Simetrías y Grupos en Física

Notas de curso

Artemio González López

Madrid, enero de 2014

Editor:
Artemio González López

Departamento de Física Teórica II
Facultad de Ciencias Físicas
Avenida Complutense s/n
Ciudad Universitaria
28040 Madrid

Impreso en Madrid

Índice general

1 Grupos Finitos **1**

1.1 Definición y propiedades elementales . . . 1
1.1.1 Ejemplos . . . 2
1.1.2 El grupo simétrico . . . 5
1.2 Subgrupos. Cosets. Teorema de Lagrange . . . 6
1.2.1 Cosets. Teorema de Lagrange . . . 8
1.3 Homomorfismos. Teorema de Cayley . . . 10
1.3.1 Teorema de Cayley . . . 13
1.4 Clases de conjugación. Subgrupos normales. Grupo cociente. Producto directo . . . 15
1.4.1 Clases de conjugación . . . 15
1.4.2 Subgrupos normales . . . 17
1.4.3 Grupo cociente . . . 20
1.4.4 Producto directo de dos subgrupos . . . 23
1.4.5 Producto directo de dos grupos . . . 25
1.5 Representaciones . . . 26
1.5.1 Representaciones lineales . . . 26
1.5.2 Representación regular . . . 28
1.5.3 Equivalencia de representaciones . . . 32
1.5.4 Representación compleja conjugada. Representaciones reales . . . 33
1.5.5 Suma y producto directos de representaciones . . . 34
1.5.6 Representaciones irreducibles . . . 38
1.5.7 Teoremas de Schur–Auerbach y de Maschke . . . 41
1.6 Lemas de Schur . . . 43
1.7 Relaciones de ortogonalidad y completitud . . . 48
1.8 Caracteres . . . 52

2 Grupos y Álgebras de Lie **59**

2.1 Espacios topológicos . . . 59
2.1.1 Preliminares . . . 59
2.1.2 Continuidad. Homeomorfismos . . . 61
2.1.3 Compacidad . . . 62
2.1.4 Conexión . . . 63
2.1.5 Conexión por arcos . . . 65
2.2 Variedades topológicas y diferenciables . . . 66
2.2.1 Variedades topológicas . . . 66
2.2.2 Variedades diferenciables . . . 68
2.2.3 Subvariedades regulares . . . 70
2.2.4 Funciones diferenciables . . . 71
2.3 Grupos topológicos y de Lie . . . 73
2.3.1 Definiciones y ejemplos . . . 73

2.3.2 Los grupos matriciales clásicos 76
2.3.3 Grupos matriciales cerrados 82
2.4 El álgebra de Lie de un grupo matricial cerrado 85
2.4.1 Espacio tangente a una subvariedad regular 85
2.4.2 Álgebras de Lie 87
2.5 Subgrupos a un parámetro 88
2.5.1 Álgebras de Lie de los grupos matriciales clásicos 92

Capítulo 1

Grupos Finitos

1.1 Definición y propiedades elementales

Definición 1.1. Un **grupo** es un conjunto G provisto de una ley de composición interna (**producto**)

$$\begin{aligned} G \times G &\to G \\ (g,h) &\mapsto gh \end{aligned}$$

que goza de las siguientes propiedades:

1. *Asociatividad:*
$$g(hk) = (gh)k\,, \qquad \forall g,h,k \in G\,.$$
2. *Elemento neutro:*
$$\exists e \in G \text{ t.q. } eg = ge = g\,, \qquad \forall g \in G\,.$$
3. *Existencia de inverso:*
$$\forall g \in G,\ \exists g^{-1} \in G \text{ t.q. } g^{-1}g = gg^{-1} = e\,.$$

Comentarios

- El elemento neutro se denomina también **unidad**.
- En la definición de grupo *no* se exige que el producto sea *conmutativo*, es decir que
$$gh = hg\,, \qquad \forall g,h \in G\,.$$
Si se verifica la propiedad anterior, se dice que G es **abeliano** (o **conmutativo)**.
- En realidad, los axiomas 2) y 3) anteriores son ligeramente *redundantes*. Más concretamente, se puede probar (ejercicio) que dichos axiomas pueden reemplazarse por los siguientes:
$$2'.\quad \exists e \in G \text{ t.q. } ge = g\,,\ \forall g \in G \qquad \text{(elemento neutro } \textit{por la derecha}\text{)}$$
$$3'.\quad \forall g \in G,\ \exists g^{-1} \in G \text{ t.q. } gg^{-1} = e \qquad \text{(inverso } \textit{por la derecha}\text{)}\,.$$
- A partir de la asociatividad del producto (primer axioma), es inmediato probar la siguiente propiedad más general: el producto $g_1 \cdots g_n$ de n elementos de un grupo G depende solo del orden en que figuran dichos elementos, y no de la forma concreta en que se efectúen los productos indicados. Por ejemplo:
$$(g_1g_2)(g_3g_4) = g_1((g_2g_3)g_4) = ((g_1g_2)g_3)g_4 = \cdots \equiv g_1g_2g_3g_4\,.$$

• De la definición de grupo se sigue inmediatamente que *el elemento neutro es único.* También es fácil probar que *el inverso de cualquier elemento $g \in G$ es único.* En efecto, si g_1 y g_2 son ambos inversos de $g \in G$, multiplicando por la derecha ambos miembros de la igualdad

$$g_1 g = g_2 g \, (= e)$$

por g_1 (o g_2) se sigue inmediatamente que $g_1 = g_2$.

• Utilizando los axiomas de grupo, se demuestran fácilmente las siguientes identidades:

$$\begin{cases} e^{-1} = e \\ (g_1 \cdots g_n)^{-1} = g_n^{-1} \cdots g_1^{-1} \, . \end{cases}$$

Notación. Si k es un entero no negativo y $g \in G$ escribiremos

$$g^k = \overbrace{g \cdots g}^{k \text{ factores}} \, , \qquad g^{-k} = \overbrace{g^{-1} \cdots g^{-1}}^{k \text{ factores}} \, ,$$

siendo $g^0 \equiv e$. Nótese que esta notación no es ambigua, debido a la asociatividad del producto de un número arbitrario de elementos de un grupo comentada anteriormente. Con la notación anterior, para todo $i, j \in \mathbb{Z}$ se verifica

$$g^i g^j = g^{i+j} \, .$$

Definición 1.2. Un grupo G es **finito** si G es un conjunto finito. El **orden** de un grupo finito G, que denotaremos por $|G|$, es el número de sus elementos.

Si $G = \{g_1, \ldots, g_n\}$ es un grupo finito, su **tabla de multiplicación** es la matriz cuyo elemento de matriz (i, j) es el producto $g_i g_j$. En la práctica, dicha tabla de multiplicación se suele representar de la forma siguiente:

	e	a	b	$\cdots$	g
e	e	a	b	$\cdots$	g
a	a	a^2	ab	$\cdots$	ag
b	b	ba	b^2	$\cdots$	bg
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
g	g	ga	gb	$\cdots$	g^2

Evidentemente, la fila o la columna de la tabla de multiplicación asociada al elemento neutro e (que generalmente es la primera) consta de los elementos $g_1, \ldots, g_n$ escritos en ese mismo orden. En general, cada fila o columna de la tabla de multiplicación está formada por una cierta *permutación* $g_{i_1}, \ldots, g_{i_n}$ de dichos elementos. En otras palabras, si g es un elemento cualquiera de G entonces los n productos gg_k o $g_k g$ (con $k = 1, \ldots, n$) son todos distintos. En efecto, si (por ejemplo) $gg_j = gg_k$ con $j \neq k$ multiplicando a la izquierda por g^{-1} llegaríamos a la contradicción de que $g_j = g_k$.

1.1.1 Ejemplos

Ejemplo 1.3. Evidentemente, el grupo más sencillo es el grupo trivial $G = \{e\}$ (único grupo de orden 1). Si $G = \{e, a\}$ es un grupo de orden 2, el único producto de elementos de G que no es evidente a priori es a^2, que en principio podría ser igual a e ó a. Sin embargo, de las igualdades $a^2 = ae = a$ se seguiría que a no posee inverso, en contradicción con los axiomas. Luego $a^2 = e$, y la tabla de multiplicación de G es como sigue:

	e	a
e	e	a
a	a	e

Por tanto hay esencialmente[1] un único grupo de orden 2, denominado *grupo cíclico de orden 2*, que se suele denotar por C_2. Dicho grupo es claramente abeliano. En particular, una realización muy sencilla de C_2 es el conjunto $\mathbb{Z}_2 = \{1, -1\}$ con la multiplicación ordinaria como producto ($e = 1$, $a = -1$).

En general, se define el **grupo cíclico de orden** n como

$$C_n = \left\{e, a, \ldots, a^{n-1}\right\},$$

con el producto[2]

$$a^i a^j = a^{i+j \bmod n}.$$

Es fácil ver que con este producto el conjunto C_n es un grupo abeliano. En particular, en C_n se verifican las igualdades

$$a^n = e\,, \qquad (a^i)^{-1} = a^{n-i}\,.$$

Definición 1.4. Se dice que un grupo G está **generado** por un subconjunto $A \subset G$ si todo elemento de G es el producto de un número finito de elementos de A.

Si denotamos

$$A^k = \left\{a_1 \cdots a_k \;\middle|\; a_i \in A\,,\ 1 \leqslant i \leqslant k\right\},$$

entonces

$$G \text{ generado por } A \iff G = \bigcup_{k=1}^{\infty} A^k\,.$$

De la definición anterior se sigue inmediatamente que C_n está generado por el elemento a. De hecho, es fácil probar que C_n puede caracterizarse como el grupo de orden n generado por un único elemento a (ejercicio). Nótese también que una realización concreta del grupo abstracto C_n es el conjunto de las *raíces n-ésimas de la unidad*

$$\left\{\mathrm{e}^{\frac{2k\pi \mathrm{i}}{n}} \;\middle|\; k = 0, 1, \ldots, n-1\right\},$$

tomando como producto el producto ordinario en $\mathbb{C}$ ($a = \mathrm{e}^{\frac{2\pi \mathrm{i}}{n}}$). Otra realización de C_n es el grupo de las rotaciones de ángulo $2k\pi/n$ ($k = 0, 1, \ldots, n-1$) alrededor de un eje cualquiera, siendo a en este caso la rotación de ángulo $2\pi/n$. □

Ejercicio 1. Si G es un grupo finito, probar que para todo $g \in G$ existe un entero $k \geqslant 1$ tal que $g^k = e$. Al menor entero que satisface esta igualdad se le denomina **orden** de g, y se le denota por $o(g)$. Demostrar que si $o(g) = m$ entonces $\{e, g, \ldots, g^{m-1}\} = C_m$.

Solución. Consideremos la serie de elementos de G

$$e, \quad g, \quad g^2, \quad \ldots \quad , g^k, \quad g^{k+1}, \quad \ldots\,.$$

Como G es finito, estos elementos no pueden ser todos distintos. Sea, por tanto, m el entero más pequeño tal que $g^m = g^j$ para $0 \leqslant j < m$. Al ser $g^{m-j} = e$, de la definición de m se sigue que $j = 0$, i.e., $g^m = e$. Además, de nuevo por definición de m, las potencias g^k con $k = 0, \ldots, m-1$ han de ser todas distintas, y por tanto

$$\{e, g, \ldots, g^{m-1}\}$$

es un conjunto de m elementos. El producto en este conjunto es el mismo que en C_m, es decir

$$g^i g^j = g^{i+j \bmod m}\,,$$

al ser $g^m = e$. □

[1] Con más precisión, todo grupo finito de orden 2 es *isomorfo* a C_2, donde el concepto de isomorfismo de grupos se define rigurosamente en la Sección 1.3.

[2] Recuérdese que si $k, n \in \mathbb{N}$ se denota por $k \bmod n$ el resto de dividir k por n. En particular, $k \bmod n \in \{0, 1, \ldots, n-1\}$.

Ejemplo 1.5. Los conjuntos $\mathbb{Z}^N$, $\mathbb{Q}^N$, $\mathbb{R}^N$ y $\mathbb{C}^N$, tomando como producto de dos elementos su *suma* (ordinaria), son claramente grupos abelianos (infinitos). En este caso, por tanto

$$a \cdot b = a + b\,, \qquad e = 0\,, \qquad a^{-1} = -a\,.$$

Nótese, sin embargo, que $(\mathbb{N}^N, +)$ *no* es grupo para ningún $N > 0$, ya que ningún elemento distinto de cero posee inverso.

Ejemplo 1.6. Los conjuntos $\mathbb{F}^* \equiv \mathbb{F} \setminus \{0\}$, donde $\mathbb{F} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$, son grupos abelianos (infinitos) con el producto habitual. (En general, si $\mathbb{F}$ es un *cuerpo* el conjunto $\mathbb{F}^*$ es un grupo abeliano.) Evidentemente, el conjunto $\mathbb{Z}^*$ (con el producto ordinario) *no* es un grupo, ya que los únicos elementos que poseen inverso son ± 1.

Ejemplo 1.7. La *circunferencia unidad* $S^1 = \left\{z \in \mathbb{C} \;\middle|\; |z| = 1\right\}$ con el producto ordinario en $\mathbb{C}$ es un grupo abeliano infinito.

Ejemplo 1.8. Sea C un conjunto arbitrario (finito o infinito). El conjunto

$$S(C) = \left\{f : C \to C \;\middle|\; f \text{ biyectiva}\right\}$$

es un grupo definiendo el producto de dos elementos $f, g \in S(C)$ como su *composición* $f \circ g$, dada por

$$(f \circ g)(x) = f\big(g(x)\big)\,.$$

En este caso la unidad es la identidad $I : C \to C$, definida por $I(x) = x$ para todo $x \in C$, y f^{-1} es la aplicación inversa de f ($f^{-1}(x) = y \iff x = f(y)$), que existe al ser f biyectiva. Nótese que $S(C)$ es finito si y solo si lo es C.

Ejemplo 1.9. Sea V un espacio vectorial (de dimensión finita o infinita, real o complejo), y definamos

$$\mathrm{GL}(V) = \left\{f : V \to V \;\middle|\; f \text{ lineal e invertible}\right\}.$$

Si se define de nuevo el producto de $f, g \in \mathrm{GL}(V)$ mediante $fg = f \circ g$, es fácil ver que $\mathrm{GL}(V)$ es un grupo (infinito, en general no abeliano) denominado **grupo general lineal** sobre V. Estrechamente relacionados con los grupos $\mathrm{GL}(\mathbb{F}^N)$ (con $\mathbb{F} = \mathbb{R}, \mathbb{C}$) están los *grupos matriciales*

$$\mathrm{GL}(N, \mathbb{F}) = \left\{A \in M_N(\mathbb{F}) \;\middle|\; A \text{ invertible}\right\},$$

donde $M_N(\mathbb{F})$ denota el conjunto de las matrices cuadradas de orden N con elementos de matriz en $\mathbb{F}$. De hecho, veremos en la sección siguiente que $\mathrm{GL}(\mathbb{F}^N)$ y $\mathrm{GL}(N, \mathbb{F})$ son *isomorfos*.

Ejemplo 1.10. Una *traslación* en $\mathbb{R}^N$ es una aplicación $\tau_a : \mathbb{R}^N \to \mathbb{R}^N$ de la forma

$$\tau_a(x) = x + a\,,$$

donde $a \in \mathbb{R}^N$ es un vector fijo. El conjunto

$$T_N = \left\{\tau_a \;\middle|\; a \in \mathbb{R}^N\right\},$$

con el producto igual a la composición, es un grupo abeliano infinito. En efecto, si $a, b \in \mathbb{R}^N$ se tiene

$$(\tau_a \tau_b)(x) = \tau_a(x + b) = x + a + b = \tau_{a+b}(x) \qquad \Longrightarrow \qquad \tau_a \tau_b = \tau_{a+b}\,.$$

En particular, el elemento neutro es la aplicación identidad τ_0, y $\tau_a^{-1} = \tau_{-a}$. Veremos más adelante que T_N es isomorfo al grupo (aditivo) $\mathbb{R}^N$.

1.1.2 El grupo simétrico

Una **permutación** de orden n es una aplicación biyectiva $\sigma : \{1,\dots,n\} \to \{1,\dots,n\}$. Si $\sigma(i) = \sigma_i$, denotaremos la permutación σ mediante el símbolo

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma_1 & \sigma_2 & \dots & \sigma_n \end{pmatrix}.$$

Denotaremos por S_n al conjunto de todas las permutaciones de orden n. Dados dos elementos $\sigma, \rho \in S_n$, se define su producto $\sigma\rho$ como la composición $\sigma \circ \rho$, que evidentemente es una nueva permutación de orden n (la composición de aplicaciones biyectivas es biyectiva). Por ejemplo,

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}\{1,2,3,4\} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}\{2,1,4,3\} = \{2,4,3,1\}\,,$$

y por tanto

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}.$$

Con esta definición el elemento neutro es la permutación identidad

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix},$$

y σ^{-1} es la permutación inversa de σ, es decir

$$\sigma^{-1} = \begin{pmatrix} \sigma_1 & \sigma_2 & \dots & \sigma_n \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

Por ejemplo,

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 4 & 2 & 1 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}.$$

El conjunto S_n con el producto definido anteriormente es claramente un grupo, llamado **grupo simétrico** (o *grupo de las permutaciones*) de n objetos[3]. Se trata de un grupo finito de orden $n!$, no abeliano para $n > 2$. (Evidentemente, el grupo S_2 coincide con el grupo cíclico de orden 2, C_2, y es por tanto abeliano.)

Dados $k > 1$ enteros distintos $i_1,\dots,i_k \in \{1,\dots,n\}$, se define el **ciclo** $(i_1 \cdots i_k) \in S_n$ como la permutación σ dada por

$$\sigma(i_1) = i_2\,,\ \ \sigma(i_2) = i_3\,,\ \ \dots\ \ ,\sigma(i_{k-1}) = i_k\,,\ \ \sigma(i_k) = i_1\,;\quad \sigma(i) = i\,,\ \ \forall i \notin \{i_1,\dots,i_k\}.$$

Es evidente que $(i_1 \cdots i_k) = (i_2\, i_3 \cdots i_k\, i_1) = \cdots = (i_k\, i_1\, i_2 \cdots i_{k-1})$, y que

$$(i_1 \cdots i_k)^{-1} = (i_k\, i_{k-1} \cdots i_2\, i_1)\,.$$

Dos ciclos $(i_1 \cdots i_k)$, $(j_1 \cdots j_l)$ son *disjuntos* si $\{i_1,\dots,i_k\} \cap \{j_1,\dots,j_l\} = \emptyset$. Evidentemente, dos ciclos disjuntos conmutan entre sí, ya que cada uno de ellos actúa sobre un subconjunto distinto de $\{1,\dots,n\}$. Por ejemplo, en S_4 se tiene

$$(143) = (431) = (314) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} = (12)(34) = (34)(12)\,.$$

En el caso especial $k = 1$, definimos el 1-ciclo (i_1) como la aplicación identidad $i_1 \mapsto i_1$. Con esta notación, por ejemplo,

$$e = (1)(2)\cdots(n)\,, \qquad (135) = (135)(2)(4) \in S_5\,.$$

Nótese que, excepto en el caso de la unidad, los ciclos de longitud 1 se suelen suprimir al expresar una permutación como producto de ciclos.

[3] Nótese que $S_n = S\big(\{1,\dots,n\}\big)$; cf. el Ejemplo 1.8.

Proposición 1.11. *Toda permutación $\sigma \in S_n$ se puede expresar como un producto de ciclos disjuntos.*

Demostración. En efecto, tomemos $m_1 \in \{1, \ldots, n\}$ y sea $p \geqslant 1$ el entero más grande tal que los números

$$m_1\,, \quad m_2 = \sigma(m_1)\,, \quad m_3 = \sigma(m_2)\,, \quad \ldots \quad , m_p = \sigma(m_{p-1})\,,$$

son todos distintos. Es fácil ver entonces que

$$\sigma(m_p) = m_1\,,$$

ya que si fuera $\sigma(m_p) = m_j$ con $2 \leqslant j \leqslant p$ se tendría

$$\sigma(m_p) = \sigma^p(m_1) = m_j = \sigma^{j-1}(m_1) \qquad \Longrightarrow \qquad \sigma^{p-j+1}(m_1) = m_{p-j+2} = m_1\,,$$

en contra de la definición de p (al ser $p - j + 2 \leqslant p$). Por tanto

$$\sigma = (m_1 \cdots m_p)\sigma'\,,$$

donde σ' es una permutación de los restantes $n - p < n$ números $\{1, \ldots, n\} \setminus \{m_1, \ldots, m_p\}$ (al ser σ biyectiva). Aplicando iterativamente este argumento se obtiene fácilmente el enunciado. □

• Es evidente que la descomposición de una permutación en ciclos disjuntos es *única salvo por el orden de los ciclos*, ya que los elementos que forman cada ciclo son subconjuntos de $\{1, \ldots, n\}$ *invariantes* bajo dicha permutación que a su vez no tienen subconjuntos propios invariantes. En particular, las *longitudes* ($\equiv$ número de elementos) de los ciclos que figuran en dicha descomposición, incluidas las de los ciclos de longitud 1, son características de la permutación.

Ejercicio 2. Probar que el orden de una permutación σ es el mínimo común múltiplo de las longitudes de los ciclos en que se descompone σ.

Ejemplo 1.12. Consideremos el grupo simétrico de 3 objetos S_3, cuyos elementos son las $3! = 6$ permutaciones

$$e, \quad (12), \quad (13), \quad (23), \quad (123), \quad (321) = (123)^{-1}\,.$$

La tabla de multiplicación de S_3 se calcula fácilmente, y resulta ser la siguiente:

	e	(12)	(13)	(23)	(123)	(321)
e	e	(12)	(13)	(23)	(123)	(321)
(12)	(12)	e	(321)	(123)	(23)	(13)
(13)	(13)	(123)	e	(321)	(12)	(23)
(23)	(23)	(321)	(123)	e	(13)	(12)
(123)	(123)	(13)	(23)	(12)	(321)	e
(321)	(321)	(23)	(12)	(13)	e	(123)

De esta tabla se sigue inmediatamente que S_3 no es abeliano, ya que dicha tabla no es simétrica respecto de la diagonal principal.

1.2 Subgrupos. Cosets. Teorema de Lagrange

Definición 1.13. Un **subgrupo** H de un grupo G es un subconjunto $H \subset G$ que es, a su vez, un grupo con el producto heredado de G.

En otras palabras, $H \subset G$ es un subgrupo de un grupo G si:

1. $\forall h, h' \in H, \;\; hh' \in H$
2. $\forall h \in H, \;\; h^{-1} \in H$

En efecto, de las dos propiedades anteriores se sigue que si $h \in H$ entonces $hh^{-1} = e \in H$. En cuanto a la asociatividad del producto, se cumple automáticamente al verificarse en G.

• Nótese que todo subgrupo ha de contener necesariamente la unidad.

• Si H_1 y H_2 son subgrupos de un grupo G, su *intersección* $H_1 \cap H_2$ es a su vez subgrupo de G. Sin embargo, la *unión* de dos subgrupos no es necesariamente un subgrupo. (Considérese, por ejemplo, el caso en que $G = S_3$, $H_1 = \{e, (12)\}$, $H_2 = \{e, (13)\}$.)

Proposición 1.14. *Un subconjunto H de un grupo G es subgrupo si es* cerrado bajo la división*, es decir si $\forall h, h' \in H$, $h'h^{-1} \in H$ (un resultado análogo es válido para $h^{-1}h'$).*

Demostración. Claramente, si H es subgrupo ha de ser cerrado bajo la división. Recíprocamente, si h es cerrado bajo la división tomando $h' = h$ en el enunciado se deduce que $e \in H$. Tomando a continuación $h' = e$ del enunciado se sigue que H contiene los inversos de todos sus elementos. Por último, de lo anterior se obtiene

$$\forall h, h' \in H\ , \quad h'h = h'(h^{-1})^{-1} \in H\ .$$

□

El criterio anterior es válido tanto para grupos finitos como infinitos. Para grupos *finitos*, dicho criterio se puede relajar sustancialmente, ya que basta con que H sea *cerrado bajo el producto*:

Proposición 1.15. *Un subconjunto H de un grupo* finito *G es subgrupo si es* cerrado bajo el producto*, es decir si $\forall h, h' \in H$, $hh' \in H$.*

Demostración. Por lo visto anteriormente, basta probar que para todo $h \in H$ el elemento inverso h^{-1} pertenece a H. Pero esto es inmediato, ya que si m es el orden de h se tiene

$$h^m = e \qquad \Longrightarrow \qquad h^{-1} = h^{m-1} \in H,$$

al ser H por hipótesis cerrado bajo el producto. □

• El resultado anterior *no* es cierto, en general, para grupos infinitos. Por ejemplo, $\mathbb{N}^N \subset \mathbb{Z}^N$ es cerrado bajo la suma, y sin embargo no es un subgrupo (no contiene el inverso de ninguno de sus elementos).

Ejemplo 1.16. Sea G un grupo finito, y sea g un elemento cualquiera de G. El conjunto

$$\{e, g, \ldots, g^{m-1}\}\,, \qquad \text{con } \ m = o(G)\,,$$

es un subgrupo de tipo C_m. Evidentemente, este es el *menor subgrupo de G que contiene a g*. Por ejemplo, en S_3 las transposiciones (12), (13) y (23) son elementos de orden 2 en virtud del Ejercicio 2, y por tanto los conjuntos

$$\{e, (12)\}\,, \qquad \{e, (13)\}\,, \qquad \{e, (23)\}$$

constituyen un subgrupo de S_3 tipo C_2. Análogamente, al ser el ciclo (123) un elemento de orden 3 el conjunto

$$\{e, (123), (123)^2\} = \{e, (123), (321)\}$$

es un subgrupo abeliano de tipo C_3.

Ejemplo 1.17. $\mathbb{Z}^N$ (con la suma vectorial como producto) es subgrupo del grupo aditivo $\mathbb{Q}^N$, que a su vez es subgrupo de $\mathbb{R}^N$, el cual es subgrupo de $\mathbb{C}^N$. $\mathbb{R}^*$ y S^1 son ambos subgrupos de $\mathbb{C}^*$ (con el producto). El conjunto de las raíces n-ésimas de la unidad es un subgrupo de S^1 (y, por tanto, de $\mathbb{C}^*$).

Ejemplo 1.18. El conjunto

$$\{e, a^2, a^4\} \subset C_6$$

es un subgrupo de C_6. En efecto, es cerrado respecto del producto (ya que $a^6 = e$ y $a^8 = a^2$ en C_6). Por la proposición anterior, ha de contener los inversos de cada uno de sus elementos. En efecto,

$$(a^2)^{-1} = a^4, \qquad (a^4)^{-1} = a^2 .$$

Evidentemente, este subgrupo (como cualquier grupo de orden 3) es una realización de C_3. En efecto, a^2 tiene orden 3 en C_6, ya que

$$(a^2)^2 = a^4 \neq e , \qquad (a^2)^3 = a^6 = e .$$

Ejemplo 1.19. Los conjuntos de matrices[4]

$$\begin{aligned} \mathrm{SL}(n,\mathbb{F}) &= \{A \in M_n(\mathbb{F}) \,|\, \det A = 1\} \\ \mathrm{O}(n,\mathbb{F}) &= \{A \in \mathrm{GL}(n,\mathbb{F}) \,|\, A^{-1} = A^{\mathsf{T}}\} \\ \mathrm{SO}(n,\mathbb{F}) &= \mathrm{O}(n,\mathbb{F}) \cap \mathrm{SL}(n,\mathbb{F}) , \end{aligned}$$

son subgrupos de $\mathrm{GL}(n,\mathbb{F})$ ($\mathbb{F} = \mathbb{R}, \mathbb{C}$), mientras que

$$\begin{aligned} \mathrm{U}(n) &= \{A \in \mathrm{GL}(n,\mathbb{C}) \,|\, A^{-1} = A^{\dagger}\} \subset \mathrm{GL}(n,\mathbb{C}) \\ \mathrm{SU}(n) &= \mathrm{U}(n) \cap \mathrm{SL}(n,\mathbb{C}) \subset \mathrm{GL}(n,\mathbb{C}) \end{aligned}$$

son subgrupos de $\mathrm{GL}(n,\mathbb{C})$. Los conjuntos anteriores se denominan, respectivamente, grupo **especial lineal**, **ortogonal**, **ortogonal especial**, **unitario** y **unitario especial**.

Ejemplo 1.20. Sea G un grupo (finito o infinito) y denotemos por $Z(G)$ a su **centro**, es decir al subconjunto de los elementos de G que conmutan con *todos* los elementos de G:

$$Z(G) = \{a \in G \,|\, ag = ga, \ \ \forall g \in G\} .$$

El conjunto $Z(G)$ es claramente un subgrupo de G, que puede reducirse al conjunto $\{e\}$ (por ejemplo, en S_3). Evidentemente, $Z(G) = G$ si y solo si G es abeliano. Si G es un grupo finito, $g_i \in Z(G)$ si y solo si la fila i de la tabla de multiplicación de G coincide con la columna i.

1.2.1 Cosets. Teorema de Lagrange

Si H es un subgrupo de un grupo G, definimos el **coset** (por la izquierda) asociado a un elemento cualquiera $a \in G$ como el conjunto

$$aH \equiv \{ah \,|\, h \in H\} .$$

El coset por la derecha Ha se define de forma análoga. Es inmediato probar que:

1. El coset aH coincide con H si y solo si $a \in H$,
2. Si $a \notin H$, la unidad e no pertenece a aH
3. Si G es finito, H y aH tienen el mismo número de elementos para todo $a \in G$, es decir[5]

$$|aH| = |H| .$$

[4] En lo que sigue denotaremos por A^{T} y $A^{\dagger}$ respectivamente la *transpuesta* y la *adjunta* (o *transpuesta conjugada*) de la matriz $A = (a_{ij})_{1 \leqslant i,j \leqslant n} \in M_n(\mathbb{C})$, definidas por

$$(A^{\mathsf{T}})_{ij} = a_{ji} , \qquad (A^{\dagger})_{ij} = \overline{a_{ji}}$$

(donde la barra denota conjugación compleja).

[5] De hecho, H y aH tienen el mismo cardinal también cuando G y H son conjuntos infinitos, ya que la aplicación $h \mapsto ah$ es claramente una biyección de H en aH.

(La última propiedad es una consecuencia inmediata del principio de simplificación.)

• Nótese, en particular, que de la segunda afirmación se sigue que el único coset de H que es subgrupo es el propio H.

Ejemplo 1.21. Sea $G = \mathbb{C}^N$, que como hemos visto es un grupo con la suma. Cualquier subespacio vectorial $H \subset G$ es claramente un subgrupo de G (aunque el recíproco es claramente falso: cf. $\mathbb{Z}^N$, $\mathbb{Q}^N$, $\mathbb{R}^N$). Los cosets (por la izquierda o la derecha, ya que G es abeliano) son en este caso los conjuntos $a + H$, con $a \in \mathbb{C}^N$ arbitrario, es decir los subespacios afines paralelos a H.

La propiedad fundamental de los cosets es la siguiente:

Proposición 1.22. *Sea H subgrupo de un grupo G, y sean a, b elementos de G. Entonces los correspondientes cosets aH, bH, o son iguales, o son disjuntos.*

Demostración. En efecto, supongamos que $aH \cap bH \neq \emptyset$, y sea $g \in aH \cap bH$. Por definición de coset, existen dos elementos $h_1, h_2 \in H$ tales que

$$g = ah_1 = bh_2 \,.$$

Entonces los cosets aH y bH son iguales, ya que

$$bH = a(h_1 h_2^{-1})H = aH$$

en virtud de la primera propiedad de los cosets (nótese que $h_1 h_2^{-1} \in H$, al ser este conjunto un subgrupo). □

En virtud de la proposición anterior, si H es un subgrupo cualquiera de un grupo (finito o infinito) G entonces G es una *unión disjunta de cosets asociados al subgrupo H*. Si G (y, por tanto, H) es finito, esto quiere decir que existen m elementos $a_1, \ldots, a_m \in G$ tales que $a_i H \cap a_j H = \emptyset$ para todo $i \neq j$ y

$$G = \bigcup_{i=1}^{m} a_i H \,.$$

Al ser los cosets $a_i H$ disjuntos y $|a_i H| = |H|$, de la ecuación anterior se sigue que

$$|G| = m|H| \,.$$

Esto demuestra el siguiente teorema, probado por vez primera por Lagrange en 1771:

Teorema de Lagrange. *Si G es un grupo finito y H es un subgrupo de G, el orden de H divide al de G.*

Definición 1.23. Si G es un grupo finito y H es un subgrupo de G, el **índice** de H es el número de cosets de H, es decir (en virtud de la demostración del teorema de Lagrange) el cociente $|G|/|H|$.

Ejercicio 3. Probar que un grupo finito G de orden *primo* n es equivalente (es decir, isomorfo; cf. la sección siguiente) a C_n.

Solución. Del teorema de Lagrange se sigue que G no posee subgrupos propios (es decir, distintos de $\{e\}$ y del propio G). Si $|G| = 1$, el resultado se cumple trivialmente. Si $|G| > 1$ y $g \neq e$ es un elemento cualquiera de G, el conjunto

$$\left\{e, g, \ldots, g^{m-1}\right\}, \qquad \text{con} \quad m = o(G) \,,$$

es un subgrupo de G de orden $m > 1$, de donde se sigue que $m = n$. Por tanto

$$G = \left\{e, g, \ldots, g^{n-1}\right\}$$

es equivalente a C_n, al ser $g^n = e$.

1.3 Homomorfismos. Teorema de Cayley

Definición 1.24. Un **homomorfismo** entre dos grupos G y G' es una aplicación $\varphi : G \to G'$ que verifica

$$\varphi(gh) = \varphi(g)\varphi(h)\,, \qquad \forall g, h \in G\,.$$

Un homomorfismo se dice **fiel** si es inyectivo. Un **isomorfismo** es un homomorfismo biyectivo, y un **automorfismo** es un isomorfismo de un grupo G en sí mismo.

En otras palabras, un homomorfismo es una aplicación $G \to G'$ que *preserva el producto*, y por tanto la estructura de grupo. En particular, dos grupos *isomorfos* son totalmente *equivalentes* desde el punto de vista de la teoría de grupos, constituyendo dos realizaciones distintas del mismo grupo abstracto. Los homomorfismos gozan de las siguientes propiedades elementales:

Proposición 1.25. *Sea $\varphi : G \to G'$ un homomorfismo. Entonces se verifica:*

1. $\varphi(e) = e'$*, siendo e' la unidad de G'*
2. $\varphi(g^{-1}) = \varphi(g)^{-1}$, $\forall g \in G$

Demostración.
1) Si $g \in G$ se tiene

$$\varphi(g) = \varphi(ge) = \varphi(g)\varphi(e) \qquad \Longrightarrow \qquad \varphi(e) = e'\,.$$

2) Utilizando el apartado anterior se tiene

$$\varphi(g)\varphi(g^{-1}) = \varphi(gg^{-1}) = \varphi(e) = e' \qquad \Longrightarrow \qquad \varphi(g^{-1}) = \varphi(g)^{-1}\,.$$

□

Ejemplo 1.26. El conjunto $\{0, 1, \ldots, n-1\}$, con el producto dado por la suma módulo n:

$$i \cdot j = i + j \mod n\,, \qquad \forall i, j = 0, 1, \ldots, n-1\,,$$

es un grupo abeliano (comprobarlo) que se suele denotar por[6] $\mathbb{Z}_n$. (La unidad es el 0, y el inverso del elemento i es $n - i$.) La aplicación

$$\begin{cases} \varphi : \mathbb{Z}_n \to C_n \equiv \{e, a, \ldots, a^{n-1}\} \\ \qquad i \mapsto a^i \end{cases}$$

es un isomorfismo de grupos. En efecto, si $i, j \in \{0, 1, \ldots, n-1\}$ entonces $i + j \mod n$ es igual a $i + j - kn$ para algún entero no negativo k (de hecho, $k = 0$ o $k = 1$), y por tanto

$$\varphi(i \cdot j) = a^{i+j-kn} = a^{i+j} = a^i a^j = \varphi(i)\varphi(j)\,.$$

Esto demuestra que φ es un homomorfismo, y por su propia definición es una aplicación suprayectiva. En cuanto a la inyectividad, si $i, j \in \{0, 1, \ldots, n-1\}$ se tiene

$$\varphi(i) = \varphi(j) \quad \Longleftrightarrow \quad a^i = a^j \quad \Longleftrightarrow \quad a^{i-j} = e \quad \Longleftrightarrow \quad i - j = 0 \mod n \quad \Longleftrightarrow \quad i = j\,,$$

ya que $|i - j| \leqslant n - 1$. Del mismo modo se demuestra que el grupo de las raíces n-ésimas de la unidad, o el de las rotaciones de ángulo $2k\pi/n$ ($k = 0, 1, \ldots, n-1$), son ambos isomorfos a C_n.

[6] En el caso $n = 2$, esta definición no coincide estrictamente con la definición de $\mathbb{Z}_2$ como el grupo multiplicativo $\{\pm 1\}$ que aparece en el ejemplo 1.3. Sin embargos, ambos grupos son isomorfos bajo la aplicación φ definida en este ejemplo, que en este caso está dada por $\varphi(i) = (-1)^i$ ($i = 0, 1$).

Ejemplo 1.27. Los grupos $\mathrm{GL}(V)$ y $\mathrm{GL}(N,\mathbb{F})$ (donde $\mathbb{F} = \mathbb{R}, \mathbb{C}$ y V es un espacio vectorial sobre $\mathbb{F}$ de dimensión N) son isomorfos. Para ello, basta introducir una base $\mathcal{B} = \{v_1, \ldots, v_N\}$ en V y definir $\varphi : \mathrm{GL}(V) \to \mathrm{GL}(N,\mathbb{F})$ mediante

$$\varphi(A) = A_{\mathcal{B}}\,,$$

donde $A_{\mathcal{B}}$ es la matriz del operador lineal $A : V \to V$ en la base $\mathcal{B}$. En efecto, es conocido del curso de Álgebra lineal que la aplicación φ es biyectiva, y que si $A, B \in \mathrm{GL}(V)$ entonces $(AB)_{\mathcal{B}} = A_{\mathcal{B}} B_{\mathcal{B}}$. Nótese, sin embargo, que el isomorfismo anterior no es *canónico*, ya que depende de la base elegida. En el caso particular $V = \mathbb{F}^N$ hay una base canónica —la formada por los vectores de la forma $(0, \ldots, 0, 1, 0 \ldots, 0)$—, y por tanto podemos identificar canónicamente $\mathrm{GL}(\mathbb{F}^N)$ y $\mathrm{GL}(N,\mathbb{F})$.

Ejemplo 1.28. El grupo abeliano T_N es claramente isomorfo al grupo aditivo $\mathbb{R}^N$, sin más que identificar $\tau_a \in T_N$ con $a \in \mathbb{R}^N$.

Ejemplo 1.29. Si C es un conjunto finito de n elementos, el grupo $S(C)$ es isomorfo a S_n. En efecto, si $C = \{c_1, \ldots, c_n\}$ la aplicación $\tilde{}: S_n \to S(C_n)$ que a cada permutación $\sigma \in S_n$ le asocia la función $\tilde{\sigma} : C \to C$ definida por $\tilde{\sigma}(c_i) = c_{\sigma_i}$ es claramente un isomorfismo. (Nótese que este isomorfismo no es canónico, ya que depende de cómo hayamos ordenado los elementos de C.)

Definición 1.30. Si $\varphi : G \to G'$, definimos su **núcleo** $\ker\varphi$ y su **imagen** $\varphi(G)$ mediante

$$\ker\varphi = \left\{g \in G \;\middle|\; \varphi(g) = e'\right\}, \qquad \varphi(G) = \left\{\varphi(g) \;\middle|\; g \in G\right\}.$$

Proposición 1.31. *Sea $\varphi : G \to G'$ un homomorfismo. Entonces se verifica:*

1. *El núcleo de φ es un* subgrupo *de G.*
2. *φ es* fiel *si y solo si* $\ker\varphi = \{e\}$.
3. *Si H es cualquier* subgrupo *de G, $\varphi(H)$ es un* subgrupo *de G'. En particular, la imagen de G es un subgrupo de G'. Si, además, φ es* fiel*, el subgrupo $\varphi(H) \subset G'$ es* isomorfo *a H.*

Demostración. Inmediata. □

Ejemplo 1.32. La aplicación

$$\begin{aligned} \varphi : \mathbb{R} &\to S^1 \\ x &\mapsto \mathrm{e}^{\mathrm{i}x} \end{aligned}$$

es un homomorfismo del grupo aditivo $\mathbb{R}$ en S^1. En este caso $\varphi(\mathbb{R}) = S^1$, y

$$\ker\varphi = 2\pi\mathbb{Z}$$

es efectivamente un subgrupo de $\mathbb{R}$.

Ejercicio 4. Probar que $\mathbb{R}$ no es isomorfo a S^1.

Solución. En S^1, dado un número natural $n > 1$ existen elementos $a \neq 1$ tales que $a^n = 1$ (se trata, evidentemente, de las $n - 1$ raíces n-ésimas de la unidad distintas de 1). Si existiera un isomorfismo $\psi : S^1 \to \mathbb{R}^n$ entonces

$$a^n = 1 \quad\Longrightarrow\quad \psi(a^n) = n\psi(a) = \psi(1) = 0 \quad\Longrightarrow\quad \psi(a) = 0 = \psi(1) \quad\Longrightarrow\quad a = 1\,,$$

en contra de la definición de a.

Ejemplo 1.33. Una **transposición** en S_n es por definición un ciclo de longitud 2. Es sabido que toda permutación $\sigma \in S_n$ se puede descomponer en un producto de transposiciones. Recordemos también que, aunque la descomposición de una permutación σ en producto de transposiciones *no* es única, sí lo es la *paridad* $(-1)^\sigma \in \{\pm 1\}$ del número de transposiciones en que se descompone dicha permutación σ. El número $(-1)^\sigma$ se denomina **signo** (o **paridad**) de σ. Por ejemplo, un ciclo $(i_1 \cdots i_l) \in S_n$ de longitud l es el producto de $l-1$ transposiciones, ya que (ejercicio)

$$(i_1 \cdots i_l) = (i_1 i_l) \cdots (i_1 i_3)(i_1 i_2) .$$

Por tanto

$$(-1)^{(i_1 \cdots i_l)} = (-1)^{l-1} .$$

Más generalmente, si $\sigma \in S_n$ se descompone en un producto de s ciclos disjuntos de longitudes $l_1, \ldots, l_s$ entonces

$$(-1)^\sigma = (-1)^{l_1 - 1} \cdots (-1)^{l_s - 1} ,$$

ya que de la definición de signo de una permutación se sigue inmediatamente que

$$(-1)^{\rho\sigma} = (-1)^\rho (-1)^\sigma .$$

En virtud de la identidad anterior, la aplicación

$$\begin{aligned} S_n &\to \mathbb{Z}_2 \\ \sigma &\mapsto (-1)^\sigma , \end{aligned}$$

donde $\mathbb{Z}_2$ denota el grupo multiplicativo $\{\pm 1\}$ (isomorfo a C_2), es un homomorfismo de grupos. El núcleo de este homomorfismo, que denotaremos por A_n, es el subconjunto de S_n formado por las permutaciones pares. Por la proposición anterior, A_n es un subgrupo de S_n. Dada una permutación *impar* cualquiera ρ, el coset ρA_n es distinto de A_n (recuérdese que si H es un subgrupo de un grupo G y $g \in G$ el coset gH coincide con H si y solo si $g \in H$), y claramente

$$S_n = A_n \cup (\rho A_n) .$$

En efecto, si $\rho' \in S_n$ es impar entonces $\rho' = \rho(\rho^{-1}\rho') \in \rho A_n$, al ser $\rho^{-1}\rho'$ par (producto de permutaciones impares). Por tanto A_n tiene índice 2 en S_n. □

Ejemplo 1.34. Un ejemplo muy importante de automorfismo es la conjugación (llamada también *automorfismo interno*). Más precisamente, dado un elemento $a \in G$ la **conjugación** bajo dicho elemento es la aplicación $\gamma_a : G \to G$ definida por

$$\gamma_a(g) = aga^{-1} , \qquad \forall g \in G .$$

Es claro que esta aplicación es un homomorfismo, ya que si $g, h \in G$ se tiene

$$\gamma_a(gh) = agha^{-1} = (aga^{-1})(aha^{-1}) = \gamma_a(g)\gamma_a(h) .$$

Es también claramente biyectiva, siendo su inversa la conjugación bajo a^{-1}. Por tanto γ_a es un automorfismo de G para todo $a \in G$. Veremos en las secciones siguientes que la conjugación es fundamental a la hora de entender la estructura de un grupo y, en particular, de estudiar sus representaciones.

Ejemplo 1.35. Si G es un grupo (finito o infinito), denotaremos por $\mathrm{Aut}(G)$ al conjunto de los automorfismos de G, que es claramente un grupo respecto de la composición. La aplicación

$$\begin{aligned} \gamma : G &\to \mathrm{Aut}(G) \\ a &\mapsto \gamma_a \end{aligned}$$

es un homomorfismo de grupos, pues

$$(\gamma_a\gamma_b)(g) = a(bgb^{-1})a^{-1} = (ab)g(ab)^{-1} = \gamma_{ab}(g)\,, \qquad \forall g \in G\,,$$

y por tanto $\gamma(a)\gamma(b) = \gamma(ab)$. El núcleo de este homomorfismo es el centro de G, ya que

$$a \in \ker\gamma \iff \gamma_a(g) = aga^{-1} = g\,,\ \forall g \in G \iff ag = ga\,,\ \forall g \in G\,.$$

En particular, de este resultado y de la Proposición 1.31 se deduce que $Z(G)$ es subgrupo de G, como ya sabíamos.

1.3.1 Teorema de Cayley

Sea $G = \{g_1, \ldots, g_n\}$ un grupo finito, y sea $g \in G$ uno cualquiera de sus elementos. Por lo visto en el comentario tras la Definición 1.2, los n productos gg_i $(i = 1, \ldots, n)$ son todos distintos. Por tanto, podemos escribir

$$gg_i = g_{\sigma(i)}\,,$$

donde $\sigma \in S_n$ es una permutación que depende de g. La aplicación

$$\begin{aligned} \varphi : G &\to S_n \\ g &\mapsto \sigma \end{aligned}$$

definida por la fórmula anterior es un homomorfismo de grupos. En efecto, sea h otro elemento de G, y sea $\rho = \varphi(h) \in S_n$ la permutación asociada a h, es decir

$$hg_i = g_{\rho(i)}\,.$$

Entonces se tiene

$$(gh)g_i = g(hg_i) = gg_{\rho(i)} = g_{\sigma(\rho(i))} \equiv g_{\sigma\circ\rho(i)}\,.$$

Por tanto $\varphi(gh) = \sigma \circ \rho \equiv \sigma\rho \equiv \varphi(g)\varphi(h)$, lo que demuestra nuestra afirmación. El homomorfismo φ es claramente inyectivo, ya que si $\varphi(g) = \varphi(g') = \sigma$ entonces

$$\forall i, \ldots, n\,, \quad gg_i = g_{\sigma(i)} = g'g_i \quad \Longrightarrow \quad g = g'\,.$$

Por tanto el grupo G es isomorfo al subgrupo $\varphi(G)$ de S_n, en virtud del tercer apartado de la Proposición 1.31. Queda así demostrado el siguiente teorema:

Teorema de Cayley. *Todo grupo finito G es isomorfo a un* subgrupo *de S_n, con $n = |G|$.*

Ejemplo 1.36. Sea D_2 el grupo de las rotaciones/reflexiones que dejan invariante un rectángulo. Si los lados de dicho rectángulo son paralelos a los ejes coordenados (cf. la Fig.1.1), este grupo consta de la identidad más los siguientes elementos:

a = reflexión respecto del eje vertical que pasa por el centro del rectángulo

b = reflexión respecto del eje horizontal que pasa por el centro del rectángulo

c = rotación de ángulo π alrededor del centro del rectángulo.

