ldots, e_n$, puestas en columna, entonces las coordenadas X' de $f(e)$ en la base $e'_1, \ldots, e'_m$ son

$$\boxed{X' = AX} \tag{3}$$

Ejemplos: Sea $A \in M_{m\times n}(K)$. La matriz de la aplicación lineal $f\colon K^n \to K^m$, $f(X) = AX$, en las bases usuales de K^n y K^m es A.

La matriz de la identidad $\mathrm{Id}\colon E \to E$, en cualquier base de E, es la matriz unidad I.

Si A y B son las matrices de dos aplicaciones lineales $f, h\colon E \to E'$ en ciertas bases de E y E', entonces la matriz de $f+h$ en tales bases es $A+B$, y la matriz de λf es λA, $\forall \lambda \in K$.

Matriz de la Composición: Dadas aplicaciones K-lineales $f\colon E \to E'$ y $g\colon E' \to E''$, si fijamos bases en E, E' y E'', tendremos las matrices A, B, C de $f, g, g\circ f$ en tales bases. Si X son las coordenadas de un vector $e \in E$, entonces AX son las coordenadas de $f(e)$, y BAX son las coordenadas de $g(f(e)) = (g\circ f)(e)$. Luego $BAX = CX$ para todo $X \in K^n$, y vemos que $C = BA$: *La matriz de la composición $g \circ f$ es el producto BA.*

Cambio de Base: Sea $e_1, \ldots, e_n$ una base de un espacio vectorial E. Si consideramos una nueva base $v_1, \ldots, v_n$ de E, tendremos escalares $b_{ij} \in K$ tales que

$$v_j = b_{1j}e_1 + \ldots + b_{nj}e_n \quad , \quad 1 \leq j \leq n, \tag{4}$$

y llamaremos **matriz de cambio de base** a la matriz $B = (b_{ij}) \in M_{n\times n}(K)$ cuyas columnas están formadas por las coordenadas de los vectores de la nueva base en la antigua.

Es decir, B es la matriz de la identidad $\mathrm{Id}\colon E \to E$ cuando en el espacio de salida se considera la nueva base $v_1, \ldots, v_n$ y en el de llegada la base antigua $e_1, \ldots, e_n$.

De acuerdo con 3, si Y son las coordenadas de un vector $e \in E$ en la nueva base, y X son las coordenadas de $\mathrm{Id}(e) = e$ en la base antigua, tendremos que $X = BY$.

Por otra parte, también tenemos una matriz de cambio de base $C \in M_{n\times n}(K)$ cuando se considera que $v_1, \ldots, v_n$ es la base inicial de E y que $e_1, \ldots, e_n$ es la nueva base, de modo que $Y = CX$. Luego $X = BCX$ y $Y = CBY$, y como estas igualdades son válidas para cualesquiera columnas X, Y, se concluye que $BC = CB = I$. Es decir, la matriz B es invertible, y su inversa es la matriz C. En resumen, la relación entre las bases $e_1, \ldots, e_n$ y $v_1, \ldots, v_n$, y las coordenadas X e Y de un mismo vector $e \in E$ en esas bases, es

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = B^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad , \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = B \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad , \quad \boxed{X = BY}\,. \tag{5}$$

Veamos ahora cómo se transforma la matriz de una aplicación lineal cuando se cambian las bases. Sea $f\colon E \to E'$ una aplicación lineal y sea $A \in M_{m\times n}(K)$ su matriz en ciertas bases $e_1, \ldots, e_n$ y $e'_1, \ldots, e'_m$ de E y E' respectivamente. Consideremos nuevas bases $v_1, \ldots, v_n$ y $v'_1, \ldots, v'_m$ de E y E' respectivamente, las correspondientes matrices de cambio de base $B \in M_{n\times n}(K)$ y $C \in M_{m\times m}(K)$, y sea $\tilde{A} \in M_{m\times n}(K)$ la matriz de f en estas nuevas bases de E y E'. Vamos a determinar $\tilde{A}$ en función de A, B y C.

Sean X e Y las coordenadas de un vector $e \in E$ en las bases $e_1, \ldots, e_n$ y $v_1, \ldots, v_n$ respectivamente, y sean X' e Y' las coordenadas de $f(e) \in E'$ en las bases $e'_1, \ldots, e'_m$ y $v'_1, \ldots, v'_m$ respectivamente. De acuerdo con 3 y 5 tendremos que

$$X' = AX \quad , \quad Y' = \tilde{A}Y \quad , \quad X = BY \quad , \quad X' = CY'.$$
$$\tilde{A}Y = Y' = C^{-1}X' = C^{-1}AX = C^{-1}ABY.$$

Como esta igualdad es válida para cualquier columna $Y \in K^n$, vemos que

$$\boxed{\tilde{A} = C^{-1}AB}\,. \tag{6}$$

3.2. Núcleo e Imagen de una Aplicación Lineal

Teorema 3.4 *Sea $f\colon E \to E'$ una aplicación lineal. Su núcleo $\operatorname{Ker} f = \{e \in E\colon f(e) = 0\}$ es un subespacio vectorial de E, y su imagen $\operatorname{Im} f = \{e' \in E'\colon e' = f(e), \exists e \in E\} = \{f(e)\}_{e\in E}$ es un subespacio vectorial de E'.*

Demostración: Como f es morfismo de grupos, de acuerdo con 1.3 tenemos que $\operatorname{Ker} f$ es un subgrupo de E, y que $\operatorname{Im} f$ es un subgrupo de E'.

Ahora, si $e \in \operatorname{Ker} f$ y λK, entonces $f(\lambda e) = \lambda f(e) = \lambda \cdot 0 = 0$, así que $\lambda e \in \operatorname{Ker} f$.

Si $e' \in \operatorname{Im} f$, existe $e \in E$ tal que $e' = f(e)$; luego $\lambda e' = \lambda f(e) = f(\lambda e) \in \operatorname{Im} f$, $\forall \lambda \in K$.

Proposición 3.5 *Si $f\colon E \to E'$ es una aplicación lineal, y $e_1, \ldots, e_n$ es una base de E,*

$$\operatorname{Im} f = \langle f(e_1), \ldots, f(e_n) \rangle.$$

Por tanto, si A es la matriz de f en ciertas bases de E y E', sus columnas son las coordenadas de un sistema de generadores de $\operatorname{Im} f$*, y se cumple que*

$$\boxed{\dim (\operatorname{Im} f) = \operatorname{rg} A.}$$

En particular f es epiyectiva si y sólo si $\operatorname{rg} A = \dim E'$.

Demostración: Como $f(\sum_i \lambda_i e_i) = \sum_i \lambda_i f(e_i)$, tenemos que $\operatorname{Im} f = \langle f(e_1), \ldots, f(e_n) \rangle$.

Ahora bien, como las columnas de A son las coordenadas de $f(e_1), \ldots, f(e_n)$ en cierta base de E', de acuerdo con 2.4 tenemos que $\dim \langle f(e_1), \ldots, f(e_n) \rangle = \operatorname{rg} A$.

Por último, la aplicación f es epiyectiva si y sólo si $\operatorname{Im} f = E'$.

Ejemplo: Sea $f\colon E \to E'$ una aplicación lineal, y sea $A = (a_{ij})$ su matriz en ciertas bases de E y E'. El núcleo de f está formado por los vectores de E cuyas coordenadas X en la base fijada cumplen que $AX = 0$,

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \text{(ecuaciones implícitas de } \operatorname{Ker} f\text{)}$$

y, de acuerdo con 3.5, si las columnas de A se denotan $A_1, \ldots, A_n$, la imagen de f está formada por los vectores de E' cuyas coordenadas X' cumplen que $X' = \lambda_1 A_1 + \ldots + \lambda_n A_n$ para ciertos escalares $\lambda_1, \ldots, \lambda_n$:

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_m \end{pmatrix} = \lambda_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + \lambda_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \qquad \text{(ecuaciones paramétricas de } \operatorname{Im} f\text{)}$$

Definición: Una aplicación K-lineal $f\colon E \to E'$ es un **isomorfismo** cuando es biyectiva, y en tal caso la aplicación inversa $f^{-1}\colon E' \to E$ también es K-lineal y por supuesto biyectiva, así que f^{-1} también es un isomorfismo.

En efecto, si $e', v' \in E'$, entonces $e' = f(e)$ y $v' = f(v)$, donde $e, v \in E$, de modo que

$$f^{-1}(e' + v') = f^{-1}\big(f(e) + f(v)\big) = f^{-1}\big(f(e+v)\big) = e + v = f^{-1}(e') + f^{-1}(v'),$$
$$f^{-1}(\lambda e') = f^{-1}\big(\lambda f(e)\big) = f^{-1}\big(f(\lambda e)\big) = \lambda e = \lambda \cdot f^{-1}(e').$$

Ejemplos:

1. Los isomorfismos transforman vectores linealmente independientes en vectores linealmente independientes, y sistemas de generadores en sistemas de generadores; luego bases en bases. Por tanto, si dos K-espacios vectoriales E y E' son isomorfos, tendremos que $\dim E = \dim E'$.

2. Si $e_1, \dots, e_n$ es una base de un K-espacio vectorial E, entonces la aplicación lineal

$$f \colon K^n \longrightarrow E \ , \quad f(x_1, \dots, x_n) = x_1 e_1 + \dots + x_n e_n,$$

es un isomorfismo. El isomorfismo inverso $f^{-1} \colon E \to K^n$ asigna a cada vector $e \in E$ sus coordenadas $(x_1, \dots, x_n)$ en la base $e_1, \dots, e_n$.

Por tanto, *todo K-espacio vectorial de dimensión n es isomorfo a K^n.*

3. Si $V_1, \dots, V_n$ son subespacios vectoriales de un espacio vectorial E, la aplicación

$$s \colon V_1 \times \dots \times V_n \longrightarrow V_1 + \dots + V_n \quad , \quad s(v_1, \dots, v_n) = v_1 + \dots + v_n \ ,$$

es lineal y epiyectiva. Además esta aplicación lineal s es inyectiva precisamente cuando la suma es directa, de modo que en tal caso $V_1 \times \dots \times V_n \simeq V_1 \oplus \dots \oplus V_n$.

Teorema de Isomorfía: *Sea $f \colon E \to E'$ una aplicación lineal, y sea V un suplementario del núcleo de f en E; es decir, $E = V \oplus (\operatorname{Ker} f)$.*

La aplicación $f|_V \colon V \to \operatorname{Im} f$, $f|_V(v) = f(v)$, es un isomorfismo lineal, y por tanto

$$\boxed{\dim E = \dim (\operatorname{Ker} f) + \dim (\operatorname{Im} f).}$$

Demostración: La aplicación lineal $f|_V$ es inyectiva porque su núcleo es nulo: Si $v \in V$ y $f(v) = 0$, entonces $v \in V \cap (\operatorname{Ker} f) = 0$.

Veamos que $f|_V$ es epiyectiva: Si $e' \in \operatorname{Im} f$, tendremos que $e' = f(e)$ para algún vector $e \in E = V + (\operatorname{Ker} f)$, de modo que $e = v + w$, donde $v \in V$, $w \in \operatorname{Ker} f$, y por tanto

$$e' = f(e) = f(v + w) = f(v) + f(w) = f(v) = f|_V(v).$$

Corolario 3.6 *Si A es la matriz de una aplicación lineal $f \colon E \to E'$ en ciertas bases,*

$$\boxed{\dim (\operatorname{Ker} f) = (\text{n}^o \text{ de columnas}) - \operatorname{rg} A.}$$

En particular f es inyectiva si y sólo si $\operatorname{rg} A = \dim E$.

Demostración: $\dim (\operatorname{Ker} f) = \dim E - \dim (\operatorname{Im} f) \overset{3.5}{=} \dim E - \operatorname{rg} A$.

Además, la aplicación f es inyectiva si y sólo si $\operatorname{Ker} f = 0$, de acuerdo con 1.4.

Corolario 3.7 *Sea $A \in M_{m\times n}(K)$. Las soluciones del sistema homogéneo $AX = 0$ forman un subespacio vectorial $V = \{X \in K^n \colon AX = 0\}$ de K^n de dimensión*

$$\boxed{\dim V = n - \operatorname{rg} A}$$

y las soluciones de un sistema no homogéneo compatible $AX = B$ forman una subvariedad lineal de K^n de dirección $AX = 0$ y dimensión $n - \operatorname{rg} A$.

Demostración: La matriz A define una aplicación lineal $f \colon K^n \to K^m$, $f(X) = AX$, y V es precisamente el núcleo de f. La matriz de f en las bases usuales de K^n y K^m es A, así que 3.6 afirma que la dimensión de V es $n - \operatorname{rg} A$.

Por último, si un sistema $AX = B$ es compatible, las soluciones se obtienen sumando a una solución particular X_0 las soluciones de la ecuación homogénea $AX = 0$; luego forman la subvariedad lineal $X_0 + V$ y, por tanto, su dimensión es $\dim V = n - \operatorname{rg} A$.

En adelante, los escalares serán $K = \mathbb{R}$ ó $\mathbb{C}$.

4. Geometría Euclídea

4.1. Producto Escalar

Definiciones: Dar un **producto escalar** en un K-espacio vectorial E es asignar a cada par de vectores $e, v \in E$ un escalar, que denotaremos $e \cdot v$ ó $\langle e|v\rangle$, de modo que

1. Es lineal a la derecha y antilineal a la izquierda:

$$\langle e|\lambda v + \mu v'\rangle = \lambda\langle e|v\rangle + \mu\langle e|v'\rangle \quad , \quad \langle \lambda e + \mu e'|v\rangle = \bar{\lambda}\langle e|v\rangle + \bar{\mu}\langle e'|v\rangle.$$

2. $\langle v|e\rangle = \overline{\langle e|v\rangle}$.

3. Es definido-positivo: $\langle e|e\rangle \geq 0$, y sólo se da la igualdad cuando $e = 0$.

y se dice que dos vectores $e, v \in E$ son **ortogonales** cuando $\langle e|v\rangle = 0$ (y por tanto $\langle v|e\rangle = 0$).

Un isomorfismo lineal $T\colon E \to E$ es una **isometría** o un **operador unitario** cuando conserva el producto escalar: $\langle e|v\rangle = \langle T(e)|T(v)\rangle$, $\forall e, v \in E$.

El **módulo** de un vector $e \in E$ es el número real $\|e\| := +\sqrt{\langle e|e\rangle} \geq 0$ (positivo cuando $e \neq 0$), de modo que $\|e\|^2 = \langle e|e\rangle$, y $\|\lambda e\| = |\lambda| \cdot \|e\|$.

La **distancia** entre dos puntos $p, q \in E$ es $\mathrm{d}(p,q) := \|q - p\|$.

Ejemplos:

1. En la Geometría Euclídea clásica, fijado un origen O, los puntos forman un espacio vectorial real de dimensión 3, dotado de un producto escalar (toda vez que se haya fijado la unidad de longitud): el producto escalar de dos vectores es el producto de sus longitudes por el coseno del ángulo que forman.

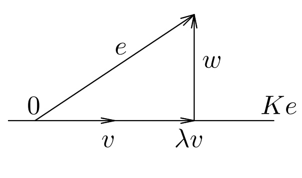
2. El producto escalar usual en K^n es $\langle (x_1, \ldots, x_n)|(y_1, \ldots, y_n)\rangle := \bar{x}_1 y_1 + \ldots + \bar{x}_n y_n$.

3. Fijado un intervalo $[a,b] \subset \mathbb{R}$, un producto escalar en el espacio vectorial complejo formado por las funciones continuas $\psi\colon [a,b] \to \mathbb{C}$, $\psi(x) = u(x) + v(x)i$, es

$$\langle \psi|\phi\rangle := \int_a^b \overline{\psi(x)}\phi(x)\,\mathrm{d}x \quad , \quad \|\psi\|^2 = \langle \psi|\psi\rangle = \int_a^b |\psi(x)|^2\,\mathrm{d}x.$$

4. En Mecánica Cuántica, el espacio de estados de un sistema cuántico es un espacio vectorial complejo E (generalmente de dimensión infinita) con un producto escalar, un estado del sistema es un subespacio vectorial unidimensional de E (normalmente representado por alguno de sus vectores de módulo 1, condición que lo determina salvo un factor $e^{i\theta}$, $\theta \in \mathbb{R}$), y la evolución temporal del estado del sistema viene dada, según la ecuación de Schrödinger (1886-1961), o de Dirac (1902-1984) en el caso relativista, por una familia de operadores unitarios $U_t\colon E \to E$, $t \in \mathbb{R}$, que cumple $U_{t+s} = U_t \circ U_s$.

Proyección Ortogonal sobre una Recta: *Fijada una recta Kv, todo vector $e \in E$ descompone en suma $e = \lambda v + w$, donde $\langle v|w\rangle = 0$ y $\lambda = (v \cdot e)/(v \cdot v)$.*



$$\boxed{\begin{bmatrix}\text{Proyecc. ortog.}\\ \text{de } e \text{ sobre } Kv\end{bmatrix} = \frac{v \cdot e}{v \cdot v}\, v}$$

Teorema de Pitágoras (s. VI a.C.)**:** $\|e+v\|^2 = \|e\|^2 + \|v\|^2$, cuando $\langle e|v\rangle = 0$.

Desigualdad de Cauchy-Schwarz (1789-1857, 1843-1921)**:** $|e \cdot v| \leq \|e\| \cdot \|v\|$.

Desigualdad triangular: $\|e+v\| \leq \|e\| + \|v\|$.

Demostración: (1) Tenemos que $v \cdot \big(e - \lambda v\big) = v \cdot e - \lambda(v \cdot v) = 0$ cuando $\lambda = \frac{v \cdot e}{v \cdot v}$.

(2) Tenemos que $\|e+v\|^2 = (e+v)\cdot(e+v) = e\cdot e + e\cdot v + v\cdot e + v\cdot v = \|e\|^2 + 0 + 0 + \|v\|^2$.

(3) Es obvia cuando $v = 0$. Cuando $v \neq 0$, ponemos $e = \lambda v + w$, con $v \cdot w = 0$.
Por el teorema de Pitágoras $\|e\|^2 = |\lambda|^2\|v\|^2 + \|w\|^2$, luego $|\lambda| \cdot \|v\| \leq \|e\|$, y

$$|e \cdot v| = |(\lambda v + w)\cdot v| = |(\lambda v)\cdot v| = |\lambda| \cdot \|v\|^2 \leq \|e\| \cdot \|v\|.$$

(4) Basta tomar raíz cuadrada en la siguiente consecuencia de la desigualdad anterior:

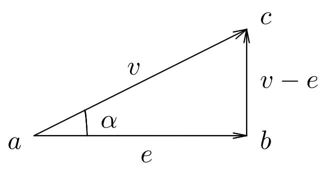
$$\begin{aligned}\|e+v\|^2 &= (e+v)\cdot(e+v) = e\cdot e + v\cdot v + e\cdot v + \overline{e\cdot v} \leq \\ &\leq e\cdot e + v\cdot v + 2|e\cdot v| \leq \|e\|^2 + \|v\|^2 + 2\|e\|\cdot\|v\| = (\|e\| + \|v\|)^2.\end{aligned}$$

Definición: En el caso real, $K = \mathbb{R}$, cuando los vectores e y v no son nulos, por la desigualdad de Cauchy-Schwarz tenemos que $-1 \leq \frac{e\cdot v}{\|e\|\cdot\|v\|} \leq 1$, y se llama medida en radianes del **ángulo** que forman e y v al único número real α entre 0 y π que cumple

$$\boxed{e \cdot v = \|e\| \cdot \|v\| \cos\alpha.}$$

Ejemplos:

1. El ángulo es $\alpha = \frac{\pi}{2}$ justo cuando e y v son ortogonales, $e \cdot v = 0$.
2. En el caso de un triángulo rectángulo, el coseno de un ángulo es el cociente del cateto contiguo por la hipotenusa:



$$0 = e\cdot(v-e) = e\cdot v - e^2\ , \quad e\cdot v = \|e\|^2$$

$$\cos\alpha = \frac{e\cdot v}{\|e\|\cdot\|v\|} = \frac{\|e\|^2}{\|e\|\cdot\|v\|} = \frac{\|e\|}{\|v\|}$$

3. Dados escalares no nulos $\lambda, \mu \in \mathbb{R}$, el ángulo β que forman λe y μv cumple que

$$\cos\beta = \frac{(\lambda e)\cdot(\mu v)}{\|\lambda e\|\cdot\|\mu v\|} = \frac{\lambda\mu(e\cdot v)}{|\lambda\mu|\|e\|\cdot\|v\|} = \begin{cases}\cos\alpha & \text{cuando } \lambda\mu > 0\\ -\cos\alpha & \text{cuando } \lambda\mu < 0\end{cases}$$

así que $\beta = \alpha$ cuando $\lambda\mu > 0$, y $\beta = \pi - \alpha$ cuando $\lambda\mu < 0$.

4.2. Espacios Vectoriales Euclídeos

Definiciones: Llamaremos **espacio vectorial euclídeo**[1], en honor de Euclides (c. 325-265 a.C.), a todo espacio vectorial E de dimensión finita dotado de un producto escalar, y diremos que una base $u_1, \ldots, u_n$ de E es **ortonormal** si los vectores de la base son de módulo 1 y mutuamente ortogonales:

$$\langle u_i|u_j\rangle = \delta_{ij} := \begin{cases}1 & \text{cuando } i = j\\ 0 & \text{cuando } i \neq j\end{cases}$$

[1]En el caso complejo se llaman **espacios vectoriales hermíticos**, en honor de Hermite (1822-1901).

En una base ortonormal el producto escalar de dos vectores $e, v \in E$ de coordenadas $X^t = (x_1, \ldots, x_n)$, $Y^t = (y_1, \ldots, y_n)$ es

$$\langle e|v\rangle = (x_1u_1 + \ldots + x_nu_n) \cdot (y_1u_1 + \ldots + y_nu_n) = \bar{x}_1y_1 + \ldots + \bar{x}_ny_n, \quad \boxed{\langle e|v\rangle = \bar{X}^tY.}$$

$$\langle e|e\rangle = \sum_i \bar{x}_ix_i = \sum_i |x_i|^2, \quad \boxed{\langle e|e\rangle = \bar{X}^tX.}$$

Ejemplos:

1. En K^n, con el producto escalar usual $(x_1, \ldots, x_n) \cdot (y_1, \ldots, y_n) = \bar{x}_1y_1 + \ldots + \bar{x}_ny_n$, la base $u_1 = (1, 0, \ldots, 0), \ldots, u_n = (0, \ldots, 0, 1)$ es ortonormal.

2. Sean $a_1, \ldots, a_n$ números enteros distintos. Las funciones $u_1 = e^{2\pi i a_1 x}, \ldots, u_n = e^{2\pi i a_n x}$ forman una base ortonormal de $\mathbb{C}e^{2\pi i a_1 x} + \ldots + \mathbb{C}e^{2\pi i a_n x}$ cuando el producto escalar es

$$\langle\psi|\phi\rangle = \int_0^1 \overline{\psi(x)}\phi(x)\,\mathrm{d}x,$$

$$\langle e^{2\pi i a_j x}|e^{2\pi i a_k x}\rangle = \int_0^1 e^{2\pi i(a_k - a_j)x}\,\mathrm{d}x = \begin{cases} \displaystyle\int_0^1 \mathrm{d}x = 1 & , \ j = k. \\ \displaystyle\left[\frac{e^{2\pi i(a_k - a_j)x}}{2\pi i(a_k - a_j)}\right]_0^1 = 0 & , \ j \neq k. \end{cases}$$

En el espacio hermítico $\mathbb{C}e^{ia_1x} + \ldots + \mathbb{C}e^{ia_nx}$, $\langle\psi|\phi\rangle = \int_0^{2\pi} \overline{\psi(x)}\phi(x)\,\mathrm{d}x$, el mismo cálculo prueba que $\langle e^{ia_ix}|e^{ia_jx}\rangle = 2\pi\delta_{ij}$, así que una base ortonormal de $\mathbb{C}e^{ia_1x} + \ldots + \mathbb{C}e^{ia_nx}$ está definida por las funciones $u_1 = \frac{1}{\sqrt{2\pi}}e^{ia_1x}, \ldots, u_n = \frac{1}{\sqrt{2\pi}}e^{ia_nx}$.

3. Un producto escalar en $M_{m\times n}(K)$ es $\langle A|B\rangle := \mathrm{tr}(\bar{A}^tB)$, y una base ortonormal de $M_{2\times 2}(K)$ está formada por las matrices

$$U_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, U_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, U_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, U_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

4. Un isomorfismo $T\colon E \to E$, de matriz A en una base ortonormal $u_1, \ldots, u_n$, es una isometría cuando $\bar{X}^tY = (\bar{A}\bar{X})^t(AY) = \bar{X}^t\bar{A}^tAY$ para todo par de columnas $X, Y \in K^n$; i.e. cuando A es una matriz **unitaria**: $\bar{A}^tA = I_n$, o lo que es lo mismo, $A^{-1} = \bar{A}^t$.

   Si $v_1, \ldots, v_n$ es otra base de E, y $B = (b_{ij})$ es la matriz de cambio de base, $v_j = \sum_i b_{ij}u_i$, entonces $(v_i \cdot v_j) = \bar{B}^tB$, así que la nueva base $v_1, \ldots, v_n$ también es ortonormal si y sólo si B es una matriz unitaria, $\bar{B}^tB = I_n$.

