



ESCUELA DE CIENCIAS
LICENCIATURA EN MATEMÁTICAS

APLICACIONES DEL ÁLGEBRA LINEAL A LA GEOMETRÍA Y A LA TOPOLOGÍA

Tesis presentada por Ixchel Dzohara Gutiérrez Rodríguez para obtener el grado de Licenciado en Matemáticas por la Universidad Autónoma "Benito Juárez" de Oaxaca.

Director de tesis:
Dr. Miguel Alejandro Xicoténcatl Merino.

*Con todo mi amor para aquellas personas
que siempre han hecho todo en la vida
para que pueda lograr mis sueños:
a mis padres y hermanos.*

Agradecimientos

“Fantasear es imaginar cosas que no tienes, puede ser malo igual que demasiada sal es mala en la sopa o incluso un poco de ajo en un pastel de chocolate. Quiero decir, que las fantasías pueden ser malas si hay demasiadas o si están en el lugar equivocado, pero pueden ser buenas por sí mismas y una gran ayuda en la vida y en la solución de problemas.”

(GEORGE PÓLYA)

Estas líneas expresan mi agradecimiento a todas aquellas personas que con su ayuda colaboraron en la realización del presente trabajo, en especial al Dr. Miguel Alejandro Xicoténcatl Merino, quién es director de esta tesis; por su orientación, seguimiento y la supervisión continua de la misma, pero sobre todo por la motivación y el apoyo que recibí durante mi estancia en el Distrito Federal.

Un especial agradecimiento por la comprensión y el ánimo recibido por parte de mi familia y amigos.

Introducción

El álgebra lineal se fundamenta en el estudio de los espacios vectoriales y transformaciones lineales. La historia del álgebra lineal moderna se remota al siglo XIX con el surgimiento del álgebra matricial a partir de la teoría de los determinantes y los métodos para resolver sistemas de ecuaciones lineales. En este sentido cabe señalar las obras de W. R. Hamilton (1805-1865) y H. G. Grassman (1809-1877). Hamilton se ocupó de los vectores y creó un sistema de números complejos extendidos que llamó *cuaterniones*. Además de Hamilton y Grassman, tres matemáticos que contribuyeron al progreso del álgebra lineal en el siglo XIX son: A. Cayley (1821-1895), J. Sylvester (1814-1897) y C. Hermite (1822-1901). Se debe a ellos la creación de la *teoría de invariantes*. Es a Sylvester a quien se deben las *matrices* y el concepto de *rango*, pero fué Cayley quien desarrolló en 1858 una nueva álgebra con el cálculo de las matrices. La introducción de las matrices es un ejemplo de que un artificio trivial en notación, es el germen de una vasta teoría con innumerables aplicaciones. El álgebra lineal tiene aplicaciones en diversas áreas, el propósito principal de esta tesis es proporcionar un texto que contenga aplicaciones a las áreas de la geometría y de la topología. Nuestro interés se centrará en el estudio de las formas bilineales, espacios con producto interno, las superficies cuádricas y algunas propiedades de los grupos de matrices; ilustraremos ejemplos importantes como son los cuaterniones y los grupos ortogonales.

En el primer capítulo introducimos lo que es una *forma bilineal* sobre un espacio vectorial, considerando los campos \mathbb{R} y \mathbb{C} . Veremos que una función cuadrática es una función de una variable para la cual existe una forma bilineal simétrica. Habiendo definido las formas bilineales sobre los espacios vectoriales, nos detendremos a estudiar un tipo muy concreto, los productos internos y con ello abordar las propiedades de los espacios euclidianos para el caso del campo \mathbb{R} y los espacios unitarios en el caso de los complejos. Hablar de espacios euclidianos nos lleva a estudiar las funciones que preservan el producto interno, ellas son las *transformaciones ortogonales*. Principalmente nos interesan los automorfismos ortogonales de \mathbb{R}^n , ya que el conjunto de todas estas transformaciones forman un grupo con la multiplicación de matrices, el grupo ortogonal $O(n)$ y nos detendremos en su estudio. De manera análoga, para el caso complejo las funciones que preservan el producto interno son las *transformaciones unitarias*; las cuales forman el grupo unitario $U(n)$. Finalmente, en este capítulo mencionaremos un teorema importante, el *teorema espectral* para el caso real y complejo.

En el segundo capítulo abordaremos el estudio de las *superficies cuádricas*, daremos su clasificación en \mathbb{R}^n . La clasificación de las superficies es una rama importante en la geometría y se hace de manera rigurosa. El criterio aquí es que dos cuádricas pertenecen

a la misma categoría cuando se puede pasar de una a la otra mediante un cambio de un sistema de coordenadas adecuado (ortogonal). Generalizaremos la noción de cuádrica en el espacio proyectivo real y complejo. Para entender una superficie cuádrica en dimensiones superiores estudiaremos sus secciones por espacios tridimensionales reales. Nuestro propósito será describir cuádricas salvo homeomorfismos, para ello estudiaremos el tipo topológico de las mismas.