Es evidente que $a^2 = b^2 = c^2 = e$ y que $c = ab = ba$, por lo que

$$D_2 = \{e, a, b, ab\}\,, \qquad \text{con } a^2 = b^2 = (ab)^2 = e\,.$$

(Nótese que de las relaciones anteriores se sigue que $ba = b^{-1}a^{-1} = (ab)^{-1} = ab$.) ¿A qué subgrupo de S_4 es isomorfo D_2? Para responder a esta pregunta, nótese que cualquiera de estas transformaciones de simetría *permuta los vértices del rectángulo* entre sí, y por tanto puede asimilarse a una permutación

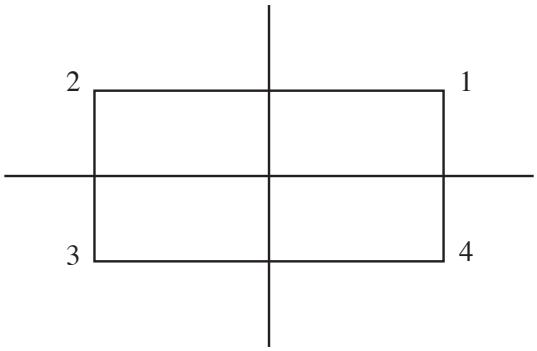

Figura 1.1: Grupo de simetrías de un rectángulo.

del conjunto $\{1, 2, 3, 4\}$. Más concretamente, si enumeramos los vértices del rectángulo tal como se indica en la Fig. 1.1 entonces

$$a \to (12)(34)\,, \qquad b \to (14)(23)\,, \qquad c = ab \to (13)(24)\,.$$

Esto también se puede comprobar (algo más laboriosamente) con el argumento utilizado para probar el teorema de Cayley. Así, por ejemplo, si llamamos

$$g_1 = e\,, \qquad g_2 = a\,, \qquad g_3 = b\,, \qquad g_4 = ab$$

entonces

$$ag_1 = a = g_2\,, \qquad ag_2 = a^2 = e = g_1\,, \qquad ag_3 = ab = g_4\,, \qquad ag_4 = a^2b = b = g_3$$

y por tanto

$$a \to \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} = (12)(34)\,.$$

Ejercicio 5. Sea G un grupo finito de orden n, y sea $g \neq e$. Probar que la descomposición en ciclos disjuntos de la permutación σ asociada a g no puede contener ningún 1-ciclo.

Solución. Esto es inmediato, ya que si σ contiene el 1-ciclo (i) entonces $\sigma(i) = i$, y por tanto

$$gg_i = g_{\sigma(i)} = g_i \quad \Longrightarrow \quad g = e\,.$$

Ejercicio 6. Sea de nuevo G un grupo finito de orden n. Probar que si σ es la permutación asociada a cualquier elemento $g \in G$, todos los ciclos que aparecen en la expresión de σ como producto de ciclos disjuntos tienen la misma longitud (cf. el ejemplo anterior).

Solución. Si $g = e$ entonces $\sigma = (1)(2)\cdots(n)$. Sea $g \neq e$, sea $\sigma \in S_n$ la permutación asociada a g, y supongamos que

$$\sigma = (i_1 \cdots i_k)(j_1 \cdots j_l)\cdots\,, \qquad \{i_1, \ldots, i_k\} \cap \{j_1, \ldots, j_l\} = \emptyset\,,$$

con $k < l$. Entonces

$$\sigma^k = (i_1)\cdots(i_k)(j_1 \cdots j_l)^k \cdots \neq e\,,$$

ya que $(j_1 \cdots j_l)^k \neq e$ si $1 \leqslant k < l$. Esto contradice el ejercicio anterior, ya que σ^k es la permutación asociada al elemento g^k, y $g^k \neq e$ al ser $\sigma^k \neq e$. □

Si G es un grupo de orden primo n, del ejercicio anterior se deduce que la permutación $\sigma \in S_n$ que representa a cualquier elemento $g \neq e$ de G ha de ser un ciclo de orden n. Por tanto g tiene orden n y $G = \{e, g, \ldots, g^{n-1}\} \approx C_n$, como ya probamos en la sección anterior utilizando el teorema de Lagrange.

1.4 Clases de conjugación. Subgrupos normales. Grupo cociente. Producto directo

1.4.1 Clases de conjugación

Dados dos elementos a, b de un grupo G, se dice que b está **conjugado** con a si b es la imagen de a bajo la conjugación por algún elemento del grupo, es decir si

$$b = gag^{-1}$$

para algún $g \in G$. Escribiremos en tal caso $b \sim a$. Es fácil ver que la relación $\sim$ definida de este modo es una *relación de equivalencia*, es decir $\forall a, b, c \in G$ se cumple:

1. $a \sim a$ (reflexividad)
2. $b \sim a \iff a \sim b$ (simetría)
3. $a \sim b,\ b \sim c \implies a \sim c$ (transitividad)

La **clase de conjugación** de $a \in G$ es la clase de equivalencia bajo la conjugación a la que pertenece dicho elemento, es decir el conjunto $[a]$ dado por

$$[a] = \left\{ gag^{-1} \mid g \in G \right\}.$$

- Nótese que, por definición,

$$b \in [a] \iff b \sim a\,.$$

- Es claro que

$$[e] = \{e\}\,,$$

y que $a \in Z(G)$ si y solo si $[a] = \{a\}$. En particular, G es *abeliano* si y solo si $[a] = \{a\}$ para todo $a \in G$ (i.e., todas las clases de conjugación de G constan de un solo elemento).

- Utilizando las propiedades (1)-(3) que definen una relación de equivalencia, es fácil ver que

$$b \in [a] \iff [a] = [b]\,,$$

y que $b, c \in G$ pertenecen a la misma clase de conjugación si y sólo si $b \sim c$.

Proposición 1.37. *Un grupo cualquiera G es la unión* disjunta *de sus clases de conjugación. En otras palabras,*

$$[a] \cap [b] \neq \emptyset \quad \implies \quad [a] = [b]\,.$$

Demostración. En efecto,

$$c \in [a] \cap [b] \implies c \sim a,\ c \sim b \iff b \sim c,\ c \sim a \implies b \sim a \iff [a] = [b]\,.$$

□

Nótese, en particular, que la única clase de conjugación que es subgrupo de G es la trivial, es decir aquella cuyo único elemento es la unidad.

Ejemplo 1.38. Si $G = \mathrm{GL}(n, \mathbb{C})$, dos matrices invertibles $A, B \in \mathrm{GL}(n, \mathbb{C})$ están conjugadas si existe una matriz invertible $C \in \mathrm{GL}(n, \mathbb{C})$ tal que $B = CAC^{-1}$. Por tanto en este caso $B \sim A$ si y solo si A y B tienen la misma *forma canónica de Jordan*, salvo por el orden de los bloques asociados a cada autovalor. (En particular, A y B han de tener los mismos autovalores, con las mismas multiplicidades.) Desde el punto de vista de la teoría de grupos, la forma canónica de Jordan de una matriz invertible es por tanto el representante "canónico" de su clase de conjugación.

Ejemplo 1.39. Sea $G = \mathrm{SO}(3,\mathbb{R})$ el grupo ortogonal especial en $\mathbb{R}^3$. Si $R \in \mathrm{SO}(3,\mathbb{R})$ es la rotación de ángulo $\alpha \in [0,\pi]$ alrededor del eje $v \in \mathbb{R}^3$ y $T \in \mathrm{SO}(3,\mathbb{R})$, entonces TRT^{-1} es la rotación de ángulo α alrededor del eje Tv. En efecto, sabemos que existe una base ortonormal positivamente orientada $\{v_1, v_2, v_3 \equiv v\}$ de $\mathbb{R}^3$ tal que

$$Rv_1 = \cos\alpha\, v_1 + \operatorname{sen}\alpha\, v_2\,, \quad Rv_2 = -\operatorname{sen}\alpha\, v_1 + \cos\alpha\, v_2\,, \quad Rv_3 = v_3\,.$$

Si $S \equiv TRT^{-1}$ entonces $S(Tx) = T(Rx)$, y por tanto

$$\begin{cases} S(Tv_1) = \cos\alpha\,(Tv_1) + \operatorname{sen}\alpha\,(Tv_2)\,, & S(Tv_2) = -\operatorname{sen}\alpha\,(Tv_1) + \cos\alpha\,(Tv_2)\,, \\ S(Tv_3) = S(Tv) = Tv_3. \end{cases}$$

Como $\{Tv_1, Tv_2, Tv_3\}$ es una base ortonormal con la misma orientación que $\{v_1, v_2, v_3\}$ (al ser T ortogonal y $\det T = 1$), de las ecuaciones anteriores se deduce que $S = TRT^{-1}$ es efectivamente una rotación de ángulo α alrededor de $Tv_3 \equiv Tv$. Por tanto, *en* $\mathrm{SO}(3,\mathbb{R})$ *las clases de conjugación están formadas por las rotaciones de un mismo ángulo* $\alpha \in [0,\pi]$ *alrededor de cualquier eje.*

Estudiemos a continuación cómo son las clases de conjugación del grupo simétrico S_n. Consideremos, para ello, dos permutaciones $\sigma, \rho \in S_n$, y sea $(i_1 \dots i_k)$ uno de los ciclos disjuntos en que se descompone σ (incluyendo los ciclos de longitud 1). En primer lugar, es inmediato que la permutación $\rho\sigma\rho^{-1}$ preserva el conjunto $\{\rho(i_1), \dots, \rho(i_k)\}$. Además, la acción de $\rho\sigma\rho^{-1}$ sobre dicho conjunto está dada por

$$\begin{aligned} \rho\sigma\rho^{-1}\{\rho(i_1), \dots, \rho(i_k)\} &= \rho\sigma\{i_1, \dots, i_k\} = \rho\{i_2, \dots, i_k, i_1\} = \{\rho(i_2), \dots, \rho(i_k), \rho(i_1)\} \\ &= \big(\rho(i_1)\cdots\rho(i_k)\big)\{\rho(i_1), \dots, \rho(i_k)\}. \end{aligned}$$

Por tanto

$$\sigma = (i_1\cdots i_k)\cdots(j_1\cdots j_l) \quad \Longrightarrow \quad \rho\sigma\rho^{-1} = \big(\rho(i_1)\cdots\rho(i_k)\big)\cdots\big(\rho(j_1)\cdots\rho(j_l)\big).$$

En particular, *si dos permutaciones están conjugadas las* longitudes *de los ciclos disjuntos en que se descomponen son las mismas*. Se dice en tal caso que ambas permutaciones tienen la misma **estructura de ciclos**. Recíprocamente, es claro que *dos permutaciones que tienen la misma estructura de ciclos están conjugadas*. Más precisamente, si

$$\sigma = (i_1\cdots i_k)\cdots(j_1\cdots j_l)\,, \qquad \sigma' = (i'_1\cdots i'_k)\cdots(j'_1\cdots j'_l)\,,$$

donde se sobreentiende que estamos escribiendo explícitamente los ciclos de longitud 1, por lo que acabamos de ver $\sigma' = \rho\sigma\rho^{-1}$, siendo (por ejemplo)

$$\rho = \begin{pmatrix} i_1 & \cdots & i_k & \cdots & j_1 & \cdots & j_l \\ i'_1 & \cdots & i'_k & \cdots & j'_1 & \cdots & j'_l \end{pmatrix}.$$

Por ejemplo, en S_6 las permutaciones

$$(235)(14) \equiv (235)(14)(6)\,, \qquad (245)(36) \equiv (245)(36)(1)$$

están conjugadas bajo

$$\rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 4 & 6 & 5 & 1 \end{pmatrix} = (1346),$$

o también bajo

$$\rho' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 2 & 4 & 3 & 5 & 1 \end{pmatrix} = (16)(34).$$

Ejemplo 1.40. Por lo visto anteriormente, las clases de conjugación de S_3 son las siguientes:

$$\{e\}\,, \qquad \big\{(12),(13),(23)\big\}\,, \qquad \big\{(123),(321)\big\}\,.$$

□

En virtud del resultado anterior, *el número de clases de conjugación del grupo simétrico S_n es igual al número de* particiones *del entero n*. En efecto, cada estructura de ciclos está determinada por sus longitudes $l_1,\ldots,l_s$ (teniendo en cuenta los posibles ciclos de longitud 1), siendo $l_1+\cdots+l_s=n$. Para n pequeño, las particiones de n pueden hallarse de manera sencilla empezando por la que consta de un sumando (es decir, la partición n) y descendiendo hasta la partición $1+\cdots+1$ con n sumandos iguales a 1.

Ejemplo 1.41. Las particiones de 4 son

$$4, \quad 3+1, \quad 2+2, \quad 2+1+1, \quad 1+1+1+1\,, \tag{1.1}$$

y por tanto S_4 tiene 5 clases de conjugación. Por ejemplo, la clase de conjugación asociada a la partición $2+2$ está formada por los elementos

$$(12)(34), \quad (13)(24), \quad (14)(23).$$

Ejercicio 7. Probar que el número de elementos de una clase de conjugación de S_n está dado por

$$\frac{n!}{\prod_{j=1}^{n}(j^{n_j}n_j!)}\,,$$

donde n_j denota el número de ciclos de longitud j. Por ejemplo, el número de elementos de las clases de conjugación de S_4 (ordenadas según la ec. (1.1)) es igual a

$$6, \quad 8, \quad 3, \quad 6, \quad 1\,.$$

Ejercicio 8. El **centralizador** $Z(a)$ del elemento $a \in G$ es el conjunto de los elementos de G que conmutan con a, es decir

$$Z(a) = \big\{g \in G \;\big|\; ag = ga\big\}\,.$$

1. Probar que $Z(a)$ es un subgrupo de G.
2. Si G es un grupo finito, probar que el cardinal de $[a]$ es igual al número de cosets de $Z(a)$. Por tanto
$$\big|[a]\big| = \frac{|G|}{|Z(a)|}.$$
En particular, *el cardinal de cualquier clase de conjugación de un grupo finito G es un divisor del orden del grupo.*

1.4.2 Subgrupos normales

Al ser la conjugación bajo un elemento cualquiera $g \in G$ un automorfismo de G, si H es un subgrupo de G su imagen $gHg^{-1} \equiv \gamma_g(H)$ es un subgrupo de G isomorfo a H (y, por tanto, con el mismo número de elementos que H si G es finito). En general, $gHg^{-1} \neq H$, ya que

$$gHg^{-1} = H \iff gH = Hg\,,$$

y en general los cosets de H por la izquierda y por la derecha no coinciden. Por ejemplo, si $G = S_3$, $H = \{e,(12)\}$ y $g = (123)$ entonces

$$gH = \big\{(123),(13)\big\} \neq Hg = \big\{(123),(23)\big\}.$$

El caso en que H es *invariante* bajo conjugación por *cualquier* elemento de G, que como veremos más adelante tiene consecuencias importantes, merece especial atención:

Definición 1.42. Un subgrupo H de un grupo G es **normal** (o *invariante*) si

$$gHg^{-1} = H\,, \qquad \forall g \in G\,.$$

Equivalentemente, H es un subgrupo normal si y solo si

$$gH = Hg\,, \qquad \forall g \in G\,.$$

Si G es un grupo finito,

$$gHg^{-1} = H \iff gHg^{-1} \subset H\,,$$

ya que $|gHg^{-1}| = |H|$. De hecho, el resultado anterior se cumple también para grupos infinitos:

Proposición 1.43. *Un subgrupo H de un grupo G es normal si y solo*

$$gHg^{-1} \subset H\,, \qquad \forall g \in G\,.$$

Solución. Basta probar que $H \subset gHg^{-1}$, para todo $g \in G$. Y, en efecto, de la condición del enunciado se sigue que

$$H \subset g^{-1}Hg\,, \qquad \forall g \in G\,,$$

de donde se deduce que $H \subset gHg^{-1}$ para todo $g \in G$.

Corolario 1.44. *Un subgrupo H de un grupo G es normal si y solo si*

$$h \in H \quad \Longrightarrow \quad [h] \in H\,.$$

En otras palabras, un subgrupo $H \subset G$ es normal si y solo si H *es la unión de clases de conjugación completas de* G.

Ejemplo 1.45. Todos los subgrupos de un grupo abeliano son necesariamente normales.

Ejemplo 1.46. En S_3, el único subgrupo propio normal es $\{e, (123), (321)\}$. En efecto, por el teorema de Lagrange los subgrupos propios de S_3 solo pueden tener orden 2 o 3. Es evidente que S_3 *no* tiene subgrupos normales de orden 2, ya que no hay ninguna clase de conjugación distinta de $\{e\}$ que conste de un solo elemento. Por otra parte, como la única clase de conjugación de S_3 que consta de dos elementos es la formada por las permutaciones cíclicas (123) y (321), el único subgrupo normal de orden 3 de S_3 es el indicado anteriormente.

Ejemplo 1.47. $\mathrm{SO}(n, \mathbb{F})$ (con $\mathbb{F} = \mathbb{R}, \mathbb{C}$) es un subgrupo de $\mathrm{SL}(n, \mathbb{F})$. Este subgrupo *no* es normal, ya que si $A \in \mathrm{SL}(n, \mathbb{F})$ y $O \in \mathrm{SO}(n, \mathbb{F})$ la matriz AOA^{-1} es ortogonal si y solo si

$$(AOA^{-1})^{\mathsf{T}} = (A^{\mathsf{T}})^{-1} O^{\mathsf{T}} A^{\mathsf{T}} = (AOA^{-1})^{-1} = AO^{-1}A^{-1} = AO^{\mathsf{T}}A^{-1}\,,$$

lo que en general solo se cumple si $A^{-1} = A^{\mathsf{T}}$ (es decir, si $A \in \mathrm{SO}(n, \mathbb{F})$). Por el contrario, $\mathrm{SU}(n)$ es un subgrupo normal de $\mathrm{U}(n)$, ya que si $A \in \mathrm{U}(n)$ y $U \in \mathrm{SU}(n)$ entonces

$$\det(AUA^{-1}) = \det U = 1 \quad \Longrightarrow \quad AUA^{-1} \in \mathrm{SU}(n)\,.$$

Ejemplo 1.48. Un *movimiento* en $\mathbb{R}^N$ es una *isometría*, es decir una aplicación $f : \mathbb{R}^N \to \mathbb{R}^N$ que verifica

$$|f(x) - f(y)| = |x - y|\,, \qquad \forall x, y \in \mathbb{R}^N\,.$$

Puede probarse que todo movimiento es de la forma $f = TR$, con $R \in \mathrm{O}(N, \mathbb{R})$ y $T \in T_N$. En otras palabras,

$$f(x) = Rx + a\,, \qquad R \in \mathrm{O}(N, \mathbb{R})\,,\ \ a \in \mathbb{R}^N\,.$$

Es claro que el conjunto E_N de todos los movimientos de $\mathbb{R}^N$ es un grupo (no abeliano, si $N > 1$), con la composición como producto. Es también evidente que los conjuntos $\mathrm{O}(N,\mathbb{R})$ y T_N (rotaciones/reflexiones y traslaciones) son subgrupos de E_N. ¿Son estos subgrupos normales? Evidentemente, como todo elemento de E_N es el producto de una rotación por una traslación, T_N es normal si y solo si

$$R\tau_a R^{-1} \in T_N\,, \qquad \forall R \in \mathrm{O}(N,\mathbb{R}), \quad a \in \mathbb{R}^N,$$

siendo τ_a la traslación $x \mapsto x + a$. Esto es obvio, ya que si $x \in \mathbb{R}^N$ se tiene

$$R\tau_a R^{-1}(x) = R(R^{-1}x + a) = x + Ra \quad \Longrightarrow \quad R\tau_a R^{-1} = \tau_{Ra} \in T_N\,.$$

Por el contrario, $O(N,\mathbb{R})$ *no* es normal, ya que

$$\tau_a R \tau_a^{-1}(x) = R(x-a) + a = Rx + (a - Ra) \quad \Longrightarrow \quad \tau_a R \tau_a^{-1} = \tau_{a-Ra} R\,,$$

y por tanto $\tau_a R \tau_a^{-1}$ no pertenece a $\mathrm{O}(N,\mathbb{R})$ si $Ra \neq a$.

Ejercicio 9. Probar que si H es un subgrupo de índice 2 de un grupo finito G entonces H es normal.

Solución. En efecto, de la demostración del teorema de Lagrange aplicada a los cosets por la derecha de un subgrupo H se sigue que el número de dichos cosets es también igual a $|G|/|H|$. Por tanto en este caso se tiene

$$G = H \cup aH = H \cup Ha\,.$$

Al ser ambas uniones *disjuntas*, se ha de verificar

$$G \setminus H = aH = Ha\,,$$

y por tanto H es normal. Un ejemplo del resultado anterior es el subgrupo $A_n \subset S_n$, necesariamente normal por lo que acabamos de ver. Por ejemplo, $\{e, (123), (321)\}$ es normal en S_3 (cf. el Ejemplo 1.46).

Proposición 1.49. *Si $f : G \to G'$ es un* homomorfismo *de grupos, el* núcleo *de f es un subgrupo normal de G.*

Demostración. Ya hemos visto que $\ker f$ es un subgrupo de G (cf. la Proposición 1.31). Basta por tanto probar que para todo $a \in \ker f$ y para todo $g \in G$ el elemento gag^{-1} pertenece a $\ker f$. Pero esto es obvio, ya que, al ser f homomorfismo se tiene

$$f(gag^{-1}) = f(g)f(a)f(g)^{-1} = f(g)e'f(g)^{-1} = e'.$$

□

Ejemplo 1.50. Ya vimos en el ejemplo anterior que A_n es un subgrupo normal de S_n. Otra forma de probar esto es que $A_n = \ker f$, siendo $f : S_n \to \mathbb{Z}_2$ el homomorfismo que a cada permutación le asocia su signo.

Ejemplo 1.51. El centro $Z(G)$ de cualquier grupo G es un subgrupo normal. Esto es inmediato a partir de la definición de $Z(G)$, y también se sigue de la proposición anterior, al ser $Z(G)$ el núcleo de la conjugación $\gamma : G \to \mathrm{Aut}(G)$.

La imagen de un subgrupo normal bajo un homomorfismo de grupos no es, en general, un subgrupo normal, a menos que f sea suprayectiva:

Proposición 1.52. *Si H es un subgrupo normal de un grupo G y $f : G \to G'$ es un homomorfismo* suprayectivo*, entonces $f(H)$ es un subgrupo normal de G'.*

Demostración. Ya hemos visto en la Proposición 1.31 que $f(H)$ es un subgrupo de G', por lo que basta probar que es normal. Esto es inmediato, ya que si $h \in H$ y $g' = f(g) \in G'$ entonces

$$g'f(h)(g')^{-1} = f(g)f(h)f(g^{-1}) = f(ghg^{-1}) \in f(H)\,,$$

al ser H por hipótesis normal en G. □

1.4.3 Grupo cociente

Si H es un subgrupo de un grupo G, parece natural definir un producto entre los cosets (por la izquierda) de H mediante la fórmula

$$aH \cdot bH = (ab)H\,. \tag{1.2}$$

¿Es esta definición correcta? Para que lo sea debemos comprobar que *depende exclusivamente de los cosets* aH y bH, y no de los elementos a y b que se han seleccionado implícitamente para describir dichos cosets. En otras palabras, debemos comprobar que

$$aH = a'H\,,\quad bH = b'H \quad\Longrightarrow\quad aH \cdot bH = a'H \cdot b'H\,.$$

o equivalentemente, teniendo en cuenta (1.2)

$$aH = a'H\,,\quad bH = b'H \quad\Longrightarrow\quad (ab)H = (a'b')H\,. \tag{1.3}$$

Para ello, nótese en primer lugar que

$$aH = a'H \iff a' = ah_1\,,\quad bH = b'H \iff b' = bh_2 \qquad \text{con } h_1, h_2 \in H\,.$$

Por tanto

$$(a'b')H = (ah_1bh_2)H\,, \tag{1.4}$$

que, como muestra el siguiente ejemplo, en general es distinto de $(ab)H$.

Ejemplo 1.53. Sean $G = S_3$, $H = \{e, (12)\}$, $a = (13)$, $b = (23)$. En este caso

$$aH = \{(13), (123)\}\,,\qquad bH = \{(23), (321)\},$$

y por tanto podemos tomar

$$a' = (123),\qquad b' = (321)\,.$$

Entonces

$$ab = (13)(23) = (321)\,,\qquad a'b' = e\,,$$

de donde se sigue que

$$(ab)H = (321)H \neq (a'b')H = H$$

(al ser $(321) \notin H$). □

Por lo que acabamos de ver, la condición necesaria y suficiente para que el producto "natural" de cosets (1.2) esté bien definido es que se cumpla la condición (1.3), es decir (en virtud de (1.4))

$$\begin{aligned}(ah_1bh_2)H = (ab)H \iff ah_1bh_2 \in abH \\ \iff b^{-1}h_1bh_2 \in H\,,\quad \forall b \in G,\ \forall h_1, h_2 \in H.\end{aligned} \tag{1.5}$$

Esta condición se cumple automáticamente si el subgrupo H es *normal*; de hecho, (1.5) es *equivalente* al carácter normal de H, como se deduce sin más que tomar $h_2 = e$. Por tanto *el producto de cosets de un subgrupo H está bien definido si y solo si H es* normal.

Supongamos, en vista de lo anterior, que el subgrupo H es *normal*, y definamos el producto de cosets de H por la ecuación (1.2). Es fácil comprobar entonces que el conjunto

$$G/H = \{aH \mid a \in G\}$$

cuyos elementos son los cosets (por la izquierda) de H es un grupo respecto del producto que acabamos de definir. En efecto:

1. La propiedad asociativa es consecuencia de la asociatividad del producto en H:

$$\begin{aligned}(aH)(bH\cdot cH) &\equiv (aH)\big((bc)H\big) = \big(a(bc)\big)H = \big((ab)c\big)H = \big((ab)H\big)\cdot cH\\ &\equiv (aH\cdot bH)(cH)\,.\end{aligned}$$

2. La unidad de G/H es el subgrupo H, ya que

$$H\cdot aH \equiv eH\cdot aH \equiv (ea)H = aH\,, \qquad \forall a\in G\,.$$

3. $(aH)^{-1} = a^{-1}H$

Definición 1.54. Si H es un subgrupo *normal* de un grupo G, el **grupo cociente** de G por H es el conjunto G/H de los cosets (por la izquierda) de H con el producto (1.2).

- Si G es un grupo finito, por el teorema de Lagrange G/H es un grupo finito de orden

$$|G/H| = |G|/|H|$$

igual al índice del subgrupo H.

Ejemplo 1.55. Sean $G = S_n$ y $H = A_n$, que como sabemos es un subgrupo normal de S_n. En este caso G/H consta de dos clases:

$$G/H = \big\{A_n, \sigma A_n\big\}\,,$$

siendo σ cualquier permutación *impar*. Al ser σ^2 una permutación par, $(\sigma A_n)^2 \equiv \sigma^2 A_n = A_n$, y por tanto la tabla de multiplicación de S_n/A_n es la siguiente:

	A_n	σA_n
A_n	A_n	σA_n
σA_n	σA_n	A_n

En particular, es evidente que $S_n/A_n \approx \mathbb{Z}_2$. De hecho, este isomorfismo es consecuencia inmediata del comentario anterior, y se extiende al caso en que H es un subgrupo normal de índice 2 (ya que el único grupo de orden 2 módulo isomorfismos es $\mathbb{Z}_2$).

Ejemplo 1.56. Hemos visto en el Ejemplo 1.47 que $\mathrm{SU}(n)$ es un subgrupo normal de $\mathrm{U}(n)$. ¿Cuál es el grupo cociente $\mathrm{U}(n)/\mathrm{SU}(n)$? Para responder a esta pregunta, consideremos una matriz cualquiera $A\in\mathrm{U}(n)$. De la condición $A^\dagger A = \mathbb{1}$ se sigue inmediatamente que $|\det A| = 1$, y por tanto

$$\det A = \mathrm{e}^{\mathrm{i}\alpha} \in S^1$$

para algún $\alpha\in\mathbb{R}$. Luego

$$\det(\mathrm{e}^{-\mathrm{i}\alpha/n}A) = 1 \quad\Longrightarrow\quad \mathrm{e}^{-\mathrm{i}\alpha/n}A\in\mathrm{SU}(n) \quad\Longrightarrow\quad A\cdot\mathrm{SU}(n) = \mathrm{e}^{\mathrm{i}\alpha/n}\mathrm{SU}(n)\,.$$

Recíprocamente, es claro que si $z\in S^1$ es un número complejo de módulo 1 la matriz $z\mathbb{1}$ pertenece a $\mathrm{U}(n)$, y define un coset $z\,\mathrm{SU}(n)\in\mathrm{U}(n)/\mathrm{SU}(n)$. Luego

$$\mathrm{U}(n)/\mathrm{SU}(n) = \big\{z\,\mathrm{SU}(n)\,\big|\,z\in S^1\big\}\,,$$

siendo por definición

$$\big(z_1\,\mathrm{SU}(n)\big)\big(z_2\,\mathrm{SU}(n)\big) = (z_1z_2)\mathrm{SU}(n)\,.$$

La aplicación $f : S^1\to\mathrm{U}(n)/\mathrm{SU}(n)$ definida por

$$z\mapsto f(z) = z\,\mathrm{SU}(n)$$

es un homomorfismo de grupos, y es suprayectiva. Si $n \geqslant 2$ *no* es, sin embargo, inyectiva, ya que

$$z \in \ker f \iff z\mathrm{SU}(n) = \mathrm{SU}(n) \iff z\mathbb{1} \in \mathrm{SU}(n) \iff z^n = 1\,,$$

y por tanto

$$\ker f = \left\{1, \omega, \ldots, \omega^{n-1}\right\} \approx C_n\,, \qquad \omega \equiv \mathrm{e}^{\frac{2\pi \mathrm{i}}{n}}\,.$$

Veremos, sin embargo, más adelante que

$$\mathrm{U}(n)/\mathrm{SU}(n) \approx S^1\,.$$

Ejemplo 1.57. Sea G un grupo, y sean H_1 y H_2 subgrupos de G tales que $H_1 \cap H_2 = \{e\}$. Si H_1 es *normal* y $G = H_1 H_2$ (es decir, todo elemento de G es el producto de un elemento de H_1 por un elemento de H_2), entonces $G/H_1 \approx H_2$. En efecto, los cosets de H_1 son de la forma

$$h_1 h_2 H_1 = h_2 (h_2^{-1} h_1 h_2) H_1 = h_2 H_1\,, \qquad h_i \in H_i\,,$$

y por tanto $G/H_1 = \left\{h_2 H_1 \mid h_2 \in H_2\right\}$. La aplicación $f : H_2 \to G/H_1$ definida por $f(h_2) = h_2 H_1$ es claramente un homomorfismo (por la definición de producto de cosets), y es suprayectiva por construcción. Por otra parte, si $h_2 \in H_2$ pertenece al núcleo de f entonces

$$h_2 H_1 = H_1 \iff h_2 \in H_1 \implies h_2 \in H_1 \cap H_2 = \{e\} \implies h_2 = e\,.$$

Luego f es inyectiva, y establece por tanto un isomorfismo entre H_2 y G/H_1.

Como se vio en el Ejemplo 1.48, el grupo E_N de los movimientos de $\mathbb{R}^N$ es el producto de sus subgrupos $H_1 = T_N$ y $H_2 = \mathrm{O}(N, \mathbb{R})$, siendo H_1 normal. Es fácil ver también que $T_N \cap \mathrm{O}(N, \mathbb{R}) = \{e\}$, ya que

$$Rx = x + a\,, \quad \forall x \in \mathbb{R}^N \iff a = 0\,, \ R = \mathbb{1}\,.$$

Por lo que acabamos de ver, $E_N / T_N \approx \mathrm{O}(N, \mathbb{R})$.

Proposición 1.58. *Si $f : G_1 \to G_2$ es un homomorfismo de grupos, entonces $G_1/\ker f \approx f(G_1)$.*

Demostración. En primer lugar, recuérdese (Proposición 1.49) que $f(G_1)$ es un subgrupo de G_2 y $\ker f$ es un subgrupo normal de G_1, por lo que el cociente $G_1/\ker f$ es un grupo con el producto (1.2). En segundo lugar, definimos la aplicación $F : G_1/\ker f \to f(G_1)$ mediante

$$F(g \ker f) = f(g)\,, \qquad \forall g \in G_1\,.$$

Esta aplicación está bien definida, ya que

$$g \ker f = g' \ker f \iff g' = gk\,, \quad k \in \ker f$$

y por tanto

$$f(g') = f(gk) = f(g) f(k) = f(g) e' = f(g)\,,$$

al ser f homomorfismo y $k \in \ker f$. La aplicación F es claramente suprayectiva. Es también un homomorfismo de grupos, ya que

$$F(g \ker f \cdot h \ker f) = F((gh) \ker f) = f(gh) = f(g) f(h) = F(g \ker f) F(h \ker f)\,.$$

Por último, F es inyectiva:

$$g \ker f \in \ker F \iff F(g \ker f) = f(g) = e' \iff g \in \ker f \iff g \ker f = \ker f$$

(recuérdese que $\ker f$ es el elemento unidad en $G/\ker f$). □

Ejemplo 1.59. Sea $G = \mathbb{R}$ (grupo aditivo), y consideremos el homomorfismo de $\mathbb{R}$ en S^1 (grupo multiplicativo) dado por $x \mapsto \mathrm{e}^{ix}$ (cf. el Ejemplo 1.32). En este caso

$$\ker f = 2\pi\mathbb{Z}\,, \qquad f(\mathbb{R}) = S^1\,,$$

y por tanto, en virtud del teorema anterior,

$$\mathbb{R}/(2\pi\mathbb{Z}) \approx S^1\,.$$

(Evidentemente, el factor 2π no es esencial, es decir $\mathbb{R}/(a\mathbb{Z})$ es también isomorfo a S_1 para todo número real $a \neq 0$.) Nótese que $\mathbb{R}/(2\pi\mathbb{Z})$ puede considerarse como el conjunto de los posibles argumentos de un número complejo. En efecto, un elemento de $\mathbb{R}/(2\pi\mathbb{Z})$ es un conjunto de la forma $\theta + 2\pi\mathbb{Z}$, con $\theta \in \mathbb{R}$ (¡recuérdese que el producto en el grupo $\mathbb{R}$ es la suma ordinaria!), y además

$$\theta + 2\pi\mathbb{Z} = \theta' + 2\pi\mathbb{Z} \iff \theta' \in \theta + 2\pi\mathbb{Z} \iff \theta' - \theta \in 2\pi\mathbb{Z}\,.$$

Ejemplo 1.60. La aplicación $\det : \mathrm{U}(n) \to S^1$ es un homomorfismo de grupos, al ser

$$\det(AB) = \det A \cdot \det B\,, \qquad \forall A, B \in M_n(\mathbb{C}).$$

El núcleo de este homomorfismo es el subgrupo $\mathrm{SU}(n)$, por lo que, en virtud de la proposición anterior,

$$\mathrm{U}(n)/\mathrm{SU}(n) \approx S^1\,.$$

Análogamente, el núcleo del homomorfismo $\det : \mathrm{O}(n,\mathbb{F}) \to \mathbb{Z}_2 = \{\pm 1\}$ es el subgrupo $\mathrm{SO}(n,\mathbb{F})$, y por tanto

$$\mathrm{O}(n,\mathbb{F})/\mathrm{SO}(n,\mathbb{F}) \approx \mathbb{Z}_2\,.$$

De la misma forma se demuestra que

$$\mathrm{GL}(n,\mathbb{F})/\mathrm{SL}(n,\mathbb{F}) \approx \mathbb{F}^*\,.$$

1.4.4 Producto directo de dos subgrupos

Definición 1.61. Sea G un grupo, y sean H_1 y H_2 dos subgrupos propios de G. Diremos que G es el **producto directo** de sus subgrupos H_1 y H_2, y escribiremos $G = H_1 \otimes H_2$, si se verifican las siguientes condiciones:

1. $G = H_1 H_2$, i.e, para todo $g \in G$ existen $h_1 \in H_1$ y $h_2 \in H_2$ tales que $g = h_1 h_2$.
2. La descomposición anterior es *única*: si $g = h_1 h_2 = h_1' h_2'$, con $h_i, h_i' \in H_i$, entonces $h_1 = h_1'$ y $h_2 = h_2'$.
3. $h_1 h_2 = h_2 h_1\,, \quad \forall h_1 \in H_1\,,\ h_2 \in H_2$.

• Evidentemente, $G = H_1 \otimes H_2$ si y solo si $G = H_2 \otimes H_1$, al ser las tres condiciones anteriores simétricas en H_1 y H_2. (Nótese, en efecto, que en virtud de la condición 3) $G = H_2 H_1$.)

• La condición (2) es equivalente a

2′) $H_1 \cap H_2 = \{e\}$.

• Nótese que la condición 3) *no* implica que G sea abeliano, ya que no se exige que H_1 o H_2 sean abelianos.

• Utilizando la propiedad 2) del producto directo, es inmediato probar que si G es finito y $G = H_1 \otimes H_2$ entonces

$$|G| = |H_1|\,|H_2|\,.$$

• Es fácil ver que las condiciones 1)-3) anteriores pueden reemplazarse 1), 2) (o 2′)) y 3′), siendo

3′) H_1 y H_2 son subgrupos *normales* de G.

En efecto, es obvio que 1) y 3) implican 3′, ya que si $g = h_1h_2$, con $h_i \in H_i$, entonces

$$gH_1g^{-1} = h_1h_2H_1h_2^{-1}h_1^{-1} = h_1H_1h_1^{-1} = H_1,$$

en virtud de la condición 3), y análogamente para gH_2g^{-1}. Para ver que 1), 2′) y 3′) implican 3), basta notar que si $h_1 \in H_1$ y $h_2 \in H_2$ entonces

$$h_1h_2(h_2h_1)^{-1} = h_1h_2h_1^{-1}h_2^{-1},$$

y (al ser H_1 y H_2 normales por hipótesis)

$$h_1h_2h_1^{-1}h_2^{-1} = h_1(h_2h_1^{-1}h_2^{-1}) \in H_1\,, \qquad h_1h_2h_1^{-1}h_2^{-1} = (h_1h_2h_1^{-1})h_2^{-1} \in H_2\,.$$

La condición 2′) implica entonces que $h_1h_2(h_2h_1)^{-1} = e$, es decir $h_1h_2 = h_2h_1$.

Ejemplo 1.62. Sea $G = C_6$, $H_1 = \{e, a^3\} \approx C_2$, $H_2 = \{e, a^2, a^4\} \approx C_3$. Al ser

$$a = a^3a^4\,, \qquad a^5 = a^3a^2\,,$$

es claro que $G = H_1H_2$, y por tanto se verifica la condición 1) de la definición del producto directo. Es evidente que se cumple también la condición 2′), y la 3′) se verifica trivialmente al ser C_6 abeliano. Por tanto $C_6 = H_1 \otimes H_2 \approx C_2 \otimes C_3$.

Ejercicio 10. Si $r, s \in \mathbb{N}$ son *primos entre sí*, probar que $C_{rs} \approx C_r \otimes C_s$. [*Ayuda:* al ser r y s primos entre sí, en virtud del algoritmo de Euclides existen 2 enteros p y q tales que $pr + qs = 1$.]

Ejemplo 1.63. Consideremos el grupo $G = D_n$, y sus subgrupos

$$H_1 = \{e, a\}\,, \qquad H_2 = \{e, b, \ldots, b^{n-1}\}\,.$$

En este caso $G = H_1H_2$, y claramente $H_1 \cap H_2 = \{e\}$. Sin embargo, el grupo D_n es el producto directo de sus subgrupos H_1 y H_2 si y solo si $n = 2$, ya que

$$ab = (ab)^{-1} = b^{-1}a^{-1} = b^{n-1}a = ba \quad \Longleftrightarrow \quad n = 2\,.$$

Ejemplo 1.64. Sean $G = S_3$, $H_1 = \{e, (12)\}$, $H_2 = \{e, (123), (321)\}$. Es fácil ver que en este caso $G = H_1H_2$, ya que

$$(12)(123) = (23)\,, \qquad (12)(321) = (13)\,,$$

y por tanto se cumple la primera condición del producto directo. Es también obvio que se verifica la condición 2′). Sin embargo, S_3 *no* es el producto directo de sus subgrupos H_1 y H_2, ya que H_1 *no* es normal, y por tanto no se verifica la condición 3′). Equivalentemente, no se cumple la condición 3), ya que (por ejemplo)

$$(12)(123) = (23)\,, \qquad (123)(12) = (13)\,.$$

Lo mismo ocurre si se toma como H_1 cualquiera de los otros subgrupos de orden 2 de S_3, ya que S_3 no posee subgrupos normales de orden 2 (cf. el Ejemplo 1.46). Por este motivo, S_3 *no* puede ser el producto directo de dos de sus subgrupos. Otra forma de probar este resultado es notar que si S_3 fuera el producto directo de dos de sus subgrupos entonces $S_3 \approx C_2 \otimes C_3$, y por tanto S_3 sería *abeliano*.

Ejemplo 1.65. Si $G = C_4$, $H_1 = \{e, a\}$, $H_2 = \{e, a^2\}$, es inmediato ver que se verifican las condiciones 1), 2′) y 3) anteriores. Sin embargo C_4 *no* es el producto directo de H_1 y H_2, ya que H_1 *no* es un subgrupo.

Ejemplo 1.66. Sean $G = E_N$, $H_1 = T_N$, $H_2 = \mathrm{O}(N)$. Entonces G *no* es el producto directo de sus subgrupos H_1 y H_2, al no ser H_2 normal.

Ejemplo 1.67. Sean $G = \mathrm{U}(n)$, $H_1 = \left\{ \mathrm{e}^{\mathrm{i}\alpha} \mathbb{1} \,\middle|\, \alpha \in \mathbb{R} \right\} \approx S^1$, $H_2 = \mathrm{SU}(n)$. Claramente, tanto H_1 como H_2 son subgrupos de G, y ambos son normales (cf. el Ejemplo 1.47). Es claro que todo elemento de $\mathrm{U}(n)$ es el producto de un elemento de H_1 y otro de H_2 (cf. el Ejemplo 1.56). Sin embargo, si $n \geqslant 2$ $\mathrm{U}(n)$ *no* es el producto directo de H_1 con H_2, ya que

$$H_1 \cap H_2 = \left\{ \mathrm{e}^{\mathrm{i}\alpha} \mathbb{1} \,\middle|\, \alpha \in \mathbb{R}\,,\ \det(\mathrm{e}^{\mathrm{i}\alpha}\mathbb{1}) = 1 \right\} = \left\{ \mathbb{1}, \omega\mathbb{1}, \ldots, \omega^{n-1}\mathbb{1} \right\}, \qquad \omega \equiv \mathrm{e}^{\frac{2\pi \mathrm{i}}{n}}\,.$$

Ejercicio 11. Estudiar si $\mathrm{O}(n) = \mathrm{SO}(n) \otimes \mathbb{Z}_2$, siendo $\mathbb{Z}_2 = \{\pm\mathbb{1}\}$.

Ejemplo 1.68. Consideremos el grupo $G = D_n$ y sus subgrupos

$$H_1 = \{e, a\}\,, \qquad H_2 = \left\{e, b, \ldots, b^{n-1}\right\},$$

que claramente satisfacen las condiciones $G = H_1 H_2$ y $H_1 \cap H_2 = \{e\}$. Por tanto G será igual al producto directo H_1 y H_2 si y solo $h_1 h_2 = h_2 h_1$, para todo $h_1 \in H_1$ y $h_2 \in H_2$. A su vez, esta última condición es claramente equivalente a la igualdad $ab = ba$. Sin embargo,

$$ab = (ab)^{-1} = b^{-1}a^{-1} = b^{n-1}a = ba \quad \Longleftrightarrow \quad b^{n-1} = b \quad \Longleftrightarrow \quad n = 2\,.$$

Por tanto $D_2 = H_1 \otimes H_2 \approx C_2 \otimes C_2$, y $D_n \neq H_1 \otimes H_2$ si $n > 2$.

Ejercicio 12. Si $G = H_1 \otimes H_2$, probar que

$$G/H_1 \approx H_2\,, \qquad G/H_2 \approx H_1\,.$$

Solución. Las aplicaciones $\pi_i : G \to H_i$ (proyecciones) definidas[7] por

$$\pi_i(h_1 h_2) = h_i$$

son claramente homomorfismos, siendo

$$\ker \pi_1 = H_2\,, \qquad \ker \pi_2 = H_1\,.$$

Por tanto, en virtud de la Proposición 1.58,

$$G/\ker \pi_1 = G/H_2 \approx \pi_1(G) = H_1\,, \qquad G/\ker \pi_2 = G/H_1 \approx \pi_2(G) = H_2\,.$$

1.4.5 Producto directo de dos grupos

Dados dos grupos G_1 y G_2, se define su **producto directo** como el *producto cartesiano*

$$G_1 \times G_2 = \left\{ (g_1, g_2) \,\middle|\, g_1 \in G_1, g_2 \in G_2 \right\},$$

con el producto natural

$$(g_1, g_2)(h_1, h_2) = (g_1 h_1, g_2 h_2)\,.$$

Es inmediato verificar que $G_1 \times G_2$ es efectivamente un grupo, siendo

$$e = (e_1, e_2)\,, \qquad (g_1, g_2)^{-1} = (g_1^{-1}, g_2^{-1})\,,$$

donde e_i denota la unidad de G_i. Los subconjuntos

$$H_1 \equiv G_1 \times \{e_2\}\,, \qquad H_2 \equiv \{e_1\} \times G_2$$

son obviamente subgrupos de $G_1 \times G_2$. De hecho, estos subgrupos son las imágenes de los grupos G_1 y G_2 bajo los *homomorfismos inyectivos* (*inyecciones canónicas*)

$$g_1 \mapsto (g_1, e_2)\,, \qquad g_2 \mapsto (e_1, g_2)\,,$$

por lo que

$$H_i \approx G_i\,, \qquad i = 1, 2\,.$$

Veamos a continuación que $G_1 \times G_2 = H_1 \otimes H_2$. Esto es inmediato, ya que:

[7] Estas aplicaciones están bien definidas porque, en virtud de las dos primeras propiedades del producto directo, todo elemento de G se escribe de manera *única* como el producto de un elemento de H_1 por un elemento de H_2.

1. $G_1 \times G_2 = H_1 H_2$, al ser
$$(g_1, g_2) = (g_1, e_2)(e_1, g_2)$$
por definición del producto en $G_1 \times G_2$.
2. $H_1 \cap H_2 = \{(e_1, e_2)\} \equiv e$.
3. $(g_1, e_2)(e_1, g_2) = (e_1, g_2)(g_1, e_2) = (g_1, g_2)$.

Normalmente G_i se identifica con H_i (recuérdese que son isomorfos, de forma natural), y se denota el producto de $G_1 \times G_2$ por $G_1 \otimes G_2$ (con un ligero abuso de notación).

• Es inmediato generalizar el producto directo al caso en que el número de factores es un entero positivo arbitrario.