5. En el espacio vectorial hermítico $E = \mathbb{C}\cos x + \mathbb{C}\mathrm{sen}\, x$, $\langle\psi|\phi\rangle = \int_0^{2\pi} \overline{\psi(x)}\phi(x)\,\mathrm{d}x$, vamos a calcular la proyección ortogonal de $\psi(x) = \cos x$ sobre la recta $\mathbb{C}\phi$ que genera la función $\phi(x) = \mathrm{sen}\, x$. Para ello fijamos en E la base ortonormal $u_1 = \frac{1}{\sqrt{2\pi}}e^{ix}$, $u_2 = \frac{1}{\sqrt{2\pi}}e^{-ix}$.

$$\psi(x) = \cos x = \tfrac{1}{2}(e^{ix} + e^{-ix}) = \tfrac{\sqrt{2\pi}}{2}u_1 + \tfrac{\sqrt{2\pi}}{2}u_2 = \left(\tfrac{\sqrt{2\pi}}{2}, \tfrac{\sqrt{2\pi}}{2}\right)$$

$$\phi(x) = \mathrm{sen}\, x = \tfrac{1}{2i}(e^{ix} - e^{-ix}) = -\tfrac{\sqrt{2\pi}}{2}iu_1 + \tfrac{\sqrt{2\pi}}{2}iu_2 = \left(-\tfrac{\sqrt{2\pi}}{2}i, \tfrac{\sqrt{2\pi}}{2}i\right)$$

$$\frac{\langle\phi|\psi\rangle}{\langle\phi|\phi\rangle}\phi = \frac{\frac{\pi}{2}i - \frac{\pi}{2}i}{\frac{\pi}{2} + \frac{\pi}{2}}\phi = 0.$$

   Es decir, las funciones $\cos x$ y $\mathrm{sen}\, x$ son ortogonales.

   Como $\|\cos x\|^2 = \langle\psi|\psi\rangle = \pi$ y $\|\mathrm{sen}\, x\|^2 = \langle\phi|\phi\rangle = \pi$, vemos que las funciones

$$v_1 = \tfrac{1}{\sqrt{\pi}}\cos x = \tfrac{1}{\sqrt{2}}u_1 + \tfrac{1}{\sqrt{2}}u_2$$

$$v_2 = \tfrac{1}{\sqrt{\pi}}\mathrm{sen}\, x = -\tfrac{i}{\sqrt{2}}u_1 + \tfrac{i}{\sqrt{2}}u_2$$

forman una base ortonormal de E; luego la matriz de cambio de base B es unitaria:

$$B = \begin{pmatrix} 1/\sqrt{2} & -i/\sqrt{2} \\ 1/\sqrt{2} & i/\sqrt{2} \end{pmatrix} \quad , \quad B^{-1} = \bar{B}^t = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ i/\sqrt{2} & -i/\sqrt{2} \end{pmatrix}.$$

Teorema 4.1 *Todo espacio vectorial euclídeo $E \neq 0$ admite bases ortonormales.*

Demostración (**Método de Gram-Schmidt** (1850-1916, 1876-1959):
Dada una base $e_1, \ldots, e_n$ de E, cada uno de los siguientes vectores v_i

$$\begin{aligned}
&v_1 = e_1 && , \quad \langle v_1 \rangle = \langle e_1 \rangle \\
&v_2 = e_2 - \frac{v_1 \cdot e_2}{v_1 \cdot v_1} v_1 && , \quad \langle v_1, v_2 \rangle = \langle e_1, e_2 \rangle \\
&v_3 = e_3 - \frac{v_1 \cdot e_3}{v_1 \cdot v_1} v_1 - \frac{v_2 \cdot e_3}{v_2 \cdot v_2} v_2 && , \quad \langle v_1, v_2, v_3 \rangle = \langle e_1, e_2, e_3 \rangle \\
&\qquad \ldots\ldots\ldots\ldots\ldots\ldots && , \qquad \ldots\ldots\ldots\ldots \\
&v_n = e_n - \frac{v_1 \cdot e_n}{v_1 \cdot v_1} v_1 - \ldots - \frac{v_{n-1} \cdot e_n}{v_{n-1} \cdot v_{n-1}} v_{n-1} && , \quad \langle v_1, \ldots, v_n \rangle = \langle e_1, \ldots, e_n \rangle
\end{aligned}$$

es no nulo, porque $e_{i+1} \notin \langle e_1, \ldots, e_i \rangle = \langle v_1, \ldots, v_i \rangle$, y ortogonal a los anteriores por construcción; es decir, $v_j \cdot v_i = 0$ cuando $j < i$.

Además $v_1, \ldots, v_n$ es una base de E porque $\langle v_1, \ldots, v_n \rangle = \langle e_1, \ldots, e_n \rangle = E$.

Luego los vectores $u_1 = \frac{v_1}{\|v_1\|}, \ldots, u_n = \frac{v_n}{\|v_n\|}$ forman una base ortonormal de E.

Ejemplo: Una hipótesis fundamental de la Mecánica Clásica es que el espacio (fijado un origen y una unidad de longitud) es un espacio vectorial euclídeo real de dimensión 3, y una base ortonormal suele denotarse $\vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}}$.

4.3. Proyección Ortogonal

Definición: El **ortogonal** de un subespacio vectorial V de un espacio euclídeo E es

$$V^\perp := \{e \in E \colon \langle v|e \rangle = 0 \text{ para todo vector } v \in V\}.$$

Teorema 4.2 *El ortogonal $V^\perp$ es un subespacio vectorial de E de dimensión*

$$\boxed{\dim V^\perp = \dim E - \dim V}$$

y cada vector de E descompone de modo único en suma de un vector de V y otro de $V^\perp$:

$$\boxed{E = V \oplus V^\perp.}$$

Demostración: Sea $v_1, \ldots, v_d$ una base de V, y consideremos el núcleo de la aplicación lineal

$$f \colon E \longrightarrow K^d \ , \quad f(e) = (v_1 \cdot e, \ldots, v_d \cdot e),$$
$$\operatorname{Ker} f = \{e \in E \colon v_1 \cdot e = \ldots = v_d \cdot e = 0\} = (Kv_1 + \ldots + Kv_d)^\perp = V^\perp.$$

En efecto, si un vector e es ortogonal a ciertos vectores, $v_1 \cdot e = \ldots = v_d \cdot e = 0$, entonces $e \in (Kv_1 + \ldots + Kv_d)^\perp$ porque también es ortogonal a todas las combinaciones lineales:

$$(\lambda_1 v_1 + \ldots + \lambda_d v_d) \cdot e = \bar{\lambda}_1 (v_1 \cdot e) + \ldots + \bar{\lambda}_d (v_d \cdot e) = 0.$$

Luego $V^\perp$ es un subespacio vectorial de E y, por el Teorema de Isomorfía,

$$\dim E = \dim V^\perp + \dim (\operatorname{Im} f) \leq \dim V^\perp + d = \dim V^\perp + \dim V, \tag{7}$$

donde la desigualdad se debe a que la imagen de f es un subespacio vectorial de $\mathbb{R}^d$.

Por otra parte, si $v \in V^\perp \cap V$, entonces $v \cdot v = 0$; luego $v = 0$, porque el producto escalar es definido-positivo, y vemos que $V^\perp \cap V = 0$. Luego la suma de V y $V^\perp$ es directa por 2.9 y, de acuerdo con 2.10, tenemos que

$$\dim V^\perp + \dim V \overset{2.7}{=} \dim (V^\perp \oplus V) \le \dim E \tag{8}$$

y comparando con (7) concluimos que $\dim (V^\perp \oplus V) = \dim V^\perp + \dim V = \dim E$.

Además 2.6 permite concluir que $V^\perp \oplus V = E$.

Corolario 4.3 $(V^\perp)^\perp = V$.

Demostración: Tenemos que $V \subseteq (V^\perp)^\perp$ por definición de $V^\perp$, y coinciden porque tienen la misma dimensión: $\dim (V^\perp)^\perp = \dim E - \dim (V^\perp) = \dim V$.

Corolario 4.4 *Dos subespacios vectoriales son iguales cuando sus ortogonales coinciden.*

Definición: Si V es un subespacio vectorial de un espacio vectorial euclídeo E, de acuerdo con 4.2 tenemos que $E = V^\perp \oplus V$. La aplicación $P_V : E \to E$, $P_V(v + w) = v$, donde $v \in V$ y $w \in V^\perp$, es lineal y se llama **proyección ortogonal** sobre V.

Es decir, cada vector $e \in E$ descompone de modo único en suma, $e = v + w$, de un vector $v \in V$ y otro $w \in V^\perp$, y por definición $P_V(e) = v$, $P_{V^\perp}(e) = w$.

Nótese que $e = P_V(e) + P_{V^\perp}(e)$, de modo que $P_{V^\perp}(e) = e - P_V(e)$, y $P_V + P_{V^\perp} = \mathrm{Id}$.

Proposición 4.5 *Si $u_1, \ldots, u_d$ es una base ortonormal de V, se cumple que*

$$P_V(e) = \langle u_1|e\rangle u_1 + \ldots + \langle u_d|e\rangle u_d.$$

Demostración: Pongamos $P_V(e) = \sum_j \lambda_j u_j$. Como $e - P_V(e) \in V^\perp$, tenemos que

$$u_i \cdot e = u_i \cdot P_V(e) = u_i \cdot \left(\sum_j \lambda_j u_j\right) = \sum_j \lambda_j (u_i \cdot u_j) = \sum_j \lambda_j \delta_{ij} = \lambda_i.$$

Ejemplo: Cuando $V = Kv$ es una recta, el vector $u = v/\|v\|$ define una base ortonormal de V, y volvemos a obtener la fórmula de la proyección ortogonal de un vector $e \in E$ sobre la recta que genera un vector v:

$$P_V(e) = (u \cdot e)u = \frac{1}{\|v\|^2}(v \cdot e)v = \frac{v \cdot e}{v \cdot v}\, v.$$

Nota: El teorema 4.2 permite dar otra demostración de la existencia de bases ortonormales:

Tomemos un vector no nulo $v \in E$ y pongamos $u = v/\|v\|$, de modo que $\|u\| = 1$.

Procedamos por inducción sobre $n = \dim E$. Si $n = 1$, tenemos que $E = Ku$, y u es una base ortonormal de E. Si $n > 1$, de acuerdo con 4.2, tenemos que $\dim (Ku)^\perp = n - 1$, y por hipótesis de inducción existe una base ortonormal $u_2, \ldots, u_n$ de $(Ku)^\perp$.

Estos vectores $u, u_2, \ldots, u_n$ son de módulo 1 y mutuamente ortogonales, y forman una base de E de acuerdo con el siguiente lema:

Lema 4.6 *Si unos vectores no nulos $v_1, \ldots, v_r$ son mutuamente ortogonales, $\langle v_i|v_j\rangle = 0$ cuando $i \neq j$, entonces son linealmente independientes.*

Demostración: Si $\sum_i \lambda_i v_i = 0$, entonces $0 = v_j \cdot (\sum_i \lambda_i v_i) = \sum_i \lambda_i (v_j \cdot v_i) = \lambda_j (v_j \cdot v_j)$ para todo índice $j = 1, \ldots, r$. Concluimos que $\lambda_j = 0$ porque $v_j \cdot v_j \neq 0$.

5. Endomorfismos

Recuérdese que los escalares son $K = \mathbb{R}$ ó $\mathbb{C}$.

Sea $p(x)$ un polinomio no constante con coeficientes complejos. En este capítulo usaremos el siguiente teorema fundamental, que se demostrará en el curso de Variable Compleja:

Teorema de D'Alembert (1717-1783)**:** *Todo polinomio no constante con coeficientes complejos admite alguna raíz compleja.*

Regla de Ruffini (1765-1822)**:** *Si α es una raíz compleja de $p(x)$, entonces*

$$p(x) = (x - \alpha)q(x).$$

Demostración: Dividiendo $p(x)$ por $x - \alpha$ tendremos que $p(x) = (x - \alpha)q(x) + r$, donde el resto r es de grado menor que 1; luego constante.

Sustituyendo $x = \alpha$ en esta igualdad obtenemos que el resto es nulo:

$$0 = p(\alpha) = (\alpha - \alpha)q(\alpha) + r = r.$$

Definición: Si α es una raíz compleja de $p(x)$, llamaremos **multiplicidad** de tal raíz al mayor número natural m tal que $(x - \alpha)^m$ divida a $p(x)$.

Las raíces de multiplicidad 1 se denominan **simples**.

Consideremos una raíz compleja α_1 de $p(x)$, que tendrá cierta multiplicidad m_1, de modo que $p(x) = (x - \alpha_1)^{m_1} q_1(x)$, donde el polinomio $q_1(x)$ ya no admite la raíz α_1. Tomemos una raíz compleja α_2 de $q_1(x)$, que tendrá cierta multiplicidad m_2, de modo que $p(x) = (x - \alpha_1)^{m_1}(x - \alpha_2)^{m_2} q_2(x)$. Por el teorema de D'Alembert, podemos proceder así hasta que el factor $q_i(x)$ sea constante, y obtenemos una descomposición

$$p(x) \; = \; c(x - \alpha_1)^{m_1}(x - \alpha_2)^{m_2} \dots (x - \alpha_r)^{m_r} \;, \tag{9}$$

donde $\alpha_1, \dots, \alpha_r$ son las raíces complejas de $p(x)$, los exponentes $m_1, \dots, m_r$ son sus respectivas multiplicidades, y c es constante.

Esta descomposición muestra que *el número de raíces complejas de un polinomio no constante, contadas con su multiplicidad, coincide siempre con el grado del polinomio.*

5.1. Valores y Vectores Propios

Definición: Los **endomorfismos** u **operadores lineales** de un K-espacio vectorial E son las aplicaciones K-lineales $T \colon E \to E$.

Si $S, T \colon E \to E$ son endomorfismos, su suma $S + T$, el producto λT por un escalar λ, y su producto $ST := S \circ T$ también son endomorfismos.

Definiciones: Dado un endomorfismo T de un K-espacio vectorial E, diremos que un escalar $\alpha \in K$ es un **valor propio** o **autovalor** de T si existe algún vector no nulo $e \in E$ tal que $T(e) = \alpha e$, en cuyo caso diremos que e es un **vector propio** o **autovector** de T de valor propio α y que el **subespacio propio** del valor propio α es el subespacio vectorial (3.4)

$$V_\alpha := \{e \in E \colon T(e) = \alpha e\} = \{e \in E \colon \alpha \mathrm{Id}(e) - T(e) = 0\} = \mathrm{Ker}\,(\alpha \mathrm{Id} - T) \neq 0.$$

El **espectro** de T es el conjunto formado por los valores propios de T (puede ser vacío).

Ejemplos:

1. Un endomorfismo T tiene el valor propio 0 si y sólo si $\mathrm{Ker}\, T \neq 0$, es decir, cuando T no es inyectivo, en cuyo caso $V_0 = \mathrm{Ker}\, T$.

2. Fijemos una base ortonormal u, v en un plano euclídeo real E.

    a) El espectro del giro de ángulo recto $G\colon E \to E$, $G(xu + yv) = -yu + xv$, es vacío.

    b) El espectro de la simetría $S\colon E \to E$, $S(xu + yv) = xu - yv$, respecto de la recta $\mathbb{R}u$ es $\{\pm 1\}$, y $V_1 = \mathbb{R}u$, $V_{-1} = \mathbb{R}v$.

    c) El espectro de la proyección ortogonal $P\colon E \to E$, $P(xu + yv) = xu$, sobre la recta $\mathbb{R}u$ es $\{0, 1\}$, y $V_0 = \mathbb{R}v$, $V_1 = \mathbb{R}u$.

    d) El espectro de la homotecia $H\colon E \to E$, $H(xu + yv) = axu + ayv$, donde $a \in \mathbb{R}$ es un número real fijado, es $\{a\}$, y $V_a = E$.

3. Sea E el espacio vectorial complejo formado por las funciones $\psi\colon \mathbb{R} \to \mathbb{C}$ infinitamente derivables. Todo número complejo α es un valor propio del operador lineal $\frac{\mathrm{d}}{\mathrm{d}x}\colon E \to E$ porque $\frac{\mathrm{d}}{\mathrm{d}x}(e^{\alpha x}) = \alpha e^{\alpha x}$. Si otra función $\psi(x) \in E$ cumple que $\psi' = \alpha\psi$, entonces

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\psi(x)}{e^{\alpha x}}\right) = \frac{\alpha\psi(x)e^{\alpha x} - \psi(x)\alpha e^{\alpha x}}{e^{2\alpha x}} = 0,$$

    y vemos que $\psi(x) = \lambda e^{\alpha x}$ para algún $\lambda \in \mathbb{C}$.

    En este caso el espectro es $\mathbb{C}$, y $V_\alpha = \mathbb{C}e^{\alpha x}$ siempre es de dimensión 1.

Fijada una base $e_1, \ldots, e_n$ de E, un endomorfismo $T\colon E \to E$ está determinado por su matriz $A = (a_{ij}) \in M_{n\times n}(K)$ en tal base, y de acuerdo con 3.6 tenemos que

$$\boxed{\dim \operatorname{Ker}(\alpha\mathrm{Id} - T) = n - \operatorname{rg}(\alpha I_n - A).}$$

Definición: Diremos que el polinomio **característico** de T es el polinomio

$$c(x) := |xI_n - A| = \begin{vmatrix} x - a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & x - a_{22} & \ldots & -a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ -a_{n1} & -a_{n2} & \ldots & x - a_{nn} \end{vmatrix} = x^n - \Big(\sum_{i=1}^{n} a_{ii}\Big)x^{n-1} + \ldots + (-1)^n|A|,$$

que sólo depende de T y no de la base fijada en el espacio vectorial E (en particular, los escalares $\operatorname{tr} T := a_{11} + \ldots + a_{nn}$ y $\det T := |A|$ no dependen de la base fijada, y reciben el nombre de **traza** y **determinante** de T), pues si se considera una nueva base en E, y B es la matriz del cambio de base, según 6 (página 22) la matriz $\tilde{A}$ de T en la nueva base es

$$\tilde{A} = B^{-1}AB \tag{10}$$

$$\begin{aligned} |xI_n - \tilde{A}| &= |xB^{-1}I_nB - B^{-1}AB| = |B^{-1}(xI_n - A)B| \\ &= |B^{-1}| \cdot |xI_n - A| \cdot |B| = |xI_n - A|. \end{aligned}$$

Teorema 5.1 *Los valores propios de un endomorfismo son las raíces en K de su polinomio característico.*

Demostración: Un escalar $\alpha \in K$ es valor propio de T cuando $\dim \operatorname{Ker}(\alpha\mathrm{Id} - T) \neq 0$.
Luego α es valor propio de T si y sólo si $\operatorname{rg}(\alpha I_n - A) < n$; i.e., $c(\alpha) = |\alpha I_n - A| = 0$.

Corolario 5.2 *El número de valores propios de un endomorfismo de un espacio vectorial de dimensión n siempre es menor o igual que n.*

Demostración: El grado del polinomio característico es n, y el número de raíces en K de un polinomio siempre está acotado por el grado del polinomio.

Corolario 5.3 *Todo endomorfismo de un espacio vectorial complejo de dimensión finita $n \geq 1$ tiene algún valor propio.*

Demostración: El polinomio característico no es constante, porque es de grado n, así que tiene alguna raíz compleja por el Teorema de D'Alembert.

Corolario 5.4 *Todo endomorfismo $T: E \to E$ de un espacio vectorial complejo de dimensión finita admite una base en que su matriz $A = (a_{ij})$ es triangular ($a_{ij} = 0$ cuando $i > j$). Es decir, para cada matriz $A \in M_{n\times n}(\mathbb{C})$ existe alguna matriz invertible $B \in M_{n\times n}(\mathbb{C})$ tal que la matriz $B^{-1}AB$ es triangular.*

Demostración: Procedemos por inducción sobre $n = \dim E$, y cuando $n = 1$, toda matriz es triangular. Cuando $n > 1$, de acuerdo con 5.3, el endomorfismo T tiene algún valor propio $\alpha \in \mathbb{C}$, y consideramos un vector propio $e_1 \in E$ de valor propio α, y una base $e_1, \dots, e_n$ en E. La matriz A de T en esta base es de la forma

$$A = \begin{pmatrix} \alpha & \dots \\ 0 & A_{n-1} \end{pmatrix} \tag{11}$$

para cierta matriz $A_{n-1} \in M_{(n-1)\times(n-1)}(\mathbb{C})$. Por hipótesis de inducción existe una matriz invertible $B_{n-1} \in M_{(n-1)\times(n-1)}(\mathbb{C})$ tal que la matriz $B_{n-1}^{-1}A_{n-1}B_{n-1}$ es triangular.

Terminamos porque también es triangular la matriz

$$B^{-1}AB = \begin{pmatrix} 1 & 0 \\ 0 & B_{n-1}^{-1} \end{pmatrix} \begin{pmatrix} \alpha & \dots \\ 0 & A_{n-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & B_{n-1} \end{pmatrix} = \begin{pmatrix} \alpha & \dots \\ 0 & B_{n-1}^{-1}A_{n-1}B_{n-1} \end{pmatrix}.$$

Teorema de Hamilton-Cayley (1805-1865 y 1821-1895)**:** *El polinomio característico $c(x) = x^n + \dots + c_1x + c_0$ de un endomorfismo T de un K-espacio vectorial de dimensión finita E siempre anula al endomorfismo: $c(T) := T^n + \dots + c_1T + c_0\mathrm{Id} = 0$.*

Demostración: Si $A \in M_{n\times n}(K)$ es la matriz de T en una base de E, la matriz del endomorfismo $c(T)$ es $c(A) = A^n + \dots + c_1A + c_0I = 0$, así que el teorema afirma que $c(A) = 0$, donde $c(x) = |xI - A|$; luego basta probarlo en el caso $K = \mathbb{C}$, y en tal caso procedemos por inducción sobre $n = \dim E$.

Si $n = 1$, entonces $A = (a)$. Luego $c(x) = x - a$ y $c(A) = A - aI = 0$.

Si $n > 1$, fijando una base adecuada podemos suponer que A es de la forma 11. Luego $c(x) = (x - \alpha)c_{n-1}(x)$, donde $c_{n-1}(x) = |xI_{n-1} - A_{n-1}|$ y, por hipótesis de inducción, $c_{n-1}(A_{n-1}) = 0$. Ahora

$$A^r = \begin{pmatrix} \alpha^r & \dots \\ 0 & A_{n-1}^r \end{pmatrix} \quad , \quad c(A) = (A - \alpha I_n)c_{n-1}(A) = \begin{pmatrix} 0 & \dots \\ 0 & C \end{pmatrix} \begin{pmatrix} c_{n-1}(\alpha) & \dots \\ 0 & 0 \end{pmatrix} = 0.$$

5.2. Diagonalización de Endomorfismos

Definición: Un endomorfismo T de un K-espacio vectorial de dimensión finita E es **diagonalizable** si existe alguna base $e_1, \dots, e_n$ de E formada por vectores propios de T; i.e., $T(e_j) = \alpha_j e_j$ para ciertos escalares $\alpha_j \in K$, lo que significa que la matriz de T en tal base es diagonal (todos sus coeficientes son nulos, salvo quizás los de la diagonal):

$$D = \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \alpha_n \end{pmatrix}$$

De acuerdo con 10, un endomorfismo T de matriz A es diagonalizable si existe alguna matriz invertible B tal que $D = B^{-1}AB$ es una matriz diagonal.

Teorema 5.5 *Sean $\alpha_1, \dots, \alpha_m$ valores propios de un endomorfismo T, distintos entre sí. La suma de los subespacios propios $V_{\alpha_1}, \dots, V_{\alpha_m}$ es directa, y por tanto*

$$\boxed{\dim V_{\alpha_1} + \dots + \dim V_{\alpha_m} = \dim (V_{\alpha_1} \oplus \dots \oplus V_{\alpha_m}) \leq \dim E.}$$

Demostración: Sea $V_{\alpha_1} \times \ldots \times V_{\alpha_m}$ el espacio vectorial formado por las sucesiones de vectores $(v_1, \ldots, v_m)$, donde $v_i \in V_{\alpha_i}$. Por definición de suma directa, hemos de probar que es biyectiva la siguiente aplicación lineal (que siempre es epiyectiva):

$$s\colon V_{\alpha_1} \times \ldots \times V_{\alpha_m} \longrightarrow V_{\alpha_1} + \ldots + V_{\alpha_m}\ , \quad s(v_1, \ldots, v_m) = v_1 + \ldots + v_m.$$

Es decir, tenemos que ver que $\operatorname{Ker} s = 0$. Procedemos por inducción sobre m.

Cuando $m = 1$, es cierto porque $s\colon V_{\alpha_1} \to V_{\alpha_1}$, $s(v_1) = v_1$, es claramente biyectiva.

Si $m > 1$ y $v_1 + \ldots + v_m = 0$, donde $v_i \in V_{\alpha_i}$, tendremos que

$$0 = T(0) = T(v_1 + \ldots + v_m) = T(v_1) + \ldots + T(v_m) = \alpha_1 v_1 + \ldots + \alpha_m v_m$$

y restando la igualdad $0 = \alpha_m(v_1 + \ldots + v_m) = \alpha_m v_1 + \ldots + \alpha_m v_m$ vemos que

$$0 = (\alpha_1 - \alpha_m)v_1 + \ldots + (\alpha_{m-1} - \alpha_m)v_{m-1}.$$

Por hipótesis de inducción, $(\alpha_1 - \alpha_m)v_1 = \ldots = (\alpha_{m-1} - \alpha_m)v_{m-1} = 0$.

Como $\alpha_i \neq \alpha_j$ cuando $i \neq j$, concluimos que $v_1 = 0, \ldots, v_{m-1} = 0$, y por tanto que también $0 = v_1 + \ldots + v_{m-1} + v_m = 0 + \ldots + 0 + v_m = v_m$.

Por último, $\dim(V_{\alpha_1} \oplus \ldots \oplus V_{\alpha_m}) = \dim V_{\alpha_1} + \ldots + \dim V_{\alpha_m}$, de acuerdo con 2.10.

Criterio de Diagonalización: *Sea $\{\alpha_1, \ldots, \alpha_r\}$ el espectro de un endomorfismo T de un K-espacio vectorial E de dimensión finita n. Las siguientes condiciones son equivalentes:*

1. *T es diagonalizable; i.e. E admite una base formada por vectores propios de T.*

2. *Todo vector no nulo de E descompone, y de modo único salvo el orden, en suma de vectores propios de valores propios distintos. Es decir, $E = V_{\alpha_1} \oplus \ldots \oplus V_{\alpha_r}$.*

3. *Todas las raíces complejas del polinomio característico $c(x)$ están en K, y la multiplicidad m_i de cada raíz $x = \alpha_i$ coincide con la dimensión d_i del subespacio propio V_{α_i}.*

Demostración: $(1 \Rightarrow 2)$ Si T es diagonalizable, por definición E admite una base $e_1, \ldots, e_n$ formada por vectores propios de T, de modo que $e_1, \ldots, e_n \in V_{\alpha_1} \oplus \ldots \oplus V_{\alpha_r}$.