Los grupos se encuentran dentro de los objetos más importantes de las matemáticas y de la física teórica, esta importancia se debe al hecho de que los grupos son una herramienta para el estudio de las simetrías. En geometría diferencial existe una clase especial de grupos, los *grupos de Lie*. Estos grupos representan los objetos en los que su aspecto geométrico y algebraico coexisten. Un objeto más simple, el álgebra de Lie, es un espacio vectorial de dimensión finita; canónicamente a cada grupo de Lie se le puede asociar su álgebra de Lie. A pesar de su simplicidad, el álgebra de Lie de un grupo codifica lo esencial de la información relativa al grupo en sí. En el tercer capítulo, abordaremos el estudio de los grupos lineales y sus álgebras de Lie, nuestro interés se centrará en estudiar a los grupos de matrices $GL(n, K)$, $SL(n, K)$, $SO(n)$, $O(n)$, $SU(n)$, $U(n)$. Así mismo, estudiaremos las descomposiciones polar y de Iwasawa. Abordaremos el estudio de la función exponencial para matrices y su inversa la función logaritmo. La función exponencial es una función que transforma elementos del álgebra de Lie en elementos del grupo de Lie. Por otro lado, la idea de la *representación* de un grupo desempeña un papel importante en el estudio de las simetrías. Formalmente, una representación de un grupo de Lie G en un espacio vectorial V es un homomorfismo de grupos $G \rightarrow Aut(V)$. Trataremos aquí la *representación adjunta* de los grupos de matrices hasta ahora mencionados.

En el último capítulo estudiaremos al grupo $SO(3)$ y su álgebra de Lie. Veremos que existe una relación entre los grupos $SU(2)$ y $SO(3)$ por sus álgebras de Lie. Mostraremos que existe un homeomorfismo entre $SO(3)$ y $\mathbb{R}P^3$, ello nos permite concluir que al mismo tiempo son grupos, espacios topológicos, variedades y grupos de Lie. Así mismo abordaremos el estudio de los cuaterniones \mathbb{H} y discutiremos sus propiedades algebraicas. Estudiaremos su geometría, la estrecha relación con las rotaciones en \mathbb{R}^3 y su relación con el grupo de Lie $SU(2)$. Estudiaremos en particular a los grupos $SU(2)$, $SO(3)$ y $SO(4)$. Y nos detendremos en el grupo de los automorfismos de \mathbb{H}^n que preservan la norma, este se denomina el *grupos simpléctico* $Sp(n)$.

Cuando la mecánica cuántica del electrón fué desarrollada en el siglo XX, se pudo observar que los cuaterniones estaban relacionados con los objetos matemáticos recientemente desarrollados llamados *espinores* y que los electrones estaban relacionados con ambos, con ello se desarrolló una notación equivalente a los cuaterniones la cuál es estándar en aplicaciones físicas, ésta se basa en matrices 2×2 conocidas como las matrices de Pauli; en este capítulo utilizaremos esta notación. En la última parte de este capítulo hablaremos de las propiedades de un grupo muy importante en física, el *grupo de Lorentz* $L = O(1, 3)$ el cual es una generalización de los grupos ortogonales. Este grupo como veremos, puede ser dotado de la estructura de grupo de Lie y está formado por cuatro componentes conexas, pero basta estudiar una componente particular para conocer las propiedades del grupo, lo veremos a detalle más adelante. Por otro lado, la descomposición de Iwasawa nos

permitirá ver al grupo de Lorentz como variedad $L \approx \mathbb{R}^3 \times \mathbb{R}P^3$. Finalmente mencionamos una importante relación entre los grupos $SL(2, \mathbb{C})$ y el grupo de Lorentz restringido L_+^\uparrow a través de la llamada *representación espinorial*, ello nos permitirá concluir que existe un isomorfismo entre el grupo de Möbius y L_+^\uparrow .

Índice general

Introducción	V
1. Formas Bilineales	1
1.1. Formas bilineales y formas cuadráticas	1
1.2. Espacios Euclidianos	4
1.3. Grupos Ortogonales	9
1.4. Transformación de los ejes principales	12
1.5. Espacios unitarios	19
2. Superficies cuádricas	27
2.1. Superficies de segundo orden	27
2.2. Secciones cónicas y superficies regladas.	34
2.3. El espacio proyectivo	37
2.4. Ecuaciones homogéneas.	42
2.5. Tipo topológico de las cuádricas	45
3. Grupos lineales y álgebras de Lie	51
3.1. Descomposición polar y de Iwasawa	51
3.2. Funciones exponencial y logaritmo	54
3.3. Álgebras de Lie	58
3.4. La representación adjunta	62
4. Cuaterniones y grupos ortogonales	67
4.1. El grupo $SO(3)$ y su álgebra de Lie	67
4.2. Cuaterniones	72
4.3. Los grupos $SU(2)$, $SO(3)$ y $SO(4)$	76
4.4. El grupo simpléctico	83
4.5. El grupo de Lorentz	87
Conclusiones	95
Bibliografía	97

Capítulo 1

Formas Bilineales

En este capítulo estudiaremos dos tipos de funciones: las formas bilineales y las formas cuadráticas; éstas son de gran interés por sus diversas aplicaciones en álgebra lineal y sobre todo en geometría. Desarrollaremos la teoría de espacios vectoriales con una forma bilineal asociada muy particular (simétrica y definida positiva), considerando los campos real y complejo. A lo largo de este capítulo V denota un espacio vectorial sobre un campo K .