Ejemplo 1.69. Sea $\mathbb{F}$ un cuerpo (y, por tanto, un grupo aditivo). Entonces $\mathbb{F}^2 = \mathbb{F} \otimes \mathbb{F}$ y, en general,
$$\mathbb{F}^N = \underbrace{\mathbb{F} \otimes \cdots \otimes \mathbb{F}}_{N} .$$

Ejemplo 1.70. Sea $G = S^1 \otimes \mathrm{SU}(n)$, y consideremos la aplicación $f : G \to \mathrm{U}(n)$ dada por
$$(\mathrm{e}^{\mathrm{i}\alpha}, A) \mapsto \mathrm{e}^{\mathrm{i}\alpha} A , \qquad \alpha \in \mathbb{R} , \quad A \in \mathrm{SU}(n) .$$
Esta aplicación es claramente suprayectiva (cf. el Ejemplo 1.56), y también es inmediato comprobar que es un homomorfismo. ¿Cuál es su núcleo? Para responder a esta pregunta, supongamos que $\mathrm{e}^{\mathrm{i}\alpha} A = \mathbb{1}$, con $\alpha \in \mathbb{R}$ y $A \in \mathrm{SU}(n)$. Tomando determinantes en ambos miembros de esta igualdad se obtiene
$$\mathrm{e}^{\mathrm{i}n\alpha} = 1 \iff \mathrm{e}^{\mathrm{i}\alpha} = \omega^k , \qquad \text{con } \omega = \mathrm{e}^{\frac{2\pi \mathrm{i}}{n}} , \quad k \in \{0, 1, \ldots, n-1\} .$$
Sustituyendo a continuación cada uno de estos valores de $\mathrm{e}^{\mathrm{i}\alpha}$ en la igualdad $\mathrm{e}^{\mathrm{i}\alpha} A = \mathbb{1}$ se obtiene $A = \omega^{-k} \mathbb{1}$, que evidentemente pertenece a $\mathrm{SU}(n)$. Por tanto
$$\ker f = \left\{ (\omega^k, \omega^{-k}\mathbb{1}) \,\middle|\, k = 0, 1, \ldots, n-1 \right\} = \left\{ (1, \mathbb{1}), (\omega, \omega^{-1}\mathbb{1}), \ldots, (\omega, \omega^{-1}\mathbb{1})^{n-1} \right\} \approx C_n .$$
De la Proposición 1.58 se sigue entonces que
$$\mathrm{U}(n) \approx \big(S^1 \otimes \mathrm{SU}(n)\big)/C_n .$$

1.5 Representaciones

1.5.1 Representaciones lineales

Definición 1.71. Sea G un grupo, y V un espacio vectorial *complejo* de dimensión *finita*. Una **representación** lineal de G en el espacio vectorial V es un *homomorfismo* $D : G \to \mathrm{GL}(V)$. La dimensión de V como espacio vectorial se denomina **dimensión** de la representación D. Una representación se dice **fiel** si es inyectiva.

En otras palabras, una representación lineal de G en el espacio vectorial V es una aplicación D que a cada elemento $g \in G$ le asocia una *aplicación lineal invertible* $D(g) : V \to V$ de forma que
$$D(gh) = D(g)D(h) , \qquad \forall g, h \in G . \tag{1.6}$$
Nótese que, por las propiedades de los homomorfismos, si D es una representación de G entonces
$$D(e) = I , \qquad D(g^{-1}) = D(g)^{-1} ,$$
siendo $I : V \to V$ la identidad.

Proposición 1.72. *Una aplicación $D : G \to \mathcal{L}(V)$ (donde $\mathcal{L}(V)$ denota el espacio de las aplicaciones lineales de V en sí mismo) es una representación si y solo si $D(e) = I$ y se verifica* (1.6)*.*

Demostración. Inmediata, ya que para todo $g \in G$ se tiene

$$D(gg^{-1}) = D(g)D(g^{-1}) = D(e) = I \quad \Longrightarrow \quad D(g)^{-1} = D(g^{-1}) \quad \Longrightarrow \quad D(g) \in \mathrm{GL}(V)\,.$$

□

Definición 1.73. Una **representaciones matricial** de un grupo G es un *homomorfismo* $D : G \to \mathrm{GL}(n, \mathbb{C})$.

Si $V = \mathbb{C}^n$ se identifica normalmente $\mathrm{GL}(\mathbb{C}^n)$ con $\mathrm{GL}(n, \mathbb{C})$, equiparando una aplicación lineal $A : \mathbb{C}^n \to \mathbb{C}^n$ con su matriz en la base canónica de $\mathbb{C}^n$. Mediante esta identificación, podemos considerar las representaciones matriciales como un caso particular de las lineales.

Ejemplo 1.74. La aplicación $D : G \to \mathrm{GL}(V)$ dada por $D(g) = I$, para todo $g \in G$, es claramente una representación de G que llamaremos **trivial**.

Ejemplo 1.75. Si $G \subset \mathrm{GL}(n, \mathbb{C})$ es un *grupo matricial*, la aplicación $D : G \to \mathrm{GL}(n, \mathbb{C})$ dada por

$$D(A) = A\,, \qquad \forall A \in G\,,$$

es claramente una representación (matricial, obviamente fiel) de G llamada **representación de definición** (*defining representation*, en inglés).

Ejemplo 1.76. Si $D : G \to \mathrm{GL}(n, \mathbb{C})$ es una representación matricial de G, la aplicación $\tilde{D} : G \to \mathrm{GL}(n, \mathbb{C})$ definida por

$$\tilde{D}(g) = \big(D(g)^{\mathsf{T}}\big)^{-1}\,,$$

es también una representación matricial de G denominada la **representación contragrediente** de D.

Ejemplo 1.77. Si $G = S^1$, la aplicación $D : S^1 \to \mathrm{GL}(2, \mathbb{C})$ definida por

$$D\big(\mathrm{e}^{\mathrm{i}\theta}\big) = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \equiv R(\theta)\,, \qquad \theta \in \mathbb{R}\,,$$

es claramente una representación. En efecto, nótese en primer lugar que D está *bien definida*, ya que

$$\Big(\mathrm{e}^{\mathrm{i}\theta} = \mathrm{e}^{\mathrm{i}\theta'} \iff \theta' = \theta + 2k\pi\,, \quad k \in \mathbb{Z}\Big) \quad \Longrightarrow \quad R(\theta) = R(\theta')\,.$$

Además, para todo $\theta_1, \theta_2 \in \mathbb{R}$ se tiene

$$D\big(\mathrm{e}^{\mathrm{i}\theta_1}\mathrm{e}^{\mathrm{i}\theta_2}\big) = D\big(\mathrm{e}^{\mathrm{i}(\theta_1+\theta_2)}\big) = R(\theta_1 + \theta_2) = R(\theta_1)R(\theta_2) = D\big(\mathrm{e}^{\mathrm{i}\theta_1}\big)D\big(\mathrm{e}^{\mathrm{i}\theta_2}\big)\,.$$

La representación D tiene dimensión 2, y es *fiel*. En efecto,

$$D\big(\mathrm{e}^{\mathrm{i}\theta_1}\big) = D\big(\mathrm{e}^{\mathrm{i}\theta_2}\big) \iff R(\theta_1) = R(\theta_2) \iff \theta_1 = \theta_2 \bmod 2\pi \iff \mathrm{e}^{\mathrm{i}\theta_1} = \mathrm{e}^{\mathrm{i}\theta_2}\,.$$

Una representación D' de G de dimensión 3 (también fiel) está dada por

$$D'\big(\mathrm{e}^{\mathrm{i}\theta}\big) = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Nótese que las matrices de esta representación dejan *invariante* el subespacio $W = \mathrm{lin}\{(0, 0, 1)\}$, es decir

$$D'\big(\mathrm{e}^{\mathrm{i}\theta}\big)W \subset W\,, \qquad \forall \theta \in \mathbb{R}\,.$$

□

De la Proposición 1.31 se sigue inmediatamente el siguiente resultado:

Proposición 1.78. *La representación D es fiel si y solo si* $\ker D = \{e\}$.

Definición 1.79. Se dice que un grupo G es **simple** si no posee subgrupos normales propios, y **semisimple** si no posee subgrupos normales *abelianos* propios.

Si D es una representación de un grupo G, al ser D un *homomorfismo* su núcleo es un subgrupo *normal* de G. Si G es simple y D es una representación de G entonces o bien $\ker D = G$, en cuyo caso D es la representación trivial, o bien $\ker D = \{e\}$, en cuyo caso D es fiel. Por tanto:

Proposición 1.80. *Toda representación no trivial de un grupo* simple *es* fiel.

Ejemplo 1.81. El grupo S_n *no* es simple, ya que posee el subgrupo normal no trivial A_n. La aplicación $D : S_n \to \mathbb{C}$ dada por $D(\sigma) = (-1)^\sigma$ es claramente una representación (de dimensión uno) no trivial de S_n. Esta representación *no* es fiel, al ser su núcleo precisamente el subgrupo normal A_n. El grupo S_3 tampoco es semisimple, ya que en este caso $A_3 = \{e, (123), (321)\}$ es un subgrupo normal abeliano.

Ejercicio 13. ¿Es S_4 semisimple?

1.5.2 Representación regular

Definición 1.82. Un **álgebra** sobre un cuerpo $\mathbb{F}$ es un *espacio vectorial* $\mathcal{A}$ sobre $\mathbb{F}$ provisto de un **producto**

$$\begin{aligned} \mathcal{A} \times \mathcal{A} &\to \mathcal{A} \\ (a, b) &\mapsto ab \end{aligned}$$

lineal en cada uno de sus argumentos, es decir tal que:

1. $a(b + c) = ab + ac\,, \quad (b + c)a = ba + ca\,, \qquad \forall a, b, c \in \mathcal{A}$ (*propiedad distributiva*).
2. $(\lambda a)b = a(\lambda b) = \lambda(ab)\,, \qquad \forall a, b \in \mathcal{A}\,,\ \forall \lambda \in \mathbb{F}\,.$

La **dimensión** del álgebra $\mathcal{A}$ es por definición su dimensión como espacio vectorial.

En lo que sigue, supondremos siempre que $\mathbb{F} = \mathbb{R}, \mathbb{C}$, y nos ocuparemos casi exclusivamente de álgebras de dimensión *finita*. Diremos que un álgebra $\mathcal{A}$ es **asociativa** o **abeliana** si su producto es respectivamente asociativo o conmutativo. Una **unidad** en un álgebra $\mathcal{A}$ es un elemento $e \in \mathcal{A}$ tal que

$$ea = ae = a\,, \qquad \forall a \in \mathcal{A}\,.$$

Es inmediato probar que en un álgebra hay a lo sumo una unidad. Si $\mathcal{A}$ es un álgebra con elemento unidad, un **inverso** de un elemento $a \in \mathcal{A}$ es cualquier elemento $b \in \mathcal{A}$ tal que

$$ab = ba = e\,.$$

En un álgebra *asociativa* $\mathcal{A}$, es fácil probar que si existe el inverso de un elemento del álgebra necesariamente ha de ser único; en tal caso, denotaremos el inverso de $a \in \mathcal{A}$ por a^{-1}. De la linealidad del producto respecto de cada uno de sus argumentos se sigue inmediatamente que

$$a0 = 0a = 0\,, \qquad \forall a \in \mathcal{A}\,,$$

por lo que 0 no puede poseer inverso. Si $\mathcal{A}$ es un álgebra asociativa con elemento unidad, y todo elemento de $\mathcal{A}$ distinto de 0 posee inverso, se dice que $\mathcal{A}$ es un **álgebra asociativa con división**. En otras palabras, un álgebra asociativa con división es un álgebra $\mathcal{A}$ en que $\mathcal{A} \setminus \{0\}$ es un *grupo* respecto de la multiplicación. Recuérdese que un **cuerpo** es un álgebra asociativa con división que es además *conmutativa*.

Ejemplo 1.83. El conjunto $M_n(\mathbb{F})$ (con $\mathbb{F} = \mathbb{R}$ o $\mathbb{C}$) es un álgebra asociativa con elemento unidad. Si $n \geqslant 2$, esta álgebra *no* es abeliana ni es un álgebra con división.

Ejercicio 14. Sea $\{e_1, e_2, e_3, e_4\}$ la base canónica de $\mathbb{R}^4$, y definamos los productos

1. $e_1 e_i = e_i e_1 = e_i$, para todo $i = 1, \ldots, 4$.
2. $e_2^2 = e_3^2 = e_4^2 = e_2 e_3 e_4 = -e_1$.

a) Demostrar que las relaciones anteriores determinan un producto *asociativo* en $\mathbb{R}^4$. b) Probar que, aunque dicho producto *no* es abeliano, $\mathbb{R}^4$ es un álgebra asociativa con división. Esta álgebra recibe el nombre de *álgebra de los cuaterniones*, y se suele denotar por $\mathbb{H}$.

Solución. a) Es fácil ver que las relaciones anteriores, junto con el requisito de que el producto sea *asociativo*, determinan los restantes productos $e_i e_j$ entre pares de elementos de la base canónica con $2 \leqslant i \neq j \leqslant 4$. En efecto,

$$e_2 e_3 e_4^2 = -e_2 e_3 = -e_1 e_4 = -e_4 \quad \Longrightarrow \quad e_2 e_3 = e_4 \, .$$

De la misma forma se demuestra que $e_3 e_4 = e_2$, y por tanto

$$e_2 e_4 = e_3 e_4^2 = -e_3 \, .$$

De estas relaciones se obtienen fácilmente los demás productos $e_i e_j$ con $i, j \geqslant 2$:

$$e_3 e_2 = e_3^2 e_4 = -e_4 \, , \qquad e_4 e_2 = -e_3 e_2^2 = e_3 \, , \qquad e_4 e_3 = e_2 e_3^2 = -e_2 \, .$$

Nótese, en particular, que

$$e_i e_j = -e_j e_i \, , \qquad 2 \leqslant i \neq j \leqslant 4 \, . \tag{1.7}$$

Extendiendo este producto por *linealidad*, es decir definiendo

$$\Big(\sum_{i=1}^{4} x_i e_i \Big) \Big(\sum_{j=1}^{4} y_j e_j \Big) = \sum_{i,j=1}^{4} x_i y_j e_i e_j \, ,$$

obtenemos un producto definido en todo $\mathbb{R}^4$, respecto del cual $\mathbb{R}^4$ es un álgebra. Nótese que el vector e_1 actúa obviamente como elemento unidad, y por tanto escribiremos a partir de ahora 1 en lugar de e_1. Se comprueba fácilmente que el producto que acabamos de definir es asociativo entre los elementos de la base canónica. En efecto, es fácil ver a partir de la ec. (1.7) y los productos 1)-2) que basta comprobar la igualdad $(e_2 e_3) e_4 = e_2 (e_3 e_4)$. Dicha igualdad se verifica, ya que ambos productos son iguales a -1. De la asociatividad del producto entre elementos de la base canónica se sigue inmediatamente que el producto es asociativo en todo $\mathbb{R}^4$, y por tanto $\mathbb{R}^4$ es un álgebra asociativa con elemento unidad. Que no es conmutativa es obvio, en virtud de la ec. (1.7). Por último, dado un elemento cualquiera $x = \sum_{i=1}^{4} x_i e_i$ y definiendo $\overline{x}$ mediante

$$\overline{x} = x_1 e_1 - \sum_{i=1}^{3} x_i e_i \, ,$$

entonces

$$x \overline{x} = x_1^2 - \sum_{i,j=2}^{4} x_i x_j e_i e_j = x_1^2 - \sum_{i=2}^{4} x_i^2 e_i^2 - \sum_{2 \leqslant i \neq j \leqslant 4} x_i x_j e_i e_j = x_1^2 + \sum_{i=2}^{4} x_i^2 = \|x\|^2 \, ,$$

en virtud de la ec. (1.7). Del mismo modo (o simplemente teniendo en cuenta que $\overline{\overline{x}} = x$) se demuestra que $\overline{x}\, x = \|x\|^2$. Por tanto todo elemento $x \neq 0$ tiene por inverso el elemento

$$x^{-1} = \frac{\overline{x}}{\|x\|^2} \, ,$$

lo que convierte a $\mathbb{R}^4$ en un *álgebra asociativa* (*no conmutativa*) *con división*. □

Todo grupo *finito* G tiene asociada de manera natural un álgebra asociativa con elemento unidad de dimensión igual al orden de G. En efecto, sea

$$G = \{g_1, \ldots, g_n\}, \qquad n \equiv |G|,$$

y denotemos por $|g_i\rangle$ el i-ésimo vector de la base canónica de $\mathbb{C}^n$. Nótese que un elemento cualquiera de $\mathbb{C}^n$ puede escribirse en la forma

$$\sum_{i=1}^{n} a_i |g_i\rangle \equiv \sum_{g\in G} a(g)|g\rangle,$$

siendo $a_i \equiv a(g_i) \in \mathbb{C}$, $i = 1, \ldots, n$. El **álgebra del grupo** G (*group algebra*, en inglés), que denotaremos por $\mathcal{A}(G)$, es el espacio vectorial $\mathbb{C}^n$ con el siguiente producto:

$$a = \sum_{g\in G} a(g)|g\rangle, \quad b = \sum_{g\in G} b(g)|g\rangle \quad \Longrightarrow \quad ab = \sum_{g,h\in G} a(g)b(h)|gh\rangle.$$

Obsérvese que podemos escribir

$$ab = \sum_{g\in G} c(g)|g\rangle,$$

siendo

$$c(g) = \sum_{\substack{h,k\in G\\ kh=g}} a(k)b(h) = \sum_{h\in G} a(gh^{-1})b(h) = \sum_{k\in G} a(k)b(k^{-1}g).$$

Es inmediato comprobar que $\mathcal{A}(G)$ es un *álgebra asociativa con elemento unidad*, siendo la unidad el vector $|e\rangle$. También es fácil probar que $\mathcal{A}(G)$ es abeliana si y solo si lo es el grupo G.

Ejemplo 1.84. Sea $G = C_2 \equiv \{e, a\}$. Entonces $\mathcal{A}(C_2) = \mathbb{C}^2$, con el producto

$$(a_0|e\rangle + a_1|a\rangle)(b_0|e\rangle + b_1|a\rangle) = (a_0b_0 + a_1b_1)|e\rangle + (a_0b_1 + a_1b_0)|a\rangle.$$

Nótese que $\mathcal{A}(C_2)$ *no* es un álgebra con división, ya que (por ejemplo)

$$(|e\rangle + |a\rangle)(|e\rangle - |a\rangle) = |e\rangle - |a^2\rangle = |e\rangle - |e\rangle = 0.$$

Por tanto los elementos $|e\rangle \pm |a\rangle$ no poseen inverso, ya que si alguno de ellos lo tuviera la igualdad anterior implicaría que el otro elemento sería nulo. □

Si G es un grupo *finito*, su **representación regular** es la aplicación $D_{\mathrm{R}} : G \to \mathcal{L}\big(\mathcal{A}(G)\big)$ dada por

$$D_{\mathrm{R}}(g)a = |g\rangle\, a, \qquad \forall a \in \mathcal{A}(G), \quad \forall g \in G.$$

En otras palabras,

$$D_{\mathrm{R}}(g)\Big(\sum_{h\in G} a(h)|h\rangle\Big) = \sum_{h\in G} a(h)|gh\rangle, \qquad \forall g \in G.$$

Es claro que $D_{\mathrm{R}}(g)$ es una aplicación *lineal* de $\mathcal{A}(G)$ en sí mismo (por la linealidad del producto de $\mathcal{A}(G)$ en cada uno de sus factores). De hecho, $D_{\mathrm{R}}(g)$ es la aplicación lineal definida en la base canónica de $\mathcal{A}(G)$ mediante

$$D_{\mathrm{R}}(g)|h\rangle = |gh\rangle, \qquad \forall h \in G,$$

y extendida por linealidad a todo $\mathcal{A}(G)$. Comprobemos que esta aplicación es efectivamente una *representación* de G:

1. En primer lugar, $D_R(g) \in \mathrm{GL}\big(\mathcal{A}(G)\big)$ para todo $g \in G$, ya que es inmediato comprobar que $D_{\mathrm{R}}(g)^{-1} = D_{\mathrm{R}}\big(g^{-1}\big)$.

2. En segundo lugar,
$$D_{\mathrm{R}}(gh) = D_{\mathrm{R}}(g)D_{\mathrm{R}}(h)\,, \qquad \forall g,h \in G\,.$$
En efecto, por linealidad basta probar que ambos miembros producen el mismo resultado cuando se aplican a los vectores de la base canónica de $\mathcal{A}(G)$. Esto se verifica claramente, ya que
$$D_{\mathrm{R}}(gh)|k\rangle = |(gh)k\rangle = |g(hk)\rangle = D_{\mathrm{R}}(g)|hk\rangle = D_{\mathrm{R}}(g)D_{\mathrm{R}}(h)|k\rangle\,, \qquad \forall k \in G\,.$$

• La representación regular de un grupo finito G tiene dimensión igual al orden de G, y es claramente *fiel*. En efecto,
$$\begin{aligned} g \in \ker D_{\mathrm{R}} &\iff D_{\mathrm{R}}(g) = I \iff D_{\mathrm{R}}(g)|h\rangle \equiv |gh\rangle = |h\rangle\,, \quad \forall h \in G \\ &\iff gh = h\,, \quad \forall h \in G \iff g = e\,. \end{aligned}$$

• Dado un elemento $g \in G$, la actuación de $D_{\mathrm{R}}(g)$ sobre la base canónica de $\mathcal{A}(G)$ está dada por
$$D_{\mathrm{R}}(g)|g_j\rangle = |gg_j\rangle = |g_{\sigma(j)}\rangle\,, \qquad j = 1,\ldots,n\,,$$
siendo $\sigma \in S_n$ una cierta permutación (dependiente de g). De esto se sigue inmediatamente que
$$[D_{\mathrm{R}}(g)]_{ij} = \begin{cases} 1, & \text{si } i = \sigma(j) \\ 0, & \text{en cualquier otro caso,} \end{cases}$$
donde hemos denotado por $[D_{\mathrm{R}}(g)]_{ij}$ el elemento de matriz (i,j) de la matriz de $D_{\mathrm{R}}(g)$ en la base canónica de $\mathcal{A}(G)$. En otras palabras, la matriz de $D_{\mathrm{R}}(g)$ tiene exactamente un 1 en cada fila y en cada columna, siendo sus demás elementos de matriz iguales a cero. Teniendo en cuenta que
$$i = \sigma(j) \iff gg_j = g_i \iff g = g_i g_j^{-1}\,,$$
otra forma más conveniente de expresar lo anterior es la siguiente:
$$[D_{\mathrm{R}}(g)]_{ij} = \begin{cases} 1, & \text{si } g = g_i g_j^{-1} \\ 0, & \text{en cualquier otro caso.} \end{cases}$$
Por tanto, para hallar la matriz de $D_{\mathrm{R}}(g)$ en la base canónica de $\mathcal{A}(G)$ basta construir la tabla de multiplicación de G colocando en las filas los elementos $g_1,\ldots,g_n$ y en las columnas sus *inversos* $g_1^{-1},\ldots,g_n^{-1}$ (en el mismo orden). En virtud del comentario anterior, la matriz de $D_{\mathrm{R}}(g)$ tendrá un 1 en las posiciones en que aparezca g en dicha tabla de multiplicación, y cero en las demás posiciones.

• De lo anterior se sigue que la representación regular de un grupo finito proporciona esencialmente la misma información que su tabla de multiplicación. De hecho, veremos más adelante que la representación regular contiene todas las representaciones irreducibles inequivalentes de un grupo finito, en un sentido que será precisado más adelante.

Ejemplo 1.85. Sea $G = C_4 = \{e, a, a^2, a^3\}$, de modo que
$$g_1^{-1} = e\,, \qquad g_2^{-1} = a^3\,, \qquad g_3^{-1} = a^2\,, \qquad g_4^{-1} = a\,.$$
La tabla de multiplicación es por tanto

	e	a^3	a^2	a
e	e	a^3	a^2	a
a	a	e	a^3	a^2
a^2	a^2	a	e	a^3
a^3	a^3	a^2	a	e

Luego (identificando $D_{\mathrm{R}}(g)$ con su matriz en la base canónica de $\mathcal{A}(G)$) $D_{\mathrm{R}}(e) = \mathbb{1}$ y

$$D_{\mathrm{R}}(a) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad D_{\mathrm{R}}(a^2) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad D_{\mathrm{R}}(a^3) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Ejercicio 15. Sea G un grupo finito. Probar que para todo $g \in G$ la matriz de $D_{\mathrm{R}}(g)$ en la base canónica de $\mathcal{A}(G)$ es *ortogonal* (y, por tanto, *unitaria*, ya que es real).

Solución. En efecto, si $g \in G$ se tiene

$$[D_{\mathrm{R}}(g)]_{ij} = 1 \iff g = g_i g_j^{-1} \iff g^{-1} = g_j g_i^{-1} \iff [D_{\mathrm{R}}(g^{-1})]_{ji} = 1\,,$$

y por tanto, al ser D_{R} representación,

$$D_{\mathrm{R}}(g)^{\mathsf{T}} = D_{\mathrm{R}}(g^{-1}) = D_{\mathrm{R}}(g)^{-1}\,.$$

1.5.3 Equivalencia de representaciones

Definición 1.86. Dadas dos representaciones $D_i : G \to \mathrm{GL}(V_i)$ $(i = 1, 2)$ de un grupo G, se dice que D_1 es **equivalente** a D_2 si existe un aplicación lineal *invertible* (es decir, un *isomorfismo lineal*) $A : V_1 \to V_2$ tal que

$$D_2(g) = A D_1(g) A^{-1}\,, \qquad \forall g \in G\,.$$

• Nótese que, al ser la aplicación lineal A en la definición anterior *invertible*, los espacios vectoriales V_1 y V_2 son necesariamente *isomorfos*, y en particular $\dim V_1 = \dim V_2$.

• La equivalencia de representaciones es claramente una *relación de equivalencia*.

• Si dos representaciones $D_i : G \to \mathrm{GL}(V_i)$ $(i = 1, 2)$ son equivalentes, *para todo $g \in G$ los operadores $D_1(g)$ y $D_2(g)$ están representados por la* misma *matriz en sendas bases de los correspondientes espacios vectoriales V_1 y V_2*. Más concretamente, sea $\mathcal{B}_1 = \{v_1, \ldots, v_n\}$ una base cualquiera de V_1, y denotemos por $\big(d_{ij}(g)\big)_{1 \leqslant i,j \leqslant n}$ la matriz de $D_1(g)$ respecto de la base $\mathcal{B}_1$, de modo que

$$D_1(g) v_j = \sum_{i=1}^{n} d_{ij}(g) v_i\,, \qquad j = 1, \ldots, n\,.$$

El conjunto $\mathcal{B}_2 = \{u_1, \ldots, u_n\}$, donde $u_i \equiv A v_i$, es una base de V_2, al ser A un isomorfismo lineal. La acción de $D_2(g)$ sobre esta base está dada por

$$D_2(g) u_j = A D_1(g) A^{-1} \cdot A v_j = A \cdot D_1(g) v_j = \sum_{i=1}^{n} d_{ij}(g) A v_i \equiv \sum_{i=1}^{n} d_{ij}(g) u_i\,.$$

Por tanto $\big(d_{ij}(g)\big)_{1 \leqslant i,j \leqslant n}$ es también la matriz de $D_2(G)$ en la base $\mathcal{B}_2$ de V_2, como habíamos afirmado.

• El objetivo fundamental de la teoría de representaciones es construir *todas* las representaciones de un grupo G *módulo equivalencia*. En otras palabras, se trata de encontrar *todas* las clases de equivalencia distintas en el conjunto de las representaciones de G, preferiblemente señalando un *representante canónico* en cada clase.

Ejemplo 1.87. Una representación de dimensión 1 solo puede ser equivalente a sí misma, ya que el producto de números complejos es conmutativo. En otras palabras, *dos representaciones distintas de dimensión* 1 *son necesariamente inequivalentes.* Lo mismo ocurre con las representaciones n-dimensionales cuyas matrices son *proporcionales a la matriz unidad.*

1.5.4 Representación compleja conjugada. Representaciones reales

Si $D : G \to \mathrm{GL}(n,\mathbb{C})$ es una representación matricial de un grupo G, definimos la aplicación $\overline{D} : G \to M_n(\mathbb{C})$ mediante

$$\overline{D}(g) = \overline{D(g)}\,, \qquad \forall g \in G\,.$$

Es fácil probar que $\overline{D}$ es una representación (matricial) de G, ya que:

1. $\det(\overline{D}(g)) \equiv \det(\overline{D(g)}) = \overline{\det(D(g))} \neq 0 \implies \overline{D}(g) \in \mathrm{GL}(n,\mathbb{C})\,, \qquad \forall g \in G\,.$
2. $\overline{D}(gh) = \overline{D(gh)} = \overline{D(g)D(h)} = \overline{D(g)}\;\overline{D(h)} \equiv \overline{D}(g)\,\overline{D}(h)\,, \qquad \forall g,h \in G\,.$

La representación $\overline{D}$ se denomina **compleja conjugada** de D. Una representación es **real** si es *equivalente* a su compleja conjugada.

Ejemplo 1.88. Consideremos la representación de definición del grupo $G = \mathrm{U}(n)$, que como hemos visto está dada por

$$D(U) = U\,, \qquad \forall U \in \mathrm{U}(n)\,.$$

¿Es esta representación real? Si lo fuera, existiría una matriz $A \in \mathrm{GL}(n,\mathbb{C})$ tal que

$$\overline{U} = AUA^{-1}\,, \qquad \forall U \in \mathrm{U}(n)\,.$$

Pero esta igualdad *no* puede ser cierta para *toda* matriz $U \in \mathrm{U}(n)$, porque tomando $U = \mathrm{i}\mathbb{1}$ se obtendría la contradicción $-\mathrm{i} = \mathrm{i}$. Por tanto la representación de definición de $\mathrm{U}(n)$ *no* es real.

Si $n \geqslant 3$, el mismo argumento vale para la representación de definición de $\mathrm{SU}(n)$, tomando en lugar de $\mathrm{i}\mathbb{1}$ la matriz $\mathrm{e}^{2\pi\mathrm{i}/n}\mathbb{1}$.

Ejercicio 16. Probar que la representación de definición de SU(2) es real.

Solución. Una matriz $U \in \mathrm{SU}(2)$ es de la forma

$$U = \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix}, \qquad \text{con } a,b \in \mathbb{C}\,, \quad |a|^2 + |b|^2 = 1 \tag{1.8}$$

(ejercicio). Necesitamos encontrar una matriz

$$C \equiv \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix}$$

tal que

$$\overline{U}C = CU \tag{1.9}$$

para *toda* matriz de la forma (1.8). Tomando $a = \mathrm{i}$, $b = 0$ en (1.8) se obtiene la matriz

$$U = \mathrm{i}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

que sustituida en (1.9) proporciona

$$-\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix} = \begin{pmatrix} -c_1 & -c_2 \\ c_3 & c_4 \end{pmatrix} = \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} c_1 & -c_2 \\ c_3 & -c_4 \end{pmatrix} \implies c_1 = c_4 = 0\,.$$

Tomando ahora $a = 0$, $b = 1$ en (1.8) se obtiene

$$U = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{1.10}$$

lo que conduce a la condición

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & c_2 \\ c_3 & 0 \end{pmatrix} = \begin{pmatrix} c_3 & 0 \\ 0 & -c_2 \end{pmatrix} = \begin{pmatrix} 0 & c_2 \\ c_3 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -c_2 & 0 \\ 0 & c_3 \end{pmatrix} \implies c_3 = -c_2\,.$$

Por tanto las únicas matrices que pueden cumplir (1.9) son las proporcionales a

$$C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{1.11}$$

Un cálculo elemental demuestra que, en efecto, la matriz (1.11) (y, por tanto, cualquier matriz proporcional a ella) satisface efectivamente (1.9) para toda matriz U de la forma (1.8). Por tanto la representación de definición de SU(2) es equivalente a su compleja conjugada, como habíamos afirmado.

Ejercicio 17. Sea $D : G \to \mathrm{GL}(n, \mathbb{C})$ una representación de un grupo G. a) Probar que si D es equivalente a una representación D' tal que $D'(g) \in \mathrm{GL}(n, \mathbb{R})$, para todo $g \in G$, entonces D es real. b) Demostrar que la representación de definición de SU(2) es real pero no cumple la condición del apartado a).

Solución. a) Por hipótesis, existen $A \in \mathrm{GL}(n, \mathbb{C})$ y una representación $D' : G \to \mathrm{GL}(n, \mathbb{C})$ con $\overline{D'} = D'$ tales que $D = AD'A^{-1}$. Pero entonces

$$\overline{D} = \overline{A}D'\overline{A^{-1}} = \overline{A}D'(\overline{A})^{-1} = (\overline{A}A^{-1})D(\overline{A}A^{-1})^{-1},$$

y por tanto D es real.
b) Si la representación de definición de SU(2) satisficiera la condición del apartado a) existiría una matriz invertible $A \in \mathrm{GL}(2, \mathbb{C})$ tal que

$$\overline{U} = (\overline{A}A^{-1})U(\overline{A}A^{-1})^{-1}, \qquad \forall U \in \mathrm{SU}(2).$$

En virtud del ejercicio anterior,

$$\overline{A}A^{-1} = \lambda\, C, \qquad \text{con } C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \lambda \in \mathbb{C},$$

lo que conduce inmediatamente a una contradicción:

$$\overline{A} = \lambda CA \quad \Longrightarrow \quad A = \overline{\lambda} C \overline{A} = |\lambda|^2 C^2 A = -|\lambda|^2 A \quad \Longrightarrow \quad |\lambda^2| = -1.$$

1.5.5 Suma y producto directos de representaciones

Suma directa de representaciones

Recordemos que un espacio vectorial V es la *suma directa* de dos de sus subespacios V_1 y V_2 si todo vector $v \in V$ se puede expresar *de manera única* como una suma $v_1 + v_2$, con $v_i \in V_i$. Escribiremos en tal caso $V = V_1 \oplus V_2$. Equivalentemente,

$$V = V_1 \oplus V_2 \quad \Longleftrightarrow \quad V = V_1 + V_2, \;\; V_1 \cap V_2 = \{0\}.$$

Si $V = V_1 \oplus V_2$ y $\mathcal{B}_i$ es una base de V_i $(i = 1, 2)$, es fácil ver que $\mathcal{B}_1 \cup \mathcal{B}_2$ es una base de V. Cuando V es de dimensión finita —como supondremos a lo largo de toda esta sección—, esto implica que

$$\dim(V_1 \oplus V_2) = \dim V_1 + \dim V_2.$$

La definición de suma directa se extiende fácilmente a un número *finito* cualquiera de subespacios. Un caso particular importante es aquél en que V es un *espacio con producto escalar* (complejo)[8] $(\cdot\,,\cdot)$. Si W es un subespacio de V, denotaremos por $W^\perp$ a su *complemento ortogonal*, definido por

$$W^\perp = \left\{ v \in V \;\middle|\; (v, w) = 0, \;\; \forall w \in W \right\}.$$

[8] Recordemos que un producto escalar $(\cdot\,,\cdot)$ en un espacio vectorial *complejo* V es una aplicación de $V \times V$ en $\mathbb{C}$ que es sesquilineal y definida positiva, es decir:

1. $(z, ax + by) = a(z, x) + b(z, y)$, $\quad (ax + by, z) = \overline{a}(x, z) + \overline{b}(y, z)$, $\quad \forall x, y, z \in V$, $\;\forall a, b \in \mathbb{C}$.
2. Para todo $z \in \mathbb{C} \setminus \{0\}$, $(z, z) \equiv \|z\|^2 > 0$.

Es fácil probar entonces que

$$V = W \oplus W^{\perp} .$$

Dados dos espacios vectoriales V_1 y V_2 sobre un mismo cuerpo $\mathbb{F}$, el producto cartesiano $V_1 \times V_2$ puede dotarse de una estructura de espacio vectorial definiendo

$$(v_1, v_2) + (w_1, w_2) = (v_1 + w_1, v_2 + w_2) , \qquad \lambda(v_1, v_2) = (\lambda v_1, \lambda v_2) ,$$

donde $v_i , w_i \in V_i$ y $\lambda \in \mathbb{F}$. Es inmediato comprobar que los conjuntos $V_1' \equiv V_1 \times \{0\}$ y $V_2' \equiv \{0\} \times V_2$ son subespacios de $V_1 \times V_2$, isomorfos a los respectivos espacios vectoriales de partida V_1 y V_2. Es también fácil demostrar que $V_1 \times V_2 = V_1' \oplus V_2'$, de donde se deduce que $V_1 \times V_2 \approx V_1 \oplus V_2$. Por tanto a partir de ahora denotaremos (con un ligero abuso de notación) $V_1 \times V_2$ por $V_1 \oplus V_2$, y lo llamaremos también la **suma directa** de los espacios vectoriales V_1 y V_2. Al igual que antes, lo anterior se extiende fácilmente a la suma directa $V_1 \oplus \cdots \oplus V_m$ de un número *finito* arbitrario de espacios vectoriales V_i.

Si V_1 y V_2 son dos espacios vectoriales sobre un mismo cuerpo $\mathbb{F}$, y $A_i : V_i \to V_i$ $(i = 1, 2)$ son sendas aplicaciones lineales, se define su **suma directa** $A_1 \oplus A_2 : V_1 \oplus V_2 \to V_1 \oplus V_2$ mediante

$$(A_1 \oplus A_2)(v_1, v_2) = (A_1 v_1, A_2 v_2) .$$

Es fácil ver que $A_1 \oplus A_2$ es una *aplicación lineal*. La extensión de la definición anterior a la suma directa de m aplicaciones lineales $A_i : V_i \to V_i$ es inmediata:

$$(A_1 \oplus \cdots \oplus A_m)(v_1, \ldots, v_m) = (A_1 v_1, \ldots, A_m v_m) .$$

Análogamente, si un espacio vectorial V es la suma directa $V_1 \oplus \cdots \oplus V_m$ de sus subespacios V_i $(i = 1, \ldots, m)$, y $A_i : V_i \to V_i$ es una aplicación lineal para todo $i = 1, \ldots, m$, la suma directa de las aplicaciones A_i es la aplicación $A_1 \oplus \cdots \oplus A_m : V \to V$ definida por

$$(A_1 \oplus \cdots \oplus A_m)(v_1 + \cdots + v_m) = A_1 v_1 + \cdots + A_m v_m .$$

(Recuérdese que si $V = V_1 \oplus \cdots \oplus V_m$ entonces todo vector $v \in V$ se escribe de manera *única* como una suma $v_1 + \cdots + v_m$, con $v_i \in V_i$.) De nuevo, es inmediato comprobar que la aplicación $A_1 \oplus \cdots \oplus A_m$ es *lineal*.

Una propiedad de la suma directa de aplicaciones lineales (en cualquiera de los dos sentidos anteriores, ligeramente distintos pero equivalentes) que utilizaremos frecuentemente es la siguiente. Sea $A = A_1 \oplus \cdots \oplus A_m$ la suma directa de m operadores lineales $A_i : V_i \to V_i$, y sea $\mathcal{B}_i$ una base de V_i para cada $i = 1, \ldots, m$. Si denotamos por $A_{\mathcal{B}_i}$ la matriz de A_i en la base $\mathcal{B}_i$, es inmediato comprobar que la matriz de A en la base $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_m$ de $V_1 \oplus \cdots \oplus V_m$ está dada por

$$A_{\mathcal{B}} = \begin{pmatrix} A_{\mathcal{B}_1} & 0 & \cdots & 0 \\ 0 & A_{\mathcal{B}_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{\mathcal{B}_m} \end{pmatrix} .$$

En particular, nótese que la matriz $A_{\mathcal{B}}$ es *diagonal por bloques*.

Si G es un grupo, y $D_i : G \to \mathrm{GL}(V_i)$ $(i = 1, 2)$ son dos representaciones de G que actúan en sendos espacios vectoriales (complejos, de dimensión finita) V_i, definimos su **suma directa** $D_1 \oplus D_2 : G \to \mathcal{L}(V_1 \oplus V_2)$ mediante

$$(D_1 \oplus D_2)(g) = D_1(g) \oplus D_2(g) , \qquad \forall g \in G .$$

En otras palabras, para cada $g \in G$ la aplicación $(D_1 \oplus D_2)(g)$ se define por

$$(D_1 \oplus D_2)(g) \cdot (v_1, v_2) = (D_1(g) v_1, D_2(g) v_2) .$$

Es fácil ver que $D_1 \oplus D_2 \equiv D$ es una *representación* de G, de dimensión

$$\dim(D_1 \oplus D_2) = \dim D_1 + \dim D_2 .$$

En efecto:

1. $D(g)$ es invertible para todo $g \in G$. En efecto, es inmediato comprobar que

$$D(g)^{-1} = D_1(g)^{-1} \oplus D_2(g)^{-1} .$$

2. Si $g, h \in G$ se tiene, al ser D_i representación:

$$\begin{aligned} D(gh)(v_1, v_2) &= (D_1(gh)v_1, D_2(gh)v_2) = (D_1(g)D_1(h)v_1, D_2(g)D_2(h)v_2) \\ &= D(g)(D_1(h)v_1, D_2(h)v_2) = D(g)D(h)(v_1, v_2) . \end{aligned}$$

Por lo visto anteriormente, si $\mathcal{B}_i$ es una base de V_i ($i = 1, 2$), y denotamos por $\mathcal{D}_i(g)$ la matriz de $D_i(g)$ respecto de la correspondiente base $\mathcal{B}_i$, entonces la matriz $\mathcal{D}(g)$ de $D_1(g) \oplus D_2(g)$ en la base $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ de $V = V_1 \oplus V_2$ es diagonal por bloques, y está dada por

$$\mathcal{D}(g) = \begin{pmatrix} \mathcal{D}_1(g) & 0 \\ 0 & \mathcal{D}_2(g) \end{pmatrix} .$$

De forma análoga se define la suma directa $D_1 \oplus \cdots \oplus D_m$ de m representaciones $D_i : G \to \mathrm{GL}(V_i)$ de G. Si $D = D_1 \oplus \cdots \oplus D_m$ y $\mathcal{B}_i$ es una base del espacio vectorial V_i para $i = 1, \ldots, m$, la matriz de $D(g)$ en la base $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_m$ de $V = V_1 \oplus \cdots \oplus V_m$ está dada por

$$\mathcal{D}(g) = \begin{pmatrix} \mathcal{D}_1(g) & 0 & \cdots & 0 \\ 0 & \mathcal{D}_2(g) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathcal{D}_m(g) \end{pmatrix} .$$

Nótese, en particular, que

$$\dim(D_1 \oplus \cdots \oplus D_m) = \dim D_1 + \cdots + \dim D_m .$$

La construcción anterior es aún más sencilla en el caso de representaciones *matriciales*. En efecto, si $D_i : G \to \mathrm{GL}(n_i, \mathbb{C})$ es una representación matricial de G para $i = 1, \ldots, m$, la suma directa $D_1 \oplus \cdots \oplus D_m$ de estas representaciones es la representación $D_1 \oplus \cdots \oplus D_m : G \to \mathrm{GL}(n, \mathbb{C})$ (con $n \equiv n_1 + \cdots + n_m$) definida por

$$(D_1 \oplus \cdots \oplus D_m)(g) = \begin{pmatrix} D_1(g) & 0 & \cdots & 0 \\ 0 & D_2(g) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_m(g) \end{pmatrix} .$$

Producto directo de representaciones

Sean, de nuevo, V_1 y V_2 dos espacios vectoriales sobre un mismo cuerpo $\mathbb{F}$, de dimensiones respectivas n_1 y n_2. Si $\mathcal{B}_1 = \{u_1, \ldots, u_{n_1}\}$ y $\mathcal{B}_2 = \{v_1, \ldots, v_{n_2}\}$ son sendas bases de V_1 y V_2, respectivamente, denotemos por

$$u_i \otimes v_j , \qquad i = 1, \ldots, n_1 , \quad j = 1, \ldots, n_2 ,$$

a los vectores de la base canónica de $\mathbb{F}^{n_1 n_2}$, ordenados (por ejemplo) siguiendo el *orden lexicográfico*:

$$u_1 \otimes v_1, \ldots, u_1 \otimes v_{n_2}, \ldots, u_{n_1} \otimes v_1, \ldots, u_{n_1} \otimes v_{n_2} . \tag{1.12}$$

Dados dos vectores

$$u = \sum_{i=1}^{n_1} a_i u_i \equiv \sum_i a_i u_i \in V_1 , \qquad v = \sum_{j=1}^{n_2} b_j u_j \equiv \sum_j b_j v_j \in V_2 ,$$

definimos su **producto directo** (**tensorial**, **de Kronecker**) $u \otimes v$ mediante la fórmula

$$u \otimes v = \sum_{i,j} a_i b_j u_i \otimes v_j \in \mathbb{F}^{n_1 n_2} .$$

El espacio $\mathbb{F}^{n_1 n_2}$ con base canónica (1.12) se denomina **producto directo** de V_1 y V_2, y se denota por $V_1 \otimes V_2$. Evidentemente la definición anterior *no* es *canónica* si V_1 y V_2 son espacios vectoriales arbitrarios sobre $\mathbb{F}$, ya que dicha definición depende de la elección de las bases en V_1 y V_2. Sí lo es (módulo la ordenación de los vectores de la base (1.12)), sin embargo, en el caso en que $V_i = \mathbb{F}^{n_i}$, $i = 1, 2$. En cualquier caso, es evidente que todos los espacios $V_1 \otimes V_2$ definidos de esta forma son *isomorfos* (tienen la misma dimensión).

Dados dos operadores lineales $A : V_1 \to V_1$ y $B : V_2 \to V_2$, definimos su **producto directo** $A \otimes B : V_1 \otimes V_2 \to V_1 \otimes V_2$ mediante la fórmula

$$(A \otimes B)(u_i \otimes v_j) = (Au_i) \otimes (Bv_j) , \qquad i = 1, \ldots, n_1 , \quad j = 1, \ldots, n_2 ,$$

extendiendo esta definición por linealidad a todo el espacio $V_1 \otimes V_2$. El producto directo $A \otimes B$ es por tanto un *operador lineal* en $V_1 \otimes V_2$. Sean

$$(a_{ik})_{i,k=1,\ldots,n_1} , \qquad (b_{jl})_{j,l=1,\ldots,n_2}$$

las respectivas matrices de los operadores A y B en las bases $\mathcal{B}_1$ y $\mathcal{B}_2$ utilizadas para definir el producto directo $V_1 \otimes V_2$. Entonces se tiene

$$(A \otimes B)(u_k \otimes v_l) = (Au_k) \otimes (Bv_l) = \Big(\sum_i a_{ik} u_i \Big) \otimes \Big(\sum_j b_{jl} v_j \Big) \equiv \sum_{i,j} a_{ik} b_{jl} u_i \otimes v_j ,$$

y por tanto los elementos de matriz de $A \otimes B$ en la base canónica (1.12) de $V_1 \otimes V_2$ están dados por

$$(A \otimes B)_{ij,kl} = a_{ik} b_{jl} . \tag{1.13}$$

Es fácil ver que, si ordenamos estos $(n_1 n_2)^2$ elementos de matriz siguiendo el orden lexicográfico $(1,1), \ldots, (1,n_2), \ldots, (n_1,1), \ldots, (n_1,n_2)$, la matriz de $A \otimes B$ en la base (1.12) se puede escribir por bloques en la forma

$$\begin{pmatrix} a_{11}B & \cdots & a_{1n_1}B \\ a_{21}B & \cdots & a_{2n_1}B \\ \vdots & \cdots & \vdots \\ a_{n_1,1}B & \cdots & a_{n_1n_1}B \end{pmatrix} , \tag{1.14}$$

donde (con un ligero abuso de notación) B denota la matriz del operador B en la base $\mathcal{B}_2$.

Si $D_i : G \to \mathrm{GL}(V_i)$ $(i = 1, 2)$ son dos representaciones de un grupo G, se define su **producto directo** como la representación $D_1 \otimes D_2 : G \to \mathrm{GL}(V_1 \otimes V_2)$ dada por

$$(D_1 \otimes D_2)(g) = D_1(g) \otimes D_2(g) , \qquad \forall g \in G .$$

En otras palabras, $(D_1 \otimes D_2)(g)$ es el operador lineal cuya actuación sobre la base canónica de $V_1 \otimes V_2$ está dada por

$$(D_1 \otimes D_2)(g) \cdot (u_i \otimes v_j) = \big(D_1(g)u_i\big) \otimes \big(D_2(g)v_j\big) .$$

Es inmediato comprobar que $D_1 \otimes D_2$ es efectivamente una representación de G de dimensión

$$\dim(D_1 \otimes D_2) = \dim D_1 \cdot \dim D_2 ,$$

ya que:

1. $D_1(g) \otimes D_2(g)$ es invertible para todo $g \in G$, ya que

$$\big[D_1(g) \otimes D_2(g)\big]^{-1} = D_1(g)^{-1} \otimes D_2(g)^{-1} .$$

2. Para todo $g, h \in G$, si $D = D_1 \otimes D_2$ se tiene (al ser D_1 y D_2 representaciones de G)

$$\begin{aligned} D(gh)(u_i \otimes v_j) &\equiv \big(D_1(gh)u_i\big) \otimes \big(D_2(gh)v_j\big) = \big(D_1(g)D_1(h)u_i\big) \otimes \big(D_2(g)D_2(h)v_j\big) \\ &\equiv D(g)\big[\big(D_1(h)u_i\big) \otimes \big(D_2(h)v_j\big)\big] \equiv D(g)D(h)(u_i \otimes v_j)\,. \end{aligned}$$

Por tanto

$$(D_1 \otimes D_2)(gh) = (D_1 \otimes D_2)(g) \cdot (D_1 \otimes D_2)(h)\,,$$

ya que ambos miembros coinciden sobre la base canónica de $V_1 \otimes V_2$.