Luego $E = \langle e_1, \ldots, e_n \rangle \subseteq V_{\alpha_1} \oplus \ldots \oplus V_{\alpha_r}$, y concluimos que $V_{\alpha_1} \oplus \ldots \oplus V_{\alpha_r} = E$.

$(2 \Rightarrow 1$ y $3)$ Si $V_{\alpha_1} \oplus \ldots \oplus V_{\alpha_r} = E$, considerando una base $e^i_1, \ldots, e^i_{d_i}$ en cada subespacio propio V_{α_i}, donde $d_i = \dim V_{\alpha_i}$, obtenemos vectores propios $e^1_1, \ldots, e^1_{d_1}, \ldots, e^r_1, \ldots, e^r_{d_r}$ que forman una base[2] de E porque generan $V_{\alpha_1} \oplus \ldots \oplus V_{\alpha_r} = E$ y

$$d_1 + \ldots + d_r = \dim V_{\alpha_1} + \ldots + \dim V_{\alpha_r} = \dim(V_{\alpha_1} \oplus \ldots \oplus V_{\alpha_r}) = \dim E.$$

Además, la matriz de T en tal base de vectores propios $e^1_1, \ldots, e^1_{d_1}, \ldots, e^r_1, \ldots, e^r_{d_r}$ es diagonal $D = \operatorname{diag}(\alpha_1, \overset{d_1}{\ldots}, \alpha_1, \ldots, \alpha_r \overset{d_r}{\ldots} \alpha_r)$, de modo que

$$c(x) = |xI_n - D| = (x - \alpha_1)^{d_1} \ldots (x - \alpha_r)^{d_r},$$

y las raíces complejas de $c(x)$ son $\alpha_1, \ldots, \alpha_r \in K$, con multiplicidades $d_1, \ldots, d_r$.

$(3 \Rightarrow 2)$ Como el número de raíces complejas de un polinomio, contadas con su multiplicidad, coincide con el grado, tenemos que

$$\dim E = \operatorname{gr} c(x) = m_1 + \ldots + m_r = \dim V_{\alpha_1} + \ldots + \dim V_{\alpha_r} = \dim(V_{\alpha_1} \oplus \ldots \oplus V_{\alpha_r}).$$

Nota: Las ecuaciones implícitas de $V_{\alpha_i} = \operatorname{Ker}(\alpha_i \mathrm{Id} - T)$ son $(\alpha_i I_n - A)X = 0$, donde A es la matriz de T en alguna base de E; así que $d_i = \dim V_{\alpha_i} = n - \operatorname{rg}(\alpha_i I_n - A)$.

[2] Por tanto, cuando T es diagonalizable, para hallar una base de E formada por vectores propios de T, basta juntar bases de los subespacios propios V_{α_i}.

Corolario 5.6 *Si todas las raíces complejas del polinomio característico $c(x)$ están en K y son simples, entonces T es diagonalizable y $\dim V_\alpha = 1$ para cualquier valor propio α.*

Demostración: Como $1 \leq \dim V_\alpha$ para todo valor propio α, siempre se cumple que

$$r \leq \dim V_{\alpha_1} + \ldots + \dim V_{\alpha_r} \leq \dim E = \operatorname{gr} c(x).$$

Cuando todas las raíces de $c(x)$ están en K y son simples, tenemos que $r = \operatorname{gr} c(x)$. Luego $\dim V_{\alpha_i} = 1$ para todo índice $i = 1, \ldots, r$, y $\dim (V_{\alpha_1} \oplus \ldots \oplus V_{\alpha_r}) = \dim E$. Concluimos que $V_{\alpha_1} \oplus \ldots \oplus V_{\alpha_r} = E$ y T es diagonalizable.

5.3. Operadores Lineales Autoadjuntos

Definición: Sea E un espacio vectorial dotado de un producto escalar. Un operador lineal $T\colon E \to E$ es **autoadjunto** o **hermítico**, en honor de Hermite (1822-1901), cuando

$$\langle T(e)|v\rangle = \langle e|T(v)\rangle \ , \quad \forall e, v \in E.$$

Si la dimensión de E es finita y A es la matriz de T en una base ortonormal de E, esto significa que $\bar{X}^t \bar{A}^t Y = \bar{X}^t A Y$ para cualesquiera columnas $X, Y \in K^n$; es decir, que la matriz A es **hermítica** o **autoadjunta**, $A = \bar{A}^t$ (**simétrica**, $A = A^t$, en el caso real).

Teorema Espectral: *Todos los valores propios de un endomorfismo autoadjunto $T\colon E \to E$ son reales, y dos vectores propios de T de valores propios distintos siempre son ortogonales.*

Si además E es de dimensión finita, admite una base ortonormal formada por vectores propios de T. Luego T es diagonalizable, y toda raíz compleja $x = \alpha$ de su polinomio característico $c(x)$ es real y de multiplicidad igual a la dimensión del subespacio propio V_α.

Demostración: Si α es un valor propio de T, entonces existe algún vector no nulo $e \in E$ tal que $T(e) = \alpha e$. Como $e \cdot e \neq 0$, concluimos que α es real (es decir, que $\bar{\alpha} = \alpha$):

$$\alpha(e \cdot e) = e \cdot (\alpha e) = e \cdot T(e) = T(e) \cdot e = (\alpha e) \cdot e = \bar{\alpha}(e \cdot e).$$

Además, si $v \in E$ es un vector propio de T de valor propio $\beta \neq \alpha$, como $\bar{\alpha} = \alpha$, tenemos que $e \cdot v = 0$ porque

$$\alpha(e \cdot v) = \bar{\alpha}(e \cdot v) = (\alpha e) \cdot v = T(e) \cdot v = e \cdot T(v) = e \cdot (\beta v) = \beta(e \cdot v).$$

Cuando $\dim E < \infty$, en el caso complejo vemos que todas las raíces complejas del polinomio característico son reales, porque son valores propios de T; es decir, si A es una matriz compleja hermítica (y las matrices reales simétricas lo son) entonces todas las raíces complejas del polinomio $c(x) = |xI - A|$ son reales.

Se sigue la existencia algún vector propio $u \in E$ (incluso en el caso real), y después de dividirlo por su módulo, podemos suponer que $\|u\| = 1$. Ahora, para probar la existencia de una base ortonormal de vectores propios, procedemos por inducción sobre $n = \dim E$.

Cuando $n = 1$, tenemos que $E = Ku$, y u es ya una base ortonormal de E.

Cuando $n > 1$, consideramos la recta Ku y su ortogonal $V := (Ku)^\perp$, que es de dimensión $n - 1$, y queda invariante por T. En efeto, si $v \in V$, entonces también $T(v) \in V$ porque

$$T(v) \cdot u = v \cdot T(u) = v \cdot (\alpha u) = \alpha(v \cdot u) = 0.$$

Luego T induce un operador lineal $T|_V\colon V \to V$, $T|_V(v) = T(v)$, que es hermítico, y por hipótesis de inducción, V admite una base ortonormal $u_1, \ldots, u_{n-1}$ formada por vectores propios de $T|_V$; luego también de T.

Los vectores propios $u_1, \ldots, u_{n-1}, u$ son de módulo 1 y mutuamente ortogonales, y forman una base de E porque son linealmente independientes: $u \notin Ku_1 + \ldots + Ku_{n-1} = V$.

Corolario 5.7 *Sea* $\{\alpha_1, \ldots, \alpha_r\}$ *el espectro de un operador hermítico* $T\colon E \to E$. *Todo vector* $\psi \in E$ *descompone, de modo único, en suma* $\psi = \psi_{\alpha_1} + \ldots + \psi_{\alpha_r}$, *donde* $\psi_{\alpha_i} \in V_{\alpha_i}$. *Además,* ψ_{α_i} *es la proyección ortogonal de* ψ *sobre* V_{α_i}.

Demostración: Tal descomposición existe y es única porque $V_{\alpha_1} \oplus \ldots \oplus V_{\alpha_r} = E$, al ser T diagonalizable, y la proyección ortogonal de ψ sobre V_{α_1} es ψ_{α_1} porque $\psi_{\alpha_2}, \ldots, \psi_{\alpha_r} \in V_{\alpha_1}^{\perp}$, de acuerdo con el Teorema Espectral.

Teorema 5.8 *Si dos operadores lineales hermíticos* S, T *de un espacio vectorial euclídeo* E *conmutan,* $ST = TS$, *entonces* E *admite una base ortonormal formada por vectores propios de* S *y de* T.

Demostración: Sea $\{\alpha_1, \ldots, \alpha_r\}$ el espectro de T, y $V_{\alpha_1}, \ldots, V_{\alpha_r}$ los correspondientes subespacios propios. Pongamos $d_i := \dim V_{\alpha_i}$.

Si $v_i \in V_{\alpha_i}$, entonces $T(v_i) = \alpha_i v_i$, de modo que

$$T(S(v_i)) = S(T(v_i)) = S(\alpha_i v_i) = \alpha_i S(v_i),$$

y vemos que también $S(v_i) \in V_{\alpha_i}$. Es decir, $S(V_{\alpha_i}) \subseteq V_{\alpha_i}$, de modo que S induce por restricción un operador hermítico $S|_V\colon V_{\alpha_i} \to V_{\alpha_i}$.

Por el Teorema Espectral, existe una base ortonormal $u_1^i, \ldots, u_{d_i}^i$ de V_{α_i} formada por vectores propios de S (y también de T porque están en V_{α_i}).

Estos vectores propios $u_1^1, \ldots, u_{d_1}^1, \ldots, u_1^r, \ldots, u_{d_r}^r$ son de módulo 1, mutuamente ortogonales por el Teorema Espectral, y, al ser T es diagonalizable, forman una base de E (ver la nota al pie de la página 33).

Ejemplos:

1. En Mecánica Cuántica, el espacio de estados de un sistema es un espacio vectorial complejo E (generalmente de dimensión infinita) con un producto escalar y los distintos estados del sistema son los subespacios vectoriales $\mathbb{C}\psi$ de dimensión 1, normalmente representados por un vector ψ de módulo 1 (esto lo determina salvo un factor $e^{i\theta}$, $\theta \in \mathbb{R}$).

   Cada magnitud escalar está descrita por un operador lineal autoadjunto $T\colon E \to E$, y al medir esa magnitud, los únicos valores que se pueden obtener son los del espectro de T, que siempre son números reales por el Teorema Espectral.

   Si el estado está representado por un vector propio de valor propio α, entonces el valor de la medición es α; pero si está representado por una suma $\psi = \psi_{\alpha_1} + \ldots + \psi_{\alpha_r}$ de vectores propios de valores propios distintos $\alpha_1, \ldots, \alpha_r$ (descomposición que es única según 5.5, y que en dimensión finita siempre existe por el Teorema Espectral) entonces los posibles valores de la medición son $\alpha_1, \ldots, \alpha_r$, y la probabilidad $\mathcal{P}(\alpha_i)$ de medir α_i es

   $$\mathcal{P}(\alpha_i) = \frac{\|\psi_{\alpha_i}\|^2}{\|\psi\|^2} = \frac{\langle \psi_{\alpha_i} | \psi_{\alpha_i} \rangle}{\langle \psi | \psi \rangle}.$$

   Si el resultado de la medición es α_i, el estado del sistema pasa a ser ψ_{α_i}, la proyección ortogonal del vector ψ sobre el subespacio propio V_{α_i}. (Aunque el representante ψ esté normalizado, $\|\psi\| = 1$, el representante ψ_{α_i} del nuevo estado no, cuando $\mathcal{P}(\alpha_i) \neq 1$).

   El teorema de Pitágoras, $\|\psi\|^2 = \sum_i \|\psi_{\alpha_i}\|^2$, afirma que $1 = \mathcal{P}(\alpha_1) + \ldots + \mathcal{P}(\alpha_r)$.

   En particular, en ese estado el valor medio de las observaciones es

   $$\sum_{i=1}^{r} \alpha_i \mathcal{P}(\alpha_i) = \sum_{i=1}^{r} \frac{\langle \psi_{\alpha_i} | \alpha_i \psi_{\alpha_i} \rangle}{\langle \psi | \psi \rangle} = \frac{\langle \sum_i \psi_{\alpha_i} | \sum_i \alpha_i \psi_{\alpha_i} \rangle}{\langle \psi | \psi \rangle} = \frac{\langle \psi | T(\psi) \rangle}{\langle \psi | \psi \rangle} = \frac{\langle T(\psi) | \psi \rangle}{\langle \psi | \psi \rangle}.$$

2. En el caso de una partícula puntual de masa m, su estado está determinado (fijado el origen del espacio, la unidad de longitud y una base ortonormal) por una función compleja $\psi\colon \mathbb{R}^3 \to \mathbb{C}$ de 3 variables reales, llamada *función de onda*, que se anula fuera de la región donde está confinada la partícula. El producto escalar de funciones que se considera es

$$\langle\psi|\phi\rangle = \int_{\mathbb{R}^3} \overline{\psi(x,y,z)}\phi(x,y,z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z\,,$$

y la probabilidad de encontrar a la partícula en cierta región Ω es

$$\frac{\displaystyle\int_\Omega |\psi(x,y,z)|^2 \mathrm{d}x\mathrm{d}y\mathrm{d}z}{\displaystyle\int_{\mathbb{R}^3} |\psi(x,y,z)|^2 \mathrm{d}x\mathrm{d}y\mathrm{d}z}\,.$$

Las coordenadas de la posición de la partícula vienen dadas por los operadores lineales $X(\psi) = x\psi$, $Y(\psi) = y\psi$, $Z(\psi) = z\psi$ (hermíticos porque $\overline{x\psi} = x\bar{\psi}$, $\overline{y\psi} = y\bar{\psi}$, $\overline{z\psi} = z\bar{\psi}$) y los momentos en las direcciones de los ejes vienen dados por los operadores lineales

$$P_x := \frac{\hbar}{i}\frac{\partial}{\partial x} = -i\hbar\frac{\partial}{\partial x} \quad , \quad P_y := \frac{\hbar}{i}\frac{\partial}{\partial y} = -i\hbar\frac{\partial}{\partial y} \quad , \quad P_z := \frac{\hbar}{i}\frac{\partial}{\partial z} = -i\hbar\frac{\partial}{\partial z}$$

que son hermíticos de acuerdo con la fórmula de integración por partes:

$$\langle P_x\psi|\phi\rangle = \int_{\mathbb{R}^3} i\hbar\frac{\partial\bar{\psi}}{\partial x}\phi\,\mathrm{d}x\mathrm{d}y\mathrm{d}z = \int_{\mathbb{R}^3} \frac{\partial(i\hbar\bar{\psi}\phi)}{\partial x}\,\mathrm{d}x\mathrm{d}y\mathrm{d}z - \int_{\mathbb{R}^3} \bar{\psi}i\hbar\frac{\partial\phi}{\partial x}\,\mathrm{d}x\mathrm{d}y\mathrm{d}z = 0 + \langle\psi|P_x\phi\rangle$$

donde el primer sumando es nulo según la regla de Barrow, pues las funciones son nulas fuera de cierta región. La energía cinética viene dada por el operador lineal hermítico

$$\frac{1}{2m}(P_x^2 + P_y^2 + P_z^2) = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)$$

y en presencia de fuerzas que se deriven de un potencial $V(x,y,z)$ (que es una función con valores reales, de modo que el operador lineal $\psi \mapsto V\psi$ es autoadjunto) la energía viene dada por el operador lineal hermítico

$$H = V(x,y,z) - \frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right).$$

La evolución temporal del estado de la partícula viene dada por ciertos operadores lineales U_t, $t \in \mathbb{R}$, donde $U_t \circ U_s = U_{t+s}$ y $U_0 = \mathrm{Id}$; y se supone que la función compleja $\Psi(t,x,y,z) := (U_t\psi)(x,y,z)$ cumple la ecuación de Schrödinger (1886-1961)

$$\frac{\partial\Psi}{\partial t} = \frac{1}{i\hbar}H(\Psi) = -\frac{i}{\hbar}H(\Psi).$$

Los operadores lineales U_t son unitarios, $\langle U_t(\psi)|U_t(\phi)\rangle = \langle U_0(\psi)|U_0(\phi)\rangle = \langle\psi|\phi\rangle$, porque

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\langle U_t\psi|U_t\phi\rangle = \langle\tfrac{\partial(U_t\psi)}{\partial t}|U_t\phi\rangle + \langle U_t\psi|\tfrac{\partial(U_t\phi)}{\partial t}\rangle = \tfrac{i}{\hbar}\langle H(U_t\psi)|U_t\phi\rangle - \tfrac{i}{\hbar}\langle U_t\psi|H(U_t\phi)\rangle = 0.$$

En particular, si $\|\psi\| = 1$, se cumple que $\|U_t\psi\| = 1$ en todo instante t.

En adelante E denotará un K-espacio vectorial de dimensión finita n (y $K = \mathbb{R}$ ó $\mathbb{C}$).

6. El Espacio Dual

Definición: El **espacio dual** de un K-espacio vectorial E es el K-espacio vectorial E^* formado (teorema 3.2) por todas las aplicaciones K-lineales $E \to K$, aplicaciones que reciben el nombre de **formas lineales** sobre E.

Teorema 6.1 *Sean $e_1, \ldots, e_n$ una base de E. Dados n escalares $a_1, \ldots, a_n \in K$, existe una única forma lineal $\omega \in E^*$ tal que $\omega(e_1) = a_1, \ldots, \omega(e_n) = a_n$, y es*

$$\omega(x_1, \ldots, x_n) := \omega(x_1 e_1 + \ldots + x_n e_n) = a_1 x_1 + \ldots + a_n x_n.$$

Demostración: Es un caso particular de 3.1, cuando $E' = K$.

Teorema 6.2 *Fijada una base $e_1, \ldots, e_n$ en E, existe una única base $\omega_1, \ldots, \omega_n$ de E^*, llamada **base dual**, tal que*

$$\omega_i(e_j) = \delta_{ij},$$

y las coordenadas de cualquier forma lineal $\omega \in E^*$ en esta base dual son $(\omega(e_1), \ldots, \omega(e_n))$.

En particular, el dual E^* es un K-espacio vectorial de dimensión

$$\boxed{\dim E^* = \dim E.}$$

Demostración: Según 6.1, existen formas lineales $\omega_1, \ldots, \omega_n$ tales que $\omega_i(e_j) = \delta_{ij}$, y son las formas lineales $\omega_i(x_1 e_1 + \ldots + x_n e_n) = x_i$, por lo que a menudo pondremos $\omega_i = \mathrm{d}x_i$, pues su valor en un vector $e = q - p$ es la diferencia entre las i-ésimas coordenadas de p y q.

Veamos que $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$ son linealmente independientes:

Si $\sum_i \lambda_i \mathrm{d}x_i = 0$, para todo índice $j = 1, \ldots, n$ tenemos que

$$0 = \big(\textstyle\sum_i \lambda_i \mathrm{d}x_i\big)(e_j) = \sum_i \lambda_i \big(\mathrm{d}x_i(e_j)\big) = \sum_i \lambda_i \delta_{ij} = \lambda_j.$$

Veamos ahora que $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$ generan el espacio dual E^*:

Dada una forma lineal $\omega \in E^*$, para todo vector $e = \sum_i x_i e_i \in E$ se cumple que

$$\omega(e) = \omega(x_1 e_1 + \ldots + x_n e_n) = x_1 \omega(e_1) + \ldots + x_n \omega(e_n),$$
$$\omega = \omega(e_1)\mathrm{d}x_1 + \ldots + \omega(e_n)\mathrm{d}x_n.$$

Proposición 6.3 *Sean $e_1, \ldots, e_n$ y $v_1, \ldots, v_n$ dos bases de E, y sean $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$ y $\mathrm{d}y_1, \ldots, \mathrm{d}y_n$ sus respectivas bases duales. Si $B = (b_{ij})$ es la matriz de cambio de base de $e_1, \ldots, e_n$ a $v_1, \ldots, v_n$, entonces B^t es la matriz de cambio de base de $\mathrm{d}y_1, \ldots, \mathrm{d}y_n$ a $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$.*

Demostración: La i-ésima coordenada del vector v_j en la base $e_1, \ldots, e_n$ es $b_{ij} = \mathrm{d}x_i(v_j)$.

Sea $C = (c_{ij})$ la matriz de cambio de base de $\mathrm{d}y_1, \ldots, \mathrm{d}y_n$ a $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$. La i-ésima coordenada de la forma lineal $\mathrm{d}x_j$ en la base $\mathrm{d}y_1, \ldots, \mathrm{d}y_n$ es $c_{ij} = \mathrm{d}x_j(v_i) = b_{ji}$.

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = B^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad , \quad \begin{pmatrix} \mathrm{d}x_1 \\ \vdots \\ \mathrm{d}x_n \end{pmatrix} = B \begin{pmatrix} \mathrm{d}y_1 \\ \vdots \\ \mathrm{d}y_n \end{pmatrix}.$$

Definición: Sea E un espacio vectorial euclídeo. El **bra** de un vector $e \in E$ es la aplicación $\langle e| \colon E \longrightarrow K$ que asigna a cada vector $v \in E$ el escalar $\langle e|v\rangle$.

Esta aplicación es lineal, porque el producto escalar es lineal por la derecha, y la base dual de una base ortonormal $u_1, \ldots, u_n$ es justamente $\langle u_1|, \ldots, \langle u_n|$; porque $\langle u_i|u_j\rangle = \delta_{ij}$.

Teorema 6.4 *Sea E un espacio vectorial euclídeo. La aplicación natural*

$$\phi\colon E \longrightarrow E^* \ , \ \phi(e) = \langle e|,$$

es un isomorfismo antilineal.

Demostración: Es antilineal porque $\langle e+e'| = \langle e| + \langle e'|$ y $\langle \lambda e| = \bar{\lambda}\langle e|$.

Es inyectiva: Si $e \in \operatorname{Ker}\phi$, entonces $0 = \phi(e) = \langle e|$. Luego $0 = \langle e|e\rangle$, y se sigue que $e = 0$. Es decir, $\operatorname{Ker}\phi = 0$, y concluimos que ϕ es inyectiva (1.4).

Ahora, por el Teorema de Isomorfía (que es válido para aplicaciones antilineales, y con la misma demostración que hemos dado), $\dim(\operatorname{Im}\phi) = \dim E = \dim E^*$; luego $\operatorname{Im}\phi = E^*$ y vemos que ϕ es epiyectiva.

Otra manera de ver que ϕ es epiyectiva es fijar una base ortonormal $u_1, \ldots, u_n$ en E y observar que toda forma lineal $\omega \in E^*$ es $\omega = \sum_i \lambda_i \langle u_i| = \sum_i \lambda_i \phi(u_i) = \phi(\sum_i \bar{\lambda}_i u_i)$.

Ejemplos:

1. Dada una función diferenciable $f\colon \mathbb{R}^n \to \mathbb{R}$ y un punto $p \in \mathbb{R}^n$, para cada vector $e \in \mathbb{R}^n$ podemos considerar la derivada en $t = 0$ de la función $h\colon \mathbb{R} \to \mathbb{R}$, $h(t) = f(p+te)$, que se obtiene al restringir f a la recta que pasa por p con dirección $\mathbb{R}e$

$$(\mathrm{d}f)(e) = \begin{bmatrix} \text{Derivada de } f \\ \text{en el punto } p \\ \text{en la dirección } e \end{bmatrix} := \lim_{t\to 0} \frac{f(p+te) - f(p)}{t}$$

de modo que, para valores pequeños ε de t se cumple que

$$\frac{f(p+\varepsilon e) - f(p)}{\varepsilon} \approx k \ , \quad f(p+\varepsilon e) \approx f(p) + k\varepsilon \ ; \qquad k := (\mathrm{d}f)(e).$$

Así, en el punto p la función f es creciente (resp. decreciente) en la dirección e cuando $(\mathrm{d}f)(e) > 0$ (resp. $(df)(e) < 0$), y los máximos y mínimos locales de f son **puntos críticos**, en el sentido de que $(\mathrm{d}f)(e) = 0$ para todo vector $e \in \mathbb{R}^n$.

Cuando el punto p no es crítico y $(\mathrm{d}f)(e) = 0$, se dice que en el punto p el vector e es tangente a la hipersuperficie de nivel $f(x_1, \ldots, x_n) = f(p)$ que pasa por p.

Tenemos así una forma lineal $\mathrm{d}f\colon \mathbb{R}^n \to \mathbb{R}$, llamada **diferencial** de la función f en el punto $p \in \mathbb{R}^n$, que da el incremento infinitesimal de f en p en cualquier dirección $e \in \mathbb{R}^n$.

Por definición $\frac{\partial f}{\partial x_i}(a_1, \ldots, a_n)$ es la derivada de la función $f(a_1, \ldots, a_i + t, \ldots, a_n)$ en $t = 0$; luego $\frac{\partial f}{\partial x_i}(p) = (\mathrm{d}f)(e_i)$, donde $e_1, \ldots, e_n$ es la base usual de $\mathbb{R}^n$. Según 6.2, en la base dual $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$ se cumple que $\mathrm{d}f = \sum_i \frac{\partial f}{\partial x_i}(p)\mathrm{d}x_i$:

$$\boxed{\mathrm{d}f = \tfrac{\partial f}{\partial x_1}\,\mathrm{d}x_1 + \ldots + \tfrac{\partial f}{\partial x_n}\,\mathrm{d}x_n.}$$

2. Cuando en $\mathbb{R}^n$ se considera el producto escalar usual, de acuerdo con 6.4, existe un único vector $\operatorname{grad} f \in \mathbb{R}^n$, llamado **gradiente** de f en p, tal que $\phi(\operatorname{grad} f) = \mathrm{d}f$; i.e.

$$\boxed{\langle \operatorname{grad} f | e\rangle = (\mathrm{d}f)(e) \ , \quad \operatorname{grad} f = \big(\tfrac{\partial f}{\partial x_1}, \ldots, \tfrac{\partial f}{\partial x_n}\big).}$$

El punto es crítico cuando $\operatorname{grad} f = 0$, y cuando el gradiente no es nulo, la dirección del hiperplano tangente en p a la hipersuperficie de nivel $f(x_1, \ldots, x_n) = f(p)$ es el ortogonal $\langle \operatorname{grad} f\rangle^{\perp}$, la dirección de la recta normal en p es $\langle \operatorname{grad} f\rangle$, y el gradiente indica la dirección (en p) de máximo crecimiento de la función f.