1.1. Formas bilineales y formas cuadráticas

Definición 1.1. Una forma bilineal γ sobre V es una función:

$$\gamma : V \times V \rightarrow K$$

que satisface:

1. $\gamma(\lambda_1 v_1 + \lambda_2 v_2, w) = \lambda_1 \gamma(v_1, w) + \lambda_2 \gamma(v_2, w)$.
2. $\gamma(v, \lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 \gamma(v, w_1) + \lambda_2 \gamma(v, w_2)$.

para todo $\lambda_1, \lambda_2 \in K$ y $v, v_1, v_2, w, w_1, w_2 \in V$

En otras palabras, γ es lineal en cada variable.

Definición 1.2. Supongamos que $\dim V = n$ y (v_1, \dots, v_n) es una base para V . Definimos la **matriz asociada** a la forma bilineal γ con respecto a la base dada como:

$$G = (g_{ij}) = (\gamma(v_i, v_j)).$$

La matriz asociada a γ sirve para calcular el valor de la forma bilineal si conocemos las coordenadas de los vectores en la base dada. De tal manera que G está determinada por γ y toda matriz G determina una forma bilineal γ por

$$\gamma \left(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j \right) = \gamma \left(\sum_{i,j=1}^n x_i y_j g_{ij} \right) = \mathbf{x}^t G \mathbf{y}.$$

Para K^n con la base canónica, la matriz asociada a la forma bilineal γ es

$$\gamma(x, y) \mapsto \mathbf{x}^t G \mathbf{y} \quad (1.1.1)$$

donde \mathbf{x}, \mathbf{y} son matrices coordenadas de x, y .

Si denotamos por $Bil(V)$ al espacio vectorial de formas bilineales sobre V con las operaciones correspondientes entonces dada una base podemos encontrar el isomorfismo:

$$\begin{aligned} M_n(K) &\xrightarrow{\cong} Bil(V) \\ G &\mapsto \gamma : V \times V \rightarrow K. \end{aligned}$$

Una pregunta natural es, ¿cómo se transforma la matriz G cuando se cambia la base de V ?. Sea T la transformación de cambio de coordenadas, se sigue que:

$$\mathbf{x} = T\tilde{\mathbf{x}}, \quad \mathbf{y} = T\tilde{\mathbf{y}},$$

por lo tanto $\gamma(T\tilde{\mathbf{x}}, T\tilde{\mathbf{y}}) = (T\tilde{\mathbf{x}})^t G T\tilde{\mathbf{y}} = \tilde{\mathbf{x}}^t (T^t G T)\tilde{\mathbf{y}}$. De tal manera que la matriz asociada a γ en las nuevas coordenadas está dada por $T^t G T$. En consecuencia, la matriz de cambio de base nos permite calcular la matriz de la forma bilineal en bases distintas.

Otra manera de ver a las formas bilineales es mediante el espacio dual. Si $\mathcal{L}_k(V, V^*)$ es el conjunto de todas las funcionales lineales de V en el espacio dual V^* , entonces:

Proposición 1.3. *Si $\Gamma : V \rightarrow V^*$ es una funcional lineal de V en el espacio dual V^* . Entonces la expresión*

$$\gamma(v, w) = \langle \Gamma(w), v \rangle \quad (1.1.2)$$

define una forma bilineal. Recíprocamente, dada una forma bilineal $\gamma : V \times V \rightarrow K$ (1.1.2) definimos la funcional lineal

$$\begin{aligned} \Gamma : V &\rightarrow V^* \\ w &\mapsto \Gamma(w) \end{aligned}$$

$\Gamma(w)(v) = \gamma(v, w)$. Por lo tanto, la ecuación (1.1.2) define un isomorfismo canónico:

$$Bil(V) \cong \mathcal{L}_k(V, V^*).$$

En notación matricial: sea (v_1^*, \dots, v_n^*) la base dual de (v_1, \dots, v_n) tal que $v_i^*(v_j) = \delta_{ij}$. Dado j fijo: $\Gamma(v_j) = \alpha_1 v_1^* + \dots + \alpha_n v_n^*$ donde $\alpha_i = \Gamma(v_j)(v_i) = \gamma(v_i, v_j) = g_{ij}$, por lo tanto

$$\Gamma(v_j) = g_{1j} v_1^* + \dots + g_{nj} v_n^* = \sum_{k=1}^n g_{kj} v_k^* \quad (1.1.3)$$

y (g_{ij}) es la matriz de Γ . En consecuencia la matriz G asociada a γ es la matriz de Γ , cuando en V^* se tiene la base dual. En K^n :

$$x^t G y = (G y)^t x$$

y la forma bilineal $(x, y) \mapsto x^t G y$ corresponde a la funcional lineal

$$\begin{aligned} K^n &\rightarrow K^{n*}, \\ y &\mapsto (G y)^t. \end{aligned}$$

Observación 1.4. *El determinante de Gram $g = \det(G)$ depende de la base pero no así el rango de la matriz, porque:*

$$\text{rango}(T^tGT) = \text{rango}(G), \quad T \in GL_n(K).$$

Por otro lado, las formas bilineales que nos interesan son las que no se alteran al intercambiar el orden de los argumentos, estas formas bilineales se llaman **simétricas**.

Definición 1.5. *Una forma bilineal $\gamma : V \times V \rightarrow K$ se dice:*

1. **Simétrica:** si $\gamma(v, w) = \gamma(w, v) \quad \forall v, w \in V$.
2. **No degenerada:** si la transformación lineal asociada $\Gamma : V \rightarrow V^*$ es un isomorfismo.

Observación 1.6.