Del mismo modo, si $A \in M_{n_1}(\mathbb{C})$ y $B \in M_{n_2}(\mathbb{C})$, definimos su **producto directo** $A \otimes B \in M_{n_1 n_2}(\mathbb{C})$ como la matriz dada por la ec. (1.14). Dadas dos representaciones *matriciales* $D_i : G \to \mathrm{GL}(n_i, \mathbb{C})$ $(i = 1, 2)$ de un grupo G, definimos su **producto directo** $D_1 \otimes D_2 : G \to \mathrm{GL}(n_1 n_2, \mathbb{C})$ mediante

$$(D_1 \otimes D_2)(g) = D_1(g) \otimes D_2(g)\,.$$

Como en el caso de las representaciones lineales, es inmediato comprobar que $D_1 \otimes D_2$ es una representación matricial de G de dimensión $n_1 n_2$.

1.5.6 Representaciones irreducibles

Si $D : G \to \mathrm{GL}(V)$ es una representación de un grupo G, diremos que un subespacio $W \subset V$ es **invariante** bajo D si

$$D(g)W \subset W\,, \qquad \forall g \in G\,,$$

o, equivalentemente,

$$D(g)w \in W\,, \qquad \forall g \in G\,, \forall w \in W\,.$$

La representación D se dice **reducible** si posee algún subespacio *propio* invariante, e **irreducible** en caso contrario.

Ejemplo 1.89. Toda representación unidimensional es irreducible, ya que un espacio vectorial de dimensión uno no posee subespacios propios.

Supongamos que $D : G \to \mathrm{GL}(V)$ es una representación reducible de un grupo G, y sea $W_1 \subset V$ un subespacio no trivial invariante bajo D. Sea W_2 un subespacio complementario de W_1, es decir un subespacio cualquiera de V tal que $V = W_1 \oplus W_2$. Si $\mathcal{B}_1$ y $\mathcal{B}_2$ son sendas bases de W_1 y W_2, y $g \in G$, la matriz de $D(g)$ en la base $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ de V es claramente de la forma

$$\mathcal{D}(g) = \begin{pmatrix} \mathcal{D}_1(g) & A(g) \\ 0 & \mathcal{D}_2(g) \end{pmatrix},$$

siendo $A(g)$ una cierta matriz $n_1 \times n_2$. Es fácil ver que $\mathcal{D}_i : G \to \mathrm{GL}(n_i, \mathbb{C})$ $(i = 1, 2)$ es una representación de G de dimensión n_i, al ser

$$\mathcal{D}(g)\mathcal{D}(h) = \begin{pmatrix} \mathcal{D}_1(g)\mathcal{D}_1(h) & \mathcal{D}_1(g)A(h) + A(g)\mathcal{D}_2(h) \\ 0 & \mathcal{D}_2(g)\mathcal{D}_2(h) \end{pmatrix}.$$

Lo anterior es cierto cualquiera que sea el subespacio complementario W_2 que tomemos. El caso más ventajoso es aquél en que W_2 es también *invariante* bajo D, ya que entonces $A(g) = 0$ para todo $g \in G$ y

$$\mathcal{D}(g) = \begin{pmatrix} \mathcal{D}_1(g) & 0 \\ 0 & \mathcal{D}_2(g) \end{pmatrix},$$

es diagonal por bloques. En tal caso $V = W_1 \oplus W_2$, con W_1 y W_2 invariantes bajo D, y $D = D_1 \oplus D_2$, siendo $D_i : G \to \mathrm{GL}(W_i)$ la restricción de D a W_i:

$$D_i(g)w_i = D(g)w_i \in W_i\,, \qquad \forall g \in G\,, \forall w_i \in W_i\,.$$

Nótese que en este caso la representación D *contiene exactamente la misma información* que las representaciones D_1 y D_2 por separado, ya que D_1 y D_2 determinan D (y viceversa). Esto *no* es así en el caso en que W_2 *no* es invariante bajo D, ya que entonces es necesario conocer las matrices $A(g)$ para reconstruir $\mathcal{D}$, además de $\mathcal{D}_1$ y $\mathcal{D}_2$.

Definición 1.90. Una representación $D : G \to \mathrm{GL}(V)$ de un grupo G es **descomponible** si V es la *suma directa* de dos subespacios propios W_i $(i = 1, 2)$ *invariantes* bajo D.

Nótese que si D es una representación descomponible entonces $D = D_1 \oplus D_2$, siendo D_i $(i = 1, 2)$ la restricción de D a cada uno de los subespacios invariantes W_i. Obviamente, toda representación descomponible es automáticamente reducible, pero el recíproco no es cierto en general (cf. el Ejemplo 1.92).

Definición 1.91. La representación D es **completamente reducible** si $D = D_1 \oplus \cdots \oplus D_m$, siendo cada D_i una representación *irreducible* de G.

En otras palabras, D *es completamente reducible si se descompone en suma directa de representaciones irreducibles.* En otras palabras, D es completamente reducible si $V = W_1 \oplus \cdots \oplus W_m$ es la suma directa de m subespacios W_i invariantes bajo la representación D, siendo la restricción de D a cada uno de estos subespacios una representación irreducible. Nótese, en particular, que *una representación irreducible es completamente reducible* (caso $m = 1$ de la definición).

Ejemplo 1.92. Sea $G = \mathbb{C}$ el grupo *aditivo* de los números complejos, y sea $D : G \to M_2(\mathbb{C})$ la aplicación definida por

$$D(z) = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}, \qquad \forall z \in \mathbb{C}\,.$$

Esta aplicación es claramente una *representación*, ya que $D(0) = \mathbb{1}$ y

$$D(z_1)D(z_2) = \begin{pmatrix} 1 & z_1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & z_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & z_1 + z_2 \\ 0 & 1 \end{pmatrix} \equiv D(z_1 \cdot z_2)\,, \qquad \forall z_1, z_2 \in \mathbb{C}$$

(recuérdese que en este caso $e = 0$ y $z_1 \cdot z_2 = z_1 + z_2$). Esta representación es claramente *reducible*, ya que $W = \mathrm{lin}\{(1, 0)\}$ es claramente invariante bajo $D(z)$ para todo $z \in \mathbb{C}$. Sin embargo, D *no* es descomponible. En efecto, si $W' \subset \mathbb{C}^2$ es un subespacio complementario de W entonces

$$W' = \mathrm{lin}\{(a, 1)\}\,, \qquad \text{con } a \in \mathbb{C}\,,$$

o equivalentemente

$$W' = \left\{(z_1, z_2) \in \mathbb{C}^2 \,\middle|\, z_1 = az_2\right\}, \qquad \text{con } a \in \mathbb{C}\,.$$

Pero en tal caso

$$D(z)(a, 1) = (a + z, 1) \in W' \iff a + z = a \iff z = 0\,.$$

□

Nótese que, al no ser la representación del ejemplo anterior descomponible, tampoco es completamente reducible. Por tanto, en general hay representaciones reducibles que *no* son completamente reducibles.

Definición 1.93. Una representación matricial $D : G \to \mathrm{GL}(n, \mathbb{C})$ es **unitaria** si

$$D(g) \in \mathrm{U}(n)\,, \qquad \forall g \in G\,.$$

Ejemplo 1.94. La representación $D : S^1 \to \mathrm{GL}(2, \mathbb{C})$ definida por

$$D(\mathrm{e}^{\mathrm{i}\theta}) = R(\theta) \equiv \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}, \qquad \forall \theta \in \mathbb{R}$$

es claramente unitaria, ya que $R(\theta) \in \mathrm{SO}(2, \mathbb{R}) \subset \mathrm{U}(2)$. La representación del Ejemplo 1.92 *no* lo es, ya que

$$\begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -z \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}^{\dagger} = \begin{pmatrix} 1 & 0 \\ \overline{z} & 1 \end{pmatrix} \iff z = 0\,.$$

Ejemplo 1.95. En virtud del Ejercicio 15, la representación regular de cualquier grupo finito es unitaria.

El concepto de representación matricial unitaria se extiende fácilmente a representaciones *lineales* $D : G \to \mathrm{GL}(V)$ si en el espacio vectorial complejo V hay definido un *producto escalar* (complejo). En efecto, recordemos que si V está dotado de un producto escalar y $A : V \to V$ es un operador lineal, se define su *adjunto* $A^\dagger : V \to V$ como el único operador lineal que satisface la relación

$$(v, Aw) = (A^\dagger v, w)\,, \qquad \forall v, w \in V\,.$$

El operador lineal A es **unitario** si $A^\dagger = A^{-1}$. Es fácil probar que

$$A \text{ unitario} \iff (Av, Aw) = (v, w)\,, \quad \forall v, w \in V. \tag{1.15}$$

(En efecto, si A es unitario sustituyendo v por Av en la igualdad $(v, Aw) = (A^{-1}v, w)$, válida para todo $v, w \in V$, se obtiene la ecuación (1.15). Recíprocamente, si se cumple (1.15) entonces A es invertible, ya que

$$Av = 0 \implies (v, w) = 0\,, \quad \forall w \in V \implies v = 0\,.$$

Sustituyendo de nuevo v por $A^{-1}v$ en (1.15) se obtiene $A^\dagger = A^{-1}$.)

Definición 1.96. Una representación $D : G \to \mathrm{GL}(V)$ de un grupo G se denomina **unitaria** si $D(g)$ es un operador lineal unitario para todo $g \in G$.

La propiedad fundamental de las representaciones unitarias es la siguiente:

Proposición 1.97. *Sea $D : G \to \mathrm{GL}(V)$ una representación* unitaria *de un grupo G. Si D es reducible, entonces D es descomponible.*

Demostración. En efecto, por hipótesis existe un subespacio propio $W \subset V$ tal que

$$D(g)W \subset W\,, \qquad \forall g \in G\,.$$

Para probar que D es descomponible basta probar que $W^\perp$ *es también invariante bajo D,* ya que $V = W \oplus W^\perp$ al ser (por hipótesis) V un espacio con producto escalar. Un vector $v \in V$ pertenece a $W^\perp$ si

$$(w, v) = 0\,, \qquad \forall w \in W\,.$$

Por tanto debemos probar que para todo $g \in G$ se tiene

$$(w, v) = 0\,, \quad \forall w \in W \implies (w, D(g)v) = 0\,, \quad \forall w \in W\,.$$

Y, en efecto, al ser la representación D unitaria (por hipótesis) se tiene

$$(w, D(g)v) = (D(g)^\dagger w, v) = (D(g)^{-1}w, v) = (D(g^{-1})w, v) = 0\,,$$

ya que $D(g^{-1})w \in W$ al ser W invariante bajo D. □

Corolario 1.98. *Toda representación* unitaria *de un grupo G es completamente reducible.*

Demostración. Sea, en efecto, $D : G \to \mathrm{GL}(V)$ una representación unitaria de un grupo G. El enunciado es cierto si la dimensión de D es igual a 1, ya que toda representación unidimensional es irreducible. Procediendo por inducción, supongamos que el enunciado es válido para representaciones unitarias de dimensión menor o igual que un cierto entero positivo n, y sea $\dim D = n + 1$. Si D es irreducible, ya hemos terminado. En caso contrario, de la proposición anterior se deduce que $D = D_1 \oplus D_2$, con $D_i : G \to \mathrm{GL}(W_i)$ $(i = 1, 2)$, $\dim D_i \leqslant n$ y $V = W_1 \oplus W_2$. Es inmediato comprobar que las representaciones D_1 y D_2 siguen siendo unitarias, ya que para todo $u_i, v_i \in W_i$ se tiene (al ser D_i la restricción de D a W_i)

$$\begin{aligned}(u_i, D_i(g)v_i) = \big(u_i, D(g)v_i\big) = \big(D(g)^{-1}u_i, v_i\big) = \big(D(g^{-1})u_i, v_i\big) = \big(D_i(g^{-1})u_i, v_i\big)\\ = \big(D_i(g)^{-1}u_i, v_i\big) \implies D_i(g)^\dagger = D_i(g)^{-1}\,.\end{aligned}$$

Por hipótesis de inducción,

$$D_i = D_{i1} \oplus \cdots \oplus D_{i,m_i}\,, \qquad i = 1, 2\,,$$

con D_{ij} irreducible. Por tanto

$$D = D_{11} \oplus \cdots \oplus D_{1,m_1} \oplus D_{21} \oplus \cdots \oplus D_{2,m_2}$$

se descompone en suma directa de representaciones irreducibles. □

1.5.7 Teoremas de Schur–Auerbach y de Maschke

El grupo G del Ejemplo (1.92) es un grupo *infinito*. Demostraremos en este apartado que *si G es un grupo* finito *entonces toda representación* (no necesariamente unitaria) *de G es completamente reducible*. Para demostrar esta propiedad fundamental de los grupos finitos, empezaremos probando el siguiente resultado previo:

Teorema de Schur–Auerbach. *Si G es un grupo* finito *y $D : G \to \mathrm{GL}(V)$ es una representación de G en un espacio vectorial V con producto escalar, entonces D es equivalente a una representación unitaria.*

Demostración. Dados dos vectores $u, v \in V$ definimos

$$\langle u, v\rangle = \frac{1}{|G|}\sum_{g\in G}\big(D(g)u, D(g)v\big)\,. \tag{1.16}$$

En otras palabras, $\langle\cdot\,,\cdot\rangle$ es el *promedio sobre el grupo* del producto escalar ordinario $(\cdot\,,\cdot)$. Es inmediato comprobar que $\langle\cdot\,,\cdot\rangle$ es un producto escalar, ya que es claramente sesquilineal y verifica

$$\begin{cases} \langle v, v\rangle = \dfrac{1}{|G|}\displaystyle\sum_{g\in G}(D(g)v, D(g)v) = \dfrac{1}{|G|}\sum_{g\in G}\|D(g)v\|^2 \geqslant 0\,, \\[2ex] \langle v, v\rangle = 0 \quad\Longleftrightarrow\quad \|D(g)v\| = 0\,, \quad \forall g \in G \quad\Longrightarrow\quad v = 0\,, \end{cases}$$

al ser $D(g)$ invertible. Es también fácil demostrar que *los operadores $D(g)$ son unitarios respecto del nuevo producto escalar* $\langle\cdot\,,\cdot\rangle$. En efecto, para todo $h \in G$ se tiene

$$\begin{aligned} \big\langle D(h)u, D(h)v\big\rangle &= \frac{1}{|G|}\sum_{g\in G}\big(D(g)D(h)u, D(g)D(h)v\big) = \frac{1}{|G|}\sum_{g\in G}\big(D(gh)u, D(gh)v\big) \\ &= \frac{1}{|G|}\sum_{k\in G}\big(D(k)u, D(k)v\big) = \langle u, v\rangle\,. \end{aligned}$$

Sean $\{u_1, \ldots, u_n\}$ y $\{v_1, \ldots, v_n\}$ sendas *bases ortonormales* de V respecto de los productos escalares $(\cdot\,,\cdot)$ y $\langle\cdot\,,\cdot\rangle$, respectivamente. Definimos entonces el operador lineal $T : V \to V$ mediante

$$Tu_i = v_i\,, \qquad i = 1, \ldots, n\,.$$

Nótese que el operador T es invertible, siendo su inverso el operador lineal $T^{-1} : V \to V$ definido por

$$T^{-1}v_i = u_i\,, \qquad i = 1, \ldots, n\,.$$

Por otra parte, si $x = \displaystyle\sum_{i=1}^{n} x_i u_i$, $y = \displaystyle\sum_{i=1}^{n} y_i u_i$ son dos vectores arbitrarios de V se tiene

$$\begin{aligned} \langle Tx, Ty\rangle &= \sum_{i,j=1}^{n}\overline{x_i}\, y_j \langle Tu_i, Tu_j\rangle = \sum_{i,j=1}^{n}\overline{x_i}\, y_j \langle v_i, v_j\rangle = \sum_{i,j=1}^{n}\overline{x_i}\, y_j \delta_{ij} \\ &= \sum_{i=1}^{n}\overline{x_i}\, y_i = (x, y)\,, \end{aligned} \tag{1.17}$$

y por tanto

$$(T^{-1}x, T^{-1}y) = \langle x, y\rangle\,. \tag{1.18}$$

Veamos, para finalizar, que la representación

$$U = T^{-1}DT\,,$$

equivalente a D por construcción, es unitaria *respecto del producto escalar de partida*. En efecto, utilizando (1.17) y (1.18) y teniendo en cuenta que $D(g)$ es unitario respecto del nuevo producto escalar $\langle\cdot\,,\cdot\rangle$ se obtiene fácilmente

$$\big(U(g)x, U(g)y\big) = \big(T^{-1}D(g)Tx, T^{-1}D(g)Ty\big) = \big\langle D(g)Tx, D(g)Ty\big\rangle = \langle Tx, Ty\rangle = (x, y)\,.$$

□

Nota. Veremos en el Capítulo 3 que el resultado anterior se extiende sin dificultad a un grupo *compacto* cualquiera G. En efecto, la clave de la demostración anterior es la posibilidad de definir un producto escalar promediado sobre el grupo via la ec. (1.16), respecto del cual los operadores $D(g)$ sean unitarios. Esto siempre es posible en un grupo compacto, dado que en tal caso existe una *medida* $\mathrm{d}\mu(g)$ definida en el grupo con la propiedad

$$\int_G f(gh)\,\mathrm{d}\mu(g) = \int_G f(g)\,\mathrm{d}\mu(g)\,, \qquad \forall h \in G\,. \tag{1.19}$$

En tal caso, se define

$$\langle u, v\rangle = \frac{1}{|G|}\int_G \big(D(g)u, D(g)v\big)\,\mathrm{d}\mu(g)\,,$$

donde

$$|G| = \int_G \mathrm{d}\mu(g)$$

es la medida total del grupo (finita, al ser G compacto). Entonces

$$\begin{aligned}\langle D(h)u, D(h)v\rangle &= \frac{1}{|G|}\int_G \big(D(g)D(h)u, D(g)D(h)v\big)\,\mathrm{d}\mu(g) = \frac{1}{|G|}\int_G \big(D(gh)u, D(gh)v\big)\,\mathrm{d}\mu(g)\\ &= \frac{1}{|G|}\int_G \big(D(g)u, D(g)v\big)\,\mathrm{d}\mu(g) = \langle u, v\rangle\,,\end{aligned}$$

donde se ha aplicado la invariancia de la medida (ec. (1.19)) en el penúltimo paso. □

A partir del teorema anterior se demuestra fácilmente el resultado fundamental de esta sección, debido a Maschke:

Teorema de Maschke. *Todas las representaciones de un grupo* finito G *son* completamente reducibles.

Demostración. Sea, en efecto, $D : G \to \mathrm{GL}(V)$ una representación de un grupo finito G. Al ser V un espacio complejo de dimensión finita, es inmediato dotarle de un producto vectorial complejo definiendo

$$(u_i, u_j) = \delta_{ij}\,, \qquad 1 \leqslant i, j \leqslant \dim V \equiv n\,,$$

donde $\{u_1, \ldots, u_n\}$ es una base cualquiera de V. En virtud de la demostración del teorema de Schur–Auerbach, la representación D es unitaria respecto del producto escalar $\langle\cdot\,,\cdot\rangle$ construido promediando $(\cdot\,,\cdot)$ sobre el grupo. Por el Corolario (1.98), la representación D es completamente reducible. □

Nota. Al ser el teorema de Schur–Auerbach válido para grupos *compactos*, también lo es (con la misma demostración) el teorema de Maschke.

1.6 Lemas de Schur

En virtud del teorema de Maschke, para clasificar las representaciones de un grupo *finito* (o *compacto*) G basta considerar sus representaciones *irreducibles*. Uno de los principales objetivos de la teoría de representaciones es, por tanto, el de determinar si una dada representación de un grupo es irreducible. Los criterios más efectivos de irreducibilidad se basan en tres resultados que veremos a continuación, probados por Issai Schur en 1905 y que se conocen en la literatura como lemas de Schur. Comenzaremos con un lema elemental previo, que es la clave para establecer los lemas de Schur:

Lema 1.99. *Sean $D : G \to \mathrm{GL}(V)$ y $D' : G \to \mathrm{GL}(V')$ dos representaciones de un grupo G, y sea $A : V \to V'$ un operador lineal que* ***entrelaza*** *ambas representaciones, es decir tal que*

$$AD(g) = D'(g)A\,, \qquad \forall g \in G\,. \tag{1.20}$$

Entonces los subespacios $\ker A$ *y* $A(V)$ *son invariantes bajo D y D', respectivamente.*

Demostración. En efecto, si $v \in \ker A$ entonces $Av = 0$, y por tanto

$$A \cdot \big(D(g)v\big) = D'(g) \cdot (Av) = 0 \quad \Longrightarrow \quad D(g)v \in \ker A\,, \quad \forall g \in G\,.$$

Análogamente, si $v' = Av \in A(V)$ se tiene

$$D'(g)v' = D'(g) \cdot (Av) = A \cdot \big(D(g)v\big) \in A(V)\,, \qquad \forall g \in G\,.$$

□

Lemas de Schur. *Sean $D : G \to \mathrm{GL}(V)$ y $D' : G \to \mathrm{GL}(V')$ sendas representaciones* irreducibles *de un grupo G entrelazadas por un operador $A : V \to V'$ (cf. la ec. (1.20)). Entonces se verifica:*

1. Si $\dim D \neq \dim D'$ entonces $A = 0$.
2. Si $\dim D = \dim D'$ entonces o bien $A = 0$, o bien D es equivalente a D'.
3. Si $D = D'$ (y, por tanto, $V = V'$) entonces existe $\lambda \in \mathbb{C}$ tal que $A = \lambda I$.

Demostración.
1) Del lema anterior y de la irreducibilidad de las representaciones D y D' se sigue que tanto $\ker A$ como $A(V)$ son triviales, y por tanto[9]

$$\dim \ker A \in \{0, \dim D\}\,, \qquad \dim A(V) \in \{0, \dim D'\}\,. \tag{1.21}$$

Pero, según un resultado bien conocido de Álgebra lineal,

$$\dim \ker A + \dim A(V) = \dim V \equiv \dim D\,. \tag{1.22}$$

Al ser $\dim D' \neq \dim D$, de las ecuaciones (1.21)-(1.22) se sigue que

$$\dim \ker A = \dim D\,, \qquad \dim A(V) = 0 \tag{1.23}$$

o, equivalentemente, $A = 0$.

2) Si $\dim D = \dim D'$, las ecuaciones (1.21)-(1.22) tienen las dos posibles soluciones

$$(\dim \ker A = \dim D\,, \quad \dim A(V) = 0)\,, \qquad (\dim \ker A = 0\,, \quad \dim A(V) = \dim D)\,.$$

[9] Recuérdese que, por definición de dimensión de una representación, $\dim V = \dim D$ y $\dim V' = \dim D'$.

En el primer caso $A = 0$, como en el apartado anterior. En el segundo, A es a la vez inyectivo ($\ker A = \{0\}$) y suprayectivo ($\dim A(V) = \dim D = \dim D' = \dim V'$). Por tanto en este último caso A es invertible, y de la ecuación de entrelazado (1.20) se sigue entonces que

$$D' = ADA^{-1}\,,$$

por lo que D' y D son equivalentes bajo A.

3) Supongamos, por último, que $D = D'$. En tal caso, de (1.20) se sigue que

$$(A-\lambda)D = D(A-\lambda)\,, \qquad \forall \lambda \in \mathbb{C}\,.$$

En virtud del apartado anterior, o bien $A - \lambda = 0$ o bien $A - \lambda$ es invertible. Por otra parte, es bien conocido que todo operador en un espacio vectorial *complejo* posee al menos un autovalor. Si λ es un autovalor de A entonces $\ker(A-\lambda) \neq \{0\}$, y por tanto $A - \lambda$ *no* es invertible, de donde se deduce que $A - \lambda = 0$. □

Ejemplo 1.100. Para la validez del tercer lema de Schur es esencial que el espacio vectorial V sea *complejo* (como, por otra parte, se exige en la Definición 1.71 de representación). En efecto, dicho lema descansa en la existencia de un autovalor para cualquier operador lineal, lo cual *no* es cierto en un espacio vectorial *real*.

Para ilustrar este punto consideremos, por ejemplo, la representación $D : S^1 \to \mathrm{GL}(2,\mathbb{C})$ del Ejemplo 1.77, definida por

$$D\big(\mathrm{e}^{\mathrm{i}\theta}\big) = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \equiv R(\theta)\,, \qquad \forall \theta \in \mathbb{R}\,.$$

Si consideramos a D como una aplicación de S^1 en $\mathrm{GL}(2,\mathbb{R})$, es claro que D no admite subespacios invariantes *reales* no triviales. Sin embargo, no se cumple en este caso la conclusión del tercer lema de Schur, ya que

$$R(\theta) = \cos\theta\,\mathbb{1} + \operatorname{sen}\theta \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \equiv \cos\theta\,\mathbb{1} + \operatorname{sen}\theta\, A \quad \Longrightarrow \quad AR(\theta) = R(\theta)A\,, \qquad \forall \theta \in \mathbb{R}\,.$$

La explicación de esta aparente contradicción es que D *no* es irreducible, ya que posee subespacios invariantes *complejos* no triviales. En efecto, es fácil ver que los autovalores de $R(\theta)$ son $\mathrm{e}^{\pm\mathrm{i}\theta}$, siendo sus respectivos subespacios propios

$$W_\pm = \mathrm{lin}\{(1, \mp\mathrm{i})\}\,.$$

Como estos subespacios *no dependen de* θ, son invariantes bajo *todos* los operadores $R(\theta) = D(\mathrm{e}^{\mathrm{i}\theta})$, y por consiguiente la representación D es *reducible*. De hecho, es claro que D es *completamente reducible*, ya que

$$D(\mathrm{e}^{\mathrm{i}\theta}) = (\mathrm{e}^{\mathrm{i}\theta} I_+) \oplus (\mathrm{e}^{-\mathrm{i}\theta} I_-)\,,$$

siendo $I_\pm$ la identidad en $W_\pm$. (Es claro que las representaciones $\mathrm{e}^{\pm\mathrm{i}\theta} I_\pm$ son *irreducibles*, al ser *unidimensionales*). En notación matricial,

$$D(\mathrm{e}^{\mathrm{i}\theta}) = A \begin{pmatrix} \mathrm{e}^{\mathrm{i}\theta} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\theta} \end{pmatrix} A^{-1}\,, \qquad A = \begin{pmatrix} 1 & 1 \\ -\mathrm{i} & \mathrm{i} \end{pmatrix}.$$

□

De los tres lemas de Schur, el tercero es el más importante desde el punto de vista práctico. Nótese que dicho lema puede formularse de la forma siguiente:

Proposición 1.101. *Si $D : G \to \mathrm{GL}(V)$ es una representación* irreducible *de un grupo G, y $A : V \to V$ es un operador que conmuta con* todos *los operadores $D(g)$, con $g \in G$, entonces A es un múltiplo de la identidad.*

Ejemplo 1.102. El tercer lema de Schur se utiliza frecuentemente para calcular el *centro* de muchos grupos matriciales interesantes. Sea, por ejemplo $G = \mathrm{U}(n)$. Veamos, en primer lugar, que la representación de definición de este grupo es irreducible. En efecto, supongamos que un subespacio $W \subset \mathbb{C}^n$ fuera invariante bajo dicha representación, es decir que

$$U \cdot W \subset W\,, \qquad \forall U \in \mathrm{U}(n)\,,$$

con $1 \leqslant \dim W \equiv k < n$. Si $\{v_1, \ldots, v_k\}$ y $\{v_{k+1}, \ldots, v_n\}$ son sendas bases ortonormales de W y $W^\perp$, respectivamente, al ser $V = W \oplus W^\perp$ el conjunto $\mathcal{B} = \{v_1, \ldots, v_n\}$ es una base ortonormal de $\mathbb{C}^n$. Sea U la matriz definida por

$$Uv_i = v_{i+1}\,, \qquad i = 1, \ldots, n\,,$$

con $v_{n+1} \equiv v_1$. Como U transforma la base ortonormal $\mathcal{B}$ en sí misma, es una matriz unitaria. Sin embargo, W *no* es invariante bajo esta matriz, ya que

$$U \cdot W = \mathrm{lin}\{v_2, \ldots, v_k, v_{k+1}\} \not\subset W$$

(recuérdese que $k < n$, y por tanto $v_{k+1} \neq v_1$).

En virtud del tercer lema de Schur, las únicas matrices $A \in M_n(\mathbb{C})$ que conmutan con todas las matrices unitarias son los múltiplos de la identidad. En particular, si $A \in Z\big(\mathrm{U}(n)\big)$ entonces

$$A = \lambda \mathbb{1}\,, \quad A^\dagger A = |\lambda|^2 \mathbb{1} = \mathbb{1} \quad \Longrightarrow \quad \lambda \in S^1\,.$$

Por tanto

$$Z\big(\mathrm{U}(n)\big) = \big\{ \mathrm{e}^{\mathrm{i}x} \mathbb{1} \;\big|\; x \in \mathbb{R} \big\} \approx S^1\,.$$

□

Ejercicio 18. Demostrar, utilizando el mismo argumento, que

$$Z\big(\mathrm{SU}(n)\big) = \big\{ \omega^k \mathbb{1} \;\big|\; k = 0, \ldots, n-1 \big\}\,, \qquad \omega \equiv \mathrm{e}^{\frac{2\pi \mathrm{i}}{n}}\,.$$

Solución. Basta redefinir $v_{n+1} = (-1)^{n-1} v_1$ para que la matriz U del ejemplo anterior pertenezca a $\mathrm{SU}(n)$. Alternativamente, si $A \in M_n(\mathbb{C})$ conmuta con todas las matrices de $\mathrm{SU}(n)$ entonces también conmuta con todas las de $\mathrm{U}(n)$ (ejercicio), y por tanto es proporcional a la matriz unidad. En particular, si $A \in Z\big(\mathrm{SU}(n)\big)$ entonces $A = \lambda \mathbb{1}$, con $\lambda^n = 1$.

Si G es un grupo *finito*, la proposición anterior tiene el siguiente recíproco:

Proposición 1.103. *Sea $D : G \to \mathrm{GL}(V)$ una representación de un grupo* finito *G, y supongamos que los únicos operadores lineales de V en V que conmutan con* todos *los operadores $D(g)$, con $g \in G$, son los múltiplos de la identidad. Entonces D es* irreducible.

Demostración. En efecto, si D fuera reducible entonces, al ser G *finito*, por el teorema de Maschke D sería descomponible. Por tanto existirían sendos subespacios $W_i \subset V$ $(i = 1, 2)$ invariantes bajo D tales que $V = W_1 \oplus W_2$ y $D = D_1 \oplus D_2$, siendo D_i la restricción de D a W_i. Si I_i denota la identidad en W_i, para todo $\lambda_1, \lambda_2 \in \mathbb{C}$ el operador $A = \lambda_1 I_1 \oplus \lambda_2 I_2$ conmuta con todos los operadores $D(g)$ (con $g \in G$), ya que

$$AD(g) \equiv (\lambda_1 I_1 \oplus \lambda_2 I_2)\big(D_1(g) \oplus D_2(g)\big) = \big(\lambda_1 D_1(g)\big) \oplus \big(\lambda_2 D_2(g)\big) \equiv D(g)A\,.$$

Esto contradice la hipótesis, ya que A no es proporcional a la identidad a menos que $\lambda_1 = \lambda_2$. □

Nota. Al ser el teorema de Maschke válido también para grupos *compactos*, el teorema anterior es válido para grupos compactos.

El siguiente resultado es una consecuencia inmediata del tercer lema de Schur:

Proposición 1.104. *Todas las representaciones irreducibles de un grupo abeliano son unidimensionales.*

Demostración. Sea, en efecto, G un grupo abeliano, y sea $D : G \to \mathrm{GL}(V)$ una representación irreducible de G. Si g es un elemento fijo de G se cumple

$$D(h)D(g) = D(hg) = D(gh) = D(g)D(h)\,, \qquad \forall h \in G\,.$$

Por el tercer lema de Schur, $D(g) = \lambda(g)I$. Como esto se cumple para $g \in G$ arbitrario, todos los operadores $D(g)$ son proporcionales a la identidad, y por tanto *cualquier* subespacio de V es invariante bajo D. Al ser esta representación irreducible, V no puede poseer subespacios no triviales, y por consiguiente $\dim V = 1$. □

Nota. Como una representación unidimensional es automáticamente irreducible, de la proposición anterior se sigue que *una representación de un grupo abeliano es irreducible si y solo si es unidimensional.*

Corolario 1.105. *Todas la representaciones irreducibles de un grupo abeliano* finito *son unitarias.*

Demostración. Por el teorema de Schur–Auerbach, cualquier representación irreducible de un grupo finito es equivalente a una representación unitaria, y dos representaciones unidimensionales son equivalentes si y solo si son idénticas. Este resultado también se puede probar de forma directa observando que, al ser G finito, todo elemento $g \in G$ satisface $g^n = e$, siendo $n = |G|$. De esto se sigue que toda representación irreducible $D : G \to \mathbb{C}$ satisface

$$D(g)^n = 1 \quad \Longrightarrow \quad |D(g)| = 1\,.$$

(De hecho, de lo anterior se deduce que $D(g)$ es una raíz n-ésima de la unidad.) □

Nota. Evidentemente, el resultado anterior es cierto también para grupos *compactos*, al ser el teorema de Schur–Auerbach válido para este tipo de grupos.

Ejemplo 1.106. *Representaciones irreducibles del grupo C_n.*
Sea D una representación irreducible de $C_n = \{e, a, \ldots, a^{n-1}\}$. Al ser este grupo abeliano, por la proposición anterior sus representaciones irreducibles son necesariamente de dimensión uno, y cualquier representación unidimensional es irreducible. Por tanto podemos identificar cada operador $D(g) : \mathbb{C} \to \mathbb{C}$ con un número complejo. Como $a^n = e$, $D(a)$ satisface

$$D(a)^n = 1 \quad \Longrightarrow \quad D(a) = \mathrm{e}^{\frac{2k\pi\mathrm{i}}{n}}\,, \qquad k \in \{0, \ldots, n-1\}\,.$$

Es inmediato comprobar que, si definimos

$$D(a^j) = D(a)^j = \mathrm{e}^{\frac{2jk\pi\mathrm{i}}{n}}\,, \qquad j = 0, \ldots, n-1\,,$$

obtenemos una representación de C_n (obviamente irreducible, al ser unidimensional). Por tanto C_n *tiene exactamente n representaciones irreducibles* $D^{(k)}$ ($k = 0, \ldots, n-1$), cada una de ellas determinada por el correspondiente entero k en la ecuación anterior. Nótese que todas estas representaciones son *unitarias* (como afirma el Corolario 1.105), ya que los números $D^{(k)}(g)$ son de módulo uno.

Ejemplo 1.107. Consideremos la *representación regular* del grupo C_3, cuya tabla de multiplicación es

	e	a^2	a
e	e	a^2	a
a	a	e	a^2
a^2	a^2	a	e

Las matrices de esta representación son, por tanto,

$$D_{\mathrm{R}}(e) = \mathbb{1}\,, \qquad D_{\mathrm{R}}(a) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad D_{\mathrm{R}}(a^2) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = D_{\mathrm{R}}(a^{-1}) = D_{\mathrm{R}}(a)^\dagger = D_{\mathrm{R}}(a)^{\mathsf{T}}\,.$$

Evidentemente, esta representación es *reducible* (al no ser unidimensional). Por el teorema de Maschke, D_{R} debe descomponerse en la suma directa de tres representaciones irreducibles (unidimensionales), que calcularemos a continuación. Para ello observamos, en primer lugar, que

$$\det\big(\lambda - D_{\mathrm{R}}(a)\big) = \begin{vmatrix} \lambda & 0 & -1 \\ -1 & \lambda & 0 \\ 0 & -1 & \lambda \end{vmatrix} = \lambda^3 - 1\,,$$

y por tanto los autovalores de $D_{\mathrm{R}}(a)$ son las tres raíces de la unidad:

$$\lambda_k = \mathrm{e}^{\frac{2k\pi\mathrm{i}}{3}}\,, \qquad k = 0, 1, 2\,.$$

Por consiguiente existe una matriz invertible $T \in \mathrm{GL}(3, \mathbb{C})$ tal que

$$D_{\mathrm{R}}(a) = T \begin{pmatrix} 1 & & \\ & \mathrm{e}^{\frac{2\pi\mathrm{i}}{3}} & \\ & & \mathrm{e}^{\frac{4\pi\mathrm{i}}{3}} \end{pmatrix} T^{-1} \tag{1.24}$$

y por tanto

$$D_{\mathrm{R}}(a^2) = D_{\mathrm{R}}(a)^2 = T \begin{pmatrix} 1 & & \\ & \mathrm{e}^{\frac{4\pi\mathrm{i}}{3}} & \\ & & \mathrm{e}^{\frac{2\pi\mathrm{i}}{3}} \end{pmatrix} T^{-1}\,. \tag{1.25}$$

La matriz T se calcula fácilmente, ya que su k-ésima columna es un autovector correspondiente al autovalor λ_k. De la expresión de $\lambda - D_{\mathrm{R}}(a)$ se deduce que si $u = (u_1, u_2, u_3)$ es un autovector de $D_{\mathrm{R}}(a)$ de autovalor λ entonces

$$u_2 = \lambda u_3\,, \qquad u_1 = \lambda u_2\,.$$

Luego u ha de ser proporcional a $(\lambda^2, \lambda, 1)$, y por tanto podemos tomar

$$T = \begin{pmatrix} 1 & \mathrm{e}^{\frac{4\pi\mathrm{i}}{3}} & \mathrm{e}^{\frac{2\pi\mathrm{i}}{3}} \\ 1 & \mathrm{e}^{\frac{2\pi\mathrm{i}}{3}} & \mathrm{e}^{\frac{4\pi\mathrm{i}}{3}} \\ 1 & 1 & 1 \end{pmatrix}.$$

Si denotamos por v_k la k-ésima columna de T, de las ecs. (1.24)-(1.25) se sigue que

$$D_{\mathrm{R}}(a^j) v_k = \mathrm{e}^{\frac{2jk\pi\mathrm{i}}{3}} v_k\,, \qquad j, k = 0, 1, 2\,.$$

Por tanto $\mathbb{C}^3 = V_0 \oplus V_1 \oplus V_2$, con $V_k = \mathrm{lin}\{v_k\}$ invariante bajo D_{R}, y

$$D_{\mathrm{R}}(a) = I_0 \oplus \left(\mathrm{e}^{\frac{2\pi\mathrm{i}}{3}} I_1\right) \oplus \left(\mathrm{e}^{\frac{4\pi\mathrm{i}}{3}} I_2\right),$$

siendo I_k la identidad en V_k. Como la representación (unidimensional) $\mathrm{e}^{\frac{2k\pi\mathrm{i}}{3}} I_k$ ($k = 0, 1, 2$) es obviamente (equivalente a) la representación irreducible $D^{(k)}$ del ejemplo anterior, hemos probado que

$$D_{\mathrm{R}} = D^{(0)} \oplus D^{(1)} \oplus D^{(2)}\,.$$

En otras palabras, *la representación regular de C_3 contiene a cualquiera de las representaciones irreducibles de dicho grupo.* De hecho, veremos al final de esta sección que una propiedad análoga se verifica para cualquier grupo *finito* G.

1.7 Relaciones de ortogonalidad y completitud

Estudiaremos en esta sección las propiedades de los elementos de matriz de las representaciones irreducibles de un grupo *finito* G. Comenzaremos con algunas cuestiones notacionales que serán útiles para enunciar dichas propiedades.

Si G es un grupo finito y $\varphi : G \to \mathbb{C}$ es una función escalar definida en G, definimos en primer lugar su **promedio** $\langle\varphi\rangle$ mediante

$$\langle\varphi\rangle = \frac{1}{|G|}\sum_{g\in G}\varphi(g)\,.$$

Nótese que una función $\varphi : G \to \mathbb{C}$ puede identificarse con el vector $\varphi \in \mathbb{C}^{|G|} \equiv \mathcal{A}(G)$ definido por

$$\varphi = \sum_{g\in G}\varphi(g)|g\rangle\,.$$

Teniendo en cuenta esta identificación, dadas dos funciones $\varphi, \psi : G \to \mathbb{C}$ es natural definir su **producto escalar** mediante

$$\langle\varphi,\psi\rangle = \frac{1}{|G|}\sum_{g\in G}\overline{\varphi(g)}\,\psi(g) \equiv \langle\overline{\varphi}\,\psi\rangle\,. \tag{1.26}$$

Si el grupo G es *compacto*, definimos de forma natural

$$\langle\varphi\rangle = \frac{1}{|G|}\int_G \varphi(g)\,\mathrm{d}\mu(g)\,,$$

y, como en el caso de un grupo finito,

$$\langle\varphi,\psi\rangle \equiv \langle\overline{\varphi}\,\psi\rangle = \frac{1}{|G|}\int_G \overline{\varphi(g)}\,\psi(g)\,\mathrm{d}\mu(g)\,.$$

Teorema 1.108. *Sean* $D : G \to \mathrm{GL}(n,\mathbb{C})$ *y* $D' : G \to \mathrm{GL}(n',\mathbb{C})$ *dos representaciones matriciales* irreducibles *de un grupo* finito (*o compacto*) G. *Entonces se verifica*

$$\left\langle (D^{-1})_{ij}\,D'_{kl}\right\rangle = \frac{\delta_{DD'}}{\dim D}\,P_{kj}(P^{-1})_{il}\,, \tag{1.27}$$

siendo

$$\delta_{DD'} = \begin{cases} 1\,, & D' = PDP^{-1}\ \ (\Longrightarrow D \approx D') \\ 0\,, & D \not\approx D'\,. \end{cases}$$

Demostración. Dada una matriz $X \in M_{n'\times n}(\mathbb{C})$, consideremos la matriz $n'\times n$ definida por

$$A(X) \equiv \frac{1}{|G|}\sum_{g\in G} D'(g)XD(g^{-1})\,.$$

Es inmediato comprobar que, cualquiera que sea la matriz X, $A(X)$ entrelaza las representaciones D yD'. En efecto:

$$\begin{aligned} D'(h)A(X) &= \frac{1}{|G|}\sum_{g\in G} D'(h)D'(g)XD(g^{-1}) = \frac{1}{|G|}\sum_{g\in G} D'(hg)XD(g^{-1}) \\ &= \frac{1}{|G|}\sum_{k\in G} D'(k)XD(k^{-1}h) = \frac{1}{|G|}\sum_{k\in G} D'(hg)XD(k^{-1})D(h) \equiv A(X)D(h)\,. \end{aligned}$$

Si D y D' *no* son equivalentes, de los dos primeros lemas de Schur se sigue que $A(X) = 0$. Por tanto el elemento de matriz (k, j) de la matriz $A(X)$:

$$a_{kj}(X) = \frac{1}{|G|}\sum_{g\in G}\sum_{l,i} D'(g)_{kl}\,x_{li}\,D(g^{-1})_{ij} \tag{1.28}$$

ha de anularse para todo $x_{li} \in \mathbb{C}$. Igualando a cero el coeficiente de x_{li} se obtiene entonces la igualdad:

$$0 = \frac{1}{|G|} \sum_{g \in G} D(g^{-1})_{ij} D'(g)_{kl} = \frac{1}{|G|} \sum_{g \in G} \left[D(g)^{-1}\right]_{ij} D'(g)_{kl} \equiv \left\langle (D^{-1})_{ij} D'_{kl} \right\rangle .$$

Supongamos, a continuación, que D y D' son equivalentes bajo una matriz P, es decir que

$$D' = PDP^{-1} . \tag{1.29}$$

Entonces la relación de entrelazado de D' y D se puede escribir en la forma

$$D'A(X) = A(X)P^{-1}D'P \iff D' \cdot A(X)P^{-1} = A(X)P^{-1} \cdot D' .$$

Por el tercer lema de Schur, la matriz $A(X)P^{-1}$ es proporcional a la identidad, es decir

$$A(X)P^{-1} = \lambda(X)\mathbb{1} , \tag{1.30}$$

con $\lambda(X) \in \mathbb{C}$. La función $\lambda(X)$ se calcula fácilmente tomando la traza de la ecuación anterior utilizando la relación de entrelazado (1.29):

$$\begin{aligned} \lambda(X) \dim D &= \operatorname{tr}\left(A(X)P^{-1}\right) = \frac{1}{|G|} \sum_{g \in G} \operatorname{tr}\left(D'(g)XD(g^{-1})P^{-1}\right) \\ &= \frac{1}{|G|} \sum_{g \in G} \operatorname{tr}\left(D'(g)XP^{-1}D'(g^{-1})\right) = \frac{1}{|G|} \sum_{g \in G} \operatorname{tr}\left(D'(g^{-1})D'(g)XP^{-1}\right) \\ &= \frac{1}{|G|} \sum_{g \in G} \operatorname{tr}\left(XP^{-1}\right) = \operatorname{tr}\left(XP^{-1}\right) . \end{aligned}$$

Sustituyendo en la ecuación (1.30) y tomando el elemento de matriz (k, j) se obtiene

$$\lambda(X)P_{kj} = \frac{P_{kj}}{\dim D} \sum_{i,l} (P^{-1})_{il} x_{li} = a_{kj}(x) .$$

Utilizando la ecuación (1.28) e igualando el coeficiente de x_{li} en ambos miembros se llega fácilmente a la ecuación

$$\frac{P_{kj}(P^{-1})_{il}}{\dim D} = \frac{1}{|G|} \sum_{g \in G} D'(g)_{kl} \left[D(g)^{-1}\right]_{ij} \equiv \left\langle (D^{-1})_{ij} D'_{kl} \right\rangle ,$$

como queríamos demostrar. □

• Si $D' = D$ entonces podemos tomar $P = P^{-1} = \mathbb{1}$, y por tanto en este caso la ecuación (1.27) se convierte en

$$\left\langle (D^{-1})_{ij} D_{kl} \right\rangle = \frac{1}{\dim D} \delta_{il}\delta_{jk} . \tag{1.31}$$

Teorema 1.109. *Sean $D : G \to \mathrm{GL}(n, \mathbb{C})$ y $D' : G \to \mathrm{GL}(n', \mathbb{C})$ dos representaciones matriciales* unitarias irreducibles *de un grupo* finito (*o compacto*) G. *Entonces se verifica*

$$\left\langle D_{ij}, D'_{kl} \right\rangle = \frac{\delta_{DD'}}{\dim D} P_{ki}(P^{-1})_{jl} , \tag{1.32}$$

En particular, si $D = D'$ se tiene

$$\left\langle D_{ij}, D_{kl} \right\rangle = \frac{1}{\dim D} \delta_{ik}\delta_{jl} . \tag{1.33}$$

Demostración. Basta probar la primera de estas fórmulas, que claramente implica la segunda. Para ello, nótese que si D es unitaria entonces

$$D(g)^{-1} = D(g)^\dagger \quad \Longrightarrow \quad \left[D(g)^{-1}\right]_{ji} = \overline{D(g)_{ij}}\,,$$

y por tanto

$$\left\langle (D^{-1})_{ji}\, D'_{kl} \right\rangle = \left\langle \overline{D_{ij}}\, D'_{kl} \right\rangle = \left\langle D_{ij}\,, D'_{kl} \right\rangle = \frac{\delta_{DD'}}{\dim D}\, P_{ki} (P^{-1})_{jl}$$

en virtud de la ecuación (1.27) con i y j intercambiados. □

En virtud del teorema de Schur–Auerbach, toda representación de un grupo finito (o compacto) es equivalente a una representación *unitaria*. A su vez, del Corolario 1.98 (y su demostración) se sigue que toda representación unitaria de un grupo finito (o compacto) se descompone en suma directa de representaciones unitarias irreducibles. Por tanto, para clasificar *todas* las representaciones de un grupo finito (o compacto) G basta hallar todas sus *representaciones unitarias irreducibles* ($\equiv$ RUI) no equivalentes. Veremos en el resto de este capítulo que esto siempre es posible para un grupo *finito*. La clave para llevar a cabo esta tarea es la proposición que acabamos de demostrar, como veremos a continuación. Evidentemente, podemos restringirnos a representaciones matriciales, ya que toda representación lineal es equivalente a una representación matricial.