3. En Física, Análisis y Geometría, casi nunca se considera un vector o una forma lineal aislada, sino campos de vectores o campos de formas lineales; es decir, en cada punto $p \in E$ se considera un vector $e \in E$ o una forma lineal $\omega \in E^*$, que depende del punto p considerado. Así, en Mecánica Clásica (fijado el origen de coordenadas y la unidad de longitud) el espacio es un espacio vectorial euclídeo real de dimensión 3, y fijada una base ortonormal $\vec{i}, \vec{j}, \vec{k}$, se identifica con $\mathbb{R}^3$. La fuerza viene dada por un campo de vectores $\vec{F} = F_1(x,y,z)\vec{i} + F_2(x,y,z)\vec{j} + F_3(x,y,z)\vec{k}$, y el trabajo es el campo de formas lineales
$$\omega_{\vec{F}} = F_1(x,y,z)\mathrm{d}x + F_2(x,y,z)\mathrm{d}y + F_3(x,y,z)\mathrm{d}z.$$
4. **Notación de Dirac:** Sea E un espacio vectorial euclídeo. El bra de un vector $e \in E$ es la forma lineal $\langle e|$, mientras que el **ket** $|e\rangle$ es el propio vector e, y dos vectores $e, v \in E$ definen un operador lineal $|e\rangle\langle v| \colon E \to E$, $(|e\rangle\langle v|)(u) = |e\rangle\langle v|u\rangle := \langle v|u\rangle e$.

6.1. Incidencia y Bidualidad

Cada vector $e \in E$ define una aplicación lineal $\psi_e \colon E^* \to K$, $\psi_e(\omega) = \omega(e)$, así que tenemos una aplicación natural
$$\Psi \colon E \longrightarrow E^{**} \ , \ \Psi(e) = \psi_e.$$
Es decir, $\Psi(e)(\omega) = \omega(e)$, $\forall e \in E, \omega \in E^*$.

Teorema de Bidualidad: *La aplicación natural $\Psi \colon E \to E^{**}$, $\Psi(e)(\omega) = \omega(e)$, es un isomorfismo lineal.*

Demostración: Veamos primero que la aplicación Ψ es lineal. Si $e, v \in E$ y $\lambda \in K$, se cumple que $\psi_{e+v} = \psi_e + \psi_v$ y $\psi_{\lambda e} = \lambda\psi_e$ porque, para toda forma lineal $\omega \in E^*$,
$$\psi_{e+v}(\omega) = \omega(e+v) = \omega(e) + \omega(v) = \psi_e(\omega) + \psi_v(\omega) = (\psi_e + \psi_v)(\omega),$$
$$\psi_{\lambda e}(\omega) = \omega(\lambda e) = \lambda\omega(e) = \lambda\big(\psi_e(\omega)\big) = \big(\lambda\psi_e\big)(\omega).$$
Veamos ahora que $\mathrm{Ker}\,\Psi = 0$.
Si $\Psi(e) = 0$, entonces $\omega(e) = 0$ para toda forma lineal $\omega \in E^*$, y de 6.2 se sigue que, en cualquier base de E, todas las coordenadas del vector e son nulas; luego $e = 0$.

Por último, del Teorema de Isomorfía se sigue que $\dim(\mathrm{Im}\,\Psi) = \dim E = \dim E^{**}$.
Luego $\mathrm{Im}\,\Psi = E^{**}$, y Ψ es epiyectiva.

Corolario 6.5 *Cada base $\omega_1, \ldots, \omega_n$ de E^* es la base dual de una única base de E.*

Demostración: Existe una única base $e_1, \ldots e_n$ de E tal que $\Psi(e_1), \ldots, \Psi(e_n)$ es la base dual de $\omega_1, \ldots, \omega_n$; es decir, tal que $\delta_{ij} = \Psi(e_i)(\omega_j) = \omega_j(e_i)$.

Definición: El **incidente** de un subespacio vectorial V de E es
$$V^{\circ} = \{\omega \in E^* \colon \omega(v) = 0, \forall v \in V\},$$
y el incidente de un subespacio vectorial W de E^*, identificando E^{**} con E, es
$$W^{\circ} = \{\psi_e \in E^{**} \colon \psi_e(\omega) = 0, \forall \omega \in W\} = \{e \in E \colon \omega(e) = 0, \forall \omega \in W\}.$$

Ejemplo: Cuando E es un espacio vectorial euclídeo, tenemos un isomorfismo antilineal $\phi \colon E \to E^*$, $\phi(e) = \langle e|$, y la condición $\phi(e) \in V^{\circ}$ significa que $e \in V^{\perp}$; i.e., $V^{\circ} = \phi(V^{\perp})$.

Teorema 6.6 *El incidente es un subespacio vectorial de dimensión*
$$\boxed{\dim V^{\circ} = \dim E - \dim V.}$$

Demostración: Ampliemos una base $e_1, \dots, e_d$ de V hasta obtener una base $e_1, \dots, e_d, \dots, e_n$ de E, y sea $\mathrm{d}x_1, \dots, \mathrm{d}x_n$ su base dual.

Una forma lineal $\omega = \sum_i \lambda_i \mathrm{d}x_i$ está en V^{o} cuando $\lambda_1 = \omega(e_1) = 0, \dots, \lambda_d = \omega(e_d) = 0$; es decir, cuando $\omega = \lambda_{d+1}\mathrm{d}x_{d+1} + \ldots + \lambda_n \mathrm{d}x_n$.

Luego $V^{\mathrm{o}} = \langle \mathrm{d}x_{d+1}, \dots, \mathrm{d}x_n \rangle$, que es un subespacio vectorial de E^*, de dimensión $n-d$ porque $\mathrm{d}x_{d+1}, \dots, \mathrm{d}x_n$ son linealmente independientes (forman parte de una base de E^*).

Corolario 6.7 $\boxed{(V^{\mathrm{o}})^{\mathrm{o}} = V.}$

Demostración: $(V^{\mathrm{o}})^{\mathrm{o}} = \{e \in E \colon \omega(e) = 0, \forall \omega \in V^{\mathrm{o}}\} \supseteq V$, y $V = (V^{\mathrm{o}})^{\mathrm{o}}$ porque

$$\dim (V^{\mathrm{o}})^{\mathrm{o}} = \dim E^* - \dim V^{\mathrm{o}} = \dim E - (\dim E - \dim V) = \dim V.$$

Corolario 6.8 *Si V y W son subespacios vectoriales de E, se cumple que:*

1. $V \subseteq W \;\Leftrightarrow\; W^{\mathrm{o}} \subseteq V^{\mathrm{o}}$ *, y en particular* $V = W \;\Leftrightarrow\; V^{\mathrm{o}} = W^{\mathrm{o}}$
2. $(V+W)^{\mathrm{o}} = V^{\mathrm{o}} \cap W^{\mathrm{o}}$
3. $(V \cap W)^{\mathrm{o}} = V^{\mathrm{o}} + W^{\mathrm{o}}$

Demostración: Si $V \subseteq W$, es claro que $W^{\mathrm{o}} \subseteq V^{\mathrm{o}}$.

Recíprocamente, si $W^{\mathrm{o}} \subseteq V^{\mathrm{o}}$, entonces $V = (V^{\mathrm{o}})^{\mathrm{o}} \subseteq (W^{\mathrm{o}})^{\mathrm{o}} = W$.

2.– Como $V \subseteq V+W$ y $W \subseteq V+W$, tenemos que $(V+W)^{\mathrm{o}} \subseteq V^{\mathrm{o}}$ y $(V+W)^{\mathrm{o}} \subseteq W^{\mathrm{o}}$; luego $(V+W)^{\mathrm{o}} \subseteq V^{\mathrm{o}} \cap W^{\mathrm{o}}$.

Además, si $\omega \in V^{\mathrm{o}} \cap W^{\mathrm{o}}$, entonces $\omega(v+w) = \omega(v) + \omega(w) = 0$ para todo vector $v + w \in V + W$. Luego $V^{\mathrm{o}} \cap W^{\mathrm{o}} \subseteq (V+W)^{\mathrm{o}}$ y concluimos que $(V+W)^{\mathrm{o}} = V^{\mathrm{o}} \cap W^{\mathrm{o}}$.

3.– Por el primer apartado, para demostrar que $(V \cap W)^{\mathrm{o}} = V^{\mathrm{o}} + W^{\mathrm{o}}$ basta ver que sus incidentes coinciden: $\big(V^{\mathrm{o}} + W^{\mathrm{o}}\big)^{\mathrm{o}} \overset{3}{=} (V^{\mathrm{o}})^{\mathrm{o}} \cap (W^{\mathrm{o}})^{\mathrm{o}} = V \cap W = \big((V\cap W)^{\mathrm{o}}\big)^{\mathrm{o}}$.

6.2. Adjunto de un Operador Lineal

Definición: Sea $T \colon E \to F$ una aplicación lineal entre espacios vectoriales euclídeos. Para cada vector $v \in F$ tenemos una aplicación $\eta_v \colon E \to K$, $\eta_v(e) = \langle v|T(e)\rangle$, que es lineal

$$\eta_v(\lambda e+\mu e') = \langle v|T(\lambda e+\mu e')\rangle = \langle v|\lambda T(e)+\mu T(e')\rangle = \lambda\langle v|T(e)\rangle+\mu\langle v|T(e')\rangle = \lambda\eta_v(e)+\mu\eta_v(e').$$

Como el producto escalar define (6.4) un isomorfismo $\phi \colon E \to E^*$, $\phi(e) = \langle e|$, existe un único vector $T^\dagger(v) \in E$ tal que $\eta_v = \langle T^\dagger(v)|$; es decir,

$$\boxed{\langle v|T(e)\rangle = \langle T^\dagger(v)|e\rangle} \qquad \forall e, v \in E.$$

Como $\langle T(e)|v\rangle = \overline{\langle v|T(e)\rangle}$ y $\langle e|T^\dagger(v)\rangle = \overline{\langle T^\dagger(v)|e\rangle}$, también tenemos que

$$\boxed{\langle T(e)|v\rangle = \langle e|T^\dagger(v)\rangle} \qquad \forall e, v \in E,$$

y en particular $\mathrm{Id}^\dagger = \mathrm{Id}$ y $(T^\dagger)^\dagger = T$.

Esta aplicación $T^\dagger \colon F \to E$ es lineal, y recibe el nombre de **adjunta**[3] de T.

En efecto, para probar que $T^\dagger(\lambda v + \mu v') = \lambda T^\dagger(v) + \mu T^\dagger(v')$, según 6.4 basta ver que ambos vectores tienen el mismo producto escalar con cualquier vector $e \in E$, y

$$\begin{aligned}\langle T^\dagger(\lambda v + \mu v')|e\rangle &= \langle \lambda v + \mu v'|T(e)\rangle = \bar\lambda\langle v|T(e)\rangle + \bar\mu\langle v'|T(e)\rangle \\ &= \bar\lambda\langle T^\dagger(v)|e\rangle + \bar\mu\langle T^\dagger(v')|e\rangle = \langle \lambda T^\dagger(v) + \mu T^\dagger(v')|e\rangle.\end{aligned}$$

[3] Un endomorfismo $T \colon E \to E$ es autoadjunto cuando $\langle T(e)|v\rangle = \langle e|T(v)\rangle$, $\forall e, v \in E$; es decir, $T = T^\dagger$.

Proposición 6.9 *Sea A la matriz de T en ciertas bases ortonormales $u_1, \ldots, u_n$ de E y $v_1, \ldots, v_m$ de F. La matriz de $T^\dagger$ en tales bases ortonormales es $\bar{A}^t$.*

Demostración: Sean $A = (a_{ij})$ y (c_{ij}) las matrices de T y $T^\dagger$ respectivamente; es decir,

$$T(u_j) = \textstyle\sum_i a_{ij} v_i \quad , \quad T^\dagger(v_j) = \sum_i c_{ij} u_i .$$

Como las bases duales de $u_1, \ldots, u_n$ y $v_1, \ldots, v_m$ son sus respectivos bras,

$$c_{ij} = \langle u_i | T^\dagger(v_j) \rangle = \langle T(u_i) | v_j \rangle = \overline{\langle v_j | T(u_i) \rangle} = \bar{a}_{ji} .$$

Teorema 6.10 $\operatorname{Ker} T^\dagger = (\operatorname{Im} T)^\perp$, $\operatorname{Im} T^\dagger = (\operatorname{Ker} T)^\perp$.

Demostración: Si $v \in F$, la condición $T^\dagger(v) = 0$ significa que $0 = \langle T^\dagger(v) | e \rangle = \langle v | T(e) \rangle$ para todo vector $e \in E$; es decir, que $v \in (\operatorname{Im} T)^\perp$.

Ahora $\operatorname{Ker} T = \operatorname{Ker} (T^\dagger)^\dagger = (\operatorname{Im} T^\dagger)^\perp$, y $(\operatorname{Ker} T)^\perp = \big((\operatorname{Im} T^\dagger)^\perp\big)^\perp = \operatorname{Im} T^\dagger$.

Corolario 6.11 $\dim (\operatorname{Im} T^\dagger) = \dim (\operatorname{Im} T)$.

Demostración: $\dim (\operatorname{Im} T^\dagger) = \dim (\operatorname{Ker} T)^\perp = \dim E - \dim (\operatorname{Ker} T) = \dim \operatorname{Im} T$.

Proposición 6.12 *Un subespacio vectorial V de E queda invariante por un endomorfismo $T \colon E \to E$, i.e. $T(V) \subseteq V$, si y sólo si su ortogonal $V^\perp$ queda invariante por $T^\dagger$.*

Demostración: Supongamos que $T(V) \subseteq V$. Si $w \in V^\perp$, se cumple que

$$\langle T^\dagger(w) | v \rangle = \langle w | T(v) \rangle = 0$$

para todo vector $v \in V$, porque $T(v) \in V$ y $w \in V^\perp$.

Luego $T^\dagger(w) \in V^\perp$, y vemos que $T^\dagger(V^\perp) \subseteq V^\perp$.

Recíprocamente, si $T^\dagger(V^\perp) \subseteq V^\perp$, entonces $(T^\dagger)^\dagger\big((V^\perp)^\perp\big) \subseteq (V^\perp)^\perp$ según acabamos de ver, y concluimos que $T(V) \subseteq V$ porque $(T^\dagger)^\dagger = T$ y $(V^\perp)^\perp = V$.

Nota: La condición de que una recta Ke quede invariante por un endomorfismo T significa que $T(e) \in Ke$; es decir, que e es un vector propio de T. Por tanto, de acuerdo con 6.12 los subespacios vectoriales de dimensión $n - 1$ invariantes por T son los ortogonales de las rectas generadas por un vector propio de $T^\dagger$.

7. Formas Cuadráticas

Definición: Llamaremos **forma cuadrática** a toda aplicación $q\colon E \to K$ que, en el sistema de coordenadas $(x_1, \ldots, x_n)$ definido por una base $e_1, \ldots, e_n$ de E, venga dada por un polinomio homogéneo de grado 2:

$$q(x_1e_1 + \ldots + x_ne_n) = \sum_{1\leq i\leq j\leq n} c_{ij}x_ix_j.$$

$$q = \sum_{1\leq i\leq j\leq n} c_{ij}\mathrm{d}x_i\mathrm{d}x_j.$$

7.1. Métricas

Definición: Dada una aplicación $S\colon E \times E \to K$, pondremos $e \cdot v := S(e, v)$, y diremos que S es una **métrica** si es bilineal y simétrica:

1. $(e + e') \cdot v = e \cdot v + e' \cdot v$, $(\lambda e) \cdot v = \lambda(e \cdot v)$; $\forall e, e', v \in E$, $\lambda \in K$.
   $v \cdot (e + e') = v \cdot e + v \cdot e'$, $v \cdot (\lambda e) = \lambda(v \cdot e)$; $\forall e, e', v \in E$, $\lambda \in K$.
2. $e \cdot v = v \cdot e$, $\forall e, v \in E$.

Matriz de una Métrica: Fijada una base $e_1, \ldots, e_n$ de E, la matriz de una métrica S es la matriz $A = (a_{ij})$, donde $a_{ij} := e_i \cdot e_j$. Es una matriz simétrica de n filas y columnas.

Fijada la base de E, esta matriz determina totalmente a la métrica S, porque el producto $e \cdot v$ del vector $e \in E$ de coordenadas $X^t = (x_1, \ldots, x_n)$ con el vector v de coordenadas $Y^t = (y_1, \ldots, y_n)$ es

$$e \cdot v = \left(\sum_{i=1}^n x_ie_i\right) \cdot \left(\sum_{j=1}^n y_je_j\right) = \sum_{i,j=1}^n a_{ij}x_iy_j = (x_1 \ldots x_n)\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn}\end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n\end{pmatrix}$$

$$\boxed{e \cdot v = X^tAY.}$$

Además, cualquier matriz simétrica A de n filas y columnas define una métrica mediante la igualdad anterior.

Forma Cuadrática Asociada a una Métrica: Cuando los vectores e y v coinciden, vemos que la función $q\colon E \to K$, $q(e) = e \cdot e = S(e, e)$, es una forma cuadrática

$$q(x_1e_1 + \ldots + x_ne_n) = X^tAX = \sum_{i=1}^n a_{ii}x_i^2 + \sum_{1\leq i<j\leq n} 2a_{ij}x_ix_j, \tag{12}$$

llamada **forma cuadrática asociada** a S. Determina totalmente a la métrica porque

$$\tfrac{1}{2}\big(q(e + v) - q(e) - q(v)\big) = \tfrac{1}{2}(e \cdot e + e \cdot v + v \cdot e + v \cdot v - e \cdot e - v \cdot v) = e \cdot v,$$

y (12) muestra que toda forma cuadrática $q\colon E \to K$ está asociada a una métrica.

En el caso real ($K = \mathbb{R}$), cuando $q(e) > 0$ ha de entenderse que $+\sqrt{q(e)} = +\sqrt{e \cdot e} \in \mathbb{R}$ es la medida del vector e con la métrica S (o con su forma cuadrática asociada q).

Ejemplos:

1. La aplicación nula $S\colon E \times E \to K$, $S(e, v) = 0$, es una métrica. Su matriz en cualquier base es la matriz 0 y la forma cuadrática asociada $q\colon E \to K$ es nula: $q(e) = 0$, $\forall e \in E$.

2. En el caso real (¡y no en el complejo!) los productos escalares son las métricas en que la forma cuadrática asociada es definido-positiva (i.e. positiva en los vectores no nulos), y una base es ortonormal justo cuando la matriz de la métrica es la matriz unidad I_n, es decir, cuando la forma cuadrática asociada es $\mathrm{d}x_1^2 + \ldots + \mathrm{d}x_n^2$.

3. Dada una forma cuadrática $q\colon E \to K$, su métrica asociada $S\colon E \times E \to K$ induce una métrica $S|_V\colon V \times V \to K$, $S|_V(v, v') = S(v, v')$, en cada subespacio vectorial $V \subseteq E$, y la forma cuadrática asociada es la restricción $q|_V\colon V \to K$, $q|_V(v) = q(v)$, de q a V.

Cambio de Base en la Matriz de una Métrica: Si $\tilde{A} = (\tilde{a}_{ij})$ es la matriz de S en otra base $\tilde{e}_1, \ldots, \tilde{e}_n$ de E, y $B = (b_{ij})$ es la matriz de cambio de base, las coordenadas $\tilde{X}$, $\tilde{Y}$ de los vectores e y v en la nueva base cumplen que $X = B\tilde{X}$, $Y = B\tilde{Y}$.

Luego $\tilde{X}^t\tilde{A}\tilde{Y} = e \cdot v = X^tAY = \tilde{X}^tB^tAB\tilde{Y}$, y la matriz de S en la nueva base es

$$\boxed{\tilde{A} = B^tAB.}$$

Lema 7.1 *En el caso real, si el determinante $|A|$ es positivo (resp. negativo, nulo), también lo es $|\tilde{A}|$. Es decir, el signo de $|A|$ no depende de la base, sólo de la métrica.*

Demostración: Tenemos que $|\tilde{A}| = |B^tAB| = |B|^2|A|$, y $|B| \neq 0$ porque B es invertible.

Teorema 7.2 *Sea S una métrica sobre un espacio vectorial real E, y sea $A = (a_{ij})$ su matriz en una base $e_1, \ldots, e_n$ de E. La métrica S es un producto escalar si y sólo si son positivos todos los* ***menores principales*** *de la matriz A*

$$a_{11}\ ,\quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}\ ,\quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}\ , \ldots ,\quad \begin{vmatrix} a_{11} & \ldots & a_{1r} \\ \ldots & \ldots & \ldots \\ a_{r1} & \ldots & a_{rr} \end{vmatrix}\ , \ldots ,\quad \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} = |A|.$$

Demostración: Según el lema, el determinante de la matriz de un producto escalar siempre es positivo, porque lo es en las bases ortonormales.

Luego, si S es un producto escalar, todo menor principal es positivo, pues es el determinante de la matriz de S en la base $e_1, \ldots, e_r$ del espacio vectorial $\mathbb{R}e_1 + \ldots + \mathbb{R}e_r$.

Recíprocamente, si todos los menores principales son positivos, procediendo por inducción sobre la dimensión n, podemos suponer que S es definido-positiva en el hiperplano $H = \mathbb{R}e_1 + \ldots + \mathbb{R}e_{n-1}$, que por tanto admite alguna base ortonormal $u_1, \ldots, u_{n-1}$.

Existe un vector no nulo $e \in E$ tal que $u_1 \cdot e = \ldots = u_{n-1} \cdot e = 0$, porque la aplicación lineal $f\colon E \to \mathbb{R}^{n-1}$, $f(e) = (u_1 \cdot e, \ldots, u_{n-1} \cdot e)$ no puede ser inyectiva.

Pero $e \notin H$, porque $e \cdot H = 0$ y S es definido-positiva en H; luego $u_1, \ldots, u_{n-1}, e$ es una base de E. La matriz de S en esta base es diagonal $\mathrm{diag}(1, \ldots, 1, a)$, de modo que la forma cuadrática es $q(x_1, \ldots, x_n) = x_1^2 + \ldots + x_{n-1}^2 + ax_n^2$.

Por el lema, el signo de a coincide con el de $|A| > 0$.

Luego $a > 0$, y la forma cuadrática $\mathrm{d}x_1^2 + \ldots + \mathrm{d}x_{n-1}^2 + a\mathrm{d}x_n^2$ es definido-positiva.

7.2. Clasificación de Métricas

Definición: Dada una métrica S en un K-espacio vectorial E, diremos que un vector $e \in E$ es **isótropo** si $q(e) = e \cdot e = 0$.

Si todos los vectores son isótropos, entonces $q = 0$, y por tanto $S = 0$.

Lema 7.3 *Si $e \in E$ no es isótropo, entonces su ortogonal $(Ke)^\perp := \{v \in E\colon e \cdot v = 0\}$ es un subespacio vectorial de dimensión $n - 1$ y $E = Ke \oplus (Ke)^\perp$.*

Demostración: Consideremos la forma lineal $\omega\colon E \to K$, $\omega(v) = e \cdot v$, que no es nula porque $\omega(e) = e \cdot e \neq 0$; luego su imagen es de dimensión 1 y, según el Teorema de Isomorfía, su núcleo $(Ke)^\perp$ es un subespacio vectorial de dimensión $n - 1$.

La suma $Ke + (Ke)^\perp$ es directa porque si $\lambda e \in (Ke)^\perp$, entonces $\lambda(e \cdot e) = 0$, y $\lambda = 0$.

Como $\dim(Ke \oplus (Ke)^\perp) = 1 + (n - 1) = \dim E$, concluimos que $Ke \oplus (Ke)^\perp = E$.

Teorema 7.4 *Toda métrica S admite una base $e_1, \dots, e_n$ en que su matriz es diagonal: $e_i \cdot e_j = 0$ cuando $i \neq j$.*

Demostración: Por inducción sobre $n = \dim E$, y es obvio cuando $n = 1$, porque toda matriz 1×1 es diagonal.

Cuando $n > 1$, es obvio si $q = 0$. Si $q \neq 0$, consideramos un vector no isótropo $e_1 \in E$.

Como $\dim (Ke_1)^\perp = n - 1$, por hipótesis de inducción existe alguna base $e_2, \dots, e_n$ de $(Ke_1)^\perp$ en que la matriz de S es diagonal.

Ahora los vectores $e_1, e_2, \dots, e_n$ cumplen que $e_1 \cdot e_j = 0$ cuando $1 < j \leq n$, porque $e_1 \cdot e_2 = \dots = e_1 \cdot e_n = 0$, y forman una base de E porque lo generan:

$$Ke_1 + Ke_2 + \dots + Ke_n = Ke_1 + (Ke_1)^\perp = E.$$

Definición: Dada una métrica $S \colon E \times E \to K$, cada vector $e \in E$ define una aplicación $\omega_e \colon E \to K$, $\omega_e(v) = e \cdot v = S(e, v)$, que es lineal porque S es lineal a la derecha.

Obtenemos así una aplicación natural $\phi \colon E \to E^*$, $\phi(e) = \omega_e$, la **polaridad** asociada a la métrica S (o a su forma cuadrática), que es lineal porque S es lineal a la izquierda:

$$\begin{aligned} \omega_{e+e'}(v) &= (e + e') \cdot v = e \cdot v + e' \cdot v = \omega_e(v) + \omega_{e'}(v) = (\omega_e + \omega_{e'})(v), \\ \omega_{\lambda e}(v) &= (\lambda e) \cdot v = \lambda(e \cdot v) = \lambda \omega_e(v) = (\lambda \omega_e)(v), \end{aligned}$$

Si $A = (a_{ij})$ es la matriz de S en una base $e_1, \dots, e_n$ de E, y consideramos la base dual $\mathrm{d}x_1, \dots, \mathrm{d}x_n$ en E^*, entonces la matriz de ϕ en estas bases también es A porque, según 6.2, la i-ésima coordenada de $\phi(e_j)$ en la base dual es $\phi(e_j)(e_i) = \omega_{e_j}(e_i) = e_j \cdot e_i = a_{ij}$.