1. *En dimensión finita, γ es simétrica si y sólo si $g_{ij} = g_{ji} \quad \forall i, j$. Es decir $G^t = G$ es simétrica.*
2. *γ es no degenerada si y sólo si G es invertible.*

Definición 1.7. *Sea $\gamma : V \times V \rightarrow K$ una forma bilineal simétrica. A la función:*

$$\begin{aligned} \gamma : V &\rightarrow K \\ v &\mapsto \gamma(v, v) \end{aligned}$$

*se le conoce como la **forma cuadrática** asociada a γ . Tal función cumple lo siguiente:*

$$\gamma(kv) = k^2v \quad \forall k \in K, \forall v \in V.$$

A partir de la forma cuadrática se puede encontrar la forma bilineal γ . Sea K un campo de característica distinta de 2. Entonces:

$$\gamma(v + w) = \gamma(v + w, v + w) = \gamma(v, v) + \gamma(w, w) + \gamma(v, w) + \gamma(w, v),$$

pero γ es simétrica, por lo tanto

$$\gamma(v, w) = \frac{1}{2}(\gamma(v + w, v + w) - \gamma(v, v) - \gamma(w, w)). \quad (1.1.4)$$

Análogamente $\gamma(v - w) = \gamma(v, v) + \gamma(w, w) - 2\gamma(v, w)$. En consecuencia:

$$\gamma(v + w, v + w) - \gamma(v - w, v - w) = 4\gamma(v, w). \quad (1.1.5)$$

Es importante notar que la forma cuadrática queda completamente determinada por la parte simétrica de la forma bilineal γ . Asimismo, la parte simétrica de una forma bilineal queda completamente determinada por la forma cuadrática correspondiente.

Definición 1.8. *Sea γ una forma bilineal. Una base (v_1, \dots, v_n) de V se dice **ortogonal**, si $\gamma(v_i, v_j) = 0$ para $i \neq j$. En otras palabras, la matriz asociada G es diagonal.*

1.2. Espacios Euclidianos

Definición 1.9. *Un espacio euclidiano es un espacio vectorial V real con una forma bilineal simétrica:*

$$\begin{aligned}\gamma : V \times V &\rightarrow \mathbb{R}, \\ (v, w) &\mapsto \langle v, w \rangle\end{aligned}$$

y un **producto interno** o **métrica** que es definida positiva, es decir, $\langle v, v \rangle > 0$ para todo $0 \neq v \in V$.

Ejemplo 1.10.

1. \mathbb{R}^n con el producto interno usual $\langle x, y \rangle = x^t y = \sum_j x_j y_j$, es un espacio euclidiano.

2. Si $A \in M_n(\mathbb{R})$ es una matriz invertible y el producto interno se define por $\gamma(x, y) := \langle Ax, Ay \rangle = x^t (A^t A) y \quad \forall x, y \in \mathbb{R}^n$, entonces se tiene un espacio euclidiano.

3. El espacio V de las funciones continuas del intervalo $f : [0, 1] \rightarrow \mathbb{R}$ con el producto interno:

$$\langle f, g \rangle := \int_0^1 f(t)g(t)dt,$$

define un espacio euclidiano.

Definición 1.11. *Sea V un espacio euclidiano en \mathbb{R}^n con una forma bilineal γ , definimos la **norma** de v por:*

$$|v| = \sqrt{\langle v, v \rangle}.$$

Observación 1.12. *Si $\dim V$ es finita entonces el producto interno es no degenerado. En efecto, si $w \neq 0$, entonces la funcional lineal*

$$\begin{aligned}\langle \cdot, w \rangle : V &\rightarrow \mathbb{R} \\ v &\mapsto \langle v, w \rangle\end{aligned}$$

no es cero ya que $\langle w, w \rangle \neq 0$.

Definición 1.13. *Sea $U \subset V$ un subespacio vectorial, definimos el **complemento ortogonal** de U como el conjunto*

$$U^\perp = \{v \in V \mid \langle v, u \rangle = 0 \quad \forall u \in U\}.$$

Claramente el complemento ortogonal U^\perp es un subespacio de V .

Definición 1.14. *Sean $(V, \langle \cdot, \cdot \rangle_V)$ y $(W, \langle \cdot, \cdot \rangle_W)$ espacios euclidianos. Una transformación lineal $\alpha : V \rightarrow W$ se dice **ortogonal** si:*

$$\langle \alpha(u), \alpha(v) \rangle_W = \langle u, v \rangle_V \quad \forall u, v \in V.$$

Definición 1.15. *Sea $\dim V = n$, una base (v_1, \dots, v_n) de V se dice **ortonormal** si:*

$$\langle v_i, v_j \rangle = \delta_{ij} \quad \forall i, j = 1, \dots, n.$$

Donde δ_{ij} es la delta de Kronecker.

Proposición 1.16 (Ortogonalización de Gram-Schmidt). *Sea v_1, \dots, v_k una familia ortonormal en V , es decir, $\langle v_i, v_j \rangle = \delta_{ij}$ y que se puede extender a una base de V por w_{k+1}, \dots, w_n . Entonces la familia se puede extender inductivamente a una base ortonormal poniendo:*

$$\begin{aligned}\tilde{v}_{k+1} &:= w_{k+1} - \sum_{i=1}^k \langle w_{k+1}, v_i \rangle \cdot v_i \\ v_{k+1} &:= \frac{\tilde{v}_{k+1}}{|\tilde{v}_{k+1}|}.\end{aligned}$$

Demostración. Note en primer lugar que los v_1, \dots, v_k son linealmente independientes pues

$$v := \sum_{j=1}^k \lambda_j v_j = 0$$

entonces $\langle v, v_j \rangle = \lambda_j = 0$.