Consideremos un conjunto finito (o compacto) cualquiera $D^{(a)}$ ($a = 1, \ldots, m$) de RUI *no equivalentes* de un grupo finito (o compacto) G. En virtud de la proposición anterior,

$$\left\langle D^{(a)}_{ij}\,, D^{(b)}_{kl} \right\rangle = \frac{1}{n_a}\, \delta_{ab} \delta_{ik} \delta_{jl}\,, \tag{1.34}$$

siendo $n_a \equiv \dim D^{(a)}$. Por tanto el conjunto

$$\left\{ D^{(a)}_{ij} \;\middle|\; 1 \leqslant i, j \leqslant n_a\,, 1 \leqslant a \leqslant m \right\} \tag{1.35}$$

es *linealmente independiente*, ya que sus elementos son ortogonales entre sí y tienen norma no nula $n_a^{-1/2}$. De esta observación se sigue inmediatamente el siguiente resultado fundamental:

Proposición 1.110. *Un grupo* finito G *posee un número* finito d *de representaciones unitarias irreducibles no equivalentes. Si* $n_1, \ldots, n_d$ *son las dimensiones de dichas representaciones,*

$$\sum_{a=1}^{d} n_a^2 \leqslant |G|\,.$$

Demostración. El número de RUI no equivalentes de G ha de ser finito, ya que el conjunto (1.35) es linealmente independiente y está contenido en el espacio vectorial de dimensión finita $\mathcal{A}(G)$. El miembro derecho de la desigualdad anterior es el cardinal del conjunto (1.35) con $m = d$, siendo d el número de RUI no equivalentes de G. Este número ha de ser menor o igual que la dimensión $|G|$ del espacio $\mathcal{A}(G)$, al ser dicho conjunto linealmente independiente. □

Nota. El resultado anterior *no* es cierto en general para un grupo *compacto*. La razón por la cual la demostración de la proposición anterior no es válida (en general) para un grupo compacto es que en este caso no es posible identificar una función de G en $\mathbb{C}$ con un vector en un espacio complejo de dimensión *finita*.

Por definición, llamaremos **conjunto completo de representaciones unitarias irreducibles** (CCRUI) de un grupo finito G a cualquier conjunto maximal (es decir, que consta de d elementos) de RUI no equivalentes de G. Nótese que *toda RUI D del grupo G es necesariamente equivalente a una representación $D^{(a)}$ del CCRUI*, ya que en caso contrario podríamos añadir D al CCRUI obteniendo un conjunto de $d + 1$ RUI no equivalentes de G. Probaremos a continuación que los elementos de matriz de un CCRUI,

además de ser linealmente independientes, forman un conjunto *completo* en $\mathcal{A}(G)$, y por tanto se verifica la importante igualdad

$$\sum_{a=1}^{d} n_a^2 = |G| \,. \tag{1.36}$$

Para ello estableceremos en primer lugar el siguiente

Lema 1.111. *Si G es un grupo finito, los elementos de matriz de su representación regular forman un conjunto completo en $\mathcal{A}(G)$.*

Demostración. En efecto, hemos visto en la Sección 1.5.2 que

$$\big[D_{\mathrm{R}}(g)\big]_{ij} = 1 \iff g = g_i g_j^{-1} \,,$$

y $\big[D_{\mathrm{R}}(g)\big]_{ij} = 0$ si $g \neq g_i g_j^{-1}$. De este hecho se sigue fácilmente que si φ es un vector cualquiera en $\mathcal{A}(G)$, es decir una función $\varphi : G \to \mathbb{C}$, entonces

$$\varphi = \frac{1}{|G|} \sum_{i,j=1}^{|G|} \varphi(g_i g_j^{-1}) (D_{\mathrm{R}})_{ij} \,.$$

En efecto,

$$\frac{1}{|G|} \sum_{i,j=1}^{|G|} \varphi(g_i g_j^{-1}) \big[D_{\mathrm{R}}(g)\big]_{ij} = \frac{1}{|G|} \sum_{\substack{1 \leqslant i,j \leqslant |G| \\ g_i g_j^{-1} = g}} \varphi(g) = \varphi(g) \,, \qquad \forall g \in G \,,$$

dado que cualquier elemento $g \in G$ aparece exactamente un vez en cada fila (y en cada columna) de la tabla de multiplicación de G. □

Proposición 1.112. *Los elementos de matriz*

$$D_{ij}^{(a)} \,, \qquad 1 \leqslant i, j \leqslant n_a \,, \quad 1 \leqslant a \leqslant d \,, \tag{1.37}$$

de un conjunto completo de representaciones unitarias irreducibles de un grupo finito G forman un conjunto completo en $\mathcal{A}(G)$.

Demostración. Por el teorema de Maschke, la representación regular de G se descompone en suma directa de representaciones irreducibles, cada una de las cuales es equivalente a una representación unitaria en virtud del teorema de Schur–Auerbach. Por tanto existe una matriz $A \in \mathrm{GL}(|G|, \mathbb{C})$ tal que

$$D_{\mathrm{R}}(g) = A \begin{pmatrix} D^{(a_1)}(g) & & & \\ & D^{(a_2)}(g) & & \\ & & \ddots & \\ & & & D^{(a_r)}(g) \end{pmatrix} A^{-1} \equiv A B(g) A^{-1} \,, \qquad \forall g \in G \,.$$

Tomando el elemento de matriz (i, j) de esta igualdad se obtiene

$$\big[D_{\mathrm{R}}(g)\big]_{ij} = \sum_{k,l} a_{ik} (A^{-1})_{lj} b_{kl}(g) \,,$$

y por tanto

$$(D_{\mathrm{R}})_{ij} \in \mathrm{lin}\big\{ b_{kl} \;\big|\; 1 \leqslant k, l \leqslant |G| \big\} \subset \mathrm{lin}\big\{ D_{ij}^{(a)} \;\big|\; 1 \leqslant i, j \leqslant n_a \,, \;\; 1 \leqslant a \leqslant d \big\} \,,$$

ya que cada b_{kl} no nulo es igual a un cierto elemento del conjunto (1.37). Como el conjunto de los elementos de matriz $(D_R)_{ij}$ es completo en $\mathcal{A}(G)$, en virtud del lema anterior, también ha de serlo el conjunto (1.37). □

1.8 Caracteres

Los elementos de matriz D_{ij} de una representación matricial $D : G \to \mathrm{GL}(n, \mathbb{C})$ evidentemente cambian bajo una transformación de semejanza $D \mapsto ADA^{-1}$. En particular, dichos elementos de matriz *no* son en general los mismos para dos representaciones equivalentes, lo que supone un serio inconveniente. Por este motivo, son particularmente importantes las combinaciones de los elementos de matriz D_{ij} que permanecen invariantes bajo transformaciones de semejanza, y por tanto no varían al pasar de una representación D a otra equivalente D'. La más sencilla de estas combinaciones es la *traza*, ya que depende linealmente de los elementos de matriz.

Definición 1.113. Dada una representación $D : G \to \mathrm{GL}(V)$ de un grupo G, se define su **carácter** como la aplicación $\chi_D : G \to \mathbb{C}$ dada por

$$\chi_D(g) = \operatorname{tr} D(g)\,, \qquad \forall g \in G\,. \tag{1.38}$$

Más explícitamente,

$$\chi_D(g) = \sum_{i=1}^{\dim D} D_{ii}(g)\,,$$

donde (con un cierto abuso de notación) denotamos por $D_{ij}(g)$ a los elementos de matriz de $D(g)$ en una base *cualquiera*[10] de V. Nótese que, en virtud del comentario al principio de la sección anterior, χ_D puede considerarse equivalentemente como un *vector* de $\mathcal{A}(G)$.

Ejemplo 1.114. Calcularemos en este ejemplo el carácter de la representación regular de un grupo finito G. Para ello, basta notar que cualquiera que sea $i \in \{1, \ldots, |G|\}$ se tiene

$$\left[D_{\mathrm{R}}(g)\right]_{ii} = 1 \iff g = g_i g_i^{-1} = e\,,$$

y $\left[D_{\mathrm{R}}(g)\right]_{ii} = 0$ si $g \neq e$. Por tanto el carácter de la representación regular está dado por

$$\chi_{D_{\mathrm{R}}}(e) = |G|\,; \qquad \chi_{D_{\mathrm{R}}}(g) = 0\,, \quad \forall g \neq e\,. \tag{1.39}$$

De la invariancia de la traza bajo transformaciones de semejanza se sigue, como acabamos de comentar, la siguiente propiedad fundamental del carácter:

Proposición 1.115. *Dos representaciones equivalentes de un grupo G tienen los mismos caracteres.*

En virtud de la proposición anterior, los caracteres de un CCRUI de un grupo finito G *no* dependen de la elección concreta de representaciones en el CCRUI. Por tanto *los caracteres de un CCRUI están unívocamente determinados por el grupo, salvo por el orden.* Veremos en el resto de esta sección que los caracteres de un CCRUI de un grupo finito G son fundamentales para el estudio de sus representaciones. Comenzaremos demostrando la siguiente *propiedad de ortonormalidad:*

Proposición 1.116. *Los caracteres*

$$\chi_a \equiv \chi_{D^{(a)}}\,, \qquad a = 1, \ldots, d\,,$$

de un conjunto completo de representaciones unitarias irreducibles de un grupo finito G forman un sistema ortonormal *en* $\mathcal{A}(G)$.

Demostración. En efecto, utilizando las relaciones de ortogonalidad (1.34) se obtiene fácilmente

$$\langle \chi_a, \chi_b \rangle = \left\langle \sum_{i=1}^{n_a} D_{ii}^{(a)}, \sum_{j=1}^{n_b} D_{jj}^{(b)} \right\rangle = \sum_{\substack{1 \leqslant i \leqslant n_a \\ 1 \leqslant j \leqslant n_b}} \langle D_{ii}^{(a)}, D_{jj}^{(b)} \rangle = \frac{\delta_{ab}}{n_a} \sum_{i,j=1}^{n_a} \delta_{ij} = \frac{\delta_{ab}}{n_a}\, n_a = \delta_{ab}\,.$$

□

[10] Recuérdese que la traza de un operador no depende de la base escogida para representarlo, en virtud de la identidad elemental $\operatorname{tr}(CAC^{-1}) = \operatorname{tr}(A)$.

Una propiedad fundamental del carácter de una representación es su *constancia sobre clases de conjugación* de G. En efecto,

$$\chi_D(hgh^{-1}) \equiv \operatorname{tr} D(hgh^{-1}) = \operatorname{tr}\left(D(h)D(g)D(h)^{-1}\right) = \operatorname{tr} D(g) \equiv \chi_D(g)\,.$$

En otras palabras, χ_D es un elemento del conjunto $\mathcal{C}(G)$ de todas las funciones $\psi : G \to \mathbb{C}$ constantes sobre las clases de conjugación de G. Este conjunto es claramente un espacio vectorial, ya que una combinación lineal de funciones constantes sobre clases de conjugación sigue siendo constante sobre dichas clases de conjugación. Utilizando la identificación de las funciones de G en $\mathbb{C}$ con el álgebra del grupo $\mathcal{A}(G)$, identificaremos a partir de ahora el espacio $\mathcal{C}(G)$ con un subespacio vectorial de $\mathcal{A}(G)$.

Lema 1.117. *La dimensión de $\mathcal{C}(G)$ es igual al número de clases de conjugación de G.*

En efecto, sean $C_1, \ldots, C_m$ las clases de conjugación de G, y denotemos por γ_i $(i = 1, \ldots, m)$ a la función definida por

$$\gamma_i(g) = \begin{cases} 1 & g \in C_i \\ 0\,, & g \notin C_i\,. \end{cases}$$

Es inmediato comprobar que las funciones γ_i pertenecen a $\mathcal{C}(G)$ y forman un conjunto linealmente independiente. Por otra parte, si $\psi \in \mathcal{C}(G)$ y denotamos por $\psi(C_i)$ al valor de ψ sobre cualquier elemento de la clase de conjugación C_i entonces

$$\psi = \sum_{i=1}^{m} \psi(C_i)\gamma_i \quad \Longrightarrow \quad \psi \in \operatorname{lin}\{\gamma_1, \ldots, \gamma_m\}\,.$$

Proposición 1.118. *Los caracteres de un conjunto completo de representaciones unitarias irreducibles de un grupo finito G son una base ortonormal de $\mathcal{C}(G)$.*

Demostración. Por la Proposición 1.116, basta demostrar que los caracteres χ_a $(a = 1, \ldots, d)$ forman un conjunto *completo* en $\mathcal{C}(G)$. Sea, por tanto, ψ un elemento de $\mathcal{C}(G)$. En virtud de la Proposición 1.112, podemos desarrollar ψ en términos de los elementos de matriz de un CCRUI de G:

$$\psi = \sum_{a,i,j} c^a_{ij} D^{(a)}_{ij}\,.$$

Al ser ψ constante sobre las clases de conjugación de G se tiene:

$$\begin{aligned}
\psi(g) &= \frac{1}{|G|}\sum_{h\in G}\psi(hgh^{-1}) = \frac{1}{|G|}\sum_{h\in G}\sum_{a,i,j} c^a_{ij} D^{(a)}_{ij}(hgh^{-1}) \\
&= \frac{1}{|G|}\sum_{h\in G}\sum_{a,i,j,k,l} c^a_{ij} D^{(a)}_{ik}(h) D^{(a)}_{kl}(g)\left[\left(D^{(a)}(h)\right)^{-1}\right]_{lj} \\
&= \frac{1}{|G|}\sum_{h\in G}\sum_{a,i,j,k,l} c^a_{ij} D^{(a)}_{ik}(h) D^{(a)}_{kl}(g)\overline{D^{(a)}(h)_{jl}}\,,
\end{aligned}$$

donde en la última igualdad hemos utilizado el carácter unitario de las representaciones $D^{(a)}$. Utilizando las relaciones (1.34) para evaluar la suma en $h \in G$ se obtiene

$$\begin{aligned}
\psi(g) &= \sum_{a,i,j,k,l} c^a_{ij} D^{(a)}_{kl}(g)\left\langle D^{(a)}_{jl}, D^{(a)}_{ik}\right\rangle \sum_{a,i,j,k,l} c^a_{ij} D^{(a)}_{kl}(g)\frac{1}{n_a}\delta_{ij}\delta_{kl} \\
&= \sum_a \left(\sum_i \frac{c^a_{ii}}{n_a}\right)\chi_a(g) \quad \Longrightarrow \quad \psi \in \operatorname{lin}\{\chi_1, \ldots, \chi_d\}\,.
\end{aligned}$$

□

De la proposición anterior y el Lema 1.117 se deduce inmediatamente el importante

Corolario 1.119. *El número de representaciones unitarias irreducibles inequivalentes de un grupo finito G coincide con el de sus clases de conjugación.*

En virtud del resultado anterior, podemos definir la **tabla de caracteres** de un grupo finito G como la matriz de orden d cuyo elemento de matriz (a, b) es el valor del carácter χ_a en un elemento cualquiera de la clase de conjugación C_b, que denotaremos por $\chi_a(C_b)$. Dicha tabla se suele representar en la forma siguiente:

	$C_1[c_1]$	$\cdots$	$C_d[c_d]$
χ_1	$\chi_1(C_1)$	$\cdots$	$\chi_1(C_d)$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
χ_d	$\chi_d(C_1)$	$\cdots$	$\chi_d(C_d)$

donde c_a denota el cardinal de la clase de conjugación C_a. Normalmente $C_1 = \{e\}$ y χ_1 es el carácter de la representación trivial de dimensión uno, por lo que la primera fila de la tabla de caracteres es simplemente $(1 \cdots 1)$. Análogamente, la primera columna de la tabla de caracteres está formada por las *dimensiones* de las RUI inequivalentes de G, al ser

$$\chi_a(e) = \operatorname{tr}\left[D^{(a)}(e)\right] = \operatorname{tr} \mathbb{1} = \dim D^{(a)} \equiv n_a \,.$$

Por tanto la tabla de caracteres del grupo tiene la forma

	$\{e\}$	$C_2[c_2]$	$\cdots$	$C_d[c_d]$
χ_1	1	1	$\cdots$	1
χ_2	n_2	$\chi_2(C_2)$	$\cdots$	$\chi_2(C_d)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
χ_d	n_d	$\chi_d(C_1)$	$\cdots$	$\chi_d(C_d)$

Utilizando la notación anterior, las relaciones de ortonormalidad de los caracteres χ_a se pueden escribir en la forma

$$\sum_{r=1}^{d} \frac{c_r}{|G|}\, \overline{\chi_a(C_r)}\, \chi_b(C_r) = \delta_{ab} \,, \qquad 1 \leqslant a, b \leqslant d \,.$$

Si definimos la matriz $X \in M_d(\mathbb{C})$ mediante

$$x_{ar} = \sqrt{\frac{c_r}{|G|}}\, \chi_a(C_r) \,,$$

las igualdades anteriores son equivalentes a la identidad

$$\sum_{r=1}^{d} \overline{x_{ar}}\, x_{br} = \delta_{ab} \,, \quad 1 \leqslant a, b \leqslant d \quad \Longleftrightarrow \quad X X^\dagger = \mathbb{1} \,.$$

Por tanto la matriz X es unitaria. En particular, de la igualdad $X^\dagger X = \mathbb{1}$ se obtiene

$$\delta_{ab} = (X^\dagger X)_{ab} = \sum_{r=1}^{d} \overline{x_{ra}}\, x_{rb} = \frac{\sqrt{c_a c_b}}{|G|} \sum_{r=1}^{d} \overline{\chi_r(C_a)} \chi_r(C_b) \,,$$

relación que puede escribirse como

$$\sum_{r=1}^{d} \overline{\chi_r(C_a)} \chi_r(C_b) = \frac{|G|}{c_a}\, \delta_{ab} \,. \tag{1.40}$$

Esto demuestra, en particular, que *las columnas de la tabla de caracteres son ortogonales entre sí.*

Ejemplo 1.120. En este ejemplo construiremos la tabla de caracteres del grupo simétrico S_3. Las clases de conjugación son en este caso

$$C_1 = \{e\}\,, \qquad C_2 = \big\{(12),(13),(23)\big\}\,, \qquad C_3 = \big\{(123),(321)\big\}\,,$$

y por tanto S_3 posee 3 RUI no equivalentes. Como la única forma de expresar $|S_3| = 6$ como la suma de 3 cuadrados es $6 = 1^2 + 1^2 + 2^2$, las 3 RUI no equivalentes de S_3 tiene dimensiones 1, 1 y 2. Una de estas RUI unidimensionales (que tomaremos como $D^{(1)}$) ha de ser la trivial, y la otra el *signo*:

$$D^{(2)}(\sigma) = (-1)^\sigma\,, \qquad \forall \sigma \in S_3\,.$$

Por tanto la tabla de caracteres de S_3 es de la forma

	$\{e\}$	$C_2[3]$	$C_3[2]$
χ_1	1	1	1
χ_2	1	-1	1
χ_3	2	x	y

Las incógnitas $x, y \in \mathbb{C}$ se hallan fácilmente imponiendo la ortogonalidad de la primera columna con las otras dos, de donde se sigue fácilmente que $x = 0$, $y = -1$. Alternativamente, imponiendo la ortogonalidad del carácter χ_3 con χ_1 y χ_2 se obtiene

$$6\langle\chi_1,\chi_3\rangle = 2 + 3x + 2y = 0\,, \quad 6\langle\chi_2,\chi_3\rangle = 2 - 3x + 2y = 0 \qquad \Longleftrightarrow \qquad x = 0\,,\ \ y = -1\,.$$

En definitiva, la tabla de caracteres de S_3 es

	$\{e\}$	$C_2[3]$	$C_3[2]$
χ_1	1	1	1
χ_2	1	-1	1
χ_3	2	0	-1

□

Hemos visto al principio de esta sección que dos representaciones equivalentes de un grupo tienen los mismos caracteres. Si G es un grupo finito, demostraremos a continuación que el recíproco de este resultado es también cierto. Por tanto *las representaciones de un grupo finito están unívocamente determinadas por sus caracteres, módulo equivalencia.* Para establecer este importante resultado, probaremos primero una serie de propiedades de los caracteres que son de gran interés en sí mismas. Comenzaremos por el siguiente lema elemental:

Lema 1.121. *Sea $D : G \to \mathrm{GL}(n,\mathbb{C})$ una representación de un grupo G. Si $D = D_1 \oplus \cdots \oplus D_m$, con $D_i : G \to \mathrm{GL}(n_i,\mathbb{C})$, entonces*

$$\chi_D = \sum_{i=1}^{m} \chi_{D_i}\,.$$

Nota. Evidentemente, el mismo resultado vale para una representación lineal $D : G \to \mathrm{GL}(V)$.

Demostración. Inmediata a partir de la definición de traza. □

De este lema y de la ortogonalidad de los caracteres de un CCRUI de un grupo finito G (Proposición 1.116) se deduce el siguiente resultado fundamental:

Proposición 1.122. *Sea $D : G \to \mathrm{GL}(n,\mathbb{C})$ una representación de un grupo finito G, y sea $D^{(a)} : G \to \mathrm{GL}(n_a,\mathbb{C})$ $(a = 1,\ldots,d)$ un conjunto completo de representaciones unitarias irreducibles de G. Entonces*

$$D \approx m_1 D^{(1)} \oplus \cdots \oplus m_d D^{(d)}\,, \tag{1.41}$$

donde los enteros no negativos m_a están dados por

$$m_a = \langle\chi_a,\chi_D\rangle\,, \qquad a = 1,\ldots,d\,. \tag{1.42}$$

Nota. Por definición,

$$m_a D^{(a)} \equiv \underbrace{D^{(a)} \oplus \cdots \oplus D^{(a)}}_{m_a} .$$

Demostración. En virtud de los teoremas de Maschke y de Schur–Auerbach, cualquier representación de un grupo finito G admite una descomposición de la forma (1.41) en términos de las representaciones $D^{(a)}$ de un CCRUI de G. Como

$$\chi_D = \sum_{a=1} m_a \chi_a ,$$

multiplicando escalarmente por χ_a y utilizando las relaciones de ortonormalidad de los caracteres se obtiene inmediatamente la relación (1.42). □

De la proposición anterior se siguen varios importantes corolarios:

Corolario 1.123. *La descomposición* (1.41) *es única, salvo por el orden.*

Demostración. En efecto, los enteros m_a están unívocamente determinados por la representación D vía la ec. (1.42) . □

Corolario 1.124. *Dos representaciones de un grupo finito G son equivalentes si y solo si tienen los mismos caracteres.*

Demostración. Ya hemos visto que dos representaciones equivalentes de cualquier grupo tienen los mismos caracteres, por lo que basta establecer el recíproco de este resultado. Sean, por tanto, dos representaciones (matriciales) D y D' de un grupo finito G, y supongamos que $\chi_D = \chi_{D'}$. De las ecs. (1.41)-(1.42) se sigue entonces que

$$D \approx m_1 D^{(1)} \oplus \cdots \oplus m_d D^{(d)} , \qquad D' \approx m'_1 D^{(1)} \oplus \cdots \oplus m'_d D^{(d)}$$

con

$$m_a = \langle \chi_a, \chi_D \rangle = \langle \chi_a, \chi_{D'} \rangle = m'_a ,$$

y por tanto

$$D \approx m_1 D^{(1)} \oplus \cdots \oplus m_d D^{(d)} \approx D' .$$

□

Corolario 1.125. *Una representación D de un grupo finito G es* irreducible *si y solo si* $\langle \chi_D, \chi_D \rangle = 1$.

Demostración.
$\Longrightarrow$) Si D es irreducible, en virtud del teorema de Schur–Auerbach $D \approx D^{(a)}$ para algún a, y por tanto

$$\chi_D = \chi_a \quad \Longrightarrow \quad \langle \chi_D, \chi_D \rangle = \langle \chi_a, \chi_a \rangle = 1 .$$

$\Longleftarrow$) Supongamos que D es una representación de G con $\langle \chi_D, \chi_D \rangle = 1$. En virtud de la Proposición 1.118 y de la ortonormalidad de los caracteres χ_a,

$$\chi_D = \sum_{a=1}^{d} m_a \chi_a \quad \Longrightarrow \quad \langle \chi_D, \chi_D \rangle = \sum_{a=1}^{d} |m_a|^2 = \sum_{a=1}^{d} m_a^2 = 1 .$$

Pero los números m_a son los *enteros no negativos* que aparecen en la descomposición (1.41) de D en términos de las RUI $D^{(a)}$. De la igualdad anterior se sigue entonces que existe $a \in \{1, \ldots, d\}$ tal que $m_a = 1$ y $m_b = 0$ para $b \neq a$. Por tanto $D \approx D^{(a)}$ es irreducible. □

Corolario 1.126. *Si $D^{(a)}$ $(a = 1, \dots, d)$ es un conjunto completo de representaciones de un grupo finito G entonces la representación regular de G admite la descomposición*

$$D_{\mathrm{R}} = \bigoplus_{a=1}^{d} n_a D^{(a)} , \qquad n_a \equiv \dim D^{(a)} .$$

Demostración. Por el Corolario 1.122, D_{R} admite una descomposición de la forma (1.41) con m_a dado por (1.42). Utilizando la ecuación (1.39) se obtiene fácilmente

$$m_a = \langle \chi_a , \chi_{D_{\mathrm{R}}} \rangle \equiv \frac{1}{|G|} \sum_{g \in G} \overline{\chi_a(g)}\, \chi_{D_{\mathrm{R}}}(g) = \frac{1}{|G|}\, \overline{\chi_a(e)}\, \chi_{D_{\mathrm{R}}}(e) = \frac{1}{|G|}\, n_a \cdot |G| = n_a .$$

□

En otras palabras, *la representación regular de un grupo finito contiene cada una de las RUI del grupo un número de veces igual a su dimensión.*

• La Proposición 1.122 se aplica a menudo al caso en que la representación D es el *producto directo* de dos representaciones D y D'. Como

$$\big[(D \otimes D')(g)\big]_{ij,kl} = D_{ik}(g) D'_{jl}(g) ,$$

el carácter de la representación $D \otimes D'$ está dado por

$$\begin{aligned} \chi_{D \otimes D'}(g) &= \sum_{i,j} [(D \otimes D')(g)]_{ij,ij} = \sum_{i,j} D_{ii}(g) D'_{jj}(g) = \Big(\sum_i D_{ii}(g) \Big) \Big(\sum_j D'_{jj}(g) \Big) \\ &= \chi_D(g) \chi_{D'}(g) . \end{aligned}$$

Por tanto *el carácter del producto tensorial de dos representaciones es el producto de sus respectivos caracteres:*

$$\chi_{D \otimes D'} = \chi_D \chi_{D'} .$$

Capítulo 2

Grupos y Álgebras de Lie

Bibliografía.

Jänich, K., *Topology*, Springer-Verlag, 1984.
Dugundji, J. *Topology*, Allyn and Bacon, 1966.

2.1 Espacios topológicos

2.1.1 Preliminares

Si M es un conjunto cualquiera, denotaremos por $\mathcal{P}(M)$ el conjunto de las *partes de* M, cuyos elementos son los subconjuntos de M:

$$\mathcal{P}(M) = \left\{ A \;\middle|\; A \subset M \right\}.$$

Definición 2.1. Una **topología** en M es un subconjunto $\mathcal{T} \subset \mathcal{P}(M)$ que verifica:

1. $\emptyset, M \in \mathcal{T}$.
2. La unión *arbitraria* de elementos de $\mathcal{T}$ pertenece a $\mathcal{T}$.
3. La intersección *finita* de elementos de $\mathcal{T}$ pertenece a $\mathcal{T}$.

Un **espacio topológico** es un conjunto M provisto de una topología $\mathcal{T}$. Los elementos de la topología se denominan conjuntos **abiertos**, y sus complementarios reciben el nombre de conjuntos **cerrados**. De esta definición se sigue inmediatamente que la unión finita y la intersección arbitraria de conjuntos cerrados es un conjunto cerrado.

En cualquier espacio topológico, el conjunto vacío y el conjunto total M son simultáneamente abiertos y cerrados. Esto sugiere la siguiente definición:

Definición 2.2. Un espacio topológico $(M, \mathcal{T})$ es **conexo** si los únicos subconjuntos de M que son a la vez abiertos y cerrados son $\emptyset$ y M.

Ejercicio 19. Probar que M es conexo si y solo si no es la unión de dos abiertos disjuntos no vacíos.

Ejemplo 2.3. Todo *espacio métrico* es un espacio topológico, siendo por definición abierto cualquier conjunto obtenido como unión de *bolas abiertas*. En particular, los espacios *normados* y *euclidianos* (reales o complejos) son espacios topológicos. Por tanto $\mathbb{R}^n$, $\mathbb{C}^n$, $M_n(\mathbb{R})$ y $M_n(\mathbb{C})$ son espacios topológicos. Más concretamente, $M_n(\mathbb{C})$ es un espacio euclidiano con el producto escalar

$$(A, B) = \sum_{i,j=1}^{n} \overline{a_{ij}}\, b_{ij} \equiv \mathrm{tr}(A^\dagger B)\,,$$

cuya norma asociada es $\|A\|^2 = \mathrm{tr}(A^\dagger A)$. En particular, la bola abierta de centro $A \equiv (a_{ij})_{1\leqslant i,j\leqslant n}$ y radio r está formada por las matrices $Z = (z_{ij})_{1\leqslant i,j\leqslant n}$ tales que

$$\sum_{i,j,=1}^{n} \left|z_{ij} - a_{ij}\right|^2 < r^2 \iff \mathrm{tr}\left[(Z-A)^\dagger (Z-A)\right] < r^2 \,.$$

Un **entorno** de un punto x en un espacio topológico M es cualquier subconjunto de M que contiene a un abierto que a su vez contiene a x. Es fácil ver, utilizando la segunda propiedad de la topología, que un subconjunto $A \subset M$ es abierto si y solo si es un entorno de cada uno de sus puntos. Por ejemplo, un subconjunto de $\mathbb{R}^n$ ($\mathbb{C}^n$, $M_n(\mathbb{R})$, $M_n(\mathbb{C})$, o de cualquier espacio métrico) es abierto si contiene una bola abierta con centro en cada uno de sus puntos.

Definición 2.4. Un espacio topológico $(M, \mathcal{T})$ es **Hausdorff** si dados dos puntos $x, y \in M$ con $x \neq y$ existen sendos entornos $U_x \ni x$, $U_y \ni y$ tales que $U_x \cap U_y = \emptyset$.

Es elemental probar que cualquier espacio métrico (en particular, normado o euclidiano) es Hausdorff. En particular, los espacios euclidianos $\mathbb{R}^n$, $\mathbb{C}^n$, $M_n(\mathbb{R})$ y $M_n(\mathbb{C})$ son Hausdorff.

Definición 2.5. Sea $(M, \mathcal{T})$ un espacio topológico. Una **base** de la topología $\mathcal{T}$ es un subconjunto $\mathcal{B} \subset \mathcal{T}$ tal que cualquier conjunto abierto es unión de elementos de $\mathcal{B}$.

Evidentemente, una topología $\mathcal{T}$ queda caracterizada por una cualquiera de sus bases.

Ejemplo 2.6. En $\mathbb{R}^n$ ($\mathbb{C}^n$, $M_n(\mathbb{R})$, $M_n(\mathbb{C})$ y, en general, en cualquier espacio métrico) una base es la formada por todas las bolas abiertas. También es base el conjunto de todas las bolas abiertas con radio racional y centro de coordenadas racionales.[1] Esta última base es un conjunto *numerable*, ya que $\mathbb{Q}^m$ es numerable. En general, se dice que un espacio topológico satisface el **segundo axioma de numerabilidad** si posee una base numerable.

Si $(M, \mathcal{T})$ es un espacio topológico y A es un subconjunto de M, la topología $\mathcal{T}_A$ **inducida** por M en A (llamada también **topología relativa**) está formada por la intersección de los abiertos de M con A:

$$\mathcal{T}_A = \left\{U \cap A \;\middle|\; U \in \mathcal{T}\right\}.$$

Es inmediato comprobar que, en efecto, $\mathcal{T}_A$ es una topología, y por tanto $(A, \mathcal{T}_A)$ es un espacio topológico. Se dice que A, con la topología relativa heredada de M, es un **subespacio** (topológico) de M. Es fácil también probar que los cerrados en la topología $\mathcal{T}_A$ son intersección de cerrados de M con A. Nótese, por último, que si el espacio topológico M es *Hausdorff* o satisface el *segundo axioma de numerabilidad* lo mismo ocurrirá con cualquier *subespacio* de M.

Ejemplo 2.7. Cualquier concepto que se define para un espacio topológico M se extiende a un subconjunto A de M sin más que considerar a dicho conjunto como un espacio topológico con la topología inducida. Así, por ejemplo, se dice que $A \subset M$ es conexo si A es conexo con la topología relativa. En otras palabras, $A \subset M$ es conexo si no es la unión disjunta de dos subconjuntos no vacíos $A_1, A_2 \subset A$ abiertos en A.

[1] En efecto, dados $x \in \mathbb{R}^n$ y $r > 0$, basta probar que existen $p \in \mathbb{Q}$ y $q \in \mathbb{Q}^n$ tales que $x \in B_p(q) \subset B_r(x)$. Como $\mathbb{Q}^n$ es denso en $\mathbb{R}^n$, hay un $q \in \mathbb{Q}^n \cap B_{r/2}(x)$. Sea $p \in \mathbb{Q}$ tal que $\|q - x\| < p \leqslant r/2$, que existe al ser $\mathbb{Q}$ denso en $\mathbb{R}$. Entonces $x \in B_p(q)$, y para todo $y \in B_p(q)$ se tiene

$$\|y - x\| \leqslant \|y - q\| + \|x - q\| < 2p \leqslant r \quad \Longrightarrow \quad y \in B_r(x)\,.$$

2.1.2 Continuidad. Homeomorfismos

Definición 2.8. Sean M_1 y M_2 dos espacios topológicos. Una función $f : M_1 \to M_2$ es **continua en un punto** x de su dominio si para todo entorno V de $f(x)$ la imagen inversa $f^{-1}(V)$ es un entorno de x. (En otras palabras, para todo entorno V de $f(x)$ existe un abierto U que contiene a x tal que de $f(U) \subset V$.) Se dice que f es **continua** (en M_1) si es continua en todos los puntos de su dominio. Equivalentemente, f es continua si para todo abierto $U_2 \subset M_2$ la imagen inversa $f^{-1}(U_2)$ es abierto en M_1 (ejercicio). Dos espacios topológicos M_1 y M_2 son **homeomorfos** si existe un **homeomorfismo** $f : M_1 \to M_2$, es decir una una *biyección* $f : M_1 \to M_2$ tal que f y f^{-1} son funciones *continuas*.

Si f es un homeomorfismo y $U \subset M$ es abierto, entonces $f(U) = \left(f^{-1}\right)^{-1}(U)$ también es abierto (por la continuidad de f^{-1}). Por tanto, U es abierto en M si y solo si $f(U)$ es abierto en M'. Dicho de otro modo,

$$\mathcal{T}_2 = f(\mathcal{T}_1) \equiv \left\{ f(U) \,\middle|\, U \in \mathcal{T}_1 \right\}$$

y, análogamente, $\mathcal{T}_1 = f^{-1}(\mathcal{T}_2)$. Por tanto, un homeomorfismo f establece una biyección no solo entre los conjuntos M y M', sino también entre sus *topologías* $\mathcal{T}_1$ y $\mathcal{T}_2$. Por esta razón, dos espacios topológicos homeomorfos pueden considerarse como *equivalentes* en sentido topológico.

Una propiedad P es **topológica** si es invariante bajo homeomorfismos; es decir, si siempre que un espacio topológico posee la propiedad P entonces cualquier otro espacio topológico homeomorfo a él también la posee. Por ejemplo, el ser de *Hausdorff* o el satisfacer el *segundo axioma de numerabilidad* son propiedades topológicas (probarlo). Una propiedad topológica P se denomina también *invariante topológico*.

Ejemplo 2.9. Es inmediato probar a partir de la definición de espacio conexo que si $f : M \to M'$ es una función continua entre dos espacios topológicos entonces M conexo implica $f(M)$ conexo. En particular, si $f : M \to M'$ es un *homeomorfismo* entonces M es conexo si y solo si M' lo es. Por tanto *la conexión es una propiedad topológica*.

Ejemplo 2.10. La aplicación $f : [0, 2\pi) \to S^1$ dada por $f(x) = \mathrm{e}^{\mathrm{i}x}$ es continua y biyectiva. Sin embargo *no* es un homeomorfismo, ya que su inversa es discontinua en el punto $1 \in S^1$. De hecho, S^1 *no* es homeomorfa a ningún intervalo, ya que S^1 menos uno cualquiera de sus puntos es *conexo*, mientras que un intervalo menos uno de sus puntos interiores es un conjunto disconexo (en efecto, los únicos conjuntos conexos de la recta real son los intervalos).

Ejemplo 2.11. Los espacios $\mathbb{C}^n$, $M_n(\mathbb{R})$ y $M_n(\mathbb{C})$ son respectivamente homeomorfos a $\mathbb{R}^{2n}$, $\mathbb{R}^{n^2}$ y $\mathbb{R}^{2n^2}$. Por ejemplo, en el caso de $M_n(\mathbb{C})$ es evidente que la aplicación $f : M_n(\mathbb{C}) \to \mathbb{R}^{2n^2}$ dada por

$$Z \equiv \left(x_{jk} + \mathrm{i}y_{jk}\right)_{1 \leqslant j,k \leqslant n} \in M_n(\mathbb{C}) \mapsto (x_{11}, y_{11}, x_{12}, y_{12}, \ldots, x_{nn}, y_{nn}) \in \mathbb{R}^{2n^2},$$

es un homeomorfismo. En efecto, f es claramente biyectiva, y la imagen bajo f de una bola abierta de radio r centrada en Z es la bola de centro $f(Z)$ y el mismo radio r.

Ejercicio 20. a) Probar que $f : M_1 \to M_2$ es continua si y solo si $f : M_1 \to f(M_1)$ es continua, donde $f(M_1)$ se considera un subespacio topológico de M_2. b) Demostrar que si $f : M_1 \to M_2$ es continua también lo es la restricción $f|_A$ de f a cualquier subconjunto $A \subset M_1$.

Demostración. La demostración de ambas afirmaciones se sigue fácilmente de las identidades

$$f^{-1}\big(V \cap f(M_1)\big) = f^{-1}(V)\,, \qquad \left(f|_A\right)^{-1}(V) = f^{-1}(V) \cap A\,.$$

□

2.1.3 Compacidad

Definición 2.12. Un espacio topológico M es **compacto** si todo recubrimiento de M por conjuntos *abiertos* contiene un subrecubrimiento *finito*.

Es inmediato comprobar que $A \subset M$ es compacto si y solo si todo recubrimiento de A por conjuntos abiertos de M contiene un subrecubrimiento finito.

Propiedades:

1. Un subconjunto compacto de un espacio topológico *Hausdorff* es *cerrado*.
2. Un subconjunto *cerrado* de un espacio topológico compacto es *compacto*.
3. En $\mathbb{R}^n$ ($\mathbb{C}^n$, $M_n(\mathbb{R})$, $M_n(\mathbb{C})$) con la topología usual, un subconjunto A es compacto si y solo si es *cerrado* y *acotado* (*teorema de Heine–Borel–Lebesgue*).
4. Si $f : M \to M'$ es continua y M es compacto entonces $f(M)$ es compacto (la demostración es inmediata a partir de la definición). En particular, la compacidad es una propiedad topológica.

   Obsérvese que este resultado implica que *una función $f : M \to \mathbb{R}$ continua en un conjunto compacto M alcanza sus valores máximo y mínimo en M*. En efecto, $f(M) \subset \mathbb{R}$ es compacto, y por tanto es cerrado y acotado. Por ser $f(M)$ acotado, existen $m_1 = \inf\{f(x) \mid x \in M\}$ y $m_2 = \sup\{f(x) \mid x \in M\}$, y por ser cerrado m_1 y m_2 pertenecen a $f(M)$.
5. *Teorema de Tychonoff*: el *producto cartesiano* (arbitrario) de espacios *compactos* es *compacto* (respecto de la *topología producto*).

Nota: se define el producto cartesiano de una familia (finita o infinita) de conjuntos $\{M_\alpha \mid \alpha \in A\}$ mediante

$$\prod_{\alpha\in A} M_\alpha = \left\{(x_\alpha)_{\alpha\in A} \;\middle|\; x_\alpha \in M_\alpha, \forall \alpha \in A\right\},$$

siendo

$$(x_\alpha)_{\alpha\in A} = (y_\alpha)_{\alpha\in A} \iff x_\alpha = y_\alpha\,, \qquad \forall \alpha \in A\,.$$

Para un producto *finito*[2] de espacios topológicos $M_1 \times M_2 \times \cdots \times M_n$, la **topología producto** tiene como base los conjuntos de la forma $U_1 \times U_2 \times \cdots \times U_n$, donde U_i es abierto en M_i para todo $1 \leqslant i \leqslant n$. Por ejemplo, la topología usual de $\mathbb{R}^n$ es la topología producto (n veces) de la topología usual de $\mathbb{R}$.

Definición 2.13. Un espacio topológico es **localmente compacto** si todo punto $x \in M$ posee un entorno compacto.

Por ejemplo, los conjuntos $\mathbb{R}^n$, $\mathbb{C}^n$, $M_n(\mathbb{R})$ o $M_n(\mathbb{C})$ no son compactos pero sí localmente compactos, mientras que $\mathbb{Q}$ (con la topología inducida por la de $\mathbb{R}$) no es compacto ni localmente compacto (ejercicio).

Definición 2.14. Una familia $\{U_\alpha\}_{\alpha\in A}$ de subconjuntos de un espacio topológico M es **localmente finita** si todo punto de M admite un entorno que tiene intersección no vacía con un número *finito* de conjuntos U_α.

Por ejemplo, $\{(n, n+2) \mid n \in \mathbb{Z}\}$ es un recubrimiento abierto localmente finito de $\mathbb{R}$. En efecto, si $k < a < b < m$, con $k, m \in \mathbb{Z}$, el intervalo (a, b) tiene intersección con un conjunto de la forma $(n, n+2)$ a lo suma para $n = k-1, k, \ldots, m-1$.

Definición 2.15. Un **refinamiento** de un recubrimiento $\{U_\alpha\}_{\alpha\in A}$ de un espacio topológico M es un recubrimiento $\{V_\beta\}_{\beta\in B}$ de M tal que para todo $\beta \in B$ existe un $\alpha \in A$ tal que $V_\beta \subset U_\alpha$.

[2] En el caso general, una base de la topología producto está formada por los conjuntos de la forma $\prod_\alpha U_\alpha$, donde U_α es abierto en M_α y $U_\alpha = M_\alpha$ excepto para un número finito de índices α.

Ejemplo 2.16. Un subrecubrimiento es un caso particular de refinamiento ($B \subset A$ y $\alpha = \beta$). El recubrimiento $\{(n, n+1] \mid n \in \mathbb{Z}\}$ de $\mathbb{R}$ es un refinamiento del recubrimiento del ejemplo anterior, ya que $(n, n+1] \subset (n, n+2)$ para todo $n \in \mathbb{Z}$.

Definición 2.17. Un espacio topológico *Hausdorff* M es **paracompacto** si todo recubrimiento abierto de M posee un refinamiento abierto localmente finito.

La paracompacidad es una generalización natural y muy útil de la compacidad: en efecto, un espacio topológico es compacto si todo recubrimiento abierto posee un *refinamiento* abierto finito. Por tanto, *todo espacio compacto y Hausdorff es paracompacto.*

Propiedades:

1. Un espacio topológico *Hausdorff*, *localmente compacto* y con una *base numerable* es *paracompacto.*

   Por ejemplo, los espacios $\mathbb{R}^n$, $\mathbb{C}^n$, $M_n(\mathbb{R})$ y $M_n(\mathbb{C})$ son paracompactos.

2. *Teorema de Stone:* un espacio topológico *metrizable* (es decir, en el que es posible definir una métrica que induce la topología) es *paracompacto*. En particular, los espacios *métricos* (normados, euclidianos) son paracompactos.

   El conjunto $\mathbb{Q}$, al ser metrizable, es paracompacto, aunque como sabemos *no* es localmente compacto.

3. *Teorema de Smirnov:* un espacio topológico es *metrizable* si y solo si es *paracompacto* y *localmente metrizable* (i.e., todo punto posee un entorno metrizable).

2.1.4 Conexión

Como ya hemos visto, un subconjunto A de un espacio topológico M es conexo si A no es la unión disjunta de dos subconjuntos abiertos (en la topología relativa) no vacíos.

Propiedades:

1. Los únicos subconjuntos conexos de $\mathbb{R}$ son los *intervalos*. En particular, $\mathbb{R}$ es conexo.

   $\mathbb{Q}$ (con la topología relativa) no es conexo. En efecto, si $x \in \mathbb{R} \setminus \mathbb{Q}$ podemos escribir $\mathbb{Q} = \big((-\infty, x) \cap \mathbb{Q}\big) \cup \big((x, \infty) \cap \mathbb{Q}\big)$. Un argumento parecido prueba que los únicos subconjuntos conexos de $\mathbb{Q}$ son los "puntos".

2. Ya se ha mencionado anteriormente que si $f : M \to M'$ es continua y M es conexo, entonces $f(M)$ es conexo (la demostración inmediata a partir de la definición).

   Este resultado implica el conocido *teorema de los valores intermedios:* si $f : M \to \mathbb{R}$ es continua en un espacio conexo M, dados dos valores $x < y$ de f para todo $z \in (x, y)$ existe $p \in M$ tal que $f(p) = z$. En efecto, $f(M)$ es conexo en $\mathbb{R}$, y por tanto ha de ser un *intervalo,* que por hipótesis contiene a x, y. Por tanto $[x, y] \subset f(M)$, como habíamos afirmado.

3. La unión $\bigcup_{\alpha \in A} C_\alpha$ de una familia cualquiera $\{C_\alpha \mid \alpha \in A\}$ de conjuntos conexos con intersección $\bigcap_{\alpha \in A} C_\alpha$ no vacía es conexa.

4. Si A es conexo y $A \subset B \subset \overline{A}$, donde

   $$\overline{A} = \bigcap_{C \supset A,\ C \text{ cerrado}} C$$

   es el **cierre** de A, entonces B es conexo. En particular, $\overline{A}$ es conexo si A es conexo.

5. El producto cartesiano $\prod_\alpha A_\alpha$ es conexo si y solo si cada A_α lo es. En particular, el espacio $\mathbb{R}^n$, y por tanto también $\mathbb{C}^n$, $M_n(\mathbb{R})$ y $M_n(\mathbb{C})$, son conexos.

Definición 2.18. Si M es un espacio topológico, la **componente conexa** $C(x)$ de un punto $x \in M$ es la unión de todos los subconjuntos conexos de M que contienen a x.

Nótese que, en virtud de la propiedad 3 de los conjuntos conexos, $C(x)$ es *conexo* para todo $x \in M$. De la definición de anterior se siguen inmediatamente las siguientes propiedades:

1. Cada componente conexa $C(x)$ es un conjunto conexo *maximal* (es decir, no propiamente contenido en ningún conjunto conexo).