El **rango** de la métrica (o de la forma cuadrática asociada) es $\mathrm{rg}\, S := \dim (\mathrm{Im}\, \phi)$.

De acuerdo con 3.5, *el rango de una métrica coincide con el rango de su matriz en cualquier base:* $\mathrm{rg}\, S = \mathrm{rg}\, A$.

La polaridad es un isomorfismo cuando la métrica es de rango máximo, $\mathrm{rg}\, S = \dim E$.

Teorema 7.5 *En el caso complejo, toda métrica de rango r admite una base $e_1, \dots, e_n$ en que su matriz es* $\mathrm{diag}(1, \overset{r}{\dots}, 1, 0, \dots, 0)$*, y su forma cuadrática es* $\mathrm{d}x_1^2 + \dots + \mathrm{d}x_r^2$.

Demostración: Según el lema anterior, en alguna base $v_1, \dots, v_n$ la matriz de la métrica es $\mathrm{diag}(a_1, \dots, a_m, 0, \dots 0)$, donde los coeficientes $a_i = v_i \cdot v_i$ no son nulos; luego $m = r$.

Sustituyendo v_i por $e_i = v_i / \sqrt{a_i}$, cuando $i = 1, \dots, r$, tendremos que $e_i \cdot e_i = 1$, y la matriz de la métrica en la base $e_1, \dots, e_r, v_{r+1}, \dots, v_n$ es $\mathrm{diag}(1, \overset{r}{\dots}, 1, 0, \dots, 0)$.

Teorema de Inercia de Sylvester (1814-1897)**:** *En el caso real, toda métrica S de rango r admite una base $e_1, \dots, e_n$ en que su matriz es* $\mathrm{diag}(1, \overset{p}{\dots}, 1, -1, \overset{q}{\dots}, -1, 0, \dots, 0)$*, con $p + q = r$, y su forma cuadrática es* $\mathrm{d}x_1^2 + \dots + \mathrm{d}x_p^2 - \mathrm{d}x_{p+1}^2 - \dots - \mathrm{d}x_{p+q}^2$.

El par (p, q) no depende de la base elegida, y decimos que la métrica S (o su forma cuadrática) es de **signatura** (p, q) o de tipo $(+, \overset{p}{\dots}, +, -, \overset{q}{\dots}, -, 0, \dots, 0)$.

Demostración: En alguna base $v_1, \dots, v_n$ la matriz de la métrica es $\mathrm{diag}(a_1, \dots, a_m, 0, \dots 0)$, donde los coeficientes $a_i = v_i \cdot v_i$ no son nulos; luego $m = r$, y sustituyendo v_i por el vector $e_i = v_i / \sqrt{|a_i|}$, cuando $i = 1, \dots, r$, tenemos que $e_i \cdot e_i = \pm 1$, y la matriz de S en la base $e_1, \dots, e_r, e_{r+1} = v_{r+1}, \dots, e_n = v_n$, después de reordenarla, tiene la forma deseada.

Por último, para ver que el número p (y por tanto $q = r - p$) no depende de la base elegida, basta probar que *p coincide con el máximo p' de las dimensiones de los subespacios vectoriales $V_+ \subseteq E$ en que la métrica S es definido-positiva*:

Si en una base la matriz de S es $\mathrm{diag}(1, \overset{p}{\dots}, 1, -1, \overset{q}{\dots}, -1, 0, \dots, 0)$, la métrica es definido-positiva en $\mathbb{R}e_1 + \dots + \mathbb{R}e_p$, así que $p \leq p'$.

Además, en $W = \mathbb{R}e_{p+1} + \dots + \mathbb{R}e_n$ tenemos que $e \cdot e \leq 0$, $\forall e \in W$; luego $W \cap V_+ = 0$, y

$$n - p + \dim V_+ = \dim W + \dim V_+ = \dim (W + V_+) \leq \dim E = n.$$

Luego $\dim V_+ \leq p$; de modo que $p' \leq p$, y concluimos que $p = p'$.

Corolario 7.6 *La signatura de una métrica $S\colon E\times E\to\mathbb{R}$ es (p,q), donde p es el máximo de las dimensiones de los subespacios vectoriales de E en que S es definido-positiva, y q es el máximo de las dimensiones de los subespacios vectoriales de E en que S es definido-negativa.*

Demostración: Sabemos que p es el máximo de las dimensiones de los subespacios vectoriales en que S es definido-positiva. Además, la signatura de $-S$ es (q,p), y $-S$ es definido-positiva en los subespacios vectoriales en que S es definido-negativa.

Mecánica Relativista

La hipótesis fundamental de la Teoría de la Relatividad es que los sucesos forman un espacio de dimensión 4, y que en cada suceso las trayectorias de la luz en el vacío definen un cono, dado por los vectores donde se anula una forma cuadrática de signatura (1,3), pues en nuestro sistema de coordenadas es $\mathrm{d}t^2-\frac{1}{c^2}(\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2)$.

En general el cono variará con el lugar y momento en que se emite la luz (Teoría de la Relatividad General); pero en regiones pequeñas es razonable suponer que el cono de luz es constante (Teoría de la Relatividad Especial):

Fijado un origen del espacio-tiempo, y unidades de longitud y tiempo (y por tanto la velocidad de la luz $c>0$), los sucesos forman un **espacio-tiempo de Minkowski** (1864-1909) (un espacio vectorial real E de dimensión 4, con una métrica g de tipo $(+,-,-,-)$, llamada **métrica de Lorentz** (1853-1928).) Las trayectorias de la luz son las rectas $p+\mathbb{R}e$ de dirección generada por un vector e de **tipo luz**, i.e. isótropo, $e\cdot e:=g(e,e)=0$.

La otra hipótesis fundamental de la Teoría de la Relatividad Especial es que los sistemas de referencia inerciales son indistinguibles, que la ecuación del cono de luz ha de ser la misma en todos ellos. Un **observador inercial** o **sistema de referencia inercial** (con el mismo origen) es una base e_0,e_1,e_2,e_3 de E en que la matriz (g_{ij}) de g es[4]

$$
\boxed{\begin{array}{c}
(g_{ij})=\begin{pmatrix}1&0&0&0\\0&-\frac{1}{c^2}&0&0\\0&0&-\frac{1}{c^2}&0\\0&0&0&-\frac{1}{c^2}\end{pmatrix}\quad,\quad\begin{cases}g_{00}=1\\ g_{11}=g_{22}=g_{33}=-c^{-2}\\ g_{ij}=0\ \text{ cuando } i\neq j\end{cases}\\
e\cdot e=t^2-c^{-2}(x^2+y^2+z^2)\quad,\quad e=te_0+xe_1+ye_2+ze_3,
\end{array}}
$$

y se entiende que las correspondientes coordenadas (t,x,y,z) de un suceso son el tiempo y la posición medidos por el observador inercial, así que $\mathbb{R}e_1$, $\mathbb{R}e_2$, $\mathbb{R}e_3$ son las direcciones de los tres ejes espaciales fijados por el observador, $\mathbb{R}e_0$ es la dirección de las trayectorias de los objetos en reposo relativo para el observador y e_0 es la velocidad 4-dimensional del observador inercial: su desplazamiento espacio-temporal por unidad de tiempo.

La métrica g recibe el nombre de **métrica del tiempo** porque un reloj que se mueve inercialmente de p a $q=p+e=p+te_0$ mide un intervalo de tiempo $|t|=\sqrt{e\cdot e}$.

El intervalo de tiempo observado entre dos sucesos p, $q=p+e$, es $e_0\cdot e$, y los sucesos son simultáneos cuando $e=q-p$ es ortogonal a e_0: los **vectores espaciales** o de **simultaneidad** para un observador inercial e_0,e_1,e_2,e_3 son los vectores de $(\mathbb{R}e_0)^\perp=\mathbb{R}e_1+\mathbb{R}e_2+\mathbb{R}e_3$, y su geometría viene dada por la forma cuadrática $\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2$, que es la restricción de la forma cuadrática $\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2-c^2\mathrm{d}t^2$ asociada a la métrica $h:=-c^2g$, y los vectores e_1,e_2,e_3 forman una base ortonormal del espacio vectorial euclídeo $\big((\mathbb{R}e_0)^\perp,h\big)$. La métrica h recibe el nombre de **métrica espacial**[5] porque mide los conceptos espaciales (distancias, ángulos, etc.) entre sucesos simultáneos para un observador inercial.

[4] Además, para cualquier par de observadores inerciales $e_0,\dots$ y $e'_0,\dots$ debe cumplirse que $e_0\cdot e'_0>0$.

[5] El cono de luz es absoluto; pero su ecuación, la forma cuadrática que lo define, sólo está bien definida salvo un factor no nulo. Con todo rigor, en un espacio-tiempo de Minkowski la métrica del tiempo g no es absoluta, sólo es absoluta la recta $Q=\mathbb{R}g$ que genera en el espacio vectorial de las métricas sobre E. Fijar una métrica $g\in Q$ de tipo $(1,3)$ significa fijar la unidad de tiempo, y fijar una métrica $h\in Q$ de tipo $(3,1)$ significa fijar la unidad de longitud. En tal caso tendremos que $h=-c^2g$ para cierta constante positiva c.

Paradoja de los Gemelos: Consideremos un sistema de referencia inercial e_0, e_1, e_2, e_3. Si un viajero parte con velocidad constante ve_1 durante un tiempo t, y luego regresa con velocidad constante $-ve_1$, el intervalo de tiempo entre la partida y la llegada es $2t$.

Calculemos el tiempo que marca el reloj de ese viajero. El trayecto de ida viene dado por el vector $e = te_0 + vte_1$, y el de vuelta por $e' = te_0 - vte_1$. Como

$$\sqrt{e \cdot e} = \sqrt{e' \cdot e'} = t\sqrt{1-(v/c)^2},$$

vemos que marca $2t\sqrt{1-(v/c)^2} < 2t$, cuando $v \neq 0$. Como $\sqrt{1-(v/c)^2} \approx 1 - \frac{1}{2}(v/c)^2$ cuando $v \ll c$, su reloj se atrasa en un factor $\approx \frac{1}{2}(v/c)^2$.

Si viaja a Marte durante 1 año a 10.000 km/h y luego regresa, $v/c \approx 10^{-5}$ y el tiempo propio es $\approx 3 \cdot 10^{-3}$ segundos menor que nuestro propio tiempo.

Contracción de Longitudes: Consideremos una varilla de longitud l, en reposo en un sistema de referencia inercial e_0, e_1, e_2, e_3, y calculemos su longitud l' para un observador inercial que se mueve con velocidad v en la dirección de la varilla (digamos e_1).

Para el nuevo observador el vector espacial que determina la varilla es $e = \lambda e_0 + le_1$ para algún $\lambda \in \mathbb{R}$, y como ha de ser ortogonal a $e_0 + ve_1$

$$0 = e \cdot (e_0 + ve_1) = \lambda - vl/c^2 \ , \ \lambda = vl/c^2 \ , \ w = l\left(\frac{v}{c^2}e_0 + e_1\right)$$
$$l' = \sqrt{h(e,e)} = l\sqrt{1-(v/c)^2} = l\big(1 - \tfrac{1}{2}(v/c)^2 - \ldots\big)$$

Las longitudes se contraen en un factor $\approx \frac{1}{2}(v/c)^2$ cuando $v \ll c$.

Como el diámetro de la Tierra es de unos 12,700 km, el viajero a Marte del ejemplo anterior observa una contracción de unos 0.6 mm en la dirección e_1.

Definición: Diremos que un vector $e \in E$ es de **tipo tiempo** cuando $e \cdot e = g(e,e) > 0$, y diremos que $\tau = \sqrt{e \cdot e}$ es el **tiempo propio** entre p y $q = p + e$.

Cuando $e \cdot e < 0$, diremos que el vector e es de **tipo espacio**, y que $\sigma = \sqrt{h(e,e)}$ es la **distancia propia** entre los sucesos p y $q = p + e$.

Cuando e es de tipo tiempo, poniendo $e_0 = \frac{1}{\tau}e$, tenemos que $e_0 \cdot e_0 = 1$. Por el Teorema de Inercia, en alguna base u_1, u_2, u_3 de $(\mathbb{R}e_0)^\perp$ la matriz de g es $\mathrm{diag}(-1,-1,-1)$, y tomando $e_i = \frac{1}{c}u_i$, obtenemos un observador inercial e_0, e_1, e_2, e_3 en el que $q - p = e = \tau e_0$; es decir, p y q ocurren en la misma posición, con un intervalo de tiempo τ.

Cuando es de tipo espacio, poniendo $e_1 = \frac{1}{\sigma}e$, tenemos que $h(e_1, e_1) = 1$. Por el Teorema de Inercia, en alguna base u_0, e_2, e_3 de $(\mathbb{R}e_1)^\perp$ la matriz de h es $\mathrm{diag}(-1,1,1)$, y tomando $e_0 = cu_0$, obtenemos un observador inercial e_0, e_1, e_2, e_3 en el que $q - p = e = \sigma e_1$; es decir, p y q ocurren al mismo tiempo, a distancia σ.

Velocidad Relativa: Para un observador inercial e_0, e_1, e_2, e_3, si un móvil puntual se mueve con velocidad relativa constante $\vec{v} = v_1e_1 + v_2e_2 + v_3e_3$ y en un instante t_0 está en la posición (x_0, y_0, z_0), entonces en el instante $t_0 + t$ está en la posición $(x_0 + v_1t, y_0 + v_2t, z_0 + v_3t)$, así que su trayectoria $(t_0 + t, x_0 + v_1t, y_0 + v_2t, z_0 + v_3t)$ en E es una recta de dirección $\mathbb{R}(e_0 + \vec{v})$: *Una velocidad aparente constante $\vec{v}$ se corresponde con una trayectoria espacio-temporal de dirección* $\langle e_0 + \vec{v}\,\rangle$, y las trayectorias de los objetos en reposo son las rectas $p + \mathbb{R}e_0$ paralelas a la trayectoria del observador inercial.

Transformaciones de Lorentz: Dado un sistema de referencia inercial e_0, e_1, e_2, e_3, si otro observador inercial se mueve con velocidad aparente ve_1, su velocidad tetradimensional será $e_0' = \gamma(e_0 + ve_1)$, $\gamma \in \mathbb{R}$, y la condición $e_0' \cdot e_0' = 1$ muestra que

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{c}{\sqrt{c^2 - v^2}} \ .$$

Un vector ortogonal a e'_0 y de módulo 1 (para $h = -c^2 g$) es $e'_1 = \gamma\big(\frac{v}{c^2}e_0 + e_1\big)$, así que

$$e'_0 = \gamma(e_0 + ve_1)\ ,\ e'_1 = \gamma\big(\tfrac{v}{c^2}e_0 + e_1\big)\ ,\ e'_2 = e_2\ ,\ e'_3 = e_3$$

es un sistema de referencia inercial que se mueve con velocidad aparente v en la dirección e_1 (y mantiene los otros dos ejes espaciales). Si (t, x, y, z) son las coordenadas de un suceso en el sistema de referencia inicial, las coordenadas (t', x', y', z') del mismo suceso para el nuevo observador vienen dadas por una transformación de Lorentz:

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\frac{v}{c^2} & 0 & 0 \\ \gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix}$$

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\frac{v}{c^2} & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}$$

Velocidad Tetradimensional: Dado un observador inercial e_0, e_1, e_2, e_3, un móvil puntual ocupará en cada instante t una posición $\big(x(t), y(t), z(t)\big)$, y su velocidad aparente es

$$\vec{v} = v_1e_1 + v_2e_2 + v_3e_3 = \big(x'(t), y'(t), z'(t)\big) = x'(t)e_1 + y'(t)e_2 + z'(t)e_3,$$

La trayectoria del móvil en E es $\big(t, x(t), y(t), z(t)\big) = te_0 + x(t)e_1 + y(t)e_2 + z(t)e_3$ y

$$\big(1, x'(t), y'(t), z'(t)\big) = e_0 + x'(t)e_1 + y'(t)e_2 + z'(t)e_3 = e_0 + \vec{v}$$

es un vector tangente a la trayectoria del móvil en E, y la dirección $\langle e_0 + \vec{v}\,\rangle$ de la recta tangente a la trayectoria es independiente del observador inercial considerado. La **velocidad** 4-dimensional del móvil es su desplazamiento infinitesimal en el espacio-tiempo por unidad de tiempo; i.e., es el vector U tangente a su trayectoria en E que cumple $U \cdot U = 1$.

Luego $U = \gamma(e_0 + \vec{v})$, y la condición $U \cdot U = 1$ muestra[6] que $\gamma = \frac{c}{\sqrt{c^2 - v_1^2 - v_2^2 - v_3^2}}$.

En particular la velocidad 4-dimensional del propio observador inercial, y de los objetos en reposo relativo, es e_0.

Unidades: He aquí las unidades de las magnitudes que hemos introducido:

c	e_0	e_1	e_2	e_3	t	x	y	z	g	h	$\vec{v}$	v_i	$v = \lvert\vec{v}\,\rvert$	U	γ
$\frac{\mathrm{m}}{\mathrm{s}}$	$\frac{1}{\mathrm{s}}$	$\frac{1}{\mathrm{m}}$	$\frac{1}{\mathrm{m}}$	$\frac{1}{\mathrm{m}}$	s	m	m	m	s^2	m^2	$\frac{1}{\mathrm{s}}$	$\frac{\mathrm{m}}{\mathrm{s}}$	$\frac{\mathrm{m}}{\mathrm{s}}$	$\frac{1}{\mathrm{s}}$	$\mathrm{s}^0\mathrm{m}^0$

7.3. Cuádricas Centrales

Definición: En un espacio euclídeo real E, las **cuádricas** (centradas en el origen) son los lugares geométricos de ecuación $q(x_1, \ldots, x_n) = $ cte. para alguna forma cuadrática no nula $q\colon E \to \mathbb{R}$, y podemos suponer que la constante es 1 ó 0 (caso en que se llama **cono**).

Sea S la métrica asociada a la forma cuadrática q, y $\phi\colon E \to E^*$ la polaridad asociada a S. Como el producto escalar define un isomorfismo natural $\psi\colon E \xrightarrow{\sim} E^*$, $\psi(e) = \langle e|$, tenemos un endomorfismo $T = \psi^{-1} \circ \phi\colon E \to E$ tal que $\phi(e) = \langle T(e)|$. Es decir,

$$e \cdot v = \phi(e)(v) = \langle T(e)|v\rangle\ , \quad \forall e, v \in E,$$

[6] La solución negativa no es razonable si queremos que los valores positivos de la coordenada t expresen sucesos futuros, como suele ser costumbre. Con todo rigor, en la definición de espacio-tiempo de Minkowski, además de fijar la métrica del tiempo g, se debe fijar, en cada recta $V = \langle e\rangle$ generada por un vector de tipo tiempo, uno de los dos vectores $\pm e_0$ que cumplen $e_0 \cdot e_0 = 1$ (la 4-velocidad de las trayectorias inerciales de dirección V) de modo que $e_0 \cdot \bar{e}_0$ sea positivo para cualquier par de velocidades $e_0, \bar{e}_0$.

y este endomorfismo T es simétrico porque la métrica S es simétrica:

$$\langle T(e)|v\rangle = e \cdot v = v \cdot e = \langle T(v)|e\rangle = \langle e|T(v)\rangle.$$

Nota: Si $A = (a_{ij})$ es la matriz de la métrica S en una base ortonormal $e_1, \ldots, e_n$ de E, también es la matriz en tal base del endomorfismo T porque $\langle e_i|T(e_j)\rangle = e_i \cdot e_j = a_{ij}$.

En particular, el polinomio característico del endomorfismo T es $c_T(x) = |xI - A|$.

Teorema 7.7 *Sea q una forma cuadrática sobre un $\mathbb{R}$-espacio vectorial euclídeo E y S su métrica. En alguna base ortonormal $u_1, \ldots, u_n$ de E la matriz de S es* $\mathrm{diag}(\alpha_1, \ldots, \alpha_n)$ *y*

$$q(x_1 u_1 + \ldots + x_n u_n) = \alpha_1 x_1^2 + \ldots + \alpha_n x_n^2,$$

donde $\alpha_1, \ldots, \alpha_n$ son las raíces del polinomio característico $c_T(x)$ del endomorfismo T asociado a S y al producto escalar, repetidas tantas veces como indique su multiplicidad.

Demostración: Por el Teorema Espectral E admite una base ortonormal $u_1, \ldots, u_n$ formada por vectores propios del endomorfismo T:

$$T(u_1) = \alpha_1 u_1 \ , \ T(u_2) = \alpha_2 u_2 \ , \ \ldots \ , \ T(u_n) = \alpha_n u_n \quad ; \quad \alpha_1, \ldots, \alpha_n \in \mathbb{R}.$$

En esta base, la matriz de T (y por tanto la matriz de S ya que la base es ortonormal) es $\mathrm{diag}(\alpha_1, \ldots, \alpha_n)$, y por tanto $c_T(x) = (x - \alpha_1) \ldots (x - \alpha_n)$.

Como la matriz de S es $\mathrm{diag}(\alpha_1, \ldots, \alpha_n)$, la forma cuadrática asociada es

$$q(x_1 u_1 + \ldots + x_n u_n) = \alpha_1 x_1^2 + \ldots + \alpha_n x_n^2.$$

Corolario 7.8 *La signatura de una métrica S es justamente (r_+, r_-), donde r_+ y r_- son el número de raíces positivas y negativas, contadas con su multiplicidad, del polinomio característico del endomorfismo asociado a S y a un producto escalar.*

Regla de Descartes (1596-1650)**:** *Si un polinomio con coeficientes reales tiene todas sus raíces reales, entonces el número de raíces positivas, contadas con su multiplicidad, coincide con el número de variaciones de signo en la sucesión de coeficientes no nulos.*

Cálculo de la Signatura de una Métrica: Sea S una métrica en un espacio vectorial real E. Si A es la matriz de S en una base $e_1, \ldots, e_n$ de E, y consideramos en E el producto escalar para el que $e_1, \ldots, e_n$ es una base ortonormal, entonces A es la matriz en tal base del endomorfismo $T \colon E \to E$ asociado a la métrica S y a ese producto escalar auxiliar y, de acuerdo con 7.8, la signatura de S es justamente (r_+, r_-), donde r_+ y r_- son el número de raíces positivas y negativas del polinomio $|xI - A|$, contadas con su multiplicidad.

Ejemplo: Se llaman **ejes** de una cuádrica central $q(x_1, \ldots, x_n) = 1$ (ó 0) a las rectas generadas por algún vector propio del endomorfismo T asociado a la correspondiente métrica y al producto escalar. Por el teorema anterior, *toda cuádrica admite un sistema de ejes perpendiculares en los que su ecuación se reduce a*

$$\frac{x_1^2}{a_1^2} + \ldots + \frac{x_p^2}{a_p^2} - \frac{y_1^2}{b_1^2} - \ldots - \frac{y_q^2}{b_q^2} = 1 \text{ (ó 0)}.$$

donde $1/a_1^2, \ldots, 1/a_p^2$ y $-1/b_1^2, \ldots, -1/b_q^2$ son las raíces positivas y negativas del polinomio característico del endomorfismo asociado a la forma cuadrática y al producto escalar, repetidas tantas veces como indique su multiplicidad.

Cuando $n = 2$, las cónicas centrales reciben los siguientes nombres:

$$
\begin{array}{rll}
\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} & = 1 & \text{Elipse} \\
\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} & = 0 & \text{Par de rectas imaginarias concurrentes} \\
\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} & = 1 & \text{Hipérbola} \\
\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} & = 0 & \text{Par de rectas concurrentes} \\
-\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} & = 1 & \text{Elipse imaginaria} \\
\dfrac{x^2}{a^2} & = 1 & \text{Par de rectas paralelas} \\
\dfrac{x^2}{a^2} & = 0 & \text{Par de rectas coincidentes} \\
-\dfrac{x^2}{a^2} & = 1 & \text{Par de rectas imaginarias paralelas}
\end{array}
$$

Cuando $n = 3$, las cuádricas centrales reciben los siguientes nombres:

$$
\begin{array}{rll}
\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} & = 1 & \text{Elipsoide} \\
\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} & = 0 & \text{Cono imaginario} \\
\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} & = 1 & \text{Hiperboloide reglado} \\
\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} & = 0 & \text{Cono} \\
\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} & = 1 & \text{Hiperboloide no reglado} \\
-\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} & = 1 & \text{Elipsoide imaginario} \\
\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} & = 1 & \text{Cilindro elíptico} \\
\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} & = 0 & \text{Par de planos imaginarios concurrentes} \\
\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} & = 1 & \text{Cilindro hiperbólico} \\
\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} & = 0 & \text{Par de planos concurrentes} \\
-\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} & = 1 & \text{Cilindro imaginario} \\
\dfrac{x^2}{a^2} & = 1 & \text{Par de planos paralelos} \\
\dfrac{x^2}{a^2} & = 0 & \text{Par de planos coincidentes} \\
-\dfrac{x^2}{a^2} & = 1 & \text{Par de planos imaginarios paralelos}
\end{array}
$$

Determinación de Máximos y Mínimos locales

En cada punto crítico $p \in \mathbb{R}^n$ de una función diferenciable $f\colon \mathbb{R}^n \to \mathbb{R}$ tenemos una forma cuadrática

$$q\colon \mathbb{R}^n \longrightarrow \mathbb{R}\ , \quad q(e) = \begin{bmatrix} \text{Segunda derivada de} \\ h(t) = f(p+te) \text{ en } t=0 \end{bmatrix}$$

y la matriz, en la base usual de $\mathbb{R}^n$, de la métrica asociada es el **hessiano** de f en p:

$$\operatorname{Hess} f = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right) .$$

Consideremos la signatura (r_+, r_-) del hessiano de f en el punto p.