Claramente $(v_1, \dots, v_k, \tilde{v}_{k+1}, w_{k+2}, \dots, w_n)$ es también una base. En particular $\tilde{v}_{k+1} \neq 0$. Por construcción

$$\begin{aligned}\langle \tilde{v}_{k+1}, v_j \rangle &= \langle w_{k+1}, v_j \rangle - \sum_i \langle w_{k+1}, v_i \rangle \langle v_i, v_j \rangle \\ &= \langle w_{k+1}, v_j \rangle - \langle w_{k+1}, v_j \rangle \\ &= 0.\end{aligned}$$

Luego \tilde{v}_{k+1} es ortogonal a v_1, \dots, v_k y por lo tanto $v_{k+1} = \frac{\tilde{v}_{k+1}}{|\tilde{v}_{k+1}|}$ es un vector unitario que también es ortogonal a v_1, \dots, v_k . \square

Definición 1.17. *Una matriz real simétrica G se dice **definida positiva** ($G > 0$), si la forma bilineal asociada $(x, y) \mapsto x^t G y$ es definida positiva.*

Corolario 1.18. *Si G es una matriz $n \times n$, real y definida positiva, entonces existe una matriz $n \times n$, real invertible tal que $G = T^t T$. Recíprocamente si T es invertible entonces $T^t T > 0$.*

Demostración. Consideremos el siguiente producto interno en \mathbb{R}^n :

$$\begin{aligned}\mathbb{R}^n \times \mathbb{R}^n &\rightarrow \mathbb{R} \\ \langle x, y \rangle_G &\mapsto x^t G y\end{aligned}$$

y sea (v_1, \dots, v_n) una base ortonormal de \mathbb{R}^n con respecto a dicho producto interno. Sea

$$\begin{aligned}T : \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ v_i &\mapsto e_i\end{aligned}$$

donde los vectores v_i se consideran con el producto interno definido anteriormente y los e_i son los vectores canónicos de \mathbb{R}^n con el producto interno usual. Entonces T es ortogonal ya que si:

$$x = \alpha_1 v_1 + \dots + \alpha_n v_n, \quad y = \beta_1 v_1 + \dots + \beta_n v_n$$

entonces $\langle x, y \rangle_G = \sum_{i,j} \alpha_i \beta_j \langle v_i, v_j \rangle = \sum_i \alpha_i \beta_i$. Por otro lado

$$\langle T(x), T(y) \rangle = \langle \alpha_1 e_1 + \dots + \alpha_n e_n, \beta_1 e_1 + \dots + \beta_n e_n \rangle = \sum_i \alpha_i \beta_i.$$

Entonces $\langle x, y \rangle_G = \langle T(x), T(y) \rangle$, por lo tanto

$$\begin{aligned} x^t G y &= (T(x))^t T(y) \quad \forall x, y \\ &= x^t (T^t T) y \quad \forall x, y \end{aligned}$$

haciendo $x = e_i, y = e_j$ obtenemos que $G = g_{ij} = (T^t T)_{ij} \quad \forall i, j = 1, \dots, n$. En consecuencia $G = T^t T$.

Es precisamente mediante la ortogonalización de Gram-Schmidt que se construye T con $T(v_i) = e_i$ una matriz triangular superior con diagonal positiva, por lo tanto $T = (t_{ij})$, $t_{ij} = 0$ para $i > j$ y $t_{ii} > 0$. \square

Si W es un espacio euclidiano y $V \subset W$ es un subespacio, entonces V hereda la estructura de W . El producto escalar de V es el inducido por la restricción de W . Si $\dim V$ es finita, entonces V es isomorfo a \mathbb{R}^n . En los espacios euclidianos se cumple lo siguiente.

Proposición 1.19 (Desigualdad de Cauchy-Schwarz). *Si V es un espacio euclidiano y $v, w \in V$ entonces se cumple:*

$$|\langle v, w \rangle| \leq |v| \cdot |w|$$

y se da la igualdad si y sólo si v, w son linealmente dependientes.

Proposición 1.20 (Desigualdad del triángulo).

$$|v + w| \leq |v| + |w|.$$

Ejemplo 1.21. *Si V es el espacio de funciones continuas $f : [0, 1] \rightarrow \mathbb{R}$ con el producto interno $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$, entonces*

$$\left(\int_0^1 f(t)g(t)dt \right)^2 \leq \left(\int_0^1 f^2(t)dt \right) \left(\int_0^1 g^2(t)dt \right).$$

Proposición 1.22 (Proyección ortogonal). *Si V un espacio euclidiano y $U \subset V$ un subespacio de dimensión finita con el producto interno inducido por V . Entonces se tiene una descomposición como suma directa $V = U \oplus U^\perp$ y una transformación lineal bien definida de V en U :*

$$p_U : V = U \oplus U^\perp \xrightarrow{pr_1} U$$

*conocida como la **proyección ortogonal**.*

Demostración. Si $v \in U \cap U^\perp$ entonces $v \in U$ y $v \in U^\perp$, por lo tanto $\langle v, v \rangle = 0$ de donde $v = 0$. En consecuencia $U \cap U^\perp = \{0\}$.