2. El conjunto de todas las componentes conexas de M forma una *partición* de dicho espacio, es decir
$$M = \bigcup_{x \in M} C(x)\,, \qquad C(x) \cap C(y) \neq \emptyset \implies C(x) = C(y)\,.$$
En efecto, si $C(x) \cap C(y) \neq \emptyset$ la propiedad 3 de los conjuntos conexos implica que $C(x) \cup C(y)$ es conexo. Por la maximalidad de $C(x)$ y $C(y)$, de esto se sigue que $C(x) \cup C(y) = C(x) = C(y)$.

3. Las componentes conexas son conjuntos *cerrados*.
En efecto, $\overline{C(x)}$ es conexo (por la propiedad 4 de la conexión) y contiene a $C(x)$, lo cual implica que $C(x) = \overline{C(x)}$ por el apartado 1.

Ejemplo 2.19. Las componentes conexas *no* tienen por qué ser abiertas. Por ejemplo, si $M = \mathbb{Q}$ entonces $C(x) = \{x\}$ para todo $x \in \mathbb{Q}$. (Un espacio topológico en el que todas las componentes conexas se reducen a puntos se denomina *totalmente disconexo*.)

El número y la estructura de las componentes conexas de un espacio topológico son *invariantes topológicos:*

Proposición 2.20. *Si $f : M \to M'$ es continua entonces la imagen de cada componente de M está contenida en una componente de M'. En particular, si f es un* homeomorfismo *entonces $f\big(C(x)\big) = C\big(f(x)\big)$, y por tanto f induce una biyección (homeomorfismo) entre las componentes de M y M'.*

Otra noción importante de conexión es la *conexión local:*

Definición 2.21. Un espacio topológico es **localmente conexo** si posee una *base* formada por conjuntos conexos.

Ejemplo 2.22. El espacio $\mathbb{R}^n$ es localmente conexo, ya que las bolas abiertas son conjuntos conexos (cf. el Ejemplo 2.26) y forman una base de la topología. Por el mismo motivo, $\mathbb{C}^n$, $M_n(\mathbb{R})$ y $M_n(\mathbb{C})$ son localmente conexos. Los conjuntos anteriores son a la vez conexos y localmente conexos. Es fácil poner ejemplos de conjuntos localmente conexos que no son conexos (por ejemplo, $\mathbb{R}^n$ menos un hiperplano cualquiera). Hay, sin embargo, ejemplos de conjuntos conexos que no son localmente conexos. (Evidentemente, también hay conjuntos que no son conexos ni localmente conexos, como los racionales.)

Proposición 2.23. *Un espacio topológico M es localmente conexo si y solo si las componentes conexas de* cualquier *abierto de M son abiertas. En particular, en un espacio localmente conexo las componentes conexas son abiertas.*

Demostración. Si las componentes conexas de cualquier abierto de M son abiertas entonces el conjunto de las componentes conexas de todos los abiertos de M es una base de la topología cuyos elementos son conjuntos conexos. Por tanto en tal caso M es localmente conexo. Recíprocamente, supongamos que M es localmente conexo, sea $U \subset M$ abierto, sea C una componente conexa de U, y sea $x \in C$. Por hipótesis, hay un abierto conexo A tal que $x \in A \subset U$. Por definición de componente conexa $A \subset C$, lo que demuestra que C es abierto (ya que es un entorno de cualquiera de sus puntos). □

2.1.5 Conexión por arcos

Un concepto más intuitivo de conexión es la conexión por arcos, que definiremos a continuación.

Un **camino** (*path*, en inglés) en un espacio topológico M es una aplicación *continua* $f : [0,1] \to M$. Los puntos $f(0), f(1) \in M$ se denominan *extremos* del camino f, y se dice que están *unidos* por dicho camino.

Definición 2.24. Un espacio topológico M es **arco-conexo** (o *conexo por arcos*) si todo par de puntos de M pueden unirse por un camino.

Es inmediato probar que M es arco-conexo si y solo si existe un punto $x_0 \in M$ tal que cualquier otro punto $x \in M$ puede unirse a x_0 por un camino. La conexión por arcos es una propiedad topológica, ya que obviamente la imagen de un espacio arco-conexo bajo una función continua es arco-conexa (ejercicio).

Proposición 2.25. *Si un espacio topológico M es arco-conexo entonces es conexo.*

Demostración. Sea $x_0 \in M$, y sea C_x la imagen de un camino que una x_0 con otro punto cualquiera $x \in M$. Entonces C_x es conexo (imagen del intervalo $[0,1]$ bajo una función continua), y se cumple

$$M = \bigcup_{x \in M} C_x \,, \qquad x_0 \in \bigcap_{x \in M} C_x \,.$$

Por la propiedad 3 de los conjuntos conexos, M es conexo. □

Ejemplo 2.26. Los conjuntos $\mathbb{R}^n$, $\mathbb{C}^n$, $M_n(\mathbb{R})$ o $M_n(\mathbb{C})$ son claramente conexos por arcos, ya que el segmento que une dos puntos $x \neq y$ del espacio es un camino ($f(t) = (1-t)x + ty$). Lo mismo ocurre, por el mismo motivo, con las bolas (abiertas o cerradas) contenidas en dichos conjuntos (o, en general, en cualquier espacio normado).

Ejemplo 2.27. Hay conjuntos conexos que *no* son arco-conexos. Un ejemplo famoso es el conjunto

$$M = \big(\{0\} \times [-1,1]\big) \cup X \,, \qquad X = \big\{(x, \operatorname{sen} x^{-1}) \;\big|\; x \in (0,1]\big\} \,,$$

con la topología relativa como subespacio de $\mathbb{R}^2$. El conjunto X es claramente conexo (imagen del intervalo $(0,1]$ bajo una función continua), y por tanto $M = \overline{X}$ es conexo en virtud de la propiedad 4 de los conjuntos conexos. Sin embargo, puede probarse que no hay ningún camino (contenido en M) que une el origen y el punto $(\pi^{-1}, 0)$. Este ejemplo demuestra también que el *cierre* de un conjunto arco-conexo no es necesariamente arco-conexo, al contrario de lo que ocurre con los conjuntos conexos (cf. la propiedad 4 de la conexión).

Proposición 2.28. *La unión arbitraria de conjuntos arco-conexos con intersección no vacía es arco-conexa.*

Demostración. La demostración es inmediata, teniendo en cuenta el comentario a continuación de la Definición 2.24). □

Definición 2.29. Si M es un espacio topológico y $x \in M$, la **componente arco-conexa** de x es la unión de todos los subconjuntos arco-conexos de M que contienen a x.

En virtud de la proposición anterior, las componentes arco-conexas de un espacio topológico son arco-conexas, y por tanto conexas. De esto se sigue que las componentes arco-conexas de M constituyen una *partición* de las componentes conexas de dicho espacio (y, por tanto, del propio espacio M). Sin embargo, como se puede ver en el ejemplo anterior, las componentes arco-conexas (a diferencia de las conexas) *no* son necesariamente cerradas.

Proposición 2.30. *En un espacio topológico M, las siguientes propiedades son equivalentes:*

1. *Cada componente arco-conexa de M es abierta* (*y, por tanto, cerrada*).
2. *Todo punto de M posee un entorno abierto arco-conexo.*

Demostración.
1) $\Rightarrow$ 2) La componente arco-conexa de $x \in M$ es un entorno abierto arco-conexo de x, al ser dicha componente abierta por hipótesis.

2) $\Rightarrow$ 1) Sea A una componente arco-conexa, sea $x \in A$, y sea U un entorno abierto arco-conexo de x. Por definición de componente arco-conexa, $U \subset A$. Por tanto A es abierto, ya que es un entorno de cualquiera de sus puntos. Nótese, para finalizar, que si todas las componentes arco-conexas son abiertas entonces también son cerradas. En efecto, el complementario de una componente arco-conexa es la unión de las demás componentes arco-conexas de M, todas ellas abiertas por hipótesis. □

Corolario 2.31. *Sea M un espacio topológico en que todo punto posee un entorno abierto arco-conexo. Entonces las componentes conexas de M coinciden con sus componentes arco-conexas, y son por tanto abiertas y cerradas en M.*

Demostración. Sea A una componente arco-conexa de M, y sea C la componente conexa que contiene a A. Por la proposición anterior, A es abierto y cerrado en M, y por tanto en C. Al ser C conexo y A no vacío, de lo anterior se deduce que $A = C$. □

La proposición anterior nos permite demostrar fácilmente el importante resultado siguiente:

Teorema 2.32. *Un espacio topológico M es arco-conexo si y solo si es conexo, y todo punto $x \in M$ posee un entorno abierto arco-conexo.*

Demostración. En virtud de la Proposición 2.25, basta probar que si M es conexo y todo $x \in M$ posee un entorno abierto arco-conexo entonces M es arco-conexo. Esto es una consecuencia inmediata de la proposición anterior, ya que al ser M conexo solo tiene una componente conexa (el propio M). □

Corolario 2.33. *Un subconjunto* abierto *de $\mathbb{R}^n$, $\mathbb{C}^n$, $M_n(\mathbb{R})$ o $M_n(\mathbb{C})$ es conexo si y solo si es arco-conexo.*

Demostración. Por la Proposición 2.25, basta probar que si U es un abierto conexo de $\mathbb{R}^n$, $\mathbb{C}^n$, $M_n(\mathbb{R})$ o $M_n(\mathbb{C})$ entonces U es arco-conexo. Y, en efecto, si $x \in U$ entonces (por definición de abierto) existe $r > 0$ tal que $B_r(x) \subset U$. Como $B_r(x)$ es abierto y arco-conexo, esto demuestra que todo punto de U posee un entorno abierto arco-conexo contenido en U. Por la proposición anterior, U es arco-conexo. □

Nota. El resultado anterior es válido en cualquier *espacio normado* (ya que en un espacio normado las bolas abiertas son conexas y forman una base de la topología).

2.2 Variedades topológicas y diferenciables

2.2.1 Variedades topológicas

Definición 2.34. Una **variedad topológica** M de **dimensión** n es un espacio topológico M *Hausdorff* con una base *numerable* de abiertos tal que todo punto de M admite un entorno abierto *homeomorfo a un abierto de* $\mathbb{R}^n$.

En otras palabras, una variedad topológica de dimensión n es un espacio topológico Hausdorff con una base numerable de abiertos que es *localmente homeomorfo* a $\mathbb{R}^n$.

• Es evidente que cualquier subconjunto abierto de $\mathbb{R}^n$ es una variedad topológica de dimensión n. En general, un subconjunto *abierto* de una variedad topológica n-dimensional (con la topología inducida) es a su vez variedad topológica de dimensión n. Por ejemplo, cualquier abierto de $\mathbb{C}^n$, $M_n(\mathbb{R})$ o $M_n(\mathbb{C})$ es una variedad topológica de dimensión respectivamente igual a $2n$, n^2 y $2n^2$.

Ejercicio 21. Comprobar que la esfera $S^n \subset \mathbb{R}^{n+1}$ es una variedad topológica de dimensión n.

Solución. Para cada $i = 1, \ldots, n$, sea $U_i^{\pm}$ el abierto de S^n definido por

$$U_i^{\pm} = \left\{ x \in \mathbb{R}^{n+1} \;\middle|\; \pm x_i > 0 \right\} \cap S^n,$$

y definamos la función $\varphi_i : S^n \to \mathbb{R}^n$ mediante

$$\varphi_i(x) = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1}), \qquad \forall x \equiv (x_1, \ldots, x_n) \in S^n .$$

(La función φ_i es la proyección sobre el plano $x_i = 0$, identificado con $\mathbb{R}^n$.) Esta función es una *biyección* de cada uno de los dos abiertos $U_i^{\pm}$ en la bola unidad $B_1(0) \subset \mathbb{R}^n$, siendo su inversa la función $\psi_i^{\pm} : B_1(0) \to U_i^{\pm}$ dada por

$$\psi_i^{\pm}(y_1, \ldots, y_n) = \left(y_1, \ldots, y_{i-1}, \pm\sqrt{1 - y^2}, y_i, \ldots, y_n \right).$$

Tanto $\varphi_i^{\pm} \equiv \varphi_i\big|_{U_i^{\pm}}$ como su inversa $\psi_i^{\pm}$ son funciones continuas. En efecto, $\varphi_i^{\pm}$ es la restricción a $U_i^{\pm}$ de una función continua (lineal) de $\mathbb{R}^{n+1}$ en $\mathbb{R}^n$, mientras que $\psi_i^{\pm}$ es continua al serlo cada una de sus componentes. Por tanto $U_i^{\pm}$ es homeomorfo a la bola abierta $B_1(0) \subset \mathbb{R}^n$ para todo $i = 1, \ldots, n+1$. Como todo punto de S^n está en uno de los abiertos $U_i^{\pm}$ para algún i (ya que un punto $x \in S^n$ necesariamente ha de tener alguna componente no nula), esto prueba que S^n es localmente homeomorfa a $\mathbb{R}^n$. Por otra parte, es claro que S^n es Hausdorff y posee una base numerable, al ser subespacio topológico de $\mathbb{R}^n$. Por tanto S^n es una variedad topológica, como habíamos afirmado.

Nótese sin embargo que S^n, a pesar de ser localmente homeomorfa a $\mathbb{R}^n$, *no* es homeomorfa a ningún abierto de $\mathbb{R}^n$. En efecto, S^n es compacta (cerrada y acotada en $\mathbb{R}^{n+1}$), y el único abierto de $\mathbb{R}^n$ que es a la vez compacto (y, por tanto, cerrado) es el conjunto vacío.

Ejercicio 22. Probar que $\mathbb{Q}$ *no* es una variedad topológica.

Definición 2.35. Un subconjunto A de un espacio topológico M es **denso** si $\overline{A} = M$, o equivalentemente si todo abierto de M contiene algún punto de A. Un espacio topológico M es **separable** si contiene un conjunto *numerable denso*.

Ejemplo 2.36. El espacio topológico $\mathbb{R}^n$ es separable, ya que $\mathbb{Q}^n$ es numerable (producto cartesiano finito de conjuntos numerables) y denso. Otro tanto ocurre con los espacios $\mathbb{C}^n$, $M_n(\mathbb{R})$ y $M_n(\mathbb{C})$. En general, un espacio topológico con una base numerable es separable (en efecto, tomando un punto en cada conjunto de la base se obtiene un conjunto numerable denso).

Proposición 2.37. *Las variedades topológicas son espacios topológicos localmente conexos, localmente compactos, paracompactos, metrizables y separables. Además, las componentes conexas de una variedad topológica coinciden con sus componentes arco-conexas, y son por tanto simultáneamente abiertas y cerradas.*

Demostración.

1. M es localmente compacta al ser localmente euclidiana. En efecto, sea $x \in M$ y sea U un entorno abierto de x isomorfo bajo $\varphi : U \to \mathbb{R}^n$ a un abierto $\varphi(U) \subset \mathbb{R}^n$. Si $r > 0$ es tal que $\overline{B_r\big(\varphi(x)\big)} \equiv K \subset \varphi(U)$, entonces $\varphi^{-1}(K)$ es un entorno compacto de x.

2. M es paracompacta en virtud del apartado anterior y de la primera propiedad de la paracompacidad.

3. M es localmente metrizable (ya que es localmente euclidiana), y por tanto es metrizable en virtud del teorema de Smirnov.

4. M es separable, al poseer una base numerable.

5. M es localmente conexa al ser localmente euclidiana. En efecto, si U es un abierto de M y $x \in U$, por definición de variedad topológica hay una abierto V que contiene a x homeomorfo bajo una aplicación $\varphi : V \to \varphi(V) \subset \mathbb{R}^n$ a un abierto $\varphi(V)$ de $\mathbb{R}^n$. Entonces $U \cap V$ contiene a x y, al ser un subconjunto abierto de V, es homeomorfo al abierto $\varphi(U \cap V) \subset \mathbb{R}^n$. Si $r > 0$ es tal que $B_r\big(\varphi(x)\big) \subset \varphi(U \cap V)$, el conjunto $\varphi^{-1}\big(B_r\big(\varphi(x)\big)\big)$ es un abierto conexo de $U \cap V \subset U$ que contiene a x. La unión de todos estos conjuntos forma por tanto una base de la topología de M.

6. Al ser M localmente euclidiana todo punto $x \in M$ posee un entorno abierto arco-conexo (imagen inversa de una bola abierta bajo un homeomorfismo), de donde se sigue la igualdad de las componentes conexas y arco-conexas de M y su carácter abierto y cerrado en virtud del Corolario 2.31.

□

Proposición 2.38. *Un subconjunto* abierto U *de una variedad topológica* M *es conexo si y solo si es arco-conexo.*

Demostración. Al ser M localmente euclidiana, todo punto $x \in U$ posee un entorno abierto arco-conexo. El enunciado se sigue entonces del Teorema 2.32. □

Proposición 2.39. *Las componentes conexas de una variedad topológica forman una familia a lo sumo* numerable.

Demostración. Sea M una variedad topológica. Por la Proposición 2.37 M es separable, es decir existe un subconjunto $A = \{a_n \mid n \in \mathbb{N}\}$ denso en M. Si C es una componente conexa de M, por la Proposición 2.37 C es abierta, y por tanto existe $n \in \mathbb{N}$ tal que $a_n \in C$. Esto implica que $C = C(a_n)$, ya que las componentes conexas forman una partición de M. Por tanto la familia de las componentes conexas de M coincide con el conjunto $\{C(a_n) \mid n \in \mathbb{N}\}$, que es a lo sumo numerable. □

2.2.2 Variedades diferenciables

Sea M una variedad topológica n-dimensional, y sea $a \in M$. Si U es un entorno abierto del punto a homeomorfo a un abierto $x(U) \subset \mathbb{R}^n$ bajo la aplicación $x : U \to x(U) \subset \mathbb{R}^n$, para cada $p \in U$ podemos considerar el vector

$$x(p) \equiv (x_1(p), \ldots, x_n(p)) \in \mathbb{R}^n$$

como las **coordenadas** del punto p en el sistema de coordenadas (U, x). En general, puede existir otro entorno abierto V de a y otro homeomorfismo $y : V \to y(V) \subset \mathbb{R}^n$, dando lugar a otro sistema de coordenadas (V, y) en el entorno abierto V. Los sistemas de coordenadas (U, x) y (V, y) están ambos definidos en el entorno abierto $U \cap V$ del punto $a \in M$, siendo en general distintos. Sin embargo, está garantizado que las *funciones cambio de coordenadas*

$$y \circ x^{-1} : x(U \cap V) \to y(U \cap V)\,, \qquad x \circ y^{-1} : y(U \cap V) \to x(U \cap V)$$

son *continuas*, al ser composición de funciones continuas. Cuando es posible conseguir (restringiendo adecuadamente los entornos y las funciones coordenadas) una familia de sistemas de coordenadas alrededor de cada punto de M de forma que los correspondientes cambios de coordenadas *sean de clase* C^k (con k un entero positivo, infinito o ω, siendo por definición C^ω la clase de las funciones *analíticas*) se dice que M es una variedad diferenciable de clase C^k.

Definición 2.40. Sea M una variedad topológica de dimensión n. Una **carta** en M es un par (U, φ), siendo U un abierto de M y $\varphi : U \to \varphi(U) \subset \mathbb{R}^n$ un homeomorfismo. Un **atlas** es una familia de cartas $\mathcal{A} = \{(U_\alpha, \varphi_\alpha) \mid \alpha \in A\}$ cuyos dominios U_α recubren M (es decir, $M = \bigcup_{\alpha \in A} U_\alpha$). Un **atlas de clase** C^k es un atlas $\mathcal{A} = \{(U_\alpha, \varphi_\alpha) \mid \alpha \in A\}$ con la siguiente propiedad: para todo $\alpha, \beta \in A$ tal que $U_\alpha \cap U_\beta \neq \emptyset$, las funciones cambio de coordenadas

$$\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)\,, \qquad \varphi_\alpha \circ \varphi_\beta^{-1} : \varphi_\beta(U_\alpha \cap U_\beta) \to \varphi_\alpha(U_\alpha \cap U_\beta)$$

son de clase C^k. Un atlas de clase C^k es **maximal** si no está propiamente contenido en otro atlas de clase C^k. Un atlas maximal de clase C^k en una variedad topológica se denomina también **estructura diferenciable de clase** C^k.

• Es fácil probar que todo atlas de clase C^k define un único atlas maximal de clase C^k que lo contiene, y que dos atlas de clase C^k definen el mismo atlas maximal si y solo si los cambios de coordenadas entre cartas de ambos atlas son de clase C^k.

Definición 2.41. Una **variedad diferenciable** de clase C^k y dimensión n es una variedad topológica M de dimensión n provista de un atlas maximal de clase C^k. Si $k = \omega$, la variedad M se denomina **analítica**.

Nótese que para definir una estructura diferenciable de clase C^k en una variedad topológica M basta con dar un atlas de clase C^k en M (ya que este atlas determina un único atlas maximal, en virtud del comentario anterior). A partir de ahora, nos ocuparemos casi exclusivamente de las variedades de clase C^∞, a las que llamaremos abreviadamente *variedades diferenciables*, o incluso, si no hay lugar a confusión, *variedades*.

• Es evidente que cualquier subconjunto *abierto* $U \subset \mathbb{R}^n$ es una variedad analítica de dimensión n. En efecto, en este caso podemos tomar un atlas $\{(U, \varphi)\}$ con una sola carta, por lo que la única función cambio de coordenadas es la identidad de U en U. Es también obvio que un abierto cualquiera U de una variedad diferenciable M es a su vez variedad diferenciable, de la misma dimensión que M. En efecto, basta tomar como atlas en U el formado por los pares de la forma $(U \cap U_\alpha, \varphi_\alpha)$, siendo $\left\{(U_\alpha, \varphi_\alpha) \;\middle|\; \alpha \in A\right\}$ un atlas de M.

Ejemplo 2.42. Los conjuntos $\mathrm{GL}(n, \mathbb{R})$ y $\mathrm{GL}(n, \mathbb{C})$ son variedades diferenciables de dimensión respectivamente igual a n^2 y $2n^2$. En efecto,

$$\mathrm{GL}(n, \mathbb{R}) = \det^{-1}(\mathbb{R} \setminus \{0\}),$$

siendo $\det : M_n(\mathbb{R}) \approx \mathbb{R}^{n^2} \to \mathbb{R}$ una función continua en todo punto $X \in M_n(\mathbb{R})$ (un *polinomio* en los elementos de matriz x_{ij}). Al ser $\mathbb{R} \setminus \{0\}$ abierto y $\det$ continua, $\mathrm{GL}(n, \mathbb{R})$ es un abierto de $M_n(\mathbb{R})$, lo cual prueba nuestra afirmación. Análogamente,

$$\mathrm{GL}(n, \mathbb{C}) = \det^{-1}(\mathbb{R}^2 \setminus \{0\}),$$

donde ahora $\det$ se considera una función de $M_n(\mathbb{C}) \approx \mathbb{R}^{2n^2}$ en $\mathbb{C} \approx \mathbb{R}^2$. Las dos componentes de esta función son la parte real y la parte imaginaria de $\det(X)$, y son por tanto funciones polinómicas de las coordenadas de la matriz X (es decir, de las partes real e imaginaria de sus elementos de matriz x_{ij}). Esto implica, de nuevo, que $\det$ es continua y que $\mathrm{GL}(n, \mathbb{C})$ es un abierto de $M_n(\mathbb{C})$, al ser la imagen inversa de un abierto de $\mathbb{R}^2$ bajo una función continua de $M_n(\mathbb{C})$ en $\mathbb{R}^2$.

Ejercicio 23. Probar que la esfera S^n es una variedad diferenciable (analítica) de dimensión n.

Solución. Tomemos como atlas en S^n el definido en el Ejercicio 21. Entonces la función cambio de coordenadas $\varphi_i \circ \varphi_j^{-1} : \varphi_j(U_i^\varepsilon \cap U_j^{\varepsilon'}) \to \varphi_i(U_i^\varepsilon \cap U_j^{\varepsilon'})$ está dada por

$$(\varphi_i \circ \varphi_j^{-1})(x_1, \ldots, x_n) = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, \underset{j}{\varepsilon' \sqrt{1 - |x|^2}}, \ldots, x_n)$$

(donde se ha supuesto que, por ejemplo, $i < j$), y análogamente para $\varphi_j \circ \varphi_i^{-1}$. Como estas aplicaciones son claramente C^∞ (analíticas, de hecho) en sus respectivos dominios de definición (en los que $|x| < 1$), el atlas considerado es C^∞ (analítico).

Proposición 2.43. *Si M y N son variedades diferenciables de dimensiones respectivas m y n, el* producto cartesiano $M \times N$ (*con la topología producto*) *es una variedad de dimensión $m + n$.*

Demostración. En primer lugar, es fácil probar que el producto cartesiano de variedades topológicas, con la topología producto, es una variedad topológica. En segundo lugar, si $\{(U_\alpha, \varphi_\alpha) \,|\, \alpha \in A\}$ y $\{(V_\beta, \psi_\beta) \,|\, \beta \in B\}$ son las estructuras diferenciables en M y N, respectivamente, entonces es fácil comprobar que los pares

$$\big(U_\alpha \times V_\beta, \varphi_\alpha \times \psi_\beta\big)\,, \qquad \alpha \in A\,,\ \ \beta \in B\,,$$

donde

$$\big(\varphi_\alpha \times \psi_\beta\big)(x, y) = \big(\varphi_\alpha(x), \psi_\beta(y)\big)\,,$$

forman un atlas C^∞ en $M \times N$, y por tanto definen una estructura diferenciable en dicho espacio. □

2.2.3 Subvariedades regulares

Definición 2.44. Sea M una variedad diferenciable de dimensión m. Una **subvariedad regular** de M de dimensión n es un subconjunto $N \subset M$ con la siguiente propiedad: para todo punto $a \in N$, existe una carta (U, φ) de M que contiene al punto a tal que $\varphi(a) = 0$ y

$$\varphi(U \cap N) = \varphi(U) \cap \big\{x \in \mathbb{R}^m \,\big|\, x_{n+1} = \cdots x_m = 0\big\}\,.$$

Se dice que la carta (U, φ) de M es una **carta adaptada** a la subvariedad N en el punto a.

Es inmediato comprobar que los pares

$$(U \cap N, \tilde{\varphi}), \qquad \text{con } \ \tilde{\varphi} \equiv (\varphi_1, \ldots, \varphi_n)\,,$$

donde (U, φ) es cualquier carta de M adaptada a N, forman una estructura diferenciable en N (con la *topología relativa*). Por tanto la subvariedad regular N es una *variedad diferenciable* de dimensión n.

El siguiente resultado permite probar de forma sencilla que un subconjunto $M \subset \mathbb{R}^p$ es una subvariedad diferenciable de $\mathbb{R}^p$, y por tanto una variedad diferenciable:

Proposición 2.45. *Sea $f : U \subset \mathbb{R}^p \to \mathbb{R}^q$ una función de clase C^∞ en un abierto U de $\mathbb{R}^p$, y sea $M = f^{-1}(0)$. Si el* rango *de $Df(x)$ es igual a k para todo $x \in M$, el conjunto M (con la* topología relativa) *es una subvariedad regular de $\mathbb{R}^p$ de dimensión $n = p - k$.*

Demostración. Es esencialmente consecuencia del *teorema de la función implícita*. □

Ejemplo 2.46. Consideremos, por ejemplo, la función $f : \mathbb{R}^{n+1} \to \mathbb{R}$ dada por $f(x) = |x|^2 - 1$. En este caso $U = \mathbb{R}^{n+1}$, $p = n + 1$, $q = 1$ y $M = S^n$, siendo

$$\nabla f(x) = 2x \neq 0\,, \qquad \forall x \in M$$

(ya que ∇f solo se anula en el origen, que no pertenece a M). Por tanto el rango de Df es igual a 1 para todo $x \in M = S^n$, lo que prueba que S^n es una variedad diferenciable (subvariedad regular de $\mathbb{R}^{n+1}$) de dimensión $n + 1 - 1 = n$.

Ejercicio 24. Probar que los conjuntos

$$\mathrm{SL}(n, \mathbb{F}) = \big\{X \in M_n(\mathbb{F}) \,\big|\, \det X = 1\big\}$$

son variedades diferenciables de dimensión $n^2 - 1$ (si $\mathbb{F} = \mathbb{R}$) o $2(n^2 - 1)$ (si $\mathbb{F} = \mathbb{C}$).

Solución. En ambos casos,

$$\mathrm{SL}(n, \mathbb{F}) = f^{-1}(0), \qquad \text{con } \ f \equiv \det - 1 : M_n(\mathbb{F}) \approx \mathbb{F}^{n^2} \to \mathbb{F}\,,$$

con f diferenciable (polinómica). Consideremos, en primer lugar, el caso $\mathbb{F} = \mathbb{R}$. De la fórmula para el desarrollo del determinante por la i-ésima fila (o la j-ésima columna) se sigue inmediatamente que

$$\frac{\partial f}{\partial x_{ij}}(X) = X_{ij}\,,$$

donde hemos denotado por X_{ij} el *adjunto* del elemento de matriz x_{ij} (es decir, $(-1)^{i+j}$ veces el menor de X obtenido suprimiendo la fila i y la columna j). Como $\det X = 1 \neq 0$ en $M = \mathrm{SL}(n, \mathbb{R})$, para todo $X \in M$ existe (i, j) tal que $\dfrac{\partial f}{\partial x_{ij}}(X) \neq 0$. Por tanto $\nabla f \neq 0$ en M, y el rango de f es igual a 1 para todo $x \in M$. De la proposición anterior se deduce entonces que M es una variedad de dimensión $n^2 - 1$.

Consideremos, a continuación, el caso $\mathbb{F} = \mathbb{C}$. Llamando $Z \equiv X + \mathrm{i}Y \in M_n(\mathbb{C})$ (siendo $X, Y \in M_n(\mathbb{R})$ las partes real e imaginaria de la matriz Z) y $f \equiv f_1 + \mathrm{i} f_2$ (con $f_1 = \mathrm{Re}\, f$, $f_2 = \mathrm{Im}\, f$), al igual que antes se tiene

$$\frac{\partial f}{\partial z_{ij}}(Z) = Z_{ij} \,.$$

Como $\det Z = 1 \neq 0$ en $M = \mathrm{SL}(n, \mathbb{C})$, para todo $Z \in M$ existe (i, j) tal que $\dfrac{\partial f}{\partial z_{ij}}(Z) \neq 0$. Utilizando las ecuaciones de Cauchy–Riemann se obtiene entonces

$$\left| \frac{\partial f}{\partial z_{ij}}(Z) \right|^2 = \left(\frac{\partial f_1}{\partial x_{ij}} \frac{\partial f_2}{\partial y_{ij}} - \frac{\partial f_1}{\partial y_{ij}} \frac{\partial f_2}{\partial x_{ij}} \right)(Z) \equiv \frac{\partial (f_1, f_2)}{\partial (x_{ij}, y_{ij})} \neq 0 \quad \Longrightarrow \quad \mathrm{rank}\, Df(Z) = 2 \,.$$

Por tanto f tiene rango constante igual a 2 en M, de donde se sigue que M es una variedad diferenciable de dimensión $2n^2 - 2$ en virtud de la proposición anterior. □

Ejercicio 25. Probar que

$$D \det(\mathbb{1}) \cdot h = \mathrm{tr}\, h \,, \qquad \forall h \in M_n(\mathbb{C}) \,. \tag{2.1}$$

Solución. Si $f = \det$ y $h \in M_n(\mathbb{C})$ se tiene:

$$Df(\mathbb{1}) \cdot h = \left. \frac{\mathrm{d}}{\mathrm{d}t} \right|_{t=0} f(\mathbb{1} + th) = \sum_{j=1}^{n} \begin{vmatrix} 1 & \cdots & h_{1j} & \cdots & 0 \\ & \ddots & & & \\ \vdots & & h_{jj} & & \vdots \\ & & & \ddots & \\ 0 & \cdots & h_{nj} & \cdots & 1 \end{vmatrix} = \sum_{j=1}^{n} h_{jj} \equiv \mathrm{tr}\, h \,.$$

Ejercicio 26. Probar que si N_1 y N_2 son subvariedades regulares de sendas variedades M_1 y M_2 entonces $N_1 \times N_2$ es una subvariedad regular de $M_1 \times M_2$, de dimensión

$$\dim(N_1 \times N_2) = \dim N_1 + \dim N_2 \,.$$

2.2.4 Funciones diferenciables

Sea $f : M \to N$ una aplicación entre dos variedades de dimensiones m y n, respectivamente, y sea $a \in M$. Si tomamos una carta (U, φ) que contenga a a y otra carta (V, ψ) (en N) que contenga a $f(a)$, es natural considerar la aplicación $\hat{f}$ entre las *coordenadas* de los puntos de U y los de V inducida por f, es decir

$$\hat{f} = \psi \circ f \circ \varphi^{-1} : \varphi\left(U \cap f^{-1}(V)\right) \subset \mathbb{R}^m \to \psi\left(f(U) \cap V\right) \subset \mathbb{R}^n \,,$$

siendo $U \cap f^{-1}(V)$ un *abierto* de $\mathbb{R}^m$. Diremos que $\hat{f}$ es la **expresión** (o el **representante**) de f en las coordenadas que hemos escogido alrededor de $a \in M$ y de $f(a) \in N$. Nótese que f es continua (en $U \cap f^{-1}(V)$) si y solo si lo es $\hat{f}$, al ser φ y ψ homeomorfismos. Las consideraciones anteriores justifican la siguiente definición:

Definición 2.47. Una función $f : M \to N$ es **diferenciable** (C^∞) en $a \in M$ si existen sendas cartas (U, φ) de M y (V, ψ) de N tales que $a \in U$, $f(a) \in V$ y $\hat{f} \equiv \psi \circ f \circ \varphi^{-1}$ es C^∞ en $\varphi(a)$. La función f es diferenciable en un *abierto* $U \subset M$ si es diferenciable en cada punto de U.

• Es importante notar que la definición *no depende de las cartas consideradas.* En efecto, si (U_1, φ_1) es otra carta tal que $a \in U \cap U_1$ y (V_1, ψ_1) es una carta en N con $f(a) \in V \cap V_1$ entonces

$$\hat{f}_1 \equiv \psi_1 \circ f \circ \varphi_1^{-1} = (\psi_1 \circ \psi^{-1}) \circ \hat{f} \circ (\varphi \circ \varphi_1^{-1})$$

es C^∞ en $\varphi_1(a)$, por serlo los cambios de coordenadas en virtud de la definición de variedad diferenciable.

Ejercicio 27. Probar que la composición de funciones diferenciables es diferenciable.

• Una **función suave** en M es una función $f : M \to \mathbb{R}$ diferenciable (C^∞) en M. Denotaremos por

$$C^\infty(M) = \left\{ f : M \to \mathbb{R} \;\middle|\; f \text{ diferenciable en } M \right\}$$

al espacio vectorial de las funciones suaves en M.

De la Proposición 2.45 se sigue el siguiente corolario, que nos será de utilidad más adelante:

Corolario 2.48. *Sea N una* subvariedad regular *de una variedad M, sea M' una variedad diferenciable, y sea $f : M \to M'$ una función diferenciable en un punto $a \in N$. Entonces la restricción $f|_N : N \to M'$ es también diferenciable en a.*

Demostración. Sean, en efecto, (U, φ) una carta de M adaptada a N en el punto a, y sea $(U \cap N, \tilde{\varphi})$ la correspondiente carta de N (recuérdese que $\tilde{\varphi} = (\varphi_1, \ldots, \varphi_n)$, siendo $n = \dim N$). Si (V, ψ) es una carta cualquiera en M' con $f(a) \in V$ y $\hat{f} = \psi \circ f \circ \varphi^{-1}$ es el representante de f respecto de las cartas (U, φ) y (V, ψ), es inmediato comprobar que

$$\tilde{\varphi}^{-1}(x_1, \ldots, x_n) = \varphi^{-1}(x_1, \ldots, x_n, 0, \ldots, 0)\,.$$

Por tanto el representante de $f|_N$ respecto de las cartas $(U \cap N, \tilde{\varphi})$ y (V, ψ) de N y M' está dado por

$$\widehat{f|_N}(x_1, \ldots, x_n) = (\psi \circ f \circ \varphi^{-1})(x_1, \ldots, x_n, 0, \ldots, 0) \equiv \hat{f}(x_1, \ldots, x_n, 0, \ldots, 0)\,.$$

Como $\hat{f}$ es por hipótesis diferenciable en $\varphi(a) = 0 \in \mathbb{R}^m$ ($m \equiv \dim M$), lo mismo ocurre con $\widehat{f|_N}$ en $\tilde{\varphi}(a) = 0 \in \mathbb{R}^n$. □

Ejercicio 28. Sea $f : M \to M'$ (siendo M, M' sendas variedades diferenciables), y supongamos que $f(M)$ está contenida en una subvariedad regular N de M'. Probar que $f : M \to M'$ es diferenciable si y solo si $f : M \to N$ lo es.

Definición 2.49. Sean M, N dos variedades diferenciables. Una aplicación $f : M \to N$ es un **difeomorfismo** si es biyectiva, y tanto f como f^{-1} son diferenciables (en M y N, respectivamente). Se dice en tal caso que las variedades diferenciables M y N son **difeomorfas**.

Dos variedades diferenciables difeomorfas son completamente *equivalentes* desde el punto de vista de la teoría de variedades diferenciables. En particular, dos variedades difeomorfas tienen la *misma dimensión*. En efecto, si $f : M \to N$ es un difeomorfismo y (U, φ), (V, ψ) son sendas cartas de M y N, respectivamente, entonces el representante $\hat{f} = \psi \circ f \circ \varphi^{-1}$ de f en dichas cartas es un difeomorfismo, y por tanto un homeomorfismo, entre los abiertos $\varphi\big(U \cap f^{-1}(V)\big) \subset \mathbb{R}^m$ y $\psi\big(f(U) \cap V\big) \subset \mathbb{R}^n$ (donde m y n denotan respectivamente las dimensiones de M y N). Por el teorema de invariancia del dominio[3], m ha de ser igual a n.

[3]Este teorema (consecuencia a su vez del teorema del punto fijo de Brouwer, según el cual toda aplicación continua de una bola cerrada en sí misma tiene un punto fijo) afirma que si $f : U \subset \mathbb{R}^n \to \mathbb{R}^n$ es inyectiva y continua, y U es abierto, entonces $f(U)$ es abierto y $f : U \to f(U)$ es un homeomorfismo. Este resultado implica que *ningún subconjunto abierto de $\mathbb{R}^m$ puede ser homeomorfo a un abierto de $\mathbb{R}^n$ si $m \neq n$*. En efecto, si $f : U \subset \mathbb{R}^m \to V \subset \mathbb{R}^n$ es un homeomorfismo entre dos conjuntos abiertos y (por ejemplo) $m > n$ podemos identificar $\mathbb{R}^n$ con el subespacio vectorial $x_{n+1} = \cdots = x_m = 0$ de $\mathbb{R}^m$, y V con un subconjunto de dicho subespacio. Esto contradice el teorema de invariancia del dominio, ya que V no es abierto en $\mathbb{R}^m$.

2.3 Grupos topológicos y de Lie

2.3.1 Definiciones y ejemplos

Sea G un grupo, y denotemos por $m : G \times G \to G$ e $i : G \to G$ la multiplicación y la inversa en G:

$$m(x, y) = xy\,, \quad i(x) = x^{-1}\,, \qquad \forall x, y \in G\,.$$

Dado un elemento $a \in G$, denotaremos respectivamente por L_a y R_a las aplicaciones de G en G dadas por

$$L_a(x) = ax\,, \quad R_a(x) = xa\,, \qquad \forall x \in G$$

(multiplicación a izquierdas o derechas por a). La aplicación

$$\mathrm{Ad}_a = L_a \circ R_{a^{-1}} = R_{a^{-1}} \circ L_a\,,$$

o equivalentemente

$$\mathrm{Ad}_a(x) = axa^{-1}\,, \qquad \forall x \in G\,,$$

recibe el nombre de **conjugación** bajo el elemento $a \in G$. Nótese que para todo $a \in G$ las aplicaciones L_a, R_a y Ad_a son claramente biyectivas, siendo

$$L_a^{-1} = L_{a^{-1}}\,, \qquad R_a^{-1} = R_{a^{-1}}\,, \qquad \mathrm{Ad}_a^{-1} = \mathrm{Ad}_{a^{-1}}\,.$$

Definición 2.50. Un **grupo topológico** es un grupo G que es además una *variedad topológica*, de forma que las aplicaciones $m : G \times G \to G$ e $i : G \to G$ son funciones *continuas*.

Nótese que si G es una variedad topológica también lo es $G \times G$ (con la topología producto).

Definición 2.51. Un **grupo de Lie** (resp. grupo de Lie *analítico*) es un grupo G que es además una *variedad diferenciable*, de forma que las aplicaciones m e i son funciones *diferenciables* (resp. *analíticas*).

• Todo grupo de Lie es obviamente un grupo topológico. De hecho, un célebre resultado de von Neumann, Gleason, Montgomery y Zippin afirma que todo grupo topológico admite una estructura de grupo de Lie analítico compatible con su topología (5º problema de Hilbert). Más aún, si G es un grupo de Lie de clase C^k, con $1 \leqslant k \leqslant \infty$, entonces el atlas diferenciable de G contiene un subatlas analítico.

• De la definición de grupo de Lie (resp. topológico) se sigue inmediatamente que para todo $a \in G$ las aplicaciones L_a, R_a y Ad_a son *difeomorfismos* (respectivamente *homeomorfismos*) de G en sí mismo.

Ejemplo 2.52. Los grupos aditivos $\mathbb{R}^n$ y $\mathbb{C}^n$ son grupos de Lie (analíticos) de dimensiones respectivas n y $2n$. Los conjuntos $\mathbb{R}^*$ y $\mathbb{C}^*$ son grupos de Lie (analíticos) de dimensiones 1 y 2, respectivamente. Por ejemplo, en el caso de $\mathbb{R}^n$ la multiplicación y la inversa están dadas por

$$m(x, y) = x + y\,, \qquad i(x) = -x\,,$$

y son por tanto funciones diferenciables al ser lineales. Análogamente, en el caso de $\mathbb{R}^*$ se tiene

$$m(x, y) = xy\,, \qquad i(x) = x^{-1}\,,$$

que de nuevo son funciones diferenciables en sus respectivos dominios.

Ejemplo 2.53. Los grupos $\mathrm{GL}(n, \mathbb{F})$ y $\mathrm{SL}(n, \mathbb{F})$ son grupos de Lie, de dimensiones respectivas n^2 y $n^2 - 1$ (si $\mathbb{F} = \mathbb{R}$) o $2n^2$ y $2(n^2 - 1)$ (si $\mathbb{F} = \mathbb{C}$). En efecto, ya hemos visto anteriormente que estos grupos son variedades diferenciables de las dimensiones indicadas (cf. el Ejemplo 2.42 y el Ejercicio 24). Basta probar, por tanto, que el producto y el inverso en dichos grupos son aplicaciones diferenciables. En primer lugar, el producto es diferenciable como aplicación de $M_n(\mathbb{F}) \times M_n(\mathbb{F}) \to M_n(\mathbb{F})$, al ser un polinomio en los elementos de matriz de sus factores. Como $\mathrm{GL}(n, \mathbb{F})$ es un *abierto* de $M_n(\mathbb{F})$, de esto se

sigue que $m : \mathrm{GL}(n,\mathbb{F}) \times \mathrm{GL}(n,\mathbb{F}) \to \mathrm{GL}(n,\mathbb{F})$ es diferenciable en $\mathrm{GL}(n,\mathbb{F}) \times \mathrm{GL}(n,\mathbb{F})$. En segundo lugar, los elementos de matriz de la inversa de una matriz $X \in \mathrm{GL}(n,\mathbb{F})$ son funciones racionales en los elementos de matriz de X cuyo denominador ($\det X$) no se anula en el abierto $\mathrm{GL}(n,\mathbb{F})$. Por tanto $i : \mathrm{GL}(n,\mathbb{F}) \to \mathrm{GL}(n,\mathbb{F})$ es también diferenciable en $\mathrm{GL}(n,\mathbb{F})$, lo que prueba que este conjunto es un grupo de Lie. Como $\mathrm{SL}(n,\mathbb{F})$ es una *subvariedad regular* de $M_n(\mathbb{F})$ (o de $\mathrm{GL}(n,\mathbb{F})$), las aplicaciones producto e inverso son también diferenciables en este grupo en virtud del Corolario 2.48. Esto demuestra que $\mathrm{SL}(n,\mathbb{F})$ es un grupo de Lie.

Probaremos a continuación que los llamados *grupos matriciales clásicos* son grupos de Lie. Para estudiar estos grupos, introduciremos la siguiente notación. En primer lugar, si $B \in \mathrm{GL}(n,\mathbb{F})$ (donde a partir de ahora $\mathbb{F} = \mathbb{R}$ o $\mathbb{C}$) es una matriz invertible fija, definimos el conjunto

$$G_1 = \left\{ X \in M_n(\mathbb{F}) \;\middle|\; X^{\mathsf{T}} B X = B \right\}. \tag{2.2}$$

Es inmediato comprobar que G_1 es un *subgrupo* de $\mathrm{GL}(n,\mathbb{F})$ (y, por lo tanto, es un grupo). En efecto, si $X, Y \in G_1$ se tiene

$$(XY)^{\mathsf{T}} B (XY) = Y^{\mathsf{T}} (X^{\mathsf{T}} B X) Y = Y^{\mathsf{T}} B Y = B \quad \Longrightarrow \quad XY \in G_1 .$$

Por otra parte, al ser B invertible $\det B \neq 0$, y por tanto para todo $X \in G_1$ se tiene

$$(\det X)^2 = 1 \quad \Longrightarrow \quad X \in \mathrm{GL}(n,\mathbb{F}) .$$

Además,

$$X^{\mathsf{T}} B X = B \quad \Longrightarrow \quad B = \left(X^{-1}\right)^{\mathsf{T}} B X^{-1} \quad \Longrightarrow \quad X^{-1} \in G_1 .$$

Si $\mathbb{F} = \mathbb{C}$ se define un segundo conjunto G_2 mediante

$$G_2 = \left\{ X \in M_n(\mathbb{C}) \;\middle|\; X^{\dagger} B X = B \right\}. \tag{2.3}$$

Al igual que antes, es inmediato comprobar que G_2 es un subgrupo de $\mathrm{GL}(n,\mathbb{C})$.