1. Cuando $r_+ = n$, el hessiano es definido positivo, la función presenta un mínimo en todas las direcciones, y tiene un mínimo local en el punto crítico p.

2. Cuando $r_- = n$, el hessiano es definido negativo, la función presenta un máximo en todas las direcciones, y tiene un máximo local en el punto crítico p.

3. Cuando $r_+ \geq 1$ y $r_- \geq 1$, la función tiene un mínimo en una dirección y un máximo en otra: el punto crítico p no es ni máximo ni mínimo local.

4. En los otros casos, el hessiano no decide el comportamiento local de la función.

8. Tensores

Definición: Dados K-espacios vectoriales $E_1, \ldots, E_r, F$, una aplicación K-**multilineal** de $E_1, \ldots, E_r$ en F es una aplicación $M\colon E_1 \times \ldots \times E_r \to F$ que es K-lineal en cada variable,

$$M(\ldots, \lambda e_i + \mu v_i, \ldots) = \lambda M(\ldots, e_i, \ldots) + \mu M(\ldots, v_i, \ldots) \ ; \quad \forall e_i, v_i \in E_i; \ \lambda, \mu \in K.$$

La aplicación nula $0\colon E_1 \times \ldots \times E_r \to F$ es multilineal, y si $M, \bar{M}\colon E_1 \times \ldots \times E_r \to F$ son aplicaciones K-multilineales, y $\lambda \in K$, entonces $M + \bar{M}$ y λM también son K-multilineales. Por tanto, las aplicaciones K-multilineales de $E_1, \ldots, E_r$ en F forman un subespacio vectorial del K-espacio vectorial formado por todas las aplicaciones de $E_1 \times \ldots \times E_r$ en F.

Si $M\colon E_1 \times \ldots \times E_r \to F$ es una aplicación multilineal, fijada una base $e_{i1}, \ldots, e_{in_i}$ en cada uno de los espacios vectoriales E_i, tendremos que

$$M\big(\sum_{j_1=1}^{n_1} \lambda_{1j_1} e_{1j_1}, \ldots, \sum_{j_r=1}^{n_r} \lambda_{rj_r} e_{rj_r}\big) = \sum_{j_1=1}^{n_1} \ldots \sum_{j_r=1}^{n_r} \lambda_{1j_1} \ldots \lambda_{rj_r} M(e_{1j_1}, \ldots, e_{rj_r}).$$

Teorema 8.1 *Fijada una base en cada uno de los espacios vectoriales $E_1, \ldots, E_r$, si dos aplicaciones multilineales $M, \bar{M}\colon E_1 \times \ldots \times E_r \to F$ coinciden en las sucesiones formadas con vectores de tales bases, entonces $M = \bar{M}$.*

Definiciones: Las aplicaciones multilineales $T\colon E \times \overset{p}{\ldots} \times E \times E^* \times \overset{q}{\ldots} \times E^* \longrightarrow K$ reciben el nombre de **tensores** de tipo (p, q) sobre el K-espacio vectorial E, y forman un K-espacio vectorial $T_p^q E$. Por convenio, los tensores de tipo (0,0) son los escalares, $T_0^0 E = K$.

Los tensores de tipo $(p, 0)$ se llaman **covariantes** de orden p, y los tensores de tipo $(0, q)$ se llaman **contravariantes** de orden q.

Ejemplos:

1. Los tensores covariantes de orden 1 son las formas lineales, y los tensores contravariantes de orden 1 son los vectores, $T_1^0 E = E^*$ y $T_0^1 E = E^{**} = E$, donde cada vector $e \in E$ define el tensor contravariante $e(\omega) := \omega(e)$.

2. Las métricas son los tensores covariantes de orden 2 simétricos.

3. Cada endomorfismo $f\colon E \to E$ define un (1,1)-tensor $T_f(e, \omega) := \omega\big(f(e)\big)$.

4. Si una métrica S es de rango máximo, $\operatorname{rg} S = \dim E$, entonces la polaridad asociada $\phi\colon E \to E^*$, $\phi(e)(v) = e \cdot v = S(e, v)$, es un isomorfismo, de modo que cada tensor covariante define un tensor contravariante y viceversa. En particular la propia métrica S define una **métrica dual**, que es un 2-tensor contravariante simétrico:

   $$S^*\colon E^* \times E^* \to K, \ S^*(\omega, \eta) = S(\phi^{-1}\omega, \phi^{-1}\eta).$$

   Si $A = (a_{ij})$ es la matriz de S en una base $e_1, \ldots, e_n$ de E, sabemos que A es también la matriz de la polaridad asociada $\phi\colon E \to E^*$ cuando en E^* se considera la base dual. Por tanto, si $\bar{X}, \bar{Y}$ son las coordenadas de ω y η en la base dual, tendremos que

   $$S^*(\omega, \eta) = (A^{-1}\bar{X})^t A (A^{-1}\bar{Y}) = \bar{X}^t A^{-1} \bar{Y},$$

   y vemos que la matriz de la métrica dual S^* en la base dada es justamente A^{-1}.

5. En un espacio de Minkowski E tenemos la métrica del tiempo g y la métrica espacial $h = -c^2 g$, que son tensores covariantes de orden 2 simétricos, y sus métricas duales g^* y h^*, que son tensores contravariantes de orden 2 simétricos. En un sistema de referencia inercial e_0, e_1, e_2, e_3, y su base dual $\mathrm{d}t, \mathrm{d}x, \mathrm{d}y, \mathrm{d}z$, sus matrices son

   $$g = \operatorname{diag}(1, -\frac{1}{c^2}, -\frac{1}{c^2}, -\frac{1}{c^2}) \quad , \quad h = -c^2 g = \operatorname{diag}(-c^2, 1, 1, 1)$$
   $$g^* = \operatorname{diag}(1, -c^2, -c^2, -c^2) \quad , \quad h^* = -\frac{1}{c^2} g^* = \operatorname{diag}(-\frac{1}{c^2}, 1, 1, 1)$$

6. En Mecánica Clásica, fijado un origen y unidades de longitud y tiempo, los sucesos forman un **espacio-tiempo de Galileo**: un espacio vectorial real E de dimensión 4, dotado de una forma lineal[7] no nula $\omega\colon E \to \mathbb{R}$ y de un producto escalar $\langle\ |\ \rangle$ definido en el subespacio vectorial $V := \operatorname{Ker}\omega$ (y $\dim V = 3$ por el Teorema de Isomorfía).

   Un **observador** o **sistema de referencia inercial** (con trayectoria a través del origen) es una base $e_0, \vec{e}_1, \vec{e}_2, \vec{e}_3$ de E tal que $\omega(e_0) = 1$ y $\vec{e}_1, \vec{e}_2, \vec{e}_3$ es una base ortonormal de V, y se entiende que las correspondientes coordenadas (t, x, y, z) son el tiempo y la posición medidos por el observador inercial: $\omega = \mathrm{d}t$, las direcciones de los ejes espaciales son $\mathbb{R}\vec{e}_1$, $\mathbb{R}\vec{e}_2$ y $\mathbb{R}\vec{e}_3$, la velocidad 4-dimensional del observador es e_0, y $\mathbb{R}e_0$ es la dirección de las trayectorias de los objetos en reposo relativo para el observador.

   La **velocidad** 4-dimensional de un móvil puntual de trayectoria $(t, x(t), y(t), z(t))$, el vector tangente U tal que $\omega(U) = 1$, es $U = (1, x'(t), y'(t), z'(t)) = e_0 + \vec{v}$.

   La métrica $g(e, v) := \omega(e)\omega(v)$ se llama **métrica del tiempo** porque el intervalo de tiempo entre dos sucesos p y $q = p + e$ es $\sqrt{g(e,e)} = |\omega(e)|$. En un sistema de referencia inercial, la matriz de g es $\operatorname{diag}(1, 0, 0, 0)$, y la forma cuadrática asociada es $\mathrm{d}t^2$.

   La polaridad $\phi\colon V \xrightarrow{\sim} V^*$ del producto escalar es un isomorfismo, y en V^* tenemos una métrica $S^*(\eta_1, \eta_2) = \langle \phi^{-1}(\eta_1) | \phi^{-1}(\eta_2)\rangle$. La **métrica espacial** dual $h^*\colon E^* \times E^* \to \mathbb{R}$, $h^*(\omega_1, \omega_2) := S^*(\omega_1|_V, \omega_2|_V)$ es un 2-tensor contravariante simétrico, y su matriz en un sistema de referencia inercial es $\operatorname{diag}(0, 1, 1, 1)$.

8.1. Producto Tensorial

Definición: Dado un (p, q)-tensor T y un (r, s)-tensor $\bar{T}$ sobre un mismo espacio vectorial E, su **producto tensorial** $T \otimes \bar{T}$ es el $(p + r, q + s)$-tensor

$$(T\otimes\bar{T})(e_1, \ldots, e_{p+r}, \omega_1, \ldots, \omega_{q+s}) = T(e_1, \ldots, e_p, \omega_1, \ldots, \omega_q)\cdot\bar{T}(e_{p+1}, \ldots, e_{p+r}, \omega_{q+1}, \ldots, \omega_{q+s})$$

y por convenio $\lambda \otimes T = T \otimes \lambda = \lambda T$ cuando $\lambda \in T^0_0 E = K$.

Veamos que $T \otimes \bar{T}$ es un tensor de tipo $(p + r, q + s)$. Para cualquier índice $1 \leq i \leq p$, poniendo $\underline{\omega} = (\omega_1, \ldots, \omega_q)$, $\underline{e}' = (e_{p+1}, \ldots, e_{p+r})$, y $\underline{\omega}' = (\omega_{q+1}, \ldots, \omega_{q+s})$, se cumple que

$$\begin{aligned}
&(T \otimes \bar{T})(\ldots, \lambda e_i + \mu v_i, \ldots, e_{p+r}, \omega_1, \ldots, \omega_{q+s}) = T(\ldots, \lambda e_i + \mu v_i, \ldots, e_p, \underline{\omega})\bar{T}(\underline{e}', \underline{\omega}') \\
&\quad = \big(\lambda T(\ldots, e_i, \ldots, e_p, \underline{\omega}) + \mu T(\ldots, v_i, \ldots, e_p, \underline{\omega})\big)\bar{T}(\underline{e}', \underline{\omega}') \\
&\quad = \lambda T(\ldots, e_i, \ldots, e_p, \omega_1, \ldots, \omega_q)\bar{T}(\underline{e}', \underline{\omega}') + \mu T(\ldots, v_i, \ldots, e_p, \omega_1, \ldots, \omega_q)\bar{T}(\underline{e}', \underline{\omega}') \\
&\quad = \lambda(T \otimes \bar{T})(\ldots, e_i, \ldots, e_{p+r}, \omega_1, \ldots, \omega_{q+s}) + \mu(T \otimes \bar{T})(\ldots, v_i, \ldots, e_{p+r}, \omega_1, \ldots, \omega_{q+s}).
\end{aligned}$$

y de modo análogo en cualquier otro índice.

Propiedades del Producto Tensorial:

1. $(\lambda T + \mu T') \otimes \bar{T} = \lambda(T \otimes \bar{T}) + \mu(T' \otimes \bar{T})$, $T \otimes (\lambda\bar{T} + \mu\bar{T}') = \lambda(T \otimes \bar{T}) + \mu(T \otimes \bar{T}')$.
2. $(T \otimes \bar{T}) \otimes T' = T \otimes (\bar{T} \otimes T')$.

Demostración: Pongamos $\underline{e} = (e_1, \ldots, e_p)$, $\underline{e}' = (e_{p+1}, \ldots, e_{p+r})$, $\underline{\omega} = (\omega_1, \ldots, \omega_q)$, $\underline{\omega}' = (\omega_{q+1}, \ldots, \omega_{q+s})$.

$$\begin{aligned}
&\big((\lambda T + \mu T') \otimes \bar{T}\big)(\underline{e}, \underline{e}', \underline{\omega}, \underline{\omega}') = (\lambda T + \mu T')(\underline{e}, \underline{\omega})\bar{T}(\underline{e}', \underline{\omega}') \\
&\qquad = \lambda T(\underline{e}, \underline{\omega})\bar{T}(\underline{e}', \underline{\omega}') + \mu T'(\underline{e}, \underline{\omega})\bar{T}(\underline{e}', \underline{\omega}') \\
&\qquad = \lambda(T \otimes \bar{T})(\underline{e}, \underline{\omega}, \underline{e}', \underline{\omega}') + \mu(T' \otimes \bar{T})(\underline{e}, \underline{\omega}, \underline{e}', \underline{\omega}') = \big(\lambda(T \otimes \bar{T}) + \mu(T' \otimes \bar{T})\big)(\underline{e}, \underline{e}', \underline{\omega}, \underline{\omega}')
\end{aligned}$$

[7] donde $\omega(e)$, con $e = q - p$, se interpreta como el intervalo de tiempo, con signo, transcurrido entre los sucesos p y q, y se entiende que el vector e es **espacial** o de **simultaneidad** cuando $\omega(e) = 0$.

e igualmente se prueba que $T\otimes(\lambda\bar{T}+\mu\bar{T}')=\lambda(T\otimes\bar{T})+\mu(T\otimes\bar{T}')$.

Para la propiedad 2, ponemos $\underline{e}''=(e_{p+r+1},\dots,e_{p+r+t})$, $\underline{\omega}''=(\omega_{q+s+1},\dots,\omega_{q+s+u})$.

$$\begin{aligned}\big((T\otimes\bar{T})\otimes T'\big)(\underline{e},\underline{e}',\underline{e}'',\underline{\omega},\underline{\omega}',\underline{\omega}'') &= (T\otimes\bar{T})(\underline{e},\underline{e}',\underline{\omega},\underline{\omega}')T'(\underline{e}'',\underline{\omega}'')\\ &= T(\underline{e},\underline{\omega})\bar{T}(\underline{e}',\underline{\omega}')T'(\underline{e}'',\underline{\omega}'')\\ &= T(\underline{e},\underline{\omega})(\bar{T}\otimes T')(\underline{e}',\underline{e}'',\underline{\omega}',\underline{\omega}'')=\big(T\otimes(\bar{T}\otimes T')\big)(\underline{e},\underline{e}',\underline{e}'',\underline{\omega},\underline{\omega}',\underline{\omega}'').\end{aligned}$$

Nota: Por la propiedad 2, en los productos tensoriales iterados pueden omitirse los paréntesis. Por ejemplo, el producto tensorial de p formas lineales $\omega_1,\dots,\omega_p\in E^*$ y q vectores $e_1,\dots,e_q\in E$ es un tensor $\omega_1\otimes\ldots\otimes\omega_p\otimes e_1\otimes\ldots\otimes e_q$ de tipo (p,q):

$$(\omega_1\otimes\ldots\otimes\omega_p\otimes e_1\otimes\ldots\otimes e_q)(e'_1,\dots,e'_p,\omega'_1,\dots,\omega'_q)=\omega_1(e'_1)\dots\omega_p(e'_p)\omega'_1(e_1)\dots\omega'_q(e_q).$$

Teorema 8.2 *Sea $e_1,\dots,e_n$ una base de E, y $\mathrm{d}x_1,\dots,\mathrm{d}x_n$ su base dual. Una base del espacio vectorial T^q_pE de los tensores de tipo (p,q) sobre E está formada por los tensores*

$$\mathrm{d}x_{i_1}\otimes\ldots\otimes\mathrm{d}x_{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}\ ,\quad 1\le i_1,\dots,i_p,j_1,\dots,j_q\le n,$$

y en particular, $\dim(T^q_pE)=(\dim E)^{p+q}$. *Además todo tensor T de tipo (p,q) sobre E es*

$$T=\sum_{i_1,\dots,i_p,j_1,\dots,j_q=1}^{n}T^{j_1\dots j_q}_{i_1\dots i_p}\,\mathrm{d}x_{i_1}\otimes\ldots\otimes\mathrm{d}x_{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q},$$

$$\boxed{T^{j_1\dots j_q}_{i_1\dots i_p}=T(e_{i_1},\dots,e_{i_p},\mathrm{d}x_{j_1},\dots,\mathrm{d}x_{j_q}).}$$

Demostración: Fijemos una sucesión $i_1,\dots,i_p,j_1,\dots,j_q$. Como $\mathrm{d}x_i(e_j)=\delta_{ij}$,

$$(\mathrm{d}x_{i_1}\otimes\ldots\otimes\mathrm{d}x_{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q})(e_{i_1},\dots,e_{i_p},\mathrm{d}x_{j_1},\dots,\mathrm{d}x_{j_q})=1,$$

y los restantes tensores de la familia dada se anulan en $(e_{i_1},\dots,e_{i_p},\mathrm{d}x_{j_1},\dots,\mathrm{d}x_{j_q})$.

Luego $\mathrm{d}x_{i_1}\otimes\ldots\otimes\mathrm{d}x_{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q}$ no es combinación lineal de los restantes tensores, y vemos que la familia considerada es linealmente independiente.

Para concluir, basta ver que cualquier tensor $T\in T^q_pE$ es

$$T=\sum_{i_1,\dots,i_p,j_1,\dots,j_q=1}^{n}T(e_{i_1},\dots,e_{i_p},\mathrm{d}x_{j_1},\dots,\mathrm{d}x_{j_q})\,\mathrm{d}x_{i_1}\otimes\ldots\otimes\mathrm{d}x_{i_p}\otimes e_{j_1}\otimes\ldots\otimes e_{j_q},$$

Ahora bien, ambos términos coinciden en cualquier sucesión $(e_{k_1},\dots,e_{k_p},\mathrm{d}x_{l_1},\dots,\mathrm{d}x_{l_q})$, pues valen $T(e_{k_1},\dots,e_{k_p},\mathrm{d}x_{l_1},\dots,\mathrm{d}x_{l_q})$; luego coinciden por 8.1.

Ejemplos:

1. El producto tensorial no es conmutativo; pero sí cumple que $e\otimes\omega=\omega\otimes e$, $\forall e\in E,\omega\in E^*$.

2. Los vectores son tensores contravariantes, así que las coordenadas de un vector $e\in E$ en una base $e_1,\dots,e_n$ de E deberían escribirse con superíndices, y no con subíndices; i.e. poner $e=x^1e_1+\ldots+x^ne_n$ en vez de $e=x_1e_1+\ldots+x_ne_n$.

   Toda forma lineal $\omega\in E^*$ es $\omega=u_1\mathrm{d}x_1+\ldots+u_n\mathrm{d}x_n$, con $u_i=\omega(e_i)$.

3. Sea $A=(a_{ij})$ la matriz de una métrica S en una base $e_1,\dots,e_n$ de E, y sea $\mathrm{d}x_1,\dots,\mathrm{d}x_n$ la base dual. El teorema afirma que $S=\sum_{i,j}S_{ij}\mathrm{d}x_i\otimes\mathrm{d}x_j$, con $S_{ij}=S(e_i,e_j)=a_{ij}$,

   $$S=\textstyle\sum_{ij}a_{ij}\mathrm{d}x_i\otimes\mathrm{d}x_j.$$

   Cuando $e_1,\dots,e_n$ es una base ortonormal de un espacio vectorial euclídeo real, las coordenadas g_{ij} del tensor covariante $g\colon E\times E\to\mathbb{R}$, $g(e,v)=\langle e|v\rangle$, que define el producto escalar son $g_{ij}=\delta_{ij}$. Es decir, $g_{11}=\ldots=g_{nn}=1$ y las restantes nulas:

   $$g=\mathrm{d}x_1\otimes\mathrm{d}x_1+\ldots+\mathrm{d}x_n\otimes\mathrm{d}x_n.$$

4. Fijado un sistema de referencia inercial e_0, e_1, e_2, e_3 en un espacio-tiempo de Minkowski E, y su base dual $\mathrm{d}t, \mathrm{d}x, \mathrm{d}y, \mathrm{d}z$, las métricas del tiempo y del espacio g, h, y sus métricas duales g^*, h^*, son (donde $i = 1, 2, 3$ y las coordenadas que se omiten son nulas):

$$
\begin{aligned}
g &= \mathrm{d}t \otimes \mathrm{d}t - c^{-2}(\mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y + \mathrm{d}z \otimes \mathrm{d}z \quad &;\quad g_{00} = 1,\ g_{ii} = -c^{-2},\\
g^* &= e_0 \otimes e_0 - c^2(e_1 \otimes e_1 + e_2 \otimes e_2 + e_3 \otimes e_3) \quad &;\quad g^{00} = 1,\ g^{ii} = -c^2,\\
h &= -c^2 \mathrm{d}t \otimes \mathrm{d}t + \mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y + \mathrm{d}z \otimes \mathrm{d}z \quad &;\quad h_{00} = -c^2,\ h_{ii} = 1,\\
h^* &= -c^{-2} e_0 \otimes e_0 + e_1 \otimes e_1 + e_2 \otimes e_2 + e_3 \otimes e_3 \quad &;\quad h^{00} = -c^{-2},\ h^{ii} = 1.
\end{aligned}
$$

5. Sea $e_0, \vec{e}_1, \vec{e}_2, \vec{e}_3$ un sistema de referencia inercial en un espacio-tiempo de Galileo $(E, \omega, \langle\, | \,\rangle)$, y $\omega = \mathrm{d}t, \mathrm{d}x, \mathrm{d}y, \mathrm{d}z$ su base dual. La métrica del tiempo g y la del espacio h^* son

$$
\begin{aligned}
g &= \mathrm{d}t \otimes \mathrm{d}t \quad &;\quad g_{00} = 1,\\
h^* &= \vec{e}_1 \otimes \vec{e}_1 + \vec{e}_2 \otimes \vec{e}_2 + \vec{e}_3 \otimes \vec{e}_3 \quad &;\quad h^{11} = h^{22} = h^{33} = 1
\end{aligned}
$$

y el **tensor de masa-momento** T de un fluido perfecto de densidad ρ, presión p y 4-velocidad media U es $T = \rho U \otimes U + p h^*$. Es un tensor simétrico, $T(\omega_1, \omega_2) = T(\omega_2, \omega_1)$.

6. En Relatividad, el **tensor de energía-impulso** T de un fluido perfecto de densidad ρ, presión p y 4-velocidad media U se define de modo que para fluidos en reposo aparente ($U = e_0$) para un observador inercial coincida con el newtoniano:

$$T = \rho e_0 \otimes e_0 + p(e_1 \otimes e_1 + e_2 \otimes e_2 + e_3 \otimes e_3) = (\rho + \tfrac{p}{c^2}) U \otimes U + p h^*.$$

7. Sea $A = (a_{ij})$ la matriz de un endomorfismo $f \colon E \to E$ en una base $e_1, \ldots, e_n$ de E, y sea $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$ la base dual. Como $a_{ij} = \mathrm{d}x_i(T(e_j)) = T_f(e_j, \mathrm{d}x_i)$, vemos que las coordenadas del tensor asociado T_f son $T_j^i = a_{ij}$.

   En el cálculo tensorial, al considerar la matriz A de un endomorfismo, es preferible poner $A = (a_j^i)$, donde el superíndice indica la fila y el subíndice la columna.

8. El tensor T asociado al operador lineal $|u\rangle\langle v|$ (notación de Dirac) es $T(e, \omega) = \omega(\langle v|e\rangle u) = \langle v|e\rangle \omega(u) = (u \otimes \langle v|)(e, \omega)$; es decir, es el tensor $|u\rangle \otimes \langle v|$.

9. Dados (p, q)-tensores T y $\bar{T}$ de coordenadas $T_{i_1 \ldots i_p}^{j_1 \ldots j_q}$ y $\bar{T}_{i_1 \ldots i_p}^{j_1 \ldots j_q}$ respectivamente, las coordenadas de los tensores $T + \bar{T}$ y λT son

$$(T + \bar{T})_{i_1 \ldots i_p}^{j_1 \ldots j_q} = T_{i_1 \ldots i_p}^{j_1 \ldots j_q} + \bar{T}_{i_1 \ldots i_p}^{j_1 \ldots j_q} \quad , \quad (\lambda T)_{i_1 \ldots i_p}^{j_1 \ldots j_q} = \lambda \cdot T_{i_1 \ldots i_p}^{j_1 \ldots j_q}.$$

10. Dado un (p, q)-tensor T, de coordenadas $T_{i_1 \ldots i_p}^{j_1 \ldots j_q}$, y un (r, s)-tensor $\bar{T}$, de coordenadas $\bar{T}_{i_1 \ldots i_r}^{j_1 \ldots j_s}$, las coordenadas $T_{i_1 \ldots i_{p+r}}^{j_1 \ldots j_{q+s}}$ del tensor $T \otimes \bar{T}$ son

$$\boxed{T_{i_1 \ldots i_{p+r}}^{j_1 \ldots j_{q+s}} = T_{i_1 \ldots i_p}^{j_1 \ldots j_q} \bar{T}_{i_{p+1} \ldots i_{p+r}}^{j_{q+1} \ldots j_{q+s}}.}$$

8.2. Contracción de Índices

Teorema 8.3 *Existe una única aplicación lineal $C_1^1 \colon T_p^q E \to T_{p-1}^{q-1} E$, $(p, q \geq 1)$, tal que, para cualesquiera formas lineales $\omega_1, \ldots, \omega_p \in E^*$ y vectores $v_1, \ldots, v_q \in E$ cumple que*

$$C_1^1(\omega_1 \otimes \ldots \otimes \omega_p \otimes v_1 \otimes \ldots \otimes v_q) = \omega_1(v_1)\, \omega_2 \otimes \ldots \otimes \omega_p \otimes v_2 \otimes \ldots \otimes v_q. \qquad (13)$$

Demostración: Fijemos una base $e_1, \ldots, e_n$ de E, y su base dual $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$.

De acuerdo con 3.1, existe una única aplicación lineal $C_1^1 \colon T_p^q E \to T_{p-1}^{q-1} E$ tal que,

$$C_1^1(\mathrm{d}x_{i_1} \otimes \ldots \otimes \mathrm{d}x_{i_p} \otimes e_{j_1} \otimes \ldots e_{j_q}) = \delta_{i_1 j_1} \mathrm{d}x_{i_2} \otimes \ldots \otimes \mathrm{d}x_{i_p} \otimes e_{j_2} \otimes \ldots \otimes e_{j_q},$$

es decir, que cumple 13 en las sucesiones $(\mathrm{d}x_{i_1},\dots,\mathrm{d}x_{i_p},e_{j_1},\dots,e_{j_q})$, y hemos de probar que lo cumple en cualquier sucesión $(\omega_1,\dots,\omega_p,v_1,\dots,v_q)$.