Sea $\beta = (u_1, \dots, u_n)$ una base ortonormal de U , entonces v se puede escribir como: $v = u + \tilde{u}$ donde

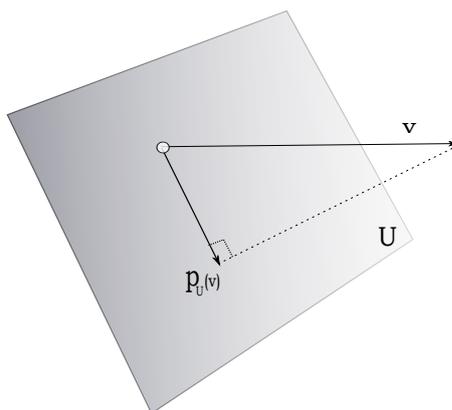


Figura 1.1: Proyección ortogonal

$$u := \sum_{j=1}^n \langle v, u_j \rangle u_j \quad \text{y} \quad \tilde{u} := v - u.$$

Claramente, $u \in U$ y $\langle \tilde{u}, u_j \rangle = \langle v, u_j \rangle - \langle u, u_j \rangle = \langle v, u_j \rangle - \langle v, u_j \rangle = 0$ para toda j . Por lo tanto $\tilde{u} \in U^\perp$. Más aún, la proyección ortogonal está dada por:

$$p_U(v) = \sum_{j=1}^n \langle v, u_j \rangle \cdot u_j.$$

□

Observación 1.23. Dado $v \in V$, para todo $u \in U$ la distancia entre $u - v$ es mayor igual a la distancia entre $p_U - v$, es decir, la menor distancia entre u y v es la proyección.

Demostración. Sea $v = p_u(v) + \tilde{u}$, donde $\tilde{u} \in U^\perp$. Entonces

$$\begin{aligned} |v - u|^2 &= |p_u(v) + \tilde{u} - u|^2 \\ &= |p_u(v) - u|^2 + |\tilde{u}|^2 \\ &\geq |\tilde{u}|^2 \\ &= |v - p_u(v)|^2 \end{aligned}$$

□

Si e_1, e_2, \dots es una familia ortonormal en V que en principio pudiera ser infinita, entonces $\langle v, e_j \rangle$ se denomina el j -ésimo **coeficiente de Fourier** de $v \in V$ con respecto a dicha familia. En este caso se cumple:

Proposición 1.24 (Desigualdad de Bessel).

$$\sum_j \langle v, e_j \rangle^2 \leq |v|^2.$$

Demostración. Sea $v = \sum_{j=1}^n \langle v, e_j \rangle e_j + w$ con $w \in L(e_1, \dots, e_n)^\perp$ por lo tanto:

$$\begin{aligned} |v|^2 &= \sum_{j=1}^n \langle v, e_j \rangle^2 + |w|^2 \\ &\geq \sum_{j=1}^n \langle v, e_j \rangle^2. \end{aligned}$$

En particular, la serie $\sum_j \langle v, e_j \rangle$ es convergente. \square

Ejemplo 1.25. Sea V el espacio de las funciones continuas $[0, 1] \rightarrow \mathbb{R}$ y sea el sistema ortonormal dado por las funciones

$$f(t) = 2 \operatorname{sen}(2\pi nt),$$

$$f(t) = 2 \operatorname{cos}(2\pi nt),$$

(con $n \in \mathbb{N}$) y la función constante 1. En este caso, los coeficientes de Fourier con respecto a esta familia ortogonal son los coeficientes de Fourier usuales.

Notemos que si (e_1, \dots, e_n) es una base ortonormal de V , entonces

$$v = \sum_j \langle v, e_j \rangle e_j,$$

es la representación de $v \in V$ como combinación lineal de los vectores de la base, porque al multiplicar una combinación lineal $v = \lambda_1 e_1 + \dots + \lambda_n e_n$ con e_k se obtiene $\langle v, e_k \rangle = \lambda_k$.

Consideremos ahora un subespacio afín A de un espacio euclidiano V de dimensión n , esto es, $A = q + U$ donde U es un subespacio vectorial de dimensión $n - k$ y $q \in A$. Elíjase una base ortonormal (c_1, \dots, c_k) de U^\perp y pongamos $\alpha_j = \langle c_j, q \rangle$. Notemos que los valores de $\alpha_1, \dots, \alpha_k$ no dependen de la elección de $q \in A$. Entonces A es el conjunto solución del siguiente sistema de ecuaciones:

Definición 1.26 (Forma normal de Hesse).

$$\langle c_j, x \rangle - \alpha_j = 0$$

con $j = 1, \dots, k$.

Observación 1.27. La distancia mínima de un punto $p \in V$ al espacio afín A es igual a $|(\lambda_1, \dots, \lambda_k)| = \sqrt{\lambda_1^2 + \dots + \lambda_k^2}$ donde $\lambda_j := \langle c_j, p \rangle - \alpha_j$, para $j = 1, \dots, k$.