Proposición 2.54. *Para toda matriz $B \in \mathrm{GL}(n,\mathbb{F})$, los grupos G_1 y G_2 son grupos de Lie.*

Demostración. Dado que el producto $m : M_n(\mathbb{F}) \times M_n(\mathbb{F}) \to M_n(\mathbb{F})$ y el inverso $i : \mathrm{GL}(n,\mathbb{F}) \to \mathrm{GL}(n,\mathbb{F})$ son aplicaciones diferenciables en sus dominios, por el Corolario 2.48 basta probar que tanto G_1 como G_2 son *subvariedades regulares* de $M_n(\mathbb{F})$ (o, equivalentemente, de $\mathrm{GL}(n,\mathbb{F})$). Consideremos, en primer lugar, el grupo G_1. Evidentemente, $G_1 = f^{-1}(0)$, siendo $f : M_n(\mathbb{F}) \to M_n(\mathbb{F})$ la aplicación dada por

$$f(X) = X^{\mathsf{T}} B X - B \equiv g(X) - B .$$

Es evidente que esta aplicación es diferenciable, dado que los elementos de matriz de f son polinomios de segundo grado en los de X. Basta probar, por tanto, que f tiene rango constante en G_1. Equivalentemente (dado que el término $-B$ es constante), debemos probar que g tiene rango constante en G_1. Para ver esto último, nótese que si $A \in M_n(\mathbb{F})$ es una matriz cualquiera entonces

$$g(XA) = A^{\mathsf{T}} g(X) A ,$$

o equivalentemente

$$g \circ R_A = \varphi_A \circ g ,$$

donde $\varphi_A : M_n(\mathbb{F}) \to M_n(\mathbb{F})$ es la aplicación definida por

$$\varphi_A(X) = A^{\mathsf{T}} X A .$$

Tomando la derivada y aplicando la regla de la cadena se obtiene

$$Dg(XA)\, DR_A(X) = D\varphi_A\big(g(X)\big)\, Dg(X) , \qquad \forall X \in M_n(\mathbb{F}) .$$

Al ser tanto R_A como φ_A aplicaciones *lineales*, $DR_A = R_A$ y $D\varphi_A = \varphi_A$ en todo punto, y por tanto

$$Dg(XA)R_A = \varphi_A Dg(X)\,, \qquad \forall A, X \in M_n(\mathbb{F})\,.$$

Como R_A y φ_A son *invertibles* si $A \in \mathrm{GL}(n,\mathbb{F})$ (en efecto, $R_A^{-1} = R_{A^{-1}}$, $\varphi_A^{-1} = \varphi_{A^{-1}}$), de lo anterior se sigue que

$$\operatorname{rank} Dg(XA) = \operatorname{rank} Dg(X)\,, \qquad \forall X \in M_n(\mathbb{F})\,,\ \ \forall A \in \mathrm{GL}(n,\mathbb{F})\,.$$

De esta relación con $X = \mathbb{1}$ y $A \in G_1$ se deduce que, efectivamente, el rango de g (y, por tanto, el de f) es constante en G_1. La demostración para G_2 es completamente análoga (basta reemplazar la transpuesta por la adjunta en todas partes). □

Ejercicio 29. Calcular la dimensión de G_1 si B es una matriz simétrica o antisimétrica.

Solución. Aunque este cálculo es sencillo utilizando el álgebra de Lie de este grupo (ver más adelante), es instructivo realizarlo directamente aplicando la Proposición 2.45. En efecto, en virtud de dicha proposición basta calcular el rango de la aplicación Df (o, equivalentemente, Dg) definida en la proposición anterior. Como dicho rango es *constante* en G_1, puede calcularse (por ejemplo) en la matriz identidad. Para ello nótese que $Dg(\mathbb{1})\cdot h$ —donde $h \in M_n(\mathbb{F})$— es el término lineal en h en el desarrollo de $g(\mathbb{1}+h)$. De esta forma se obtiene la expresión

$$Dg(\mathbb{1})\cdot h = h^{\mathsf{T}}B + Bh\,, \qquad \forall h \in M_n(\mathbb{F})\,,$$

y por tanto

$$\operatorname{rank} Dg(\mathbb{1}) = \dim\left\{h^{\mathsf{T}}B + Bh \;\middle|\; h \in M_n(\mathbb{F})\right\}.$$

Para evaluar el miembro derecho, nótese que

$$B^{\mathsf{T}} = \pm B \quad\Longrightarrow\quad (h^{\mathsf{T}}B + Bh)^{\mathsf{T}} = \pm(h^{\mathsf{T}}B + Bh)\,,$$

y por consiguiente

$$\left\{h^{\mathsf{T}}B + Bh \;\middle|\; h \in M_n(\mathbb{F})\right\} \subset S_\pm\,,$$

donde S_+ (resp. S_-) denota el subespacio de las matrices simétricas (resp. antisimétricas) en $M_n(\mathbb{F})$. De hecho, ambos miembros de la ecuación anterior son *iguales*, ya que si $A \in S_\pm$ y $h = \frac{1}{2}B^{-1}A$ entonces

$$h^{\mathsf{T}}B + Bh = \frac{1}{2}(\pm A)(\pm B^{-1})B + \frac{1}{2}A = A\,.$$

En definitiva,

$$\operatorname{rank} Dg(\mathbb{1}) = \dim S_\pm = \begin{cases} \frac{1}{2}\,n(n\pm 1)\,, & \mathbb{F} = \mathbb{R} \\ n(n\pm 1)\,, & \mathbb{F} = \mathbb{C}\,, \end{cases}$$

y por tanto, en virtud de la Proposición 2.45:

$$\dim G_1 = \begin{cases} \frac{1}{2}\,n(n\mp 1)\,, & \mathbb{F} = \mathbb{R} \\ n(n\mp 1)\,, & \mathbb{F} = \mathbb{C}\,, \end{cases}$$

donde el signo "−" corresponde al caso en que B es simétrica y el "+" a aquél en que es antisimétrica. Nótese que en este último caso

$$B^{\mathsf{T}} = -B \quad\Longrightarrow\quad \det B = (-1)^n \det B\,,$$

y por tanto (al ser B invertible por hipótesis) n ha de ser necesariamente *par*.

Ejercicio 30. Hallar la dimensión de G_2 si B es una matriz autoadjunta.

Solución. Procediendo como en el ejercicio anterior se llega a las identidades

$$\operatorname{rank} Dg(\mathbb{1}) = \dim\{h^\dagger B + Bh \mid h \in M_n(\mathbb{C})\} = \dim S\,,$$

donde ahora S denota el subespacio de las matrices autoadjuntas de orden n. La dimensión *real* de dicho espacio es

$$n + 2\cdot\frac{1}{2}n(n-1) = n^2\,,$$

por lo que, en virtud de la Proposición 2.45:

$$\dim G_2 = 2n^2 - n^2 = n^2\,.$$

2.3.2 Los grupos matriciales clásicos

Ya hemos visto (cf. el Ejemplo 2.53) que los grupos matriciales $\mathrm{GL}(n,\mathbb{F})$ y $\mathrm{SL}(n,\mathbb{F})$ son grupos de Lie. Consideremos a continuación los siguientes conjuntos de matrices:

$$\begin{aligned}
\mathrm{O}(n,\mathbb{F}) &= \{X \in M_n(\mathbb{F}) \mid X^\mathsf{T}X = \mathbb{1}\}\,, & \mathrm{O}(n) &\equiv \mathrm{O}(n,\mathbb{R})\,,\\
\mathrm{SO}(n,\mathbb{F}) &= \mathrm{O}(n,\mathbb{F}) \cap \mathrm{SL}(n,\mathbb{F})\,, & \mathrm{SO}(n) &\equiv \mathrm{SO}(n,\mathbb{R})\,,\\
\mathrm{O}(p,q) &= \{X \in M_{p+q}(\mathbb{R}) \mid X^\mathsf{T}BX = B\}\,, & B &\equiv \begin{pmatrix}\mathbb{1}_p & 0\\ 0 & -\mathbb{1}_q\end{pmatrix} \quad (p,q \geqslant 1)\,,\\
\mathrm{SO}(p,q) &= \mathrm{O}(p,q) \cap \mathrm{SL}(p+q,\mathbb{R})\,,\\
\mathrm{U}(n) &= \{X \in M_n(\mathbb{F}) \mid X^\dagger X = \mathbb{1}\}\,, & \mathrm{SU}(n) &= \mathrm{U}(n) \cap \mathrm{SL}(n,\mathbb{C})\,,\\
\mathrm{SP}(n,\mathbb{F}) &= \{X \in M_{2n}(\mathbb{F}) \mid X^\mathsf{T}JX = J\}\,, & J &\equiv \begin{pmatrix}0 & \mathbb{1}_n\\ -\mathbb{1}_n & 0\end{pmatrix}\,,\\
\mathrm{SP}(n) &= \mathrm{SP}(n,\mathbb{C}) \cap \mathrm{U}(2n)\,.
\end{aligned}$$

De la Proposición 2.54 se deduce que $\mathrm{O}(n,\mathbb{F})$, $\mathrm{O}(p,q)$ y $\mathrm{SP}(n,\mathbb{F})$ son grupos de Lie, denominados respectivamente **grupo ortogonal**, **grupo pseudoortogonal** (de tipo (p,q)) y **grupo simpléctico**. Las dimensiones de estos grupos son

$$\begin{aligned}
&\dim \mathrm{O}(n,\mathbb{C}) = n(n-1)\,, \qquad \dim \mathrm{O}(n) = \dim \mathrm{O}(p,q) = \frac{1}{2}n(n-1) \quad (p+q=n)\,,\\
&\dim \mathrm{SP}(n,\mathbb{R}) = n(2n+1)\,, \quad \dim \mathrm{SP}(n,\mathbb{C}) = 2n(2n+1)
\end{aligned}$$

(véase el Ejercicio 29). Nótese también que si $X \in \mathrm{O}(p,q)$ entonces $\det X = \pm 1$, y la continuidad de la función det implica que $\mathrm{SO}(p,q)$ es un subconjunto *abierto* de $\mathrm{O}(p,q)$. Por tanto $\mathrm{SO}(p,q)$ es también un grupo de Lie, de dimensión igual a la de $\mathrm{O}(p,q)$. Análogamente, $\mathrm{SO}(n)$ es abierto en $\mathrm{O}(n)$ y, por consiguiente, un grupo de Lie de dimensión igual a la de $\mathrm{O}(n)$. De hecho, $\mathrm{SO}(n)$ es la componente conexa de la identidad en $\mathrm{O}(n)$ (véase el Ejercicio 32).

Análogamente, del último ejercicio (con $B = \mathbb{1}$) se deduce que $\mathrm{U}(n)$ es un grupo de Lie, de dimensión

$$\dim U(n) = n^2\,.$$

Mediante una ligera modificación de este ejercicio basada en la identidad (2.1) (cf. el Ejercicio 31), se demuestra que $\mathrm{SU}(n)$ es un grupo de Lie (subvariedad regular de $M_n(\mathbb{C})$) de dimensión

$$\dim \mathrm{SU}(n) = n^2 - 1\,.$$

Esta última relación se puede entender de forma heurística teniendo en cuenta que toda matriz $X \in \mathrm{U}(n)$ tiene determinante de módulo 1, por lo que la restricción adicional $\det X = 1$ simplemente fija el argumento de $\det X$, y por tanto añade una sola condición real independiente a las que definen $\mathrm{U}(n)$. Las consideraciones anteriores se extienden fácilmente a los *grupos pseudounitarios*

$$\mathrm{U}(p,q) = \{X \in M_{p+q}(\mathbb{C}) \mid X^\dagger BX = B\}\,, \qquad \mathrm{SU}(p,q) = \mathrm{U}(p,q) \cap \mathrm{SL}(p+q,\mathbb{C})\,,$$

con

$$B \equiv \begin{pmatrix} \mathbb{1}_p & 0 \\ 0 & -\mathbb{1}_q \end{pmatrix},$$

cuyo interés práctico es sin embargo mucho menor que el de los correspondientes grupos pseudoortogonales.

Consideremos, finalmente, el conjunto $\mathrm{SP}(n)$, que claramente es un grupo al ser la intersección de dos subgrupos de $M_{2n}(\mathbb{C})$. Teniendo en cuenta que $\mathrm{SP}(n) = f^{-1}(0)$, con $f : M_{2n}(\mathbb{C}) \to M_{2n}(\mathbb{C}) \times M_{2n}(\mathbb{C})$ dada por

$$f(X) = (X^{\mathsf{T}} J X - J, X^{\dagger} X - \mathbb{1}) = (X^{\mathsf{T}} J X, X^{\dagger} X) - (J, \mathbb{1}) \equiv g(X) - (J, \mathbb{1}) .$$

Procediendo de forma análoga a los dos ejercicios anteriores es fácil probar que

$$g \circ R_A = \psi_A \circ g ,$$

donde ahora $\psi_A : M_{2n}(\mathbb{C}) \times M_{2n}(\mathbb{C}) \to M_{2n}(\mathbb{C}) \times M_{2n}(\mathbb{C})$ está dada por

$$\psi_A(X, Y) = (A^{\mathsf{T}} X A, A^{\dagger} Y A) .$$

Teniendo en cuenta que $Dg = Df$, y que tanto R_A como ψ_A son aplicaciones lineales, de esta igualdad se deduce que

$$Df(XA) R_A = \psi_A Df(X) , \qquad \forall A, X \in M_{2n}(\mathbb{C}) ,$$

Como la aplicación ψ_A es invertible si A lo es ($\psi_A^{-1} = \psi_{A^{-1}}$), el rango de f es constante en $\mathrm{SP}(n)$. De la Proposición 2.45 se deduce entonces que $\mathrm{SP}(n)$ es una subvariedad regular de $M_{2n}(\mathbb{C})$. Esto demuestra, al igual que antes, que $\mathrm{SP}(n)$ es un grupo de Lie. Para calcular su dimensión, en este caso es más sencillo utilizar la identidad

$$\dim \mathrm{SP}(n) = \dim M_{2n}(\mathbb{C}) - \operatorname{rank} Df(\mathbb{1}) = \dim \ker Df(\mathbb{1}) .$$

De la definición de f se sigue fácilmente que

$$Df(\mathbb{1}) \cdot h = \left(h^{\mathsf{T}} J + J h, h^{\dagger} + h\right) ,$$

y por tanto

$$h \in \ker Df(\mathbb{1}) \quad \Longrightarrow \quad h^{\dagger} = -h .$$

Sea $h = h_1 + \mathrm{i} h_2$, con $h_{1,2} \in M_{2n}(\mathbb{R})$ y

$$h_1^{\mathsf{T}} = -h_1 , \qquad h_2^{\mathsf{T}} = h_2$$

al ser h antihermítica. Entonces

$$h^{\mathsf{T}} J + J h = (J h_1 - h_1 J) + \mathrm{i}(J h_2 + h_2 J) = 0 \quad \Longleftrightarrow \quad J h_1 - h_1 J = J h_2 + h_2 J = 0 .$$

Estas dos últimas ecuaciones se resuelven fácilmente representando h_1 y h_2 como supermatrices 2×2 con bloques en $M_n(\mathbb{R})$. Se obtiene de esta forma

$$h_1 = \begin{pmatrix} A_1 & B_1 \\ -B_1 & A_1 \end{pmatrix}, \qquad h_2 = \begin{pmatrix} A_2 & B_2 \\ B_2 & -A_2 \end{pmatrix}$$

con $A_{1,2}, B_{1,2} \in M_n(\mathbb{C})$ matrices tales que

$$A_1^{\mathsf{T}} = -A_1 , \qquad A_2^{\mathsf{T}} = A_2 , \qquad B_{1,2}^{\mathsf{T}} = B_{1,2} .$$

Por tanto

$$\dim \mathrm{SP}(n) = n^2 + n(n+1) = n(2n+1) .$$

Nota. Estrictamente hablando, los grupos matriciales clásicos son $\mathrm{SU}(n)$, $\mathrm{SO}(n)$ y $\mathrm{SP}(n)$.

Definición 2.55. Sean G y G' dos grupos de Lie. Una aplicación $f : G \to G'$ es un **homomorfismo de grupos de Lie** si es un homomorfismo algebraico, es decir si

$$f(gh) = f(g)f(h)\,, \qquad \forall g, h \in G\,,$$

y es además una aplicación *diferenciable*. Un **isomorfismo de grupos de Lie** es un homomorfismo de grupos de Lie $f : G \to G'$ que es además un *difeomorfismo*. Dos grupos de Lie G y G' son **isomorfos** (como grupos de Lie) si existe un isomorfismo de grupos de Lie $f : G \to G'$.

Nótese que dos grupos de Lie isomorfos son isomorfos también en sentido algebraico (cf. la Definición 1.24), ya que los difeomorfismos son aplicaciones *biyectivas*.

Ejemplo 2.56. Consideremos dos grupos G y G' de tipo G_1 definidos por sendas matrices B y B', es decir

$$G = \left\{X \in M_n(\mathbb{F}) \;\middle|\; X^{\mathsf{T}} B X\right\}, \qquad G' = \left\{X \in M_n(\mathbb{F}) \;\middle|\; X^{\mathsf{T}} B' X\right\},$$

y supongamos que $B' = C^{\mathsf{T}} B C$ con $C \in \mathrm{GL}(n, \mathbb{F})$. En tal caso

$$\begin{aligned} X' \in G' &\iff (X')^{\mathsf{T}} C^{\mathsf{T}} B C X' = B' = C^{\mathsf{T}} B C \iff (CX'C^{-1})^{\mathsf{T}} B (CX'C^{-1}) = B \\ &\iff CX'C^{-1} \in G\,. \end{aligned}$$

Podemos entonces definir un aplicación $\varphi : G \to G'$ dada por $\varphi(X) = C^{-1} X C$, que es claramente lineal y biyectiva ($\varphi^{-1}(X') = CX'C^{-1}$). Además, tanto φ como φ^{-1} son diferenciables, al ser restricción de aplicaciones lineales de $M_n(\mathbb{F})$ en $M_n(\mathbb{F})$. Por último, la aplicación φ es un homomorfismo algebraico, ya que

$$\varphi(XY) = C^{-1} X Y C = \big(C^{-1} X C\big)\big(C^{-1} Y C\big) = \varphi(X)\varphi(Y)\,.$$

Luego φ es un isomorfismo de grupos de Lie, y los grupos G y G' son por tanto isomorfos.

Sea $\mathbb{F} = \mathbb{R}$, y sean $B, B' \in \mathrm{GL}(n, \mathbb{R})$ dos matrices *simétricas* invertibles. Es sabido entonces que la condición necesaria y suficiente para que exista una matriz invertible $C \in \mathrm{GL}(n, \mathbb{R})$ tal que $B' = C^{\mathsf{T}} B C$ es que B y B' tengan la misma *signatura*, es decir, el mismo número de autovalores positivos y negativos. Por lo que acabamos de ver, en tal caso los correspondientes grupos G y G' son isomorfos. En consecuencia, si $\mathbb{F} = \mathbb{R}$ y B es una matriz simétrica invertible entonces el correspondiente grupo G_1 es isomorfo a $\mathrm{O}(p, q)$ para ciertos $p, q \in \mathbb{N}$. Si $\mathbb{F} = \mathbb{C}$, sin embargo, dadas dos matrices simétricas $B, B' \in \mathrm{GL}(n, \mathbb{C})$ siempre existe una matriz invertible C tal que $B' = C^{\mathsf{T}} B C$. Por tanto en el caso complejo todos los grupos de tipo G_1 con B simétrica son isomorfos a $\mathrm{O}(n, \mathbb{C})$. En particular, en este caso no tiene ningún interés definir $\mathrm{O}(p, q; \mathbb{C})$.

Ejercicio 31. Probar que $\mathrm{SU}(n)$ es un grupo de Lie de dimensión $n^2 - 1$.

Solución. Consideremos la aplicación $f : M_n(\mathbb{C}) \to M_n(\mathbb{C}) \times \mathbb{C}$ definida por

$$f(X) = \big(X^\dagger X, \det X\big) - (\mathbb{1}, 1)\,,$$

de modo que $\mathrm{SU}(n) = f^{-1}(0)$. Procediendo como en el ejemplo anterior, se demuestra que f es diferenciable y el rango de Df es constante en $\mathrm{SU}(n)$, siendo

$$Df(\mathbb{1}) \cdot h = \big(h^\dagger + h, \operatorname{tr} h\big)\,, \qquad \forall h \in M_n(\mathbb{C})\,.$$

Es fácil probar que el miembro derecho de esta ecuación es igual al subespacio

$$\left\{(A, \tfrac{1}{2} \operatorname{tr} A + \mathrm{i}b) \;\middle|\; A \in S,\ b \in \mathbb{R}\right\},$$

siendo de nuevo S el espacio de las matrices autoadjuntas de orden n. (En efecto, si $A = h + h^\dagger$ se tiene

$$\operatorname{tr} h = \frac{1}{2} \operatorname{tr} A + \frac{1}{2} \operatorname{tr}(h - h^\dagger)\,,$$

donde el último término es imaginario puro.) Por tanto

$$\operatorname{rank} Df(\mathbb{1}) = \dim S + 1 = n^2 + 1\,,$$

y de la Proposición 2.45 se deduce entonces que $\mathrm{SU}(n)$ es una subvariedad regular de $M_n(\mathbb{C})$ de dimensión

$$2n^2 - (n^2 + 1) = n^2 - 1\,.$$

Ejercicio 32. Probar que $\mathrm{SO}(n)$ es la componente conexa de la identidad en $\mathrm{O}(n)$.

Solución. Sea

$$R(\theta_1, \ldots, \theta_m) = \begin{pmatrix} R(\theta_1) & & \\ & \ddots & \\ & & R(\theta_m) \end{pmatrix},$$

si $n = 2m$, o

$$R(\theta_1, \ldots, \theta_m) = \begin{pmatrix} R(\theta_1) & & & \\ & \ddots & & \\ & & R(\theta_m) & \\ & & & 1 \end{pmatrix},$$

si $n = 2m + 1$, donde $\theta_i \in \mathbb{R}$ y

$$R(\theta) \equiv \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

Nótese que $R(\theta_1, \ldots, \theta_m) \in \mathrm{SO}(n)$ para todo $\theta_1, \ldots \theta_n \in \mathbb{R}$. Toda matriz ortogonal $X \in \mathrm{SO}(n)$ es de la forma

$$X = CR(\theta_1, \ldots, \theta_m)C^{-1}\,,$$

con C *ortogonal*. Un camino $\gamma : [0, 1] \to \mathrm{SO}(n)$ que une la identidad con la matriz X está dado entonces por

$$\gamma(t) = CR(t\theta_1, \ldots, t\theta_m)C^{-1}\,, \qquad 0 \leqslant t \leqslant 1\,.$$

(Nótese que $\gamma(t)$ es una matriz ortogonal para todo $0 \leqslant t \leqslant 1$ al serlo C y $R(t\theta_1, \ldots, t\theta_m)$.) Esto prueba que $\mathrm{SO}(n)$ es *conexo*, y por tanto ha de estar contenido en la componente conexa de la identidad en $\mathrm{O}(n)$. Por otra parte, dicha componente no puede contener ninguna matriz con determinante -1. En efecto, la imagen de la componente conexa de la identidad bajo la función $\det : \mathrm{O}(n) \to \mathbb{R}$, continua en $\mathrm{O}(n)$ en virtud de su continuidad en $M_n(\mathbb{R})$, ha de ser un conjunto conexo. Como dicha imagen está contenida en el conjunto $\{-1, 1\}$ solo puede contener uno de los dos elementos ± 1, que ha de ser obviamente el 1 dado que $\det \mathbb{1} = 1$.

Ejercicio 33. Probar que los grupos $\mathrm{U}(n)$, $\mathrm{SU}(n)$, $\mathrm{O}(n)$, $\mathrm{SO}(n)$ y $\mathrm{SP}(n)$ son *compactos*, mientras que $\mathrm{GL}(n, \mathbb{F})$, $\mathrm{SL}(n, \mathbb{F})$, $\mathrm{O}(n, \mathbb{C})$, $\mathrm{SO}(n, \mathbb{C})$, $\mathrm{O}(p, q)$, $\mathrm{SO}(p, q)$ y $\mathrm{SP}(n, \mathbb{F})$ no lo son (si $n \geqslant 2$, o $n \geqslant 1$ en el caso de $\mathrm{GL}(n, \mathbb{F})$ y $\mathrm{SP}(n, \mathbb{F})$).

Solución. Dado que todos los grupos reseñados son *subespacios topológicos* de $M_m(\mathbb{R}) \approx \mathbb{R}^{m^2}$ para un número natural m apropiado, por el teorema de Heine–Borel–Lebesgue son compactos si y solo si son cerrados y acotados en $M_m(\mathbb{R})$. En particular, $\mathrm{GL}(n, \mathbb{F})$ *no* es compacto al no ser cerrado (de hecho, se demuestra fácilmente que tampoco es acotado). En cuanto a los demás grupos matriciales considerados en el ejercicio, es inmediato comprobar que todos ellos son *cerrados*, ya que son las imágenes inversas de conjuntos cerrados adecuados bajo funciones continuas. Por tanto, dichos grupos son compactos si y solo si son subconjuntos *acotados* de $M_m(\mathbb{R})$. Nótese que $\mathrm{SU}(n)$, $\mathrm{O}(n)$ y $\mathrm{SO}(n)$ son todos ellos subgrupos de $\mathrm{U}(n)$ y $\mathrm{SP}(n)$ lo es de $\mathrm{U}(2n)$, por lo que para demostrar que los conjuntos anteriores son acotados basta probar este resultado para $\mathrm{U}(n)$ con n arbitrario. Esto último es evidente, ya que si $X \in \mathrm{U}(n)$ entonces

$$\|X\|^2 = \operatorname{tr}(X^\dagger X) = n\,.$$

Del mismo modo, como $\mathrm{SL}(n,\mathbb{R})$, $\mathrm{SP}(n,\mathbb{R})$, $\mathrm{SO}(n,\mathbb{C})$ y $\mathrm{SO}(p,q)$ son respectivamente subgrupos de $\mathrm{SL}(n,\mathbb{C})$, $\mathrm{SP}(n,\mathbb{C})$, $\mathrm{O}(n,\mathbb{C})$ y $\mathrm{O}(p,q)$, para demostrar la no compacidad de todos estos grupos basta probar que los cuatro primeros son no acotados. En el caso de $\mathrm{SL}(n,\mathbb{R})$ esto es evidente, ya que dicho grupo contiene matrices de la forma

$$X = \mathrm{diag}(\lambda, \lambda^{-1}, 1, \ldots, 1)\,, \qquad \lambda \in \mathbb{R} \setminus \{0\}\,,$$

con $\|X\|^2 = \lambda^2 + \lambda^{-2} + n - 2$ no acotado para $\lambda \to 0$ o $\lambda \to \pm\infty$. Análogamente, la matriz

$$X = \begin{pmatrix} 0 & \lambda \mathbb{1}_n \\ -\lambda^{-1}\mathbb{1}_n & 0 \end{pmatrix}, \qquad \lambda \in \mathbb{R} \setminus \{0\}\,,$$

pertenece a $\mathrm{SP}(n,\mathbb{R})$ y tiene norma

$$\|X\|^2 = n(\lambda^2 + \lambda^{-2})$$

arbitrariamente grande para $\lambda \to 0$ o $\lambda \to \pm\infty$. Por otra parte, la matriz

$$X = \left(\begin{array}{cc|c} \cosh t & -\mathrm{i}\,\mathrm{senh}\, t & \\ \mathrm{i}\,\mathrm{senh}\, t & \cosh t & 0 \\ \hline & 0 & \mathbb{1}_{n-2} \end{array}\right), \qquad t \in \mathbb{R}\,,$$

pertenece a $\mathrm{SO}(n,\mathbb{C})$ y

$$\|X\|^2 = 2(\cosh^2 t + \sinh^2 t) + n - 2 = 4\sinh^2 t + n \underset{t\to\pm\infty}{\longrightarrow} \infty\,.$$

Finalmente, la matriz

$$X = \left(\begin{array}{ccc|c} \cosh t & \cdots & \mathrm{senh}\, t & \\ \vdots & \ddots & \vdots & 0 \\ \mathrm{senh}\, t & \cdots & \cosh t & \\ \hline & 0 & & \mathbb{1}_{q-1} \end{array}\right)$$

(donde los elementos de matriz diagonales no indicados son iguales a 1, y los extradiagonales son iguales a cero) pertenece a $\mathrm{SO}(p,q)$ y tiene norma

$$\|X\|^2 = 2(\cosh^2 t + \sinh^2 t) + p + q - 2 = 4\sinh^2 t + p + q \underset{t\to\pm\infty}{\longrightarrow} \infty\,.$$

Ejercicio 34. Probar que los grupos $\mathrm{GL}(n,\mathbb{C})$, $\mathrm{SL}(n,\mathbb{F})$, $\mathrm{U}(n)$ y $\mathrm{SU}(n)$ son conexos. ¿Cuántas componentes conexas tiene $\mathrm{GL}(n,\mathbb{R})$? [*Ayuda:* en el caso de $\mathrm{SL}(n,\mathbb{R})$, se puede utilizar el hecho de que toda matriz $X \in \mathrm{SL}(n,\mathbb{R})$ es el producto de una matriz real simétrica definida positiva de determinante 1 y una matriz de $\mathrm{SO}(n)$ (*descomposición polar*).]

Solución. Nótese antes de nada que, al ser un grupo de Lie G una variedad topológica, G es conexo si y solo si G es arco-conexo (Proposición 2.38). Consideremos, en primer lugar, $\mathrm{GL}(n,\mathbb{C})$. Toda matriz $X \in \mathrm{GL}(n,\mathbb{C})$ es semejante a una matriz triangular superior, de modo que podemos escribir

$$X = C \begin{pmatrix} \lambda_1 & & \\ \vdots & \ddots & A \\ 0 & \cdots & \lambda_n \end{pmatrix} C^{-1}\,,$$

donde $\lambda_i \neq 0$ para todo i al ser $\det X \neq 0$. Sea $\gamma(t)$ la matriz obtenida sustituyendo λ_i por $\lambda_i(t)$ $(i = 1, \ldots, n)$ y A por tA, donde $\lambda_i(t)$ es cualquier camino en $\mathbb{C} \setminus \{0\}$ que une 1 con λ_i (nótese que tal camino existe, al ser $\mathbb{C} \setminus \{0\}$ conexo). Entonces γ es continua en $[0,1]$,

$$\det \gamma(t) = \lambda_1(t) \cdots \lambda_n(t) \neq 0 \quad \Longrightarrow \quad \gamma(t) \in \mathrm{GL}(n,\mathbb{C})\,,$$

$\gamma(0) = \mathbb{1}$ y $\gamma(1) = X$, por lo que γ es un camino en $\mathrm{GL}(n,\mathbb{C})$ que une $\mathbb{1}$ con X. El mismo razonamiento vale para probar la conexión de $\mathrm{SL}(n,\mathbb{C})$, teniendo en cuenta que en este caso $\lambda_n = (\lambda_1 \cdots \lambda_{n-1})^{-1}$ y tomando en consecuencia $\lambda_n(t) = \big(\lambda_1(t) \cdots \lambda_{n-1}(t)\big)^{-1}$.

En el caso de $\mathrm{SL}(n,\mathbb{R})$, si $X = PR$ es la *descomposición polar* de una matriz $X \in \mathrm{SL}(n,\mathbb{R})$ entonces, al ser P simétrica, definida positiva y de determinante 1, existe $C \in \mathrm{GL}(n,\mathbb{R})$ tal que

$$P = C \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} C^{-1} ,$$

con $\lambda_i > 0$ para todo i y $\lambda_1 \cdots \lambda_n = 1$. Si definimos

$$\gamma(t) = C \begin{pmatrix} \lambda_1(t) & & \\ & \ddots & \\ & & \lambda_n(t) \end{pmatrix} C^{-1} R(t) ,$$

donde[4]

$$\lambda_i(t) = 1 + t(\lambda_i - 1) , \quad i = 1, \ldots, n-1; \qquad \lambda_n(t) = \prod_{i=1}^{n-1} \lambda_i(t)^{-1}$$

y $t \mapsto R(t)$ es cualquier camino en $\mathrm{SO}(n)$ que une $\mathbb{1}$ con R, es fácil ver que γ es un camino en $\mathrm{SL}(n,\mathbb{R})$ que une $\mathbb{1}$ con X.

Consideremos, a continuación, el grupo $\mathrm{U}(n)$. En este caso basta tener en cuenta que toda matriz unitaria X es de la forma

$$X = U \begin{pmatrix} \mathrm{e}^{\mathrm{i}\theta_1} & & \\ & \ddots & \\ & & \mathrm{e}^{\mathrm{i}\theta_n} \end{pmatrix} U^{-1} ,$$

con U unitaria y $\theta_i \in \mathbb{R}$. Por tanto

$$\gamma(t) = U \begin{pmatrix} \mathrm{e}^{\mathrm{i}t\theta_1} & & \\ & \ddots & \\ & & \mathrm{e}^{\mathrm{i}t\theta_n} \end{pmatrix} U^{-1} , \qquad 0 \leqslant t \leqslant 1 ,$$

es claramente un camino en $\mathrm{U}(n)$ que une $\mathbb{1}$ con X. Finalmente, si $X \in \mathrm{SU}(n)$ entonces vale la demostración anterior, teniendo en cuenta que ahora podemos tomar $\theta_n = -(\theta_1 + \cdots + \theta_{n-1})$.

Veamos, por ultimo, que el grupo $\mathrm{GL}(n,\mathbb{R})$ tiene dos componentes conexas dadas por

$$\mathrm{GL}(n,\mathbb{R})_\pm = \big\{ X \in M_n(\mathbb{F}) \;\big|\; \pm \det X > 0 \big\} .$$

En efecto, es claro que $\mathrm{GL}(n,\mathbb{R})$ es la unión disjunta de $\mathrm{GL}(n,\mathbb{R})_+$ y $\mathrm{GL}(n,\mathbb{R})_-$, y que dos matrices $X_\pm \in \mathrm{GL}(n,\mathbb{R})_\pm$ no pueden unirse por ningún camino en $\mathrm{GL}(n,\mathbb{R})$ (por la continuidad de det y el teorema de los valores intermedios). Basta, por tanto, probar que ambos conjuntos $\mathrm{GL}(n,\mathbb{R})_\pm$ son conexos. Esto es, a su vez, equivalente a la conexión de $\mathrm{GL}(n,\mathbb{R})_+$, ya que obviamente $\mathrm{GL}(n,\mathbb{R})_- = L_A \mathrm{GL}(n,\mathbb{R})_+$ con A cualquier matriz de determinante negativo. Sea, por tanto, $X \in \mathrm{GL}(n,\mathbb{R})_+$, y denotemos por $\delta = (\det X)^{1/n} \in \mathbb{R}_+$ la raíz n-ésima positiva del determinante de X. Entonces

$$X = \delta A , \qquad \text{con } \; A \in \mathrm{SL}(n,\mathbb{R}) .$$

Por ser $A \in \mathrm{SL}(n,\mathbb{R})$, existe un camino $\gamma_1 : [0,1] \to \mathrm{SL}(n,\mathbb{R}) \subset \mathrm{GL}(n,\mathbb{R})$ que une $\mathbb{1}$ con A. Análogamente, al ser $\mathbb{R}_+$ conexo existe un segundo camino $\gamma_2 : [0,1] \to \mathbb{R}_+$ que une 1 con δ. Entonces la curva $\gamma = \gamma_1 \gamma_2$ es un camino en $\mathrm{GL}(n,\mathbb{R})_+$ que une $\mathbb{1}$ con X. En efecto, γ es continua al serlo γ_1 y γ_2,

$$\det \gamma(t) = \gamma_2(t)^n > 0 , \qquad \forall t \in [0,1] ,$$

[4] Nótese que si $1 \leqslant i \leqslant n-1$ entonces $\lambda_i(t) = 1 - t + t\lambda_i > 0$ para todo $t \in [0,1]$.

y claramente

$$\gamma(0) = 1 \cdot \mathbb{1} = \mathbb{1}\,, \qquad \gamma(1) = \delta A = X\,.$$

□

Nota. Los grupos $\mathrm{SP}(n,\mathbb{F})$ y $\mathrm{SP}(n)$ son también conexos. En el caso de $\mathrm{SP}(n)$, esto se sigue fácilmente del siguiente resultado: una matriz $X \in \mathrm{SP}(n)$ es de la forma

$$X = S \begin{pmatrix} \mathrm{e}^{\mathrm{i}\theta_1} & & & & & \\ & \ddots & & & & \\ & & \mathrm{e}^{\mathrm{i}\theta_n} & & & \\ & & & \mathrm{e}^{-\mathrm{i}\theta_1} & & \\ & & & & \ddots & \\ & & & & & \mathrm{e}^{-\mathrm{i}\theta_n} \end{pmatrix} S^{-1}\,,$$

con $\theta_i \in \mathbb{R}$ y $S \in \mathrm{SP}(n)$. La demostración del carácter conexo de $\mathrm{SP}(n,\mathbb{F})$, que omitiremos, es algo más complicada (puede hacerse, por ejemplo, por inducción sobre n, ya que $\mathrm{SP}(1,\mathbb{F}) = \mathrm{SL}(1,\mathbb{F})$ es conexo). Nótese, en particular, que de lo anterior se deduce que las matrices simplécticas (reales o complejas) tienen determinante 1. En efecto,

$$X^J X = J \quad \Longrightarrow \quad \det X = \pm 1\,,$$

de donde se sigue que $\det X = 1$ al ser det continua y $\mathrm{SP}(n,\mathbb{C})$ conexo.

2.3.3 Grupos matriciales cerrados

Como acabamos de ver, todos los subgrupos propios de $\mathrm{GL}(n,\mathbb{F})$ estudiados en el apartado anterior son subconjuntos *cerrados* de $M_n(\mathbb{F})$, y por tanto son también cerrados en $\mathrm{GL}(n,\mathbb{F})$. Esto motiva la siguiente definición:

Definición 2.57. Un **grupo matricial** (o **lineal**) **cerrado** es un subgrupo *cerrado* de $\mathrm{GL}(n,\mathbb{F})$.

• Como hemos comentado más arriba, si G es un subgrupo de $\mathrm{GL}(n,\mathbb{F})$ y G es cerrado en $M_n(\mathbb{F})$ entonces G es automáticamente un grupo matricial cerrado.

Ejercicio 35. Probar que un subconjunto $G \subset \mathrm{GL}(n,\mathbb{F})$ es cerrado en $\mathrm{GL}(n,\mathbb{F})$ si y solo si para cualquier sucesión $(X_k)_{k\in\mathbb{N}}$ de matrices $X_k \in G$ convergente a una matriz $X \in M_n(\mathbb{F})$ se verifica que, o bien $X \in G$, o bien $\det X = 0$.

Solución. En general, un subconjunto de un espacio topológico M es cerrado si y solo si contiene todos sus *puntos de acumulación* (por definición, $x \in M$ es punto de acumulación de $A \subset M$ si y solo si todo entorno de x contiene algún elemento de A). Si M es un *espacio métrico*, $x \in M$ es punto de acumulación de A si y solo si hay una sucesión de elementos de A que converge a x. Por tanto, un subconjunto A de un espacio métrico M es cerrado si y solo si el límite de toda sucesión convergente de elementos de A pertenece a A. En particular, $G \subset \mathrm{GL}(n,\mathbb{F})$ es cerrado en $\mathrm{GL}(n,\mathbb{F})$ si dada cualquier sucesión $\{X_k\}_{k\in\mathbb{N}}$ de matrices $X_k \in G$ convergente en $\mathrm{GL}(n,\mathbb{F})$, es decir tal que

$$\exists \lim_{k\to\infty} X_k \equiv X \in \mathrm{GL}(n,\mathbb{F})\,,$$

entonces $X \in G$. En otras palabras, G es cerrado en $\mathrm{GL}(n,\mathbb{F})$ si dada cualquier sucesión $\{X_k\}_{k\in\mathbb{N}}$, con $X_k \in G$ para todo $k \in \mathbb{N}$, o bien $\lim_{k\to\infty} X_k$ no existe, o bien existe pero no pertenece a $\mathrm{GL}(n,\mathbb{F})$ (es decir, tiene determinante nulo), o bien existe y pertenece a G.

Los grupos matriciales que acabamos de estudiar son también *subvariedades regulares* de $M_n(\mathbb{F})$, y por tanto de $\mathrm{GL}(n,\mathbb{F})$ (ya que $\mathrm{GL}(n,\mathbb{F})$ es abierto en $M_n(\mathbb{F})$). Esto motiva la siguiente definición general:

Definición 2.58. Sea G un grupo de Lie, y sea $H \subset G$. Se dice que H es un **subgrupo de Lie *regular*** de G si: 1) H es *subgrupo* de G, y 2) H es *subvariedad regular* de G.

No es difícil demostrar (esencialmente en virtud del Corolario 2.48) que si H es un subgrupo de Lie regular de un grupo de Lie G entonces H, con la *topología relativa* y la estructura diferenciable que posee al ser subvariedad regular, es también un grupo de Lie. El siguiente resultado, fundamental en la teoría de grupos de Lie y altamente no trivial, caracteriza los subgrupos de Lie *regulares* de un grupo de Lie G de manera sorprendentemente sencilla:

Teorema del subgrupo cerrado. *Sea H un subgrupo de un grupo de Lie G. Entonces H es un subgrupo de Lie regular de G si y solo si H es* cerrado *en G.*

Corolario 2.59. *Sea $G \subset \mathrm{GL}(n,\mathbb{F})$ un grupo matricial cerrado. Entonces G es un subgrupo de Lie regular de $\mathrm{GL}(n,\mathbb{F})$ y, por tanto, un* grupo de Lie.

Ejemplo 2.60. El grupo $\mathrm{GL}(n,\mathbb{R})$ es un subgrupo cerrado de $\mathrm{GL}(n,\mathbb{C})$ (aunque, evidentemente, $\mathrm{GL}(n,\mathbb{R})$ no es cerrado en $M_n(\mathbb{C})$). En efecto,

$$\mathrm{GL}(n,\mathbb{R}) = \mathrm{GL}(n,\mathbb{C}) \cap M_n(\mathbb{R}) \,,$$

siendo $M_n(\mathbb{R})$ cerrado en $M_n(\mathbb{C})$. Para convencerse de esto último, basta notar que si $\{X_n\}_{n\in\mathbb{N}}$ es una sucesión convergente de matrices reales su límite es también una matriz real. En particular, todo subgrupo cerrado de $\mathrm{GL}(n,\mathbb{R})$ es también un subgrupo cerrado de $\mathrm{GL}(n,\mathbb{C})$. Por tanto *un subconjunto $G \subset \mathrm{GL}(n,\mathbb{F})$ es un grupo matricial cerrado si y solo si G es un subgrupo cerrado de* $\mathrm{GL}(n,\mathbb{C})$.

Ejercicio 36. Considérese el grupo matricial

$$G = \left\{ \begin{pmatrix} \mathrm{e}^{\mathrm{i}t} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\alpha t} \end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\} \,,$$

donde $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ es un número *irracional*. Probar que G es un grupo de Lie (isomorfo al grupo aditivo $\mathbb{R}$), pero *no* es cerrado en $\mathrm{GL}(2,\mathbb{C})$. Por tanto *$G$ no es un grupo matricial cerrado, ni un subgrupo de Lie regular de* $\mathrm{GL}(2,\mathbb{C})$. [*Ayuda:* puede probarse que, al ser α irracional, para todo $\lambda \in \mathbb{R}$ existe una sucesión $(n_k)_{k\in\mathbb{N}}$ de enteros positivos tal que $\lim_{k\to\infty} \mathrm{e}^{2\pi\mathrm{i}n_k\alpha} = \mathrm{e}^{\mathrm{i}\lambda}$.]

Solución. Es fácil ver que la aplicación $g : \mathbb{R} \to G$ dada por

$$g(t) = \begin{pmatrix} \mathrm{e}^{\mathrm{i}t} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\alpha t} \end{pmatrix}$$

es biyectiva, al ser α irracional. Podemos dotar a G de una topología estableciendo que un conjunto $A \subset G$ es abierto si es la imagen de un abierto de $\mathbb{R}$ bajo g. Es evidente que g es un homeomorfismo respecto de la topología así definida. En particular, al ser G homeomorfo a $\mathbb{R}$, G es una variedad topológica. También es fácil dotar a G de una estructura diferenciable de una sola carta $\left(G, g^{-1}\right)$, en la cual la coordenada de una matriz $g(t) \in G$ es el número real $t \in \mathbb{R}$. Por tanto G es una variedad diferenciable. Como

$$g(t+s) = g(t)g(s) \,, \qquad g(t)^{-1} = g(-t) \,,$$

en la carta que acabamos de definir el producto y la inversa están representados por las aplicaciones diferenciables

$$(t,s) \mapsto t+s \,, \qquad t \mapsto -t \,.$$

Por tanto G es grupo de Lie, claramente isomorfo al grupo aditivo de los reales bajo la aplicación g.

Veamos a continuación que G no es cerrado en $\mathrm{GL}(2,\mathbb{C})$. Como α es irracional, para todo $\lambda \in \mathbb{R}$ existe una sucesión de enteros positivos n_k tal que

$$\lim_{k\to\infty} \mathrm{e}^{2\pi\mathrm{i}n_k\alpha} = \mathrm{e}^{\mathrm{i}\lambda} \,,$$

y por tanto

$$\lim_{k\to\infty} g(2\pi n_k) = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\lambda} \end{pmatrix} \in \mathrm{GL}(2,\mathbb{C})\,.$$

Sin embargo, es fácil ver que la matriz del miembro derecho solo pertenece a G si y solo si

$$\lambda = 2\pi(n\alpha + m)\,, \qquad \text{con } n, m \in \mathbb{Z}\,;$$

en particular,

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \notin G\,.$$

Por tanto G *no* es cerrado en $\mathrm{GL}(2,\mathbb{C})$, como habíamos afirmado.

Ejemplo 2.61. Consideremos el grupo E_n de los movimientos de $\mathbb{R}^n$, cuyos elementos son las aplicaciones afines $f : \mathbb{R}^n \to \mathbb{R}^n$ de la forma

$$f(x) = Rx + a\,, \qquad R \in \mathrm{O}(n)\,,\ \ a \in \mathbb{R}^n\,.$$

Este grupo *no* es un grupo matricial. Sin embargo, si identificamos $\mathbb{R}^n$ con el hiperplano afín de $\mathbb{R}^{n+1}$ de ecuación $x_{n+1} = 1$ entonces podemos escribir

$$f \cdot \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} Rx + a \\ 1 \end{pmatrix} \equiv X_f \begin{pmatrix} x \\ 1 \end{pmatrix},$$

donde $X_f \in M_{n+1}(\mathbb{R})$ es la matriz dada por

$$X_f = \left(\begin{array}{ccc|c} & & & a_1 \\ & R & & \vdots \\ & & & a_n \\ \hline 0 & \cdots & 0 & 1 \end{array}\right) \equiv \begin{pmatrix} R & a \\ 0 & 1 \end{pmatrix}.$$

De esto se sigue inmediatamente que $X_{fg} = X_f X_g$, y por tanto E_n es isomorfo al subgrupo $\tilde{E}_n$ de $M_{n+1}(\mathbb{R})$ definido por

$$\tilde{E}_n = \left\{ \begin{pmatrix} R & a \\ 0 & 1 \end{pmatrix} \,\middle|\, R \in \mathrm{O}(N)\,,\ a \in \mathbb{R}^n \right\}\,.$$

Este grupo es claramente un *grupo matricial cerrado*, ya que $\tilde{E}_n$ es de hecho cerrado en $M_{n+1}(\mathbb{R})$. En efecto, si la sucesión

$$X_k \equiv \begin{pmatrix} R_k & a_k \\ 0 & 1 \end{pmatrix}, \qquad R_k \in \mathrm{O}(n)\,,\ \ a_k \in \mathbb{R}^n\,,$$

converge a una matriz $X \in M_{n+1}(\mathbb{R})$ entonces

$$X = \begin{pmatrix} R & a \\ 0 & 1 \end{pmatrix}, \qquad \text{con} \quad R = \lim_{n\to\infty} R_k\,,\ \ a = \lim_{n\to\infty} a_k\,.$$

Como el límite de una sucesión de matrices ortogonales es una matriz ortogonal —ya que $\mathrm{O}(n)$ es cerrado—, $X \in \tilde{E}_n$, lo que demuestra que $\tilde{E}_n$ es cerrado. Por tanto $\tilde{E}_n$ es un *grupo de Lie* (subvariedad regular de $\mathrm{GL}(n,\mathbb{R})$). Además, la aplicación $\varphi : \tilde{E}_n \to \mathrm{O}(n) \times \mathbb{R}^n$ definida por

$$\varphi\big(\left(\begin{smallmatrix} R & a \\ 0 & 1 \end{smallmatrix}\right)\big) = (R, a)$$

es un homeomorfismo, ya que es biyectiva y tanto φ como φ^{-1} son claramente continuas[5]. Por tanto $\tilde{E}_n \approx \mathrm{O}(n) \times \mathbb{R}^n$ como variedades topológicas. En particular, la dimensión de $\tilde{E}_n$ está dada por

$$\dim \tilde{E}_n = \dim \mathrm{O}(n) + n = \frac{1}{2}\, n(n-1) + n = \frac{1}{2}\, n(n+1)\,.$$

A partir de ahora, identificaremos habitualmente los grupos E_n y $\tilde{E}_n$.