Ahora bien, los dos términos de la igualdad (13) definen sendas aplicaciones

$$M,\bar{M}\colon E^*\times \overset{p}{\dots}\times E^*\times E\times \overset{q}{\dots}\times E\longrightarrow T^{q-1}_{p-1}E,$$

que son multilineales porque el producto tensorial es lineal en cada factor y la aplicación $C_1 1$ es lineal. Como M y $\bar{M}$ coinciden en las sucesiones formadas con vectores de la base dada y su base dual, de acuerdo con 8.1 coinciden en todas las sucesiones. q.e.d.

Dado un tensor T de coordenadas $T^{j_1\dots j_q}_{i_1\dots i_p}$ en una base $e_1,\dots,e_n$, tenemos que

$$\begin{aligned}
C^1_1 T &= C^1_1\Big(\sum_{i_1,\dots,j_q=1}^{n} T^{j_1\dots j_q}_{i_1\dots i_p}\mathrm{d}x_{i_1}\otimes\dots\otimes\mathrm{d}x_{i_p}\otimes e_{j_1}\otimes\dots e_{j_q}\Big)\\
&= \sum_{i_1,\dots,j_q=1}^{n} T^{j_1\dots j_q}_{i_1\dots i_p} C^1_1(\mathrm{d}x_{i_1}\otimes\dots\otimes\mathrm{d}x_{i_p}\otimes e_{j_1}\otimes\dots e_{j_q})\\
&= \sum_{i_1,\dots,j_q=1}^{n} T^{j_1\dots j_q}_{i_1\dots i_p}\delta_{i_1 j_1}\mathrm{d}x_{i_2}\otimes\dots\otimes\mathrm{d}x_{i_p}\otimes e_{j_2}\otimes\dots e_{j_q},
\end{aligned}$$

así que las coordenadas del tensor $C^1_1 T$ son $T^{j_2\dots j_q}_{i_2\dots i_p}=\sum\limits_{i_1,j_1=1}^{n} T^{j_1\dots j_q}_{i_1\dots i_p}\,\delta_{i_1 j_1}=\sum\limits_{a=1}^{n} T^{a j_2\dots j_q}_{a i_2\dots i_p}$,

$$\boxed{T^{j_2\dots j_q}_{i_2\dots i_p}=T^{a j_2\dots j_q}_{a i_2\dots i_p}}$$

donde se adopta el **convenio de Einstein**: *hay una suma extendida a todos los posibles valores de los índices repetidos.*

Ejemplos:

1. Análogamente se define la contracción $C^j_i\colon T^q_p E\to T^{q-1}_{p-1}E$ del índice covariante i con el índice contravariante j, la contracción C^{kl}_{ij} del índice covariante i con el índice contravariante k y del índice covariante j con el índice contravariante l, etc.

   Dado un tensor T de coordenadas T^{mnp}_{ijkl}, las coordenadas del tensor $C^3_1 T$ son $T^{mn}_{jkl}=T^{mna}_{ajkl}$ y las coordenadas del tensor $C^{32}_{13}T$ son $T^{m}_{jl}=T^{mba}_{ajbl}$.

2. Con el convenio de Einstein, si $A=(a_{ij})$ es una matriz $m\times n$ y $B=(b_{ij})$ es una matriz $n\times r$, los coeficientes c_{ij} de la matriz AB son $c_{ij}=a_{ik}b_{kj}$.

3. Dada una forma lineal $\omega\in E^*$ de coordenadas u_i y un vector $e\in E$ de coordenadas x^j, las coordenadas del tensor $\omega\otimes e$ son $T^j_i=u_i x^j$, de modo que $\omega(e)=C^1_1(\omega\otimes e)=T^i_i=u_i x^i$.

4. Dada una métrica $S\colon E\times E\to K$ de matriz (a_{ij}), y vectores $e,v\in E$ de coordenadas x^i e y^j respectivamente, las coordenadas del tensor $S\otimes e\otimes v$ son $T^{kl}_{ij}=a_{ij}x^k y^l$, de modo que $C^{12}_{12}(S\otimes e\otimes v)=T^{ij}_{ij}=a_{ij}x^i y^j=S(e,v)$.

5. Sea $f\colon E\to E$ un endomorfismo de matriz $A=(a_{ij})$ y sea T_f el tensor asociado, de coordenadas $T^j_i=a_{ji}$. Se cumple que $C^1_1(T_f)=T^i_i=a_{ii}=\operatorname{tr}A$.

   Además, dado un vector $e\in E$ de coordenadas x^i, las coordenadas del tensor $T_f\otimes e$ son $T^{jk}_i=T^j_i x^k=a_{ji}x^k$, y las coordenadas del vector $C^2_1(T_f\otimes e)$ son $T^i=T^{ij}_j=a_{ij}x^j$. Vemos que $f(e)=C^2_1(T_f\otimes e)$. Igualmente se ve que $f(e)=C^1_1(e\otimes T_f)$.

Subida y Bajada de Índices: Fijada una métrica, de coordenadas g_{ij}, los índices contravariantes pueden transformarse en índices covariantes; y fijada una métrica en el dual, de coordenadas g^{ij}, los índices covariantes pueden transformarse en índices contravariantes:

$$T^{\dots\dots}_{\dots i\dots} = g_{ia}T^{\dots a\dots}_{\dots\dots} = g_{ai}T^{\dots a\dots}_{\dots\dots} \quad , \quad T^{\dots j\dots}_{\dots\dots} = g^{ja}T^{\dots\dots}_{\dots a\dots} = g^{aj}T^{\dots\dots}_{\dots a\dots} \ .$$

En particular, cada vector x^i define una forma lineal $u_i = g_{ia}x^a$, y cada forma lineal u_i define un vector $x^i = g^{ia}u_a$.

Cuando una métrica g, de coordenadas g_{ij}, es de rango máximo, $\operatorname{rg}(g_{ij}) = \dim E$, podemos bajar índices con g, y subir índices con la métrica dual g^*, cuyas coordenadas g^{ij} son los coeficientes de la matriz inversa, $(g^{ij}) = (g_{ij})^{-1}$, así que

$$\boxed{g_{ik}g^{kj} = g^{ik}g_{kj} = \delta_{ij}.}$$

Ejemplos: Fijemos una métrica, de coordenadas g_{ij} y una métrica contravariante de coordenadas g^{ij}. Dado un tensor T de tipo (2,2), de coordenadas T^{kl}_{ij}, decimos que:

1. El tensor de coordenadas $T^l_{ijk} = g_{ia}T^{al}_{jk}$ (o bien $T^l_{kij} = g_{ka}T^{al}_{ij}$) se obtiene del tensor T bajando el primer índice contravariante como primer índice covariante.

2. El tensor de coordenadas $T^l_{ijk} = g_{ka}T^{al}_{ij}$ se obtiene del tensor T bajando el primer índice contravariante como tercer índice covariante.

3. El tensor de coordenadas $T^l_{ijk} = g_{ja}T^{la}_{ik}$ (o bien $T^k_{ilj} = g_{la}T^{ka}_{ij}$) se obtiene del tensor T bajando el segundo índice contravariante como segundo índice covariante.

4. El tensor de coordenadas $T_{ijkl} = g_{ka}g_{lb}T^{ab}_{ij}$ se obtiene del tensor T bajando los dos índices contravariantes como los dos últimos índices covariante.

5. El tensor de coordenadas $T^{jkl}_i = g^{la}T^{jk}_{ia}$ (o bien $T^{klj}_i = g^{ja}T^{jk}_{ia}$) se obtiene del tensor T subiendo el segundo índice covariante como tercer índice contravariante.

6. El tensor de coordenadas $T^{ijkl} = g^{ia}g^{jb}T^{kl}_{ab}$ se obtiene del tensor T subiendo los dos índices covariantes como los dos primeros índices contravariantes.

Definición: Dado un (p,q)-tensor T, con $p \geq 1$, su **contracción interior** con un vector $e \in E$ es el $(p-1,q)$-tensor $(i_eT)(e_2,\dots,e_p,\omega_1,\dots,\omega_q) = T(e,e_2,\dots,e_p,\omega_1,\dots,\omega_q)$.

Sea $e_1,\dots,e_n$ una base de E y $\mathrm{d}x_1,\dots,\mathrm{d}x_n$ su base dual. Si las coordenadas de T y e en esta base son $T^{j_1\dots j_q}_{i_1\dots i_p}$ y x^i respectivamente, entonces las coordenadas $T^{j_1\dots j_q}_{i_2\dots i_p}$ de i_eT son

$$\begin{aligned} T^{j_1\dots j_q}_{i_2\dots i_p} &= (i_eT)(e_{i_2},\dots,e_{i_p},\mathrm{d}x_{j_1},\dots,\mathrm{d}x_{j_q}) = T(e,e_{i_2},\dots,e_{i_p},\mathrm{d}x_{j_1},\dots,\mathrm{d}x_{j_q}) \\ &= T(x^ae_a,e_{i_2},\dots,e_{i_p},\mathrm{d}x_{j_1},\dots,\mathrm{d}x_{j_q}) \\ &= x^aT(e_a,e_{i_2},\dots,e_{i_p},\mathrm{d}x_{j_1},\dots,\mathrm{d}x_{j_q}) = x^aT^{j_1\dots j_q}_{ai_2\dots i_p}. \end{aligned}$$

Ejemplos:

1. Este cálculo de las coordenadas muestra que $i_eT = C^1_1(e\otimes T)$.

2. La aplicación $E\times T^q_pE \to T^q_{p-1}E$, $(e,T)\mapsto i_eT$, es bilineal; es decir,
$$i_{\lambda e+\mu v}T = \lambda i_eT + \mu i_vT \quad , \quad i_e(\lambda T + \mu\bar{T}) = \lambda i_eT + \mu i_e\bar{T}.$$

3. Dada una métrica $S\colon E\times E\to K$ y un vector $e\in E$, la forma lineal $\omega_e\colon E\to K$, $\omega_e(v) = S(e,v)$, es precisamente $\omega_e = i_eS$.

9. Tensores Alternados

Definición: Dado un tensor covariante $T \in T_p^0 E$, para cada permutación $\sigma \in S_p$ ponemos

$$(\sigma T)(e_1, \dots, e_p) = T(e_{\sigma(1)}, \dots, e_{\sigma(p)}).$$

y esta acción del grupo S_p sobre los p-tensores covariantes tiene las siguientes propiedades:

1. *La aplicación* $\sigma\colon T_p^0 E \longrightarrow T_p^0 E$ *es lineal.*
2. $\tau(\sigma T) = (\tau\sigma)T$; $\forall \sigma, \tau \in S_p$.
3. $\sigma(\omega_1 \otimes \dots \otimes \omega_p) = \omega_{\sigma^{-1}(1)} \otimes \dots \otimes \omega_{\sigma^{-1}(p)}$; $\forall \omega_1, \dots, \omega_p \in E^*$.

Demostración: La primera se deja como ejercicio. Veamos las otras dos:

$$\begin{aligned}
\big(\tau(\sigma T)\big)(e_1, \dots, e_p) &= (\sigma T)(e_{\tau(1)}, \dots, e_{\tau(p)}) = T(e_{\tau(\sigma 1)}, \dots, e_{\tau(\sigma p)}) \\
&= T(e_{(\tau\sigma)(1)}, \dots, e_{(\tau\sigma)(p)}) = \big((\tau\sigma)T\big)(e_1, \dots, e_p), \\
(\sigma(\omega_1 \otimes \dots \otimes \omega_p))(e_1, \dots, e_p) &= \omega_1(e_{\sigma(1)}) \cdot \dots \cdot \omega_p(e_{\sigma(p)}) \\
&= \omega_{\sigma^{-1}(1)}(e_1) \cdot \dots \cdot \omega_{\sigma^{-1}(p)}(e_p) = (\omega_{\sigma^{-1}(1)} \otimes \dots \otimes \omega_{\sigma^{-1}(p)})(e_1, \dots, e_p).
\end{aligned}$$

Definición: Un p-tensor covariante S es **simétrico** cuando $\sigma S = S$ para toda permutación $\sigma \in S_p$; y un p-tensor covariante Ω es **hemisimétrico** o **alternado**, (es una **p-forma lineal**) cuando $\sigma\Omega = (\operatorname{sgn}\sigma)\Omega$ para toda permutación $\sigma \in S_p$:

$$\Omega(e_{\sigma(1)}, \dots, e_{\sigma(p)}) = (\operatorname{sgn}\sigma)\Omega(e_1, \dots, e_p),$$

y convenimos que las 1-formas lineales son las formas lineales $\omega \in E^*$.

Las p-formas lineales forman un subespacio vectorial de $T_p^0 E$.

Ejemplo: En el caso de un 2-tensor covariante Ω, la condición de que sea alternado significa que $\Omega(e, v) = -\Omega(v, e)$, $\forall e, v \in E$, y en particular $\Omega(e, e) = 0$, $\forall e \in E$.

De hecho, esta última condición equivale a la de ser alternado, pues implica que

$$0 = \Omega(e + v, e + v) = \Omega(e, e) + \Omega(e, v) + \Omega(v, e) + \Omega(v, v) = \Omega(e, v) + \Omega(v, e).$$

Si Ω_{ij} son las coordenadas de Ω en una base de E, la condición de ser alternado significa que $\Omega_{ij} = -\Omega_{ji}$, y en particular $\Omega_{ij} = 0$ cuando $i = j$.

Ejemplo: En un espacio-tiempo de Minkowski E las fuerzas electromagnéticas vienen representadas por un endomorfismo $\tilde{F}\colon E \to E$, donde se interpreta que para un observador inercial de 4-velocidad e_0, la fuerza que actúa sobre la unidad de carga en reposo es $\tilde{F}(e_0)$. Consideremos el 2-tensor covariante $F(e, v) := h(\tilde{F}(e), v)$, donde h es la métrica espacial. La fuerza que mide un observador ha de ser un vector espacial, así que $F(e_0, e_0) = 0$, y es natural suponer que *el tensor F del campo electromagnético es una 2-forma lineal.*

Fijado un observador inercial e_0, e_1, e_2, e_3, el vector $E := \tilde{F}(e_0) = E_1 e_1 + E_2 e_2 + E_3 e_3$ recibe el nombre de **campo eléctrico**, de modo que

$$E_i = h(E, e_i) = h\big(\tilde{F}(e_0), e_i\big) = F(e_0, e_i) = F_{0i} = -F_{i0}.$$

Además $F_{00} = F_{11} = F_{22} = F_{33} = 0$, y las restante coordenadas F_{ij} de F se denotan

$$F_{12} = -F_{21} = -\tfrac{1}{c}B_3\ ,\ F_{13} = -F_{31} = \tfrac{1}{c}B_2\ ,\ F_{23} = -F_{32} = -\tfrac{1}{c}B_1\ ,$$

y se dice que el vector espacial $B := B_1 e_1 + B_2 e_2 + B_3 e_3$ es el **campo magnético** que mide el observador inercial considerado.

Consideremos las coordenadas F_i^j del tensor asociado a $\tilde{F}$, de modo que $\tilde{F}(e_i) = F_i^a e_a$.

Como $F(e,v) = h(\tilde{F}(e),v)$, las coordenadas F_{ij} del tensor F son

$$F_{ij} = F(e_i,e_j) = h\big(\tilde{F}(e_i),e_j\big) = h\big(F_i^a e_a,e_j\big) = F_i^a h(e_a,e_j) = F_i^a h_{aj},$$

así que el tensor F se obtiene al bajar, con la métrica h, el índice contravariante del tensor asociado a $\tilde{F}$. Veamos que el tensor F_i^j se obtiene al subir con la métrica dual h^{ij} el segundo índice covariante del tensor F_{ij}:

$$F_{ia}h^{aj} = F_i^b h_{ba}h^{aj} = F_i^b \delta_{bj} = F_i^j,$$

porque $h_{ik}h^{kj} = \delta_{ij}$. Si $A = (a_{ij})$ es la matriz de $\tilde{F}$, tenemos que $a_{ij} = F_j^i = F_{ja}h^{ai}$:

$$A = \begin{pmatrix} 0 & \frac{1}{c^2}E_1 & \frac{1}{c^2}E_2 & \frac{1}{c^2}E_3 \\ E_1 & 0 & \frac{1}{c}B_3 & -\frac{1}{c}B_2 \\ E_2 & -\frac{1}{c}B_3 & 0 & \frac{1}{c}B_1 \\ E_3 & \frac{1}{c}B_2 & -\frac{1}{c}B_1 & 0 \end{pmatrix} \;,\; |\lambda I - A| = \lambda^4 - \tfrac{1}{c^2}(\|E\|^2 - \|B\|^2)\lambda^2 - \tfrac{1}{c^4}\langle E|B\rangle^2,$$

y vemos que los siguientes escalares no dependen del observador inercial fijado, son invariantes escalares del campo electromagnético:

$$\begin{aligned} \|E\|^2 - \|B\|^2 &= h(E,E) - h(B,B) = E_1^2 + E_2^2 + E_3^2 - B_1^2 - B_2^2 - B_3^2, \\ \langle E|B\rangle^2 &= h(E,B) = (E_1B_1 + E_2B_2 + E_3B_3)^2. \end{aligned}$$

Definición: La **hemisimetrización** de un p-tensor covariante $T \in T_p^0E$ es el tensor

$$h(T) = \sum_{\sigma \in S_p} (\operatorname{sgn}\sigma)(\sigma T) \in T_p^0 E,$$

que ya es alternado, porque para toda permutación $\tau \in S_p$ se cumple que

$$\tau\Big(\sum_{\sigma\in S_p}(\operatorname{sgn}\sigma)(\sigma T)\Big) = (\operatorname{sgn}\tau)\sum_{\sigma\in S_p}\operatorname{sgn}(\tau\sigma)(\tau\sigma T) = (\operatorname{sgn}\tau)h(T).$$

Ejemplos:

1. Si ω es una forma lineal, tenemos que $h(\omega) = \omega$.
2. Si $\omega_1,\omega_2 \in E^*$, y $\sigma = (12)$, tenemos que $\sigma(\omega_1 \otimes \omega_2) = \omega_2 \otimes \omega_1$. Luego $h(\omega_1 \otimes \omega_2) = \omega_1 \otimes \omega_2 - \omega_2 \otimes \omega_1$.
3. Tenemos que $h(\Omega) = p!\,\Omega$, para toda p-forma lineal Ω, porque $\sigma\Omega = (\operatorname{sgn}\sigma)\Omega$, $\forall \sigma \in S_p$, de modo que $h(\Omega) = \sum_{\sigma\in S_p}(\operatorname{sgn}\sigma)(\sigma\Omega) = \sum_{\sigma\in S_p}(\operatorname{sgn}\sigma)^2\Omega = p!\,\Omega$.

Lema 9.1 *Sea H un subgrupo de un grupo G, y $a,b \in G$. Si $b \in aH$, entonces $aH = bH$. Por tanto, los conjuntos aH y bH o son disjuntos o coinciden,*

Demostración: Si $b \in aH$, entonces $b = ah$ para algún $h \in H$, y $bH = ahH \subseteq aH$.

Como $a = bh^{-1} \in bH$, también tenemos que $aH \subseteq bH$. Luego $aH = bH$.

Por último, si $c \in (aH) \cap (bH)$, tenemos que $aH = cH = bH$.

Lema 9.2 *La aplicación $h\colon T_p^0 E \longrightarrow T_p^0 E$ es lineal, su imagen está formada por las p-formas lineales, y si $h(T) = 0$, entonces para todo tensor $\bar{T} \in T_q^0 E$ se cumple que*

$$h(T \otimes \bar{T}) = h(\bar{T} \otimes T) = 0.$$

Demostración: Las aplicaciones $\sigma\colon T_p^0 E \to T_p^0 E$ son lineales; luego también $h = \sum_\sigma (\operatorname{sgn}\sigma)\sigma$.

Además para toda p-forma lineal Ω tenemos que $\Omega = h\big(\frac{1}{p!}\Omega\big) \in \operatorname{Im} h$.

Finalmente, identifiquemos cada permutación $\sigma \in S_p$ con una permutación $\sigma \in S_{p+q}$ que deja fijos los números $p+1, \dots, p+q$. Se cumple que

$$\sum_{\sigma\in S_p} (\operatorname{sgn}\sigma)\sigma(T\otimes\bar{T}) = h(T)\otimes\bar{T} = 0,$$
$$\sum_{\sigma\in S_p} (\operatorname{sgn}\tau\sigma)\big(\tau\sigma(T\otimes\bar{T})\big) = 0, \text{ para toda } \tau \in S_{p+q},$$

y vemos que en $h(T\otimes\bar{T})$ la suma correspondiente a los elementos de τS_p es nula.

Como S_p es un subgrupo del grupo S_{p+q}, por el lema S_{p+q} puede descomponerse en unión disjunta $S_{p+q} = \tau_1 S_p \cup \ldots \cup \tau_r S_p$, y vemos que $h(T\otimes\bar{T}) = 0$.

Igualmente se prueba que $h(\bar{T}\otimes T) = 0$.

9.1. Producto Exterior

Definición: El **producto exterior** $\Omega\wedge\bar{\Omega}$ de una p-forma lineal Ω con una q-forma lineal $\bar{\Omega}$ es la $(p+q)$-forma lineal

$$\Omega\wedge\bar{\Omega} = h(T)\wedge h(\bar{T}) := h(T\otimes\bar{T}),$$

y no depende de los tensores T y $\bar{T}$ elegidos: si $\Omega = h(T')$, entonces $T' = T+S$ con $h(S) = 0$, y por el lema anterior $h(T'\otimes\bar{T}) = h(T\otimes\bar{T} + S\otimes\bar{T}) = h(T\otimes\bar{T})$.

Cuando $p = 0$ (ó $q = 0$), convenimos que $\lambda\wedge\Omega = \Omega\wedge\lambda = \lambda\Omega$.

El subespacio vectorial de $T_p^0 E$ formado por las p-formas lineales se denota $\Lambda^p E^*$.

Ejemplos:

1. Dadas formas lineales $\omega, \omega' \in E^*$, tenemos que
$$\omega\wedge\omega' = h(\omega)\wedge h(\omega') = h(\omega\otimes\omega') = \omega\otimes\omega' - \omega'\otimes\omega.$$

2. El producto exterior de p formas lineales $\omega_1, \dots, \omega_p \in E^*$ es:
$$\omega_1\wedge\ldots\wedge\omega_p = h(\omega_1)\wedge\ldots\wedge h(\omega_p) = h(\omega_1\otimes\ldots\otimes\omega_p) = \sum_{\sigma\in S_p} (\operatorname{sgn}\sigma)\sigma(\omega_1\otimes\ldots\otimes\omega_p)$$
$$(\omega_1\wedge\ldots\wedge\omega_p)(e_1,\dots,e_p) = \sum_{\sigma\in S_p} (\operatorname{sgn}\sigma)\omega_1(e_{\sigma(1)})\ldots\omega_p(e_{\sigma(p)}) = \begin{vmatrix} \omega_1(e_1) & \ldots & \omega_1(e_p) \\ \ldots & \ldots & \ldots \\ \omega_p(e_1) & \ldots & \omega_p(e_p) \end{vmatrix}.$$

3. Sea $e_1, \dots, e_n$ una base de E, y $\mathrm{d}x_1, \dots, \mathrm{d}x_n$ la base dual. Si A es la matriz cuyas columnas son las coordenadas de ciertos vectores $v_1, \dots, v_n \in E$ en la base dada, se cumple que
$$(\mathrm{d}x_1\wedge\ldots\wedge\mathrm{d}x_n)(v_1,\dots,v_n) = |A|.$$
En particular $(\mathrm{d}x_1\wedge\ldots\wedge\mathrm{d}x_n)(e_1,\dots,e_n) = 1$.

Propiedades del Producto Exterior:

1. *Es bilineal*: $(\Omega+\Omega')\wedge\bar{\Omega} = \Omega\wedge\bar{\Omega} + \Omega'\wedge\bar{\Omega} \quad , \quad (\lambda\Omega)\wedge\bar{\Omega} = \lambda(\Omega\wedge\bar{\Omega})$.
$\Omega\wedge(\bar{\Omega}+\bar{\Omega}') = \Omega\wedge\bar{\Omega} + \Omega\wedge\bar{\Omega}' \quad , \quad \Omega\wedge(\lambda\bar{\Omega}) = \lambda(\Omega\wedge\bar{\Omega})$.

2. *Asociativo*: $(\Omega\wedge\bar{\Omega})\wedge\Omega' = \Omega\wedge(\bar{\Omega}\wedge\Omega')$.

3. $\omega\wedge\omega' = -\omega'\wedge\omega$, *y por tanto* $\omega\wedge\omega = 0$, $\forall\omega,\omega' \in E^*$.

4. *Sea* $e_1, \ldots, e_n$ *una base de* E, *y* $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$ *su base dual. Las* p*-formas*

$$\mathrm{d}x_{i_1} \wedge \ldots \wedge \mathrm{d}x_{i_p} \ , \quad 1 \leq i_1 < \ldots < i_p \leq n,$$

definen una base de $\Lambda^p E^*$, *y para toda* p*-forma* Ω *se cumple que*

$$\Omega = \sum_{1 \leq i_1 < \ldots < i_p \leq n} \Omega_{i_1 \ldots i_p} \, \mathrm{d}x_{i_1} \wedge \ldots \wedge \mathrm{d}x_{i_p} \ , \quad \Omega_{i_1 \ldots i_p} := \Omega(e_{i_1}, \ldots, e_{i_p}).$$

Por tanto $\dim \Lambda^p E^* = \binom{n}{p}$, *y en particular* $\dim \Lambda^n E^* = 1$ *y* $\Lambda^p E^* = 0$ *cuando* $p > n$.

5. *Anticonmutativo*: $\Omega_p \wedge \Omega_q = (-1)^{pq} \Omega_q \wedge \Omega_p, \ \ \forall \Omega_p \in \Lambda^p E^*, \, \Omega_q \in \Lambda^q E^*$.

6. $i_e(\Omega_p \wedge \Omega_q) = (i_e \Omega_p) \wedge \Omega_q + (-1)^p \Omega_p \wedge (i_e \Omega_q), \ \ \forall \Omega_p \in \Lambda^p E^*, \, \Omega_q \in \Lambda^q E^*$.

Demostración: Las propiedades (1) y (2) se siguen de la definición, de la linealidad de h y de las correspondientes propiedades del producto tensorial.
En cuanto a (3), tenemos que $\omega \wedge \omega' = \omega \otimes \omega' - \omega' \otimes \omega = -\omega' \wedge \omega$.

(4) Como la aplicación lineal $h \colon T_p^0 E \to \Lambda^p E^*$ es epiyectiva, y los tensores $\mathrm{d}x_{i_1} \otimes \ldots \otimes \mathrm{d}x_{i_p}$ generan $T_p^0 E$, se sigue que las p-formas $h(\mathrm{d}x_{i_1} \otimes \ldots \otimes \mathrm{d}x_{i_p}) = \mathrm{d}x_{i_1} \wedge \ldots \wedge \mathrm{d}x_{i_p}$ generan $\Lambda^p E^*$.
Como $\mathrm{d}x_{i_1} \wedge \ldots \wedge \mathrm{d}x_{i_p} = 0$ cuando algún índice está repetido, y

$$\mathrm{d}x_{i_{\sigma(1)}} \wedge \ldots \wedge \mathrm{d}x_{i_{\sigma(p)}} = \pm \mathrm{d}x_{i_1} \wedge \ldots \wedge \mathrm{d}x_{i_p}$$

para toda permutación σ, vemos que, dada una p-forma Ω, hay escalares $\Omega_{i_1 \ldots i_p}$ tales que

$$\Omega = \sum_{1 \leq i_1 < \ldots < i_p \leq n} \Omega_{i_1 \ldots i_p} \, \mathrm{d}x_{i_1} \wedge \ldots \wedge \mathrm{d}x_{i_p}.$$

Ahora, fijada una sucesión de índices $1 \leq j_1 < \ldots < j_p \leq n$, como se cumple que $(\mathrm{d}x_{i_1} \wedge \ldots \wedge \mathrm{d}x_{i_p})(e_{j_1}, \ldots, e_{j_p}) = 0$, salvo cuando la sucesión $i_1 < \ldots < i_p$ coincide con la sucesión $j_1 < \ldots < j_p$, en cuyo caso el valor es 1, vemos que

$$\Omega_{j_1 \ldots j_p} = \Omega(e_{j_1}, \ldots, e_{j_p}).$$

Cuando $\Omega = 0$, concluimos que las p-formas $\mathrm{d}x_{i_1} \wedge \ldots \wedge \mathrm{d}x_{i_p}$, $1 \leq i_1 < \ldots < i_p \leq n$, son linealmente independientes.

(5) El caso $p = q = 1$ es (3). El caso $\Omega_p = \omega_1 \wedge \ldots \wedge \omega_p$, $\Omega_q = \theta_1 \wedge \ldots \wedge \theta_q$ se sigue directamente, y el caso general se sigue de la bilinealidad del producto exterior.

(6) Si $e = 0$, es ambos términos son nulos.
Si $e \neq 0$, fijemos una base $e = e_1, \ldots, e_n$ y su base dual $\omega_1, \ldots, \omega_n$.
Podemos suponer que $\Omega_p = \omega_{i_1} \wedge \ldots \wedge \omega_{i_p} = \omega_{i_1} \wedge \omega_I$, y $\Omega_q = \omega_{j_1} \wedge \ldots \wedge \omega_{j_q} = \omega_{j_1} \wedge \omega_J$.
Cuando $i_1 > 1$ y $j_1 > 1$, ambos términos son nulos. Cuando $i_1 = 1$ y $j_1 > 1$,

$$\begin{aligned} i_e(\Omega_p \wedge \Omega_q) &= i_e(\omega_1 \wedge \omega_I \wedge \Omega_q) = \omega_I \wedge \Omega_q, \\ (i_e \Omega_p) \wedge \Omega_q + (-1)^p \Omega_p \wedge (i_e \Omega_q) &= \omega_I \wedge \Omega_q + (-1)^p \Omega_p \wedge 0 = \omega_I \wedge \Omega_q, \end{aligned}$$

y el caso $i_1 > 1$, $j_1 = 1$ es similar. Cuando $i_1 = 1$ y $j_1 = 1$,

$$\begin{aligned} i_e(\Omega_p \wedge \Omega_q) &= i_e(\omega_1 \wedge \omega_I \wedge \omega_1 \wedge \omega_J) = i_e(0) = 0, \\ (i_e \Omega_p) \wedge \Omega_q + (-1)^p \Omega_p \wedge (i_e \Omega_q) &= \omega_I \wedge \Omega_q + (-1)^p \omega_1 \wedge \omega_I \wedge \omega_J \\ &= \omega_I \wedge \Omega_q + (-1)^p (-1)^{p-1} \omega_I \wedge \omega_1 \wedge \omega_J = \omega_I \wedge \Omega_q - \omega_I \wedge \Omega_q = 0. \end{aligned}$$

Ejemplo: Fijado un sistema de referencia inercial e_0, e_1, e_2, e_3 en un espacio-tiempo de Minkowski, y su base dual $\mathrm{d}t, \mathrm{d}x, \mathrm{d}y, \mathrm{d}z$, el tensor F del campo electromagnético es

$$F = E_1 \mathrm{d}t \wedge \mathrm{d}x + E_2 \mathrm{d}t \wedge \mathrm{d}y + E_3 \mathrm{d}t \wedge \mathrm{d}z - \tfrac{1}{c}(B_3 \mathrm{d}x \wedge \mathrm{d}y - B_2 \mathrm{d}x \wedge \mathrm{d}z + B_1 \mathrm{d}y \wedge \mathrm{d}z).$$

Nota: Sólo hemos estudiado tensores alternados covariantes, porque el estudio de los tensores contravariantes alternados se reduce inmediatamente a ellos, pues los tensores contravariantes sobre un espacio vectorial E son los tensores covariantes sobre su dual E^*.

Si $e_1, \ldots, e_n$ es una base de E, los p-vectores $e_{i_1} \wedge \ldots \wedge e_{i_p}$, $1 \leq i_1 < \ldots < i_p \leq n$, definen una base del espacio vectorial $\Lambda^p E$ formado por los tensores contravariantes alternados de orden p sobre E. Por otra parte, también podemos considerar los tensores contravariantes de orden p simétricos, que forman un subespacio vectorial $S^p E$ de $T_0^p E$, y tenemos un operador de **simetrización** $s\colon T_0^p E \to S^p E$, $s(T) = \sum_{\sigma \in S_p} \sigma T$.

En Mecánica Cuántica, si E es el espacio de estados de cierto tipo de partícula, entonces los posibles estados de un sistema formado por p partículas de dicho tipo vienen dados por tensores contravariantes de orden p, entendiendo que $e_1 \otimes e_2 \otimes \ldots \otimes e_p$ representa el estado del sistema cuando el estado de la primera partícula es $\mathbb{C}e_1$, el de la segunda es $\mathbb{C}e_2$, etc.

El caso es que las partículas elementales se dividen en bosones y fermiones. En los bosones el espacio de estados es $S^p E$, y en los fermiones el espacio de estados es $\Lambda^p E$. Fijada una base $\psi_1, \ldots, \psi_n$ de E, los textos de Mecánica Cuántica escriben

$$|a_1, \ldots, a_n\rangle = \begin{cases} s(\psi_1 \otimes \overset{a_1}{\ldots} \otimes \psi_1 \otimes \ldots \otimes \psi_n \otimes \overset{a_n}{\ldots} \otimes \psi_n) \in S^p E & \text{en los bosones} \\ h(\psi_1 \otimes \overset{a_1}{\ldots} \otimes \psi_1 \otimes \ldots \otimes \psi_n \otimes \overset{a_n}{\ldots} \otimes \psi_n) \in \Lambda^p E & \text{en los fermiones} \end{cases}$$

donde los números naturales $a_1, \ldots, a_n$ suman p y se llaman números de ocupación, pues a_i indica el número de partículas que están en el estado ψ_i; aunque, al ser tensores simétricos o hemisimétricos, nunca pueda decirse qué partícula está en el estado ψ_i (salvo cuando $a_i = p$, en el caso de los bosones). En el caso de los bosones, los números a_i son arbitrarios, con la única condición de que $a_1 + \ldots + a_n = p$; pero en el caso de los fermiones los números a_i sólo pueden tomar los valores 0 y 1 (Principio de Exclusión de Pauli) porque

$$|a_1, \ldots, a_n\rangle = h(\psi_1 \otimes \overset{a_1}{\ldots} \otimes \psi_1 \ldots \psi_n \otimes \overset{a_n}{\ldots} \otimes \psi_n) = \psi_1 \wedge \overset{a_1}{\ldots} \wedge \psi_1 \wedge \ldots \wedge \psi_n \wedge \overset{a_n}{\ldots} \wedge \psi_n \neq 0.$$

9.2. Formas de Volumen y Orientaciones

Teorema 9.3 *Sea E un $\mathbb{R}$-espacio vectorial euclídeo E de dimensión n. Si una n-forma Ω cumple que $|\Omega(u_1, \ldots, u_n)| = 1$ en alguna base ortonormal $u_1, \ldots, u_n$ de E, entonces $|\Omega(v_1, \ldots, v_n)| = 1$ en toda base ortonormal $v_1, \ldots, v_n$ de E.*

Demostración: Consideremos la base dual $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$ de $u_1, \ldots, u_n$.

Tenemos que $\Omega = \Omega(u_1, \ldots, u_n)\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n = \pm\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n$.

Si $B = (b_{ij})$ es la matriz de cambio de base, $v_j = \sum_i b_{ij} u_i$, tenemos que

$$(\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n)(v_1, \ldots, v_n) = \det\bigl(\mathrm{d}x_i(v_j)\bigr) = \det(b_{ij}) = |B|.$$

La matriz del producto escalar en la nueva base es $B^t I B$ (porque $K = \mathbb{R}$); luego $B^t B = I$ porque la base $v_1, \ldots, v_n$ es ortonormal, así que $|B|^2 = |B^t B| = 1$, y $|B| = \pm 1$.

Definición: Sea E un $\mathbb{R}$-espacio vectorial euclídeo de dimensión n. Las **formas de volumen** son las n-formas $\pm\Omega_E$ que cumplan $|\Omega_E(u_1, \ldots, u_n)| = 1$ en alguna (luego en toda) base ortonormal $u_1, \ldots, u_n$ de E, y diremos que $|\Omega_E(e_1, \ldots, e_n)|$ es el **volumen** del paralelepípedo (área del paralelogramo, cuando $n = 2$) determinado por los vectores $e_1, \ldots, e_n \in E$.

Si $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$ la base dual de una base ortonormal, las formas de volumen son

$$\boxed{\Omega_E = \pm\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n}$$

y el volumen del paralelepípedo que determinan unos vectores $e_1, \ldots, e_n$ es el valor absoluto del determinante de la matriz cuyas columnas son las coordenadas de $e_1, \ldots, e_n$ en alguna base ortonormal de E (luego nulo si y sólo si $e_1, \ldots, e_n$ son linealmente dependientes).

Ejemplo: En un plano euclídeo real E, dado un paralelogramo de lados e, v, podemos fijar una base ortonormal u_1, u_2 tal que $e = bu_1$, $v = au_1 + hu_2$, donde b es la base y h es la altura, y vemos que el área $S = |\Omega_E(e,v)|$ del paralelogramo es la base por la altura:

$$S = |\Omega_E(e,v)| = |\Omega_E(bu_1, au_1 + hu_2)| = |ba\Omega_E(u_1,u_1) + bh\Omega_E(u_1,u_2)| = bh,$$

y el área de un triángulo[8] de vértices de coordenadas (x_0,y_0), (x_1,y_1), (x_2,y_2) en una base ortonormal es

$$\frac{S}{2} = \pm\frac{1}{2}\begin{vmatrix} x_1 - x_0 & x_2 - x_0 \\ y_1 - y_0 & y_2 - y_0 \end{vmatrix}.$$

Proposición 9.4 *Sea $T\colon E \to E$ un endomorfismo de un $\mathbb{R}$-espacio vectorial euclídeo E de dimensión n, y sea $|A|$ el valor absoluto del determinante de la matriz de T en cualquier base de E. Se cumple que*

$$\begin{bmatrix}\text{Volumen del paralelepípedo} \\ \text{de aristas } T(e_1), \ldots, T(e_n)\end{bmatrix} = |A| \begin{bmatrix}\text{Vol. del paralelepípedo} \\ \text{de aristas } e_1, \ldots, e_n\end{bmatrix}.$$

Demostración: Si $e_1, \ldots, e_n$ son linealmente dependientes, $\sum \lambda_i e_i = 0$, también tendremos que $0 = T(\sum_i \lambda_i e_i) = \sum_i \lambda_i T(e_i)$, y ambos términos son nulos.

Si $e_1, \ldots, e_n$ forman una base de E, consideramos la base dual $\omega_1, \ldots, \omega_n$, y tendremos que $\Omega_E = \lambda\omega_1 \wedge \ldots \wedge \omega_n$, donde $\lambda = \Omega_E(e_1, \ldots, e_n)$. Consideramos también la matriz $A = (a_{ij})$ de T en tal base, $T(e_j) = \sum_i a_{ij} e_i$, donde $a_{ij} = \omega_i(T(e_j))$, y terminamos:

$$\Omega_E\big(T(e_1), \ldots, T(e_n)\big) = \lambda(\omega_1 \wedge \ldots \wedge \omega_n)\big(T(e_1), \ldots, T(e_n)\big) = \lambda \det\big(\omega_i(Te_j)\big) = \lambda \det(a_{ij}).$$

Proposición 9.5 *Sea $e_1, \ldots, e_n$ una base de E, y $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$ la base dual. Si (g_{ij}) denota la matriz del producto escalar en esta base, $g_{ij} = e_i \cdot e_j$, las formas de volumen de E son*

$$\Omega_E = \pm\sqrt{|g_{ij}|}\, \mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n.$$

Demostración: Tenemos que $\Omega_E = \pm\lambda \mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n$, con $\lambda = \pm\Omega_E(e_1, \ldots, e_n) = \pm|B|$, donde B es la matriz de cambio de base de una base ortonormal $u_1, \ldots, u_n$ a la base $e_1, \ldots, e_n$.

Como $(g_{ij}) = B^t I B$, concluimos que $|g_{ij}| = |B|^2$.

Ejemplo: En un plano, el área S del paralelogramo que determinan dos vectores e, v es

$$S = \sqrt{\begin{vmatrix} e\cdot e & e \cdot v \\ v \cdot e & v \cdot v \end{vmatrix}} = \sqrt{(e\cdot e)(v\cdot v) - (e\cdot v)^2}.$$

Como $a^2 = e \cdot e$, $b^2 = v \cdot v$, $c^2 = (e - v)\cdot(e - v) = a^2 + b^2 - 2(e\cdot v)$, donde a, b, c son las longitudes de los lados del triángulo que determinan e y v, obtenemos la **fórmula de Herón** (s. I d.C.) para el área $S' = \frac{1}{2}S$ de un triángulo:

$$\begin{aligned} S' &= \tfrac{1}{2}\sqrt{a^2b^2 - \tfrac{1}{4}(a^2 + b^2 - c^2)^2} = \tfrac{1}{4}\sqrt{2(a^2b^2 + a^2c^2 + b^2c^2) - a^4 - b^4 - c^4} \\ &= \tfrac{1}{4}\sqrt{(a+b+c)(a+b-c)(a-b+c)(-a+b+c)} \\ &= \sqrt{s(s-a)(s-b)(s-c)}\ , \qquad s = \tfrac{1}{2}(a+b+c). \end{aligned}$$

Definiciones: Dar una **orientación** en un espacio vectorial euclídeo real E es dar una (de las dos posibles) forma de volumen Ω_E, y diremos que una base $e_1, \ldots, e_n$ es **directa** o que está bien orientada cuando $\Omega_E(e_1, \ldots, e_n) > 0$.

[8]que se define como la mitad del área del paralelogramo; y en dimensión 3, el volumen del tetraedro que determinan tres vectores e, v, w es $1/6$ del volumen $\Omega_E(e,v,w)$ del paralelepípedo que determinan e, v, w.

Sea Ω_E la forma de volumen de un espacio vectorial euclídeo real orientado E de dimensión 3. Si $e, v \in E$, entonces $i_v(i_e\Omega_E)$ es una forma lineal, y de acuerdo con 6.4 existe un único vector $e \times v \in E$, llamado **producto vectorial** de e y v, tal que

$$(e \times v) \cdot w = \Omega_E(e, v, w) \ , \quad \forall w \in E.$$

Sea u_1, u_2, u_3 una base ortonormal directa de E, y pongamos $e = \sum_i x_i u_i$, $v = \sum_i y_i u_i$. Las coordenadas de $e \times v$ son $(e \times v) \cdot u_i = \Omega_E(e, v, u_i)$, $i = 1, 2, 3$; luego

$$e \times v = \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} u_1 - \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} u_2 + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} u_3 = \begin{vmatrix} x_1 & y_1 & u_1 \\ x_2 & y_2 & u_2 \\ x_3 & y_3 & u_3 \end{vmatrix}.$$

En particular $u_1 \times u_2 = u_3$, $u_2 \times u_3 = u_1$, $u_3 \times u_1 = u_2$.

Propiedades del Producto Vectorial:

1. *Es bilineal*: $(\lambda e + \mu e') \times v = \lambda(e \times v) + \mu(e' \times v)$,
   $e \times (\lambda v + \mu v') = \lambda(e \times v) + \mu(e \times v')$.

2. *Es anticonmutativo*: $e \times v = -v \times e$, $e \times e = 0$.

3. *El producto vectorial $e \times v$ es ortogonal a ambos factores, y $e \times v \neq 0$ si y sólo si los vectores e y v son linealmente independientes. En tal caso su módulo es el área del paralelogramo determinado por e y v, y la base $e, v, e \times v$ es directa.*

Demostración: Para ver que dos vectores de E coinciden, basta ver que tienen el mismo producto escalar con cualquier vector $w \in E$. Ahora bien,

$$\big((\lambda e + \mu e') \times v\big) \cdot w = \Omega_E(\lambda e + \mu e', v, w) = \lambda\Omega_E(e, v, w) + \mu\Omega_E(e', v, w),$$
$$\big(\lambda(e \times v) + \mu(e' \times v)\big) \cdot w = \lambda(e \times v) \cdot w + \mu(e' \times v) \cdot w = \lambda\Omega_E(e, v, w) + \mu\Omega_E(e', v, w).$$
$$(e \times v) \cdot w = \Omega_E(e, v, w) = -\Omega_E(v, e, w) = -(v \times e) \cdot w,$$

y la linealidad por la derecha se sigue de la linealidad por la izquierda y el carácter alternado.

Veamos también que $e \times v$ es ortogonal a ambos factores:

$$(e \times v) \cdot e = \Omega_E(e, v, e) = 0 \quad , \quad (e \times v) \cdot v = \Omega_E(e, v, v) = 0.$$

Si e y v son linealmente dependientes, digamos $v = \lambda e$, entonces $e \times v = \lambda e \times e = 0$.

Si e y v son linealmente independientes, los ampliamos hasta obtener una base e, v, w de E. Como $(e \times v) \cdot w = \Omega_E(e, v, w) \neq 0$, vemos que $e \times v \neq 0$.

Además, si tomamos un vector u de módulo 1 ortogonal al plano $\mathbb{R}e + \mathbb{R}v$ tendremos que $i_u\Omega_E$ es una formas de área de tal plano: Si u_1, u_2 es una base ortonormal de $\mathbb{R}e + \mathbb{R}v$, entonces u, u_1, u_2 es una base ortonormal en E, así que $(i_u\Omega_E)(u_1, u_2) = \Omega_E(u, u_1, u_2) = \pm 1$.

Ahora $e \times v = \lambda u$, donde $|\lambda| = \|e \times v\|$; luego

$$\begin{bmatrix} \text{Área del paralelo–} \\ \text{gramo de aristas } e, v \end{bmatrix} = |(i_u\Omega_E)(e, v)| = |\Omega_E(u, e, v)| = |(e \times v) \cdot u| = |\lambda u \cdot u| = \|e \times v\|.$$

Por último, la base $e, v, e \times v$ es directa porque $\Omega_E(e, v, e \times v) = (e \times v) \cdot (e \times v) > 0$.

9.3. Determinantes

Teorema 9.6 *Unos vectores $e_1, \ldots, e_p \in E$ son linealmente independientes si y sólo si existe alguna p-forma lineal Ω tal que $\Omega(e_1, \ldots, e_p) \neq 0$.*

En particular, fijada una n-forma no nula Ω en un espacio vectorial E de dimensión n, unos vectores $e_1, \ldots, e_n$ forman una base de E si y sólo si $\Omega(e_1, \ldots, e_n) \neq 0$.

Demostración: Si son linealmente independientes, los ampliamos hasta obtener una base $e_1, \ldots, e_p, \ldots, e_n$ de E. Sea $\mathrm{d}x_1, \ldots, \mathrm{d}x_p, \ldots, \mathrm{d}x_n$ la base dual. Tenemos que

$$(\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_p)(e_1, \ldots, e_p) = \det\big(\mathrm{d}x_i(e_j)\big) = \det(\delta_{ij}) = 1 \neq 0.$$

Si son linealmente dependientes, alguno (digamos e_1) es combinación lineal de los restantes, $e_1 = \lambda_2 e_2 + \ldots + \lambda_p e_p$, y para toda p-forma Ω tenemos que

$$\Omega(e_1, e_2, \ldots, e_p) = \lambda_2 \Omega(e_2, e_2, \ldots, e_p) + \ldots + \lambda_p \Omega(e_p, e_2, \ldots, e_p) = 0.$$

Nota: Fijada una base $e_1, \ldots, e_n$ en un espacio vectorial E, y su base dual $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$, como las formas $\mathrm{d}x_{i_1} \wedge \ldots \wedge \mathrm{d}x_{i_p}$, $1 \leq i_1 < \ldots < i_p \leq n$, definen una base de las p-formas, vemos que p vectores $v_j = \sum_i a_{ij} e_i$, $j = 1, \ldots, p$ son linealmente independientes si y sólo si para alguna sucesión $1 \leq i_1 < \ldots < i_p \leq n$ se cumple que

$$0 \neq (\mathrm{d}x_{i_1} \wedge \ldots \wedge \mathrm{d}x_{i_p})(v_1, \ldots, v_p) = \begin{vmatrix} \mathrm{d}x_{i_1}(v_1) & \ldots & \mathrm{d}x_{i_1}(v_p) \\ \ldots & \ldots & \ldots \\ \mathrm{d}x_{i_p}(v_1) & \ldots & \mathrm{d}x_{i_p}(v_p) \end{vmatrix} = \begin{vmatrix} a_{i_1 1} & \ldots & a_{i_1 p} \\ \ldots & \ldots & \ldots \\ a_{i_p 1} & \ldots & a_{i_p p} \end{vmatrix},$$

es decir, que la matriz (a_{ij}) tiene algún menor de orden p no nulo.

Vemos así que el teorema anterior es una expresión intrínseca del Teorema del Rango.

Definición: Sea E un espacio vectorial de dimensión n. Como $\dim \Lambda^n E^* = \binom{n}{n} = 1$, todo endomorfismo $\Lambda^n E^* \to \Lambda^n E^*$ es la multiplicación por cierto escalar. Ahora bien, cada endomorfismo $T \colon E \to E$ induce un endomorfismo

$$T^* \colon \Lambda^n E^* \to \Lambda^n E^* \ , \ (T^*\Omega)(e_1, \ldots, e_n) = \Omega(T(e_1), \ldots, T(e_n)),$$

que ha de ser la multiplicación por un escalar $\det(T)$, llamado **determinante** de T,

$$\Omega\big(T(e_1), \ldots, T(e_n)\big) = (T^*\Omega)(e_1, \ldots, e_n) = (\det T)\, \Omega(e_1, \ldots, e_n) \ , \quad \Omega \in \Lambda^n E^*. \qquad (14)$$

Proposición 9.7 *El determinante de un endomorfismo $T \colon E \to E$ coincide con el determinante de su matriz $A = (a_{ij})$ en cualquier base,* $\det(T) = |A|$.

Demostración: Sea $e_1, \ldots, e_n$ una base de E, y $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$ su base dual. La matriz (a_{ij}) de T cumple que $a_{ij} = \mathrm{d}x_i(T(e_j))$, y como $(\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n)(e_1, \ldots, e_n) = 1$, concluimos:

$$\begin{aligned} \det T &= (\det T)(\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n)(e_1, \ldots, e_n) = \big(T^*(\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n)\big)(e_1, \ldots, e_n) \\ &= (\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n)\big(T(e_1), \ldots, T(e_n)\big) = |\mathrm{d}x_i(T(e_j))| = |a_{ij}|. \end{aligned}$$

Teorema 9.8 $\det(T \circ S) = (\det T)(\det S)$.

Demostración: $(TS)^* = S^* T^*$.

Teorema 9.9 *Un endomorfismo T es un isomorfismo si y sólo si* $\det(T) \neq 0$.

Demostración: Fijada una base $e_1, \ldots, e_n$ en E, y una n-forma $0 \neq \Omega \in \Lambda^n E^*$, se cumple

$$\Omega\big(T(e_1), \ldots, T(e_n)\big) = (\det T)\, \Omega(e_1, \ldots, e_n),$$

de acuerdo con 14. Por 9.6, tenemos que $\det(T) \neq 0$ si y sólo si $T(e_1), \ldots, T(e_n)$ es una base de E, lo que equivale a que T sea un isomorfismo.

Nota: El teorema 9.8 afirma que el determinante de un producto de matrices es el producto de sus determinantes, y 9.9 que las matrices invertibles son las de determinante no nulo.

Índice alfabético