Demostración. Notemos que la distancia de un punto $q \in A$ a p es mínima precisamente cuando $p - q \in U^\perp$ y por lo tanto $p - q = \sum_{j=1}^k \lambda_j c_j$. En tal caso

$$\lambda_j = \langle c_j, p \rangle - \langle c_j, q \rangle = \langle c_j, p \rangle - \alpha_j$$

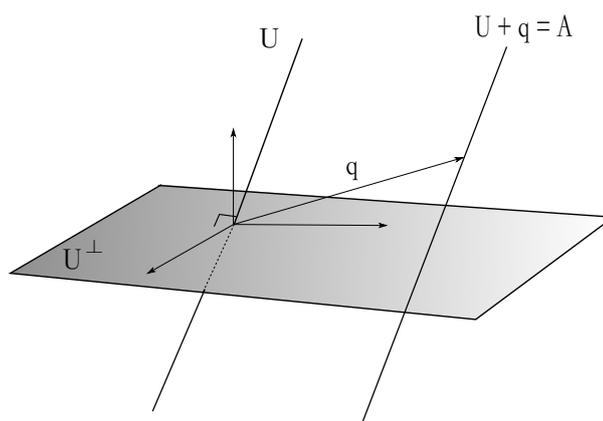


Figura 1.2: Observación 1.27

y la distancia buscada es:

$$\begin{aligned}
 |p - q| &= \left| \sum_{i=1}^k \lambda_i c_i \right| \\
 &= |(\lambda_1, \dots, \lambda_k)| \\
 &= \sqrt{\lambda_1^2 + \dots + \lambda_k^2}
 \end{aligned}$$

□

Para un hiperplano A la forma normal de Hesse consiste de una sola ecuación $\langle c, x \rangle - \alpha = 0$, con $|c| = 1$. En este caso la distancia de un punto $p \in V$ a A está dada por:

$$|\langle c, p \rangle - \alpha|.$$

1.3. Grupos Ortogonales

Naturalmente no sólo estudiaremos a los espacios euclidianos, sino más bien la categoría de espacios euclidianos y transformaciones ortogonales: la transformación identidad entre espacios euclidianos es ortogonal y la composición de transformaciones ortogonales es ortogonal. Una transformación ortogonal $\alpha : V \rightarrow W$ es siempre inyectiva ya que si $\alpha(v) = 0$ entonces $|v| = |\alpha(v)| = 0$ por lo que $v = 0$. Si las dimensiones de los espacios son iguales, entonces la transformación es un isomorfismo. Nos interesan principalmente los isomorfismos ortogonales de \mathbb{R}^n en sí mismo, para ello identificaremos $L(\mathbb{R}^n, \mathbb{R}^n)$ con $M_n(\mathbb{R})$ el espacio vectorial de las matrices $n \times n$. Es fácil ver que el conjunto de todas las transformaciones ortogonales es un grupo con la multiplicación de matrices, como veremos a continuación.

Teorema 1.28. *El grupo de las transformaciones lineales ortogonales $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ con el producto interno usual se denomina el **grupo ortogonal** en dimensión n y se denota por $O(n)$. Además:*

1. Una matriz $A \in M_n(\mathbb{R})$ pertenece a $O(n)$ si y sólo si $A^t \cdot A = I$, es decir, si las columnas de A forman una base ortonormal de \mathbb{R}^n .

2. Si $A \in O(n)$ entonces $\det(A) = \pm 1$.

El subgrupo $SO(n) = \{A \in O(n) \mid \det(A) = 1\}$ se conoce como el **grupo ortogonal especial**

Demostración. Notemos que:

$$\begin{aligned} A \in O(n) &\iff (Ax)^t(Ay) = x^t y \quad \forall x, y \in \mathbb{R}^n \\ &\iff x^t A^t A y = x^t I y \\ &\iff A^t \cdot A = I, \end{aligned}$$

la última equivalencia se obtiene tomando $x = e_i, y = e_j, \forall i, j \in \{1, \dots, n\}$.

Por otro lado, si $A \in O(n)$ entonces:

$$\begin{aligned} (\det(A))^2 &= \det(A)^t \cdot \det(A) \\ &= \det(A^t \cdot A) \\ &= \det(I) \\ &= 1. \end{aligned}$$

□

Notemos también que $A \cdot A^t = I$ porque $A^t = A^{-1}$. Por lo tanto A^t es ortogonal, las filas de A forman una base ortonormal y A^{-1} es ortogonal.

Recordemos que una base ordenada (v_1, \dots, v_n) de \mathbb{R}^n se dice positiva si $\det(v_1, \dots, v_n) > 0$. El grupo $SO(n)$ es el subgrupo de las transformaciones ortogonales que preservan orientación, es decir, $A \in SO(n)$ manda bases positivas en bases positivas.

Veamos que sucede con los grupos $O(n)$ y $SO(n)$ en dimensiones bajas:

Ejemplo 1.29.

1. Notemos que $O(1) = \{id, -id\} \cong \mathbb{Z}/2, SO(1) = \{id\} = 1$.

2. El grupo $SO(2)$ es el grupo de rotaciones en \mathbb{R}^2 .

En efecto, si $A \in SO(2)$ entonces

$$Ae_1 = \begin{pmatrix} a \\ b \end{pmatrix}, \quad a^2 + b^2 = 1$$

y por lo tanto $a = \cos \varphi, b = \sin \varphi$ para algún φ . Como Ae_2 es ortogonal a Ae_1 y tiene norma 1. Ae_2 tiene únicamente dos posibilidades $\begin{pmatrix} -b \\ a \end{pmatrix}$ ó $\begin{pmatrix} b \\ -a \end{pmatrix}$, así:

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

por lo tanto $SO(2) \cong S^1$. El polinomio característico de A es

$$\chi_A(t) = (t - \cos \varphi)^2 + \sin^2 \varphi$$

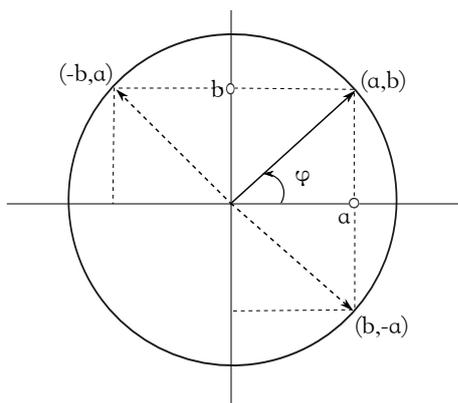


Figura 1.3: Ejemplo 1.29 2.

y cuando $\sin \varphi = 0$, $\chi_A(t)$ tiene raíces reales, es decir $A = \pm I$. En los complejos, $\chi_A(t)$ tiene las raíces $\cos \varphi \pm i \sin \varphi$. Las matrices ortogonales de la forma

$$B = \begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix}, \quad \det(B) = -1$$

consisten precisamente de las matrices en $O(2) \setminus SO(2)$. En este caso, el polinomio característico es $\chi_B(t) = t^2 - 1 = (t + 1) \cdot (t - 1)$, es decir, B deja un vector fijo y manda los vectores ortogonales en sus negativos. Esto es, B es una **reflexión**.

Sea $v(\varphi) = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$ y notemos que $B(\varphi) \cdot v(\varphi/2) = v(\varphi/2)$. Luego $B(\varphi)$ es la reflexión a la recta $L(v(\varphi/2))$. Notemos que $v(\pi) = -v(0)$, pero que $L(v(\pi)) = L(v(0))$. Tomemos a φ como parámetro de S^1 y denotemos por $H(\varphi) = L(v(\varphi/2))$ a la recta que deja fija $B(\varphi)$, es decir, el subespacio propio $B(\varphi)$ de valor propio 1 y consideremos la superficie:

$$M = \{(\varphi, x) \in S^1 \times \mathbb{R}^2 \mid x \in H(\varphi)\}.$$

Es fácil ver que M es una **banda de Möbius** (no compacta).

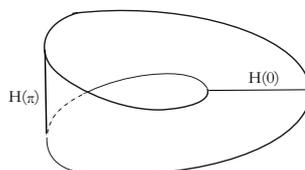


Figura 1.4: Banda de Möbius.

Ejemplo 1.30. *El grupo $SO(3)$,*

El polinomio característico $\chi_A(t)$ de una matriz $A \in SO(3)$ es un polinomio mónico de grado 3, por lo que tiene al menos una raíz real. Entonces existe $\lambda \in \mathbb{R}$ y un vector propio unitario $v \in \mathbb{R}$ tal que $Av = \lambda v$. Notemos que $|\lambda v| = |Av| = |v| = 1$ de tal forma que $|\lambda| = 1$. En consecuencia $\lambda = \pm 1$.

Al considerar a A como matriz sobre los complejos, obtenemos tres valores propios correspondientes a las tres raíces complejas $\lambda_1, \lambda_2, \lambda_3$ de $\chi_A(t)$, los cuales satisfacen que $\lambda_1 \cdot \lambda_2 \cdot \lambda_3 = \det(A) = 1$. Pero como $\chi_A(t)$ es un polinomio con coeficientes reales, sus raíces complejas aparecen conjugadas a pares. De manera que, o bien todos los λ_i son reales y $\lambda_i = \pm 1$ de donde se sigue que alguno de ellos es 1 o alguna de las raíces λ es real y las dos restantes son complejas conjugadas digamos $\mu, \bar{\mu}$. Pero como $\lambda \cdot \mu \cdot \bar{\mu} = 1$ y $\mu \cdot \bar{\mu} = |\mu|^2 > 0$ entonces $\lambda = 1$.

Observación 1.31. *Notemos que toda $A \in SO(3)$ deja fija a una recta $L_A = \{r \cdot v \mid r \in \mathbb{R}\}$ el eje de rotación de A . Por lo que si $U_A = L_A^\perp$ es ortogonal entonces la restricción $A|_{U_A} : U_A \rightarrow U_A$ es una rotación del plano U_A .*

Ejemplo 1.32. *El grupo $O(3)$.*

Si $B \in O(3)$, entonces $\det(B) \cdot B = A \in SO(3)$. Por lo tanto B es una rotación L_A compuesta de una reflexión en el plano U_A (en caso de que $\det(B) = -1$). Mas aún, tenemos un isomorfismo de grupos:

$$SO(3) \times \mathbb{Z}/2 \rightarrow O(3)$$

$$(A, [n]) \mapsto (-1)^n A$$

cuya inversa está dada por $B \mapsto (\det(B) \cdot B, [n])$, donde $(-1)^n = \det(B)$.

1.4. Transformación de los ejes principales

Sea V un espacio euclidiano de dimensión n , con producto interno $\langle \cdot, \cdot \rangle$ y una base ortonormal (v_1, \dots, v_n) . Identificamos V con V^* mediante el isomorfismo

$$V \rightarrow V^*$$

$$w \mapsto \langle \cdot, w \rangle.$$

Tenemos lo siguiente:

1. A todo endomorfismo Γ de V se le asigna una forma bilineal γ dada por:

$$\gamma