[5]Recuérdese que si M_1 y M_2 son dos *espacios métricos* una aplicación $f : M_1 \to M_2$ es *continua* si y solo si para toda sucesión $\{x_k\}_{k\in\mathbb{N}} \subset M_1$ que converge a un punto $x \in M_1$ la sucesión $\{f(x_k)\}_{k\in\mathbb{N}}$ converge a $f(x)$ en M_2.

2.4 El álgebra de Lie de un grupo matricial cerrado

2.4.1 Espacio tangente a una subvariedad regular

Si M es una variedad diferenciable (C^∞), una **curva suave** en M es una aplicación diferenciable (C^∞) $\gamma : I \to M$, donde I es un intervalo abierto de la recta real. En particular, si M es una *subvariedad regular* de $\mathbb{R}^n$ toda curva suave en M es una curva suave en $\mathbb{R}^n$. En efecto, si $t_0 \in I$, $a = \gamma(t_0)$ y (φ, U) es una carta de $\mathbb{R}^n$ adaptada a M en a entonces en un entorno abierto suficientemente pequeño de t_0 se tiene

$$(\varphi \circ \gamma)(t) = \big((\tilde{\varphi} \circ \gamma)(t), 0 \ldots, 0\big)\,,$$

con $\tilde{\varphi} \circ \gamma$ diferenciable. Esto implica que $\varphi \circ \gamma$ es diferenciable en un entorno de t_0, de donde se sigue (al ser φ un difeomorfismo) que $\gamma : I \to \mathbb{R}^n$ es diferenciable en un entorno de t_0.

Sea $M \subset \mathbb{R}^n$ una subvariedad regular, sea $a \in M$, y sea $\gamma : I \to M$ una curva suave en M tal que $\varphi(t_0) = a$. El **vector tangente** a γ en a es el vector

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t}(t_0) \equiv \gamma'(t_0) \in \mathbb{R}^n\,,$$

que por el comentario anterior está bien definido. Diremos que un vector $v \in \mathbb{R}^n$ es **tangente** a M en a si v es el vector tangente en el punto a a una curva suave γ contenida en M.

Definición 2.62. Sea M una subvariedad regular de $\mathbb{R}^n$, y sea $a \in M$. El **espacio tangente** a M en a es el conjunto $T_a M$ de todos los vectores tangentes a M en a.

Proposición 2.63. *El espacio tangente a una subvariedad regular $M \subset \mathbb{R}^n$ en un punto $a \in M$ es un espacio vectorial real de dimensión* $\dim M$.

Demostración. Sea $v = \gamma'(t_0)$ el vector tangente en $a = \gamma(t_0)$ a una curva suave γ contenida en M. Utilizando la notación del comentario al principio de esta sección podemos escribir

$$\varphi\big(\gamma(t)\big) = (\hat{\gamma}(t), 0, \ldots, 0)\,,$$

con $\hat{\gamma} \equiv \tilde{\varphi} \circ \gamma$ diferenciable en t_0. Derivando respecto de t y haciendo $t = t_0$ se obtiene

$$D\varphi(a) \cdot v = \big(\hat{\gamma}'(t_0), 0, \ldots, 0\big) \quad \Longrightarrow \quad v = D\varphi(a)^{-1}\big(\hat{\gamma}'(t_0), 0, \ldots, 0\big)\,,$$

donde la invertibilidad de $D\varphi(a)$ está garantizada al ser φ un difeomorfismo. Por tanto

$$T_a M \subset D\varphi(a)^{-1} \cdot \big(\mathbb{R}^m \times \{0, \ldots, 0\}\big)\,, \qquad m \equiv \dim M\,.$$

Recíprocamente, si $v = D\varphi(a)^{-1}(w, 0, \ldots, 0)$, con $w \in \mathbb{R}^m$, es inmediato comprobar que v es tangente a la curva suave en M

$$t \mapsto \varphi^{-1}(tw, 0, \ldots, 0)$$

en el punto $a = \varphi(0)$. Por tanto

$$T_a M = D\varphi(a)^{-1} \cdot \big(\mathbb{R}^m \times \{0, \ldots, 0\}\big)$$

es un espacio vectorial de dimensión $m = \dim M$, como habíamos afirmado. □

Ejemplo 2.64. Sea $f : \mathbb{R}^p \to \mathbb{R}^q$ una aplicación diferenciable tal que Df tiene rango constante k en $M = f^{-1}(0)$. Por la Proposición 2.45, el conjunto M es una subvariedad regular de $\mathbb{R}^p$ de dimensión $m = p - k$. Si $\gamma : I \to M$ es una curva diferenciable y $\gamma(t_0) = a \in M$ entonces

$$f\big(\gamma(t)\big) = 0\,, \qquad \forall t \in I\,.$$

Derivando esta identidad en $t = t_0$ obtenemos

$$Df(a) \cdot \gamma'(t_0) = 0\,,$$

y por tanto

$$T_a M \subset \ker Df(a)\,.$$

Pero, al ser

$$\dim \ker Df(a) = p - \operatorname{rank} Df(a) = p - k = \dim M = \dim T_a M\,,$$

el contenido anterior es de hecho una igualdad. En otras palabras, el espacio tangente a $M = f^{-1}(0)$ en un punto cualquiera $a \in M$ está dado por

$$T_a M = \ker Df(a)\,. \tag{2.4}$$

□

Si $G \subset \mathrm{GL}(n,\mathbb{F})$ es un grupo matricial cerrado, por el Corolario 2.59 G es una subvariedad regular de $M_n(\mathbb{F}) \approx \mathbb{R}^{d_{\mathbb{F}} n^2}$ (siendo $d_{\mathbb{F}} \equiv \dim_{\mathbb{R}} \mathbb{F}$). Por tanto, su espacio tangente en cada punto $g \in G$ está bien definido y es un espacio vectorial *real* de dimensión igual a la del propio G. ($\mathbb{F} = \mathbb{C}$). En particular:

Definición 2.65. Si $G \subset \mathrm{GL}(n,\mathbb{F})$ es un grupo matricial cerrado, denotaremos por $\mathrm{Lie}(G)$ el espacio tangente a G en la identidad.

Nótese, en particular, que $\dim \mathrm{Lie}(G) = \dim G$.

Notación. Si $G \subset \mathrm{GL}(n,\mathbb{F})$ es un grupo matricial cerrado y $X \in \mathrm{Lie}(G) \equiv T_e G$, denotaremos por $g_X : I_X \to G$ (con I_X un intervalo abierto) cualquier curva suave en G cuyo vector tangente en la identidad es X. Supondremos siempre, por sencillez (y sin pérdida de generalidad), que la curva está parametrizada de forma que $g_X(0) = 0$, y por tanto $g'_X(0) = X$.

Lema 2.66. *Sea $A : I \to \mathrm{GL}(n,\mathbb{F})$ una curva suave. Entonces $t \mapsto A(t)^{-1}$ es una curva suave, con derivada*

$$\frac{\mathrm{d}}{\mathrm{d}t} A(t)^{-1} = -A(t)^{-1} A'(t) A(t)^{-1}\,.$$

Demostración. Que $A(t)^{-1}$ es diferenciable (C^∞) en I es consecuencia inmediata de que sus elementos de matriz son funciones racionales de los de $A(t)$ con denominador no nulo. La fórmula para su derivada se obtiene fácilmente derivando la identidad

$$A(t) A(t)^{-1} = \mathbb{1}\,, \qquad \forall t \in I\,.$$

□

Proposición 2.67. *Si $G \subset \mathrm{GL}(n,\mathbb{F})$ es un grupo matricial cerrado y $X \in \mathrm{Lie}(G)$, para todo $A \in G$ la matriz AXA^{-1} pertenece a* $\mathrm{Lie}(G)$. *En otras palabras,* $\mathrm{Lie}(G)$ *es invariante bajo conjugación por elementos de G.*

Demostración. La matriz AXA^{-1} es el vector tangente en e a la curva suave $\gamma : I_X \to G$ definida por

$$\gamma(t) = A g_X(t) A^{-1}\,.$$

□

Corolario 2.68. *Sea $G \subset \mathrm{GL}(n,\mathbb{F})$ un grupo matricial cerrado, y sean $X, Y \in \mathrm{Lie}(G)$. Entonces el conmutador*

$$[X,Y] \equiv XY - YX$$

pertenece a $\mathrm{Lie}(G)$.

Demostración. Dados dos elementos $X, Y \in \mathrm{Lie}(G)$, la aplicación

$$t \mapsto \gamma(t) = g_X(t) Y g_X(t)^{-1}\,, \qquad t \in I_X\,,$$

es diferenciable en virtud del Lema 2.66. Por la proposición anterior, la curva γ está contenida en el *subespacio vectorial* $\mathrm{Lie}(G) \subset M_n(\mathbb{F})$, y por tanto su vector tangente en cualquier punto pertenece a $\mathrm{Lie}(G)$. En particular, evaluando dicho vector tangente en $t = 0$ mediante el Lema 2.66 obtenemos fácilmente

$$\gamma'(0) = g_X'(0) Y - Y g_X'(0) = XY - YX \in \mathrm{Lie}(G)\,.$$

□

2.4.2 Álgebras de Lie

Definición 2.69. Un **álgebra de Lie** sobre un cuerpo $\mathbb{F}$ es un álgebra $\mathfrak{g}$ sobre $\mathbb{F}$ cuyo producto, que denotaremos por $[\,\cdot\,,\cdot\,]$, verifica las siguientes propiedades:

1. *Anticonmutatividad:* $[x, y] = -[y, x]\,, \qquad \forall x, y \in \mathfrak{g}$
2. *Identidad de Jacobi:*

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0\,, \qquad \forall x, y, z \in \mathfrak{g}\,.$$

La **dimensión** de $\mathfrak{g}$ como álgebra de Lie es su dimensión como espacio vectorial sobre $\mathbb{F}$. Un álgebra de Lie $\mathfrak{g}$ es **conmutativa** (o abeliana) si $[x, y] = 0$ para todo $x, y \in \mathfrak{g}$.

Comentarios.

- Por definición de álgebra, el producto $[\,\cdot\,,\cdot\,]$, que se denomina habitualmente **corchete de Lie**, es lineal en cada uno de sus argumentos.
- La anticonmutatividad implica que

$$[x, x] = 0\,, \qquad \forall x \in \mathfrak{g}\,.$$

  Esta última propiedad es de hecho *equivalente* a la anticonmutatividad, ya que

$$[x + y, x + y] = 0 = [x, y] + [y, x]\,.$$

- La identidad de Jacobi prescribe como difieren los elementos $[[x, y], z]$ y $[x, [y, z]]$, que en un álgebra asociativa serían idénticos:

$$[[x, y], z] - [x, [y, z]] = [[x, z], y]\,.$$

Definición 2.70. Si $\mathfrak{g}_1$ y $\mathfrak{g}_2$ son dos álgebras de Lie, una aplicación $f : \mathfrak{g}_1 \to \mathfrak{g}_2$ es un **homomorfismo** si es *lineal* y satisface

$$f([x, y]) = [f(x), f(y)]\,, \qquad \forall x, y \in \mathfrak{g}_1\,.$$

Un **isomorfismo** de álgebras de Lie es un homomorfismo biyectivo. Las álgebras de Lie $\mathfrak{g}_1$ y $\mathfrak{g}_2$ son **isomorfas** si existe un isomorfismo de álgebras de Lie $f : \mathfrak{g}_1 \to \mathfrak{g}_2$.

Nótese que dos álgebras de Lie isomorfas son también isomorfas como espacios vectoriales, ya que los isomorfismos de álgebras de Lie son isomorfismos lineales.

Si A es un álgebra **asociativa** sobre un cuerpo $\mathbb{F}$, y definimos el corchete de Lie mediante

$$[a, b] = ab - ba\,, \qquad \forall a, b \in A\,,$$

entonces $\big(A, [\,\cdot\,,\cdot\,]\big) \equiv \mathfrak{g}_A$ es un álgebra de Lie sobre $\mathbb{F}$ denominada **álgebra de los conmutadores** de A. En efecto, es claro que $[\,\cdot\,,\cdot\,]$ es lineal en cada componente, y por tanto $\mathfrak{g}_A$ es un álgebra. Además, el producto es claramente anticonmutativo. Por último, la identidad de Jacobi se sigue del siguiente cálculo elemental:

$$[a,[b,c]] + \text{perm. cicl.} = (abc - bca) + (cba - acb) + \text{perm. cicl.} = 0\,.$$

Evidentemente, $\mathfrak{g}_A$ es conmutativa si y solo si lo es A. En particular, $M_n(\mathbb{F})$ con el conmutador de matrices como corchete de Lie es un álgebra de Lie, real o compleja según sea $\mathbb{F} = \mathbb{R}$ o $\mathbb{F} = \mathbb{C}$. En este último caso ($\mathbb{F} = \mathbb{C}$), podemos considerar a $M_n(\mathbb{C})$ como un álgebra de Lie *real* de dimensión $2n^2$.

Definición 2.71. Si $\mathfrak{g}$ es un álgebra de Lie sobre un cuerpo $\mathbb{F}$, una **subálgebra** de $\mathfrak{g}$ es un subespacio $\mathfrak{h} \subset \mathfrak{g}$ que es a su vez un álgebra de Lie con el corchete heredado de $\mathfrak{g}$.

Es evidente que un subespacio $\mathfrak{h} \subset \mathfrak{g}$ es una subálgebra de Lie si y solo si

$$x, y \in \mathfrak{h} \quad \Longrightarrow \quad [x,y] \in \mathfrak{h}\,.$$

En virtud del Corolario 2.68, si G es un grupo matricial cerrado su espacio tangente en la identidad $\mathrm{Lie}(G)$ es una subálgebra de Lie de $M_n(\mathbb{F})$, considerada como álgebra de Lie *real*. En particular:

Proposición 2.72. *Si G es un grupo matricial cerrado, el conjunto* $\mathrm{Lie}(G)$ *con el conmutador de matrices como corchete de Lie es un álgebra de Lie* real *denominada* ***álgebra de Lie*** *del grupo G.*

Nota. Si G es un grupo de Lie abstracto, es posible también definir (de forma algo distinta, pero equivalente, a la anterior) un álgebra de Lie real $\mathrm{Lie}(G)$ asociada a G, de la misma dimensión que G.

2.5 Subgrupos a un parámetro

La herramienta fundamental para estudiar el álgebra de Lie de un grupo de Lie es el concepto de subgrupo a un parámetro, que definiremos a continuación.

Definición 2.73. Un **subgrupo a un parámetro** de un grupo de Lie G es un homomorfismo del grupo aditivo de los números reales en G.

En otras palabras, un subgrupo a un parámetro de un grupo de Lie G es una aplicación diferenciable $g : \mathbb{R} \to G$ tal que

$$g(t)g(s) = g(t+s)\,, \qquad \forall s,t \in \mathbb{R}\,.$$

Nótese que si $f : G_1 \to G_2$ es un homomorfismo entonces

$$f(e) = e\,, \qquad f(g^{-1}) = f(g)^{-1}$$

(cf. la Proposición 1.25). En particular, si $g : \mathbb{R} \to G$ es un subgrupo a un parámetro se verifica

$$g(0) = \mathbb{1}\,, \qquad g(t)^{-1} = g(-t)\,.$$

Obsérvese también que el conjunto $g(\mathbb{R})$, al ser la imagen de un grupo ($\mathbb{R}$) bajo un homomorfismo, es efectivamente un subgrupo de G (Proposición 1.31). (De ahí la denominación de *subgrupo a un parámetro*.)

Nota. Se puede probar que un homomorfismo *continuo* $g : \mathbb{R} \to G$ (donde G es un grupo de Lie) es automáticamente *diferenciable*.

Estudiaremos a continuación la estructura de los subgrupos a un parámetro de un *grupo matricial cerrado* G. Para ello será fundamental la *función exponencial matricial* $\exp : M_n(\mathbb{C}) \to M_n(\mathbb{C})$, definida por

$$\exp(A) \equiv \mathrm{e}^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}\,, \qquad \forall A \in M_n(\mathbb{C})\,,$$

donde $A^0 \equiv \mathbb{1}$. Se demuestra (basta aplicar el criterio M de Weierstrass) que la serie anterior converge absolutamente para toda matriz $A \in M_n(\mathbb{C})$. La exponencial matricial tiene las siguientes propiedades, que enunciaremos sin demostración:

1. $\exp : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ es una aplicación C^∞ (donde $M_n(\mathbb{C})$ se identifica con $\mathbb{R}^{2n^2}$).
2. $\dfrac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{e}^{tA} = A\mathrm{e}^{tA} = \mathrm{e}^{tA}A\,, \quad \forall t \in \mathbb{R}\,,\ \ A \in M_n(\mathbb{C})$.
3. Para todo $t, s \in \mathbb{R}$ y para toda matriz $A \in M_n(\mathbb{C})$ se verifica
$$\mathrm{e}^{tA}\mathrm{e}^{sA} = \mathrm{e}^{(t+s)A}\,.$$
En particular, e^A es invertible para toda matriz $A \in M_n(\mathbb{C})$, siendo
$$\left(\mathrm{e}^A\right)^{-1} = \mathrm{e}^{-A}\,.$$
4. $[A, B] = 0 \implies \left[\mathrm{e}^A, \mathrm{e}^B\right] = 0\,, \quad \forall A, B \in M_n(\mathbb{C})\,.$
5. $\det(\mathrm{e}^A) = \mathrm{e}^{\mathrm{tr}\,A}\,, \quad \forall A \in M_n(\mathbb{C})\,.$
6. Para todo $A \in M_n(\mathbb{C})$ y $C \in \mathrm{GL}(n, \mathbb{C})$ se verifica
$$(\mathrm{e}^A)^{\mathsf{T}} = \mathrm{e}^{A^{\mathsf{T}}}\,, \qquad (\mathrm{e}^A)^{\dagger} = \mathrm{e}^{A^{\dagger}}\,, \qquad C\mathrm{e}^A C^{-1} = \mathrm{e}^{CAC^{-1}}\,.$$

Ejercicio 37. Utilizando la desigualdad de Cauchy–Schwarz, probar que

$$\|AB\| \leqslant \|A\|\|B\|\,, \qquad \forall A, B \in M_n(\mathbb{C})\,.$$

Deducir que $\|A^n\| \leqslant \|A\|^n$, para todo $n \in \mathbb{N}$, y por tanto

$$\left\|\mathrm{e}^A\right\| \leqslant \mathrm{e}^{\|A\|}\,, \qquad \forall A \in M_n(\mathbb{C})\,.$$

Ejercicio 38. Probar que $D\exp(0)$ es la identidad. Deducir que hay sendos entornos abiertos U de 0 y V de $\mathbb{1}$ en $M_n(\mathbb{C})$ tales que $\exp : U \to V$ es biyectiva, con inversa $\exp^{-1} \equiv \log : V \to U$ derivable en V, siendo $D\log(\mathbb{1})$ la identidad. *Ayuda:* si $f : \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $a \in \mathbb{R}^n$ y $h \in \mathbb{R}^n$,

$$Df(a) \cdot h = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f(a + th)\,.$$

Solución. Si $h \in M_n(\mathbb{C})$, utilizando la fórmula de la ayuda se obtiene

$$D\exp(0) \cdot h = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{e}^{th} = h\mathrm{e}^{th}\Big|_{t=0} = h\,,$$

y por tanto $D\exp(0)$ es la identidad $I : M_n(\mathbb{C}) \to M_n(\mathbb{C})$. Por el teorema de la función inversa, existen sendos entornos abiertos U de 0 y V de $\mathrm{e}^0 = \mathbb{1}$ en $M_n(\mathbb{C})$ tales que $\exp : U \to V$ es biyectiva, con inversa $\exp^{-1} \equiv \log : V \to U$ derivable en V. Además, la derivada de log en una matriz $A \in V$ está dada por

$$D\log(A) = [(D\exp)(\log A)]^{-1} \quad \implies \quad D\log(\mathbb{1}) = [D\exp(0)]^{-1} = I\,.$$

Nota. Puede probarse que la componente conexa de la unidad en un grupo de Lie G está generada por un entorno cualquiera de la unidad. Como $\mathrm{GL}(n,\mathbb{C})$ es conexo, aplicando este resultado al entorno V del ejercicio anterior se deduce inmediatamente que toda matriz $X \in \mathrm{GL}(n,\mathbb{C})$ se puede expresar en la forma

$$X = \mathrm{e}^{A_1} \cdots \mathrm{e}^{A_m} ,$$

con $A_i \in M_n(\mathbb{C})$ para $i = 1, \ldots, m$ (por supuesto, m depende de la matriz X).

Proposición 2.74. *Sea $g : \mathbb{R} \to G$ un subgrupo a un parámetro de un grupo matricial cerrado G. Entonces*

$$g(t) = \mathrm{e}^{tA} , \qquad \textit{con} \quad A = g'(0) \in \mathrm{Lie}(G) .$$

Demostración. Al ser g subgrupo a un parámetro, para todo $s, t \in \mathbb{R}$ se tiene

$$g(t+s) = g(s)g(t) .$$

Derivando esta relación respecto de s y haciendo $s = 0$ se obtiene

$$g'(t) = Ag(t) , \qquad \text{con} \;\; A \equiv g'(0) .$$

La ecuación anterior es un sistema lineal de primer orden en la matriz $g(t)$ con coeficientes constantes. Por tanto

$$g(t) = \mathrm{e}^{tA} g(0) = \mathrm{e}^{tA} ,$$

ya que $g(0) = \mathbb{1}$. Nótese, para finalizar, que $A = g'(0)$ pertenece a $\mathrm{Lie}(G)$, al ser el vector tangente a la curva suave $g : \mathbb{R} \to G$ en la identidad. □

Una consecuencia inmediata de la segunda propiedad de la exponencial matricial es que si $\mathrm{e}^{tA} \in G$ para todo $t \in \mathbb{R}$ entonces $A \in \mathrm{Lie}(G)$; en otras palabras,

$$\left\{ A \in M_n(\mathbb{F}) \;\middle|\; \mathrm{e}^{tA} \in G , \; \forall t \in \mathbb{R} \right\} \subset \mathrm{Lie}(G) . \tag{2.5}$$

De hecho, el contenido anterior es una *igualdad*. Para establecer este importante resultado, necesitaremos probar dos lemas previos.

Lema 2.75. *Sea $A : \mathbb{N} \to M_n(\mathbb{C})$ tal que $A(k) = O(k^{-2})$ para $k \to \infty$. Entonces*

$$\lim_{k\to\infty} \big(\mathbb{1} + A(k)\big)^k = \mathbb{1} .$$

Demostración. Utilizando la fórmula del binomio de Newton se obtiene

$$(\mathbb{1} + A)^k - \mathbb{1} = \sum_{l=1}^{k} k(k-1)\cdots(k-l+1)\frac{A^l}{l!} \quad \Longrightarrow \quad \big\|(\mathbb{1}+A)^k - \mathbb{1}\big\| \leqslant \sum_{l=1}^{k} \frac{(k\|A\|)^l}{l!} ,$$

donde hemos tenido en cuenta que $\|A^l\| \leqslant \|A\|^l$. Al ser $A(k) = O(k^{-2})$, existe $M > 0$ tal que

$$k\|A\| \leqslant \frac{M}{k}$$

para k suficientemente grande, y por tanto

$$\big\|(\mathbb{1}+A)^k - \mathbb{1}\big\| \leqslant \sum_{l=1}^{k} \frac{(M/k)^l}{l!} \leqslant \mathrm{e}^{M/k} - 1 \underset{k\to\infty}{\longrightarrow} 0 .$$

□

Lema 2.76. *Sea $g : \mathbb{R} \to M_n(\mathbb{C})$ una curva suave tal que $g(0) = \mathbb{1}$. Entonces para todo $t \in \mathbb{R}$ se tiene*

$$\lim_{\substack{k\to\infty \\ k\in\mathbb{N}}} g(t/k)^k = \mathrm{e}^{tA}\,, \qquad A \equiv g'(0)\,.$$

Demostración. Al ser $A = g'(0)$ se tiene

$$g(s) = \mathrm{e}^{sA} + O(s^2)\,,$$

y por tanto para todo $t \in \mathbb{R}$ *fijo* y $k \in \mathbb{N}$ se verifica

$$g(t/k)^k = \left[\mathrm{e}^{tA/k} + O(k^{-2})\right]^k = \mathrm{e}^{tA}\left[\mathbb{1} + \mathrm{e}^{-tA/k}\,O(k^{-2})\right]^k. \tag{2.6}$$

Como

$$\lim_{k\to\infty} \mathrm{e}^{-tA/k} = \mathbb{1}$$

por la continuidad de la función exponencial, $\left\|\mathrm{e}^{-tA/k}\right\|$ está acotada para $k \to \infty$. Por tanto

$$\mathrm{e}^{-tA/k}\,O(k^{-2}) = O(k^{-2})\,,$$

y el resultado anunciado se sigue de (2.6) y del lema anterior. □

Corolario 2.77. *Si $G \subset M_n(\mathbb{F})$ es un grupo matricial cerrado entonces*

$$\mathrm{Lie}(G) = \left\{A \in M_n(\mathbb{F}) \,\middle|\, \mathrm{e}^{tA} \in G\,,\ \forall t \in \mathbb{R}\right\}. \tag{2.7}$$

Demostración. En virtud de (2.5), basta probar que

$$\mathrm{Lie}(G) \subset \left\{A \in M_n(\mathbb{F}) \,\middle|\, \mathrm{e}^{tA} \in G\,,\ \forall t \in \mathbb{R}\right\}.$$

Sea, por tanto, $A \in \mathrm{Lie}(G)$, y consideremos una curva suave $g_A : I_A \to G$ tal que $g_A(0) = \mathbb{1}$ y $g_A'(0) = A$. Como $g_A(t/k)^k \in G$ para todo $t \in I_A$ y para todo $k \in \mathbb{N}$, por el lema anterior

$$\lim_{k\to\infty} g_A(t/k)^k = \mathrm{e}^{tA} \in \mathrm{GL}(n,\mathbb{F})\,.$$

Como G es *cerrado* en $\mathrm{GL}(n,\mathbb{F})$, e^{tA} ha de pertenecer a G para todo $t \in I_A$. Por otra parte, al ser I_A abierto y $0 \in I_A$ existe $\varepsilon > 0$ tal que $(-\varepsilon,\varepsilon) \subset I_A$. Si t es cualquier número real, hay un número natural N tal que $t/N \in (-\varepsilon,\varepsilon)$, y por tanto

$$\frac{t}{N} \in (-\varepsilon,\varepsilon) \subset I_A \quad\Longrightarrow\quad \mathrm{e}^{tA/N} \in G \quad\Longrightarrow\quad \mathrm{e}^{tA} = (\mathrm{e}^{tA/N})^N \in G\,.$$

□

• Es importante notar que e^A puede pertenecer a un grupo matricial cerrado G aun cuando A no pertenezca a $\mathrm{Lie}(G)$. Así, por ejemplo, si $k \in \mathbb{Z} \setminus \{0\}$ la matriz $A = 2\pi\mathrm{i}k\mathbb{1}$ satisface $\mathrm{e}^A = \mathbb{1} \in G$ para cualquier grupo $G \subset \mathrm{GL}(n,\mathbb{C})$. Sin embargo, es fácil ver que A no pertenece al álgebra de Lie de los grupos $\mathrm{SL}(n,\mathbb{C})$, $\mathrm{O}(n,\mathbb{C})$, $\mathrm{SU}(n)$, $\mathrm{SP}(n,\mathbb{F})$ o $\mathrm{SP}(n)$ (cf. la sección siguiente).

Del Corolario (2.77) y de la tercera propiedad de la exponencial matricial se deduce que si $A \in \mathrm{Lie}(G)$ entonces la aplicación $t \mapsto \mathrm{e}^{tA}$ es un subgrupo a un parámetro del grupo G. Combinando esta observación con la Proposición 2.74 se obtiene el siguiente resultado:

Proposición 2.78. *Una aplicación $g : \mathbb{R} \to G$ es un subgrupo a un parámetro de un grupo matricial cerrado G si y solo si $g(t) = \mathrm{e}^{tA}$, con $A \in \mathrm{Lie}(G)$.*

2.5.1 Álgebras de Lie de los grupos matriciales clásicos

El Corolario anterior permite calcular de forma muy sencilla el álgebra de Lie de los grupos matriciales clásicos estudiados en la Sección 2.3.2.

Consideremos, en primer lugar, el grupo $\mathrm{GL}(n,\mathbb{F})$. En este caso

$$\mathfrak{gl}(n,\mathbb{F}) \equiv \mathrm{Lie}\left(\mathrm{GL}(n,\mathbb{F})\right) = \left\{A \in M_n(\mathbb{F}) \,\middle|\, \mathrm{e}^{tA} \in \mathrm{GL}(n,\mathbb{F})\,,\ \ \forall t \in \mathbb{R}\right\} = M_n(\mathbb{F})\,,$$

ya que la exponencial de una matriz es siempre invertible (cf. la tercera propiedad de la exponencial matricial). En el caso de $\mathrm{SL}(n,\mathbb{R})$, al ser

$$\det \mathrm{e}^{tA} = \mathrm{e}^{t\,\mathrm{tr}\,A} = 1\,, \quad \forall t \in \mathbb{R} \quad \Longleftrightarrow \quad \mathrm{tr}\,A = 0 \qquad (A \in M_n(\mathbb{R}))$$

se tiene

$$\mathfrak{sl}(n,\mathbb{R}) \equiv \mathrm{Lie}\left(\mathrm{SL}(n,\mathbb{R})\right) = \left\{A \in M_n(\mathbb{R}) \,\middle|\, \mathrm{tr}\,A = 0\right\}.$$

El caso complejo es similar, ya que si $A \in M_n(\mathbb{C})$ se tiene

$$\det \mathrm{e}^{tA} = \mathrm{e}^{t\,\mathrm{tr}\,A} = 0\,, \quad \forall t \in \mathbb{R} \quad \Longleftrightarrow \quad t\,\mathrm{tr}\,A \in 2\pi\mathrm{i}\mathbb{Z}\,,\ \ \forall t \quad \Longleftrightarrow \quad \mathrm{tr}\,A = 0\,.$$

Por tanto

$$\mathfrak{sl}(n,\mathbb{C}) \equiv \mathrm{Lie}\left(\mathrm{SL}(n,\mathbb{C})\right) = \left\{A \in M_n(\mathbb{C}) \,\middle|\, \mathrm{tr}\,A = 0\right\}.$$

Calculemos a continuación el álgebra de Lie del grupo G_1 definido por la ec. (2.2), que de acuerdo con el Corolario 2.77 está dado por

$$\mathfrak{g}_1 \equiv \mathrm{Lie}(G_1) = \left\{A \in M_n(\mathbb{F}) \,\middle|\, \mathrm{e}^{tA^{\mathsf{T}}} B \mathrm{e}^{tA} = B\,,\ \ \forall t \in \mathbb{R}\right\}. \tag{2.8}$$

Derivando respecto de t la relación

$$\mathrm{e}^{tA^{\mathsf{T}}} B \mathrm{e}^{tA} = B$$

y haciendo $t = 0$ se obtiene fácilmente

$$A^{\mathsf{T}}B + BA = 0\,. \tag{2.9}$$

Por tanto

$$\mathfrak{g}_1 \subset \left\{A \in M_n(\mathbb{F}) \,\middle|\, A^{\mathsf{T}}B + BA = 0\right\}.$$

Recíprocamente, si una matriz $A \in M_n(\mathbb{F})$ satisface la relación (2.9) entonces para todo $t \in \mathbb{R}$ se tiene

$$A^{\mathsf{T}} = -BAB^{-1} \quad \Longrightarrow \quad \mathrm{e}^{tA^{\mathsf{T}}} B \mathrm{e}^{tA} = \mathrm{e}^{-tBAB^{-1}} B \mathrm{e}^{tA} = B \mathrm{e}^{-tA} B^{-1} B \mathrm{e}^{tA} = B \mathrm{e}^{-tA} \mathrm{e}^{tA} = B\,,$$

donde hemos utilizado las propiedades 3 y 6 de la exponencial matricial. Por tanto

$$A^{\mathsf{T}}B + BA = 0 \quad \Longrightarrow \quad A \in \mathfrak{g}_1\,,$$

de donde se sigue que

$$\mathfrak{g}_1 = \left\{A \in M_n(\mathbb{F}) \,\middle|\, A^{\mathsf{T}}B + BA = 0\right\}. \tag{2.10}$$

En particular, las álgebras de Lie de los grupos $\mathrm{O}(n,\mathbb{F})$, $\mathrm{O}(p,q)$ y $\mathrm{SP}(n,\mathbb{F})$ están dadas respectivamente por

$$\begin{aligned}
\mathrm{Lie}\left(\mathrm{O}(n,\mathbb{F})\right) &= \left\{A \in M_n(\mathbb{F}) \,\middle|\, A^{\mathsf{T}} = -A\right\} \equiv \mathfrak{o}(n)\,,\\
\mathrm{Lie}\left(\mathrm{O}(p,q)\right) &= \left\{A \in M_n(\mathbb{F}) \,\middle|\, A^{\mathsf{T}}B + AB = 0\right\} \equiv \mathfrak{o}(p,q)\,, \qquad B = \begin{pmatrix} \mathbb{1}_p & 0 \\ 0 & -\mathbb{1}_q \end{pmatrix},\\
\mathrm{Lie}\left(\mathrm{SP}(n,\mathbb{F})\right) &= \left\{A \in M_n(\mathbb{F}) \,\middle|\, A^{\mathsf{T}}J + AJ = 0\right\} \equiv \mathfrak{sp}(n,\mathbb{F})\,, \qquad J = \begin{pmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix}.
\end{aligned}$$

Como $\mathrm{SO}(n,\mathbb{F})$ y $\mathrm{SO}(p,q)$ son subgrupos abiertos respectivamente de $\mathrm{O}(n,\mathbb{F})$ y $\mathrm{O}(p,q)$, sus álgebras de Lie $\mathfrak{so}(n,\mathbb{F})$ y $\mathfrak{so}(p,q)$ coinciden respectivamente con $\mathfrak{o}(n,\mathbb{F})$ y $\mathfrak{o}(p,q)$, es decir

$$\mathfrak{so}(n,\mathbb{F}) = \mathfrak{o}(n,\mathbb{F})\,, \qquad \mathfrak{so}(p,q) = \mathfrak{o}(p,q)\,.$$

Análogamente, el álgebra de Lie del grupo G_2 definido por (2.3) está dada por

$$\mathfrak{g}_2 \equiv \mathrm{Lie}(G_2) = \left\{A \in M_n(\mathbb{C}) \;\middle|\; A^\dagger B + BA = 0\right\}.$$

En particular, el álgebra de Lie $\mathfrak{u}(n)$ del grupo unitario $\mathrm{U}(n)$ es el conjunto

$$\mathfrak{u}(n) = \left\{A \in M_n(\mathbb{C}) \;\middle|\; A^\dagger = -A\right\}.$$

Consideremos, por último, los grupos $\mathrm{SU}(n)$ y $\mathrm{SP}(n)$.

Lema 2.79. *Sean G_1 y G_2 sendos grupos matriciales cerrados. Entonces $G = G_1 \cap G_2$ es un grupo matricial cerrado, y su álgebra de Lie está dada por*

$$\mathrm{Lie}(G) = \mathrm{Lie}(G_1) \cap \mathrm{Lie}(G_2)\,.$$

Demostración. En efecto, al ser G_1 y G_2 cerrados en $\mathrm{GL}(n,\mathbb{F})$ también lo es $G = G_1 \cap G_2$. Por tanto G es un grupo matricial cerrado. Por otra parte,

$$\begin{aligned}\mathrm{Lie}(G) &= \left\{A \in M_n(\mathbb{C}) \;\middle|\; \mathrm{e}^{tA} \in G_1 \cap G_2\,,\ \ \forall t \in \mathbb{R}\right\}\\ &= \left\{A \in M_n(\mathbb{C}) \;\middle|\; \mathrm{e}^{tA} \in G_1\,,\ \ \forall t \in \mathbb{R}\right\} \cap \left\{A \in M_n(\mathbb{C}) \;\middle|\; \mathrm{e}^{tA} \in G_2\,,\ \ \forall t \in \mathbb{R}\right\}\\ &= \mathrm{Lie}(G_1) \cap \mathrm{Lie}(G_2)\,.\end{aligned}$$

□

Al ser $\mathrm{SU}(n) = \mathrm{U}(n) \cap \mathrm{SL}(n,\mathbb{C})$ y $\mathrm{SP}(n) = \mathrm{SP}(n,\mathbb{C}) \cap \mathrm{U}(2n)$, en virtud del lema anterior sus respectivas álgebras de Lie $\mathfrak{su}(n)$ y $\mathfrak{sp}(n)$ están dadas por

$$\begin{aligned}\mathfrak{su}(n) &= \mathfrak{u}(n) \cap \mathfrak{sl}(n,\mathbb{C}) = \left\{A \in M_n(\mathbb{C}) \;\middle|\; A^\dagger = -A\,,\ \ \mathrm{tr}\,A = 0\right\},\\ \mathfrak{sp}(n) &= \mathfrak{sp}(n,\mathbb{C}) \cap \mathfrak{u}(n) = \left\{A \in M_n(\mathbb{C}) \;\middle|\; A^{\mathsf{T}} J + JA = 0\,,\ \ A^\dagger = -A\right\}.\end{aligned}$$

Las caracterizaciones anteriores de las álgebras de Lie de los grupos clásicos se pueden utilizar para calcular fácilmente su dimensión. Por ejemplo, la dimensión del grupo $\mathrm{SU}(n)$ coincide con la del espacio $\mathfrak{su}(n)$ de las matrices antiautoadjuntas de traza nula. Dicha dimensión se calcula fácilmente observando que toda matriz antiautoadjunta $A \in M_n(\mathbb{C})$ se puede escribir en la forma

$$A = \sum_{1\leqslant j<k\leqslant n} \left[a_{jk}(E_{jk} - E_{kj}) + \mathrm{i}b_{jk}(E_{jk} + E_{kj})\right] + \mathrm{i}\sum_{j=1}^{n} c_j E_{jj}\,, \qquad a_{jk}\,, b_{jk}\,, c_j \in \mathbb{R}\,.$$

(En la expresión anterior, E_{ij} denota la matriz de orden n con todos los elementos nulos excepto un 1 en la intersección de la fila i con la columna j.) Imponiendo la condición de traza nula se obtiene

$$c_{nn} = -\sum_{j=1}^{n-1} c_{jj}\,,$$

y en consecuencia

$$A = \sum_{1\leqslant j<k\leqslant n} \left[a_{jk}(E_{jk} - E_{kj}) + \mathrm{i}b_{jk}(E_{jk} + E_{kj})\right] + \mathrm{i}\sum_{j=1}^{n-1} c_j(E_{jj} - E_{nn})\,.$$

Por tanto $\mathfrak{su}(n)$ está generado por las n^2-1 matrices

$$E_{jk}-E_{kj}\,,\quad \mathrm{i}(E_{jk}+E_{kj})\,,\quad E_{ll}-E_{nn}\,,\qquad 1\leqslant j<k\leqslant n\,,\ \ 1\leqslant l\leqslant n-1\,.$$

Es fácil ver que estas matrices son linealmente independientes, y por tanto forman una base de $\mathfrak{su}(n)$. Por consiguiente

$$\dim \mathrm{SU}(n)=\dim \mathfrak{su}(n)=n^2-1\,,$$

como ya sabíamos.

Ejemplo 2.80. Consideremos el conjunto G' de las transformaciones $f:\mathbb{R}^2\to\mathbb{R}^2$ de la forma

$$f(t,x)=(t+a,\lambda x+vt+b)\,,\qquad a,b,v,\in\mathbb{R}\,,\quad \lambda\in\mathbb{R}^*\,.$$

Si representamos los puntos de $\mathbb{R}^2$ mediante vectores columna $\left(\begin{smallmatrix}t\\x\\1\end{smallmatrix}\right)\in\mathbb{R}^3$, la aplicación f se identifica con la transformación

$$\begin{pmatrix}t\\x\\1\end{pmatrix}\mapsto\begin{pmatrix}1&0&a\\v&\lambda&b\\0&0&1\end{pmatrix}\begin{pmatrix}t\\x\\1\end{pmatrix}\equiv A(a,b,v,\lambda)\begin{pmatrix}t\\x\\1\end{pmatrix},$$

con $A(a,b,v,\lambda)\in M_3(\mathbb{R})$. Por tanto podemos identificar G' con el conjunto

$$G=\big\{A(a,b,v,\lambda)\,\big|\,(a,b,v,\lambda)\in\mathbb{R}^3\times\mathbb{R}^*\big\}\subset M_3(\mathbb{R})\,.$$

Este conjunto es un subgrupo de $\mathrm{GL}(3,\mathbb{R})$, ya que

$$\begin{pmatrix}1&0&a\\v&\lambda&b\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&a'\\v'&\lambda'&b'\\0&0&1\end{pmatrix}=\begin{pmatrix}1&0&a+a'\\v+\lambda v'&\lambda\lambda'&va'+\lambda b'+b\\0&0&1\end{pmatrix},$$

y por tanto

$$\begin{cases}A(a,b,v,\lambda)A(a',b',v',\lambda')=A(a+a',b+\lambda b'+va',v+\lambda v',\lambda\lambda')\in G\,,\\ A(a,b,v,\lambda)^{-1}=A\Big(-a,\dfrac{-b+va}{\lambda},-\dfrac{v}{\lambda},\dfrac{1}{\lambda}\Big)\in G\,.\end{cases}$$

Como G es claramente cerrado en $M_3(\mathbb{R})$, es un grupo matricial cerrado y, por tanto, un grupo de Lie. El grupo G es la imagen de $\mathbb{R}^3\times\mathbb{R}^*$ bajo la aplicación $\varphi:\mathbb{R}^3\times\mathbb{R}^*\to G$ definida por

$$\varphi(a,b,v,\lambda)=A(a,b,v,\lambda)\,,$$

cuya inversa $\varphi^{-1}:G\to\mathbb{R}^3\times\mathbb{R}^*$ está dada por

$$\varphi^{-1}\big(A(a,b,v,\lambda)\big)=(a,b,v,\lambda)\,.$$

Las aplicaciones φ y φ^{-1} son ambas continuas. En efecto, si la sucesión $\{(a_n,b_n,v_n,\lambda_n)\}_{n\in\mathbb{N}}\subset\mathbb{R}^3\times\mathbb{R}^*$ converge a $(a,b,v,\lambda)\in\mathbb{R}^3\times\mathbb{R}^*$ es evidente que $A(a_n,b_n,v_n,\lambda_n)\to A(a,b,v,\lambda)$, por lo que φ es continua. Análogamente, si $\{A(a_n,b_n,v_n,\lambda_n)\}_{n\in\mathbb{N}}\subset G$ converge a una matriz $A(a,b,v,\lambda)\in G$ entonces

$$a=\lim_{n\to\infty}a_n\,,\quad b=\lim_{n\to\infty}b_n\,,\quad v=\lim_{n\to\infty}v_n\,,\quad \lambda=\lim_{n\to\infty}\lambda_n\neq 0\,,$$

y en consecuencia

$$\lim_{n\to\infty}\varphi^{-1}\big(A(a_n,b_n,v_n,\lambda_n)\big)=\varphi^{-1}\big(A(a,b,v,\lambda)\big)\,.$$

Luego φ^{-1} también es continua, y φ es un homeomorfismo. El grupo G es por tanto homeomorfo al abierto $\mathbb{R}^3\times\mathbb{R}^*\subset\mathbb{R}^4$, de donde se sigue que

$$\dim G=4\,.$$

Como $\mathbb{R}^3 \times \mathbb{R}^*$ es no compacto y tiene dos componentes conexas ($\mathbb{R}^3 \times \mathbb{R}_+$ y $\mathbb{R}^3 \times \mathbb{R}_-$), G es no compacto y tiene dos componentes conexas $G_\pm$ dadas por

$$G_\pm = \big\{A(a,b,v,\lambda) \;\big|\; \pm\lambda > 0\big\} = \big\{A \in G \;\big|\; \pm \det A > 0\big\}\,.$$

Evidentemente G_+, la componente conexa de la identidad, es un subconjunto abierto de G y por tanto es un grupo de Lie de dimensión 4.

Hallemos a continuación el álgebra de Lie de G. En primer lugar, sea

$$t \mapsto A\big(a(t), b(t), v(t), \lambda(t)\big)$$

cualquier curva en G que pase por la identidad en $t = 0$ (es decir, $a(0) = b(0) = v(0) = 0$, $\lambda(0) = 1$), y llamemos

$$\alpha = a'(0)\,, \quad \beta = b'(0)\,, \quad \nu = v'(0)\,, \quad \mu = \lambda'(0)\,.$$

Entonces

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} A\big(a(t), b(t), v(t), \lambda(t)\big) \equiv \begin{pmatrix} 0 & 0 & \alpha \\ \nu & \mu & \beta \\ 0 & 0 & 0 \end{pmatrix} \in \mathrm{Lie}(G)\,,$$

y por tanto

$$\mathrm{Lie}(G) \subset \left\{ \begin{pmatrix} 0 & 0 & \alpha \\ \nu & \mu & \beta \\ 0 & 0 & 0 \end{pmatrix} \;\middle|\; (\alpha, \beta, \nu, \mu) \in \mathbb{R}^4 \right\}\,.$$

Recíprocamente, dados $(\alpha, \beta, \nu, \mu) \in \mathbb{R}^4$ la aplicación

$$t \mapsto A\big(\alpha t, \beta t, \nu t, \mathrm{e}^{\mu t}\big)$$

es una curva en G que pasa por la identidad en $t = 0$, y su vector tangente en la identidad es la matriz $\begin{pmatrix} 0 & 0 & \alpha \\ \nu & \mu & \beta \\ 0 & 0 & 0 \end{pmatrix}$. Por tanto

$$\mathrm{Lie}(G) = \left\{ \begin{pmatrix} 0 & 0 & \alpha \\ \nu & \mu & \beta \\ 0 & 0 & 0 \end{pmatrix} \;\middle|\; (\alpha, \beta, \nu, \mu) \in \mathbb{R}^4 \right\} = \mathrm{lin}\{E_{13}, E_{21}, E_{22}, E_{23}\}\,.$$

Una base de $\mathrm{Lie}(G)$ es, por tanto, la formada por las matrices

$$T = E_{13}\,, \quad K = E_{21}\,, \quad D = E_{22}\,, \quad P = E_{23}\,. \tag{2.11}$$

Las relaciones de conmutación de los elementos de esta base se calculan fácilmente teniendo en cuenta que

$$E_{ij} E_{kl} = \delta_{jk} E_{il}\,.$$

De esta forma se comprueba que los únicos conmutadores no nulos están dados por

$$[T, K] = -P\,, \qquad [K, D] = -K\,, \qquad [D, P] = P$$

y sus opuestos.

Estudiemos, a continuación, los grupos a un parámetro generados por los elementos de la base (2.11). Al ser

$$T^2 = K^2 = P^2 = 0\,,$$

de la definición de la exponencial se sigue que

$$\mathrm{e}^{tT} = \mathbb{1} + tT = A(t,0,0,1)\,, \quad \mathrm{e}^{tK} = \mathbb{1} + tK = A(0,0,t,1)\,, \quad \mathrm{e}^{tP} = \mathbb{1} + tP = A(0,t,0,1)\,. \tag{2.12}$$

Por otra parte, al ser D diagonal se tiene

$$e^{tD} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^t & 0 \\ 0 & 0 & 1 \end{pmatrix} = A(0,0,0,e^t)\,. \tag{2.13}$$

Es inmediato comprobar (utilizando, por ejemplo, la identificación de G con G') que

$$A(a,b,v,\lambda) = A(a,0,0,1)A(b,0,0,1)A(0,0,v,1)A(0,0,0,\lambda)\,.$$

De las ecs. (2.12)-(2.13) se sigue entonces que todo elemento de la componente conexa con la identidad de G se puede expresar en la forma

$$A(a,b,v,\lambda) = e^{aT}e^{bP}e^{vK}e^{(\log\lambda)D}\,.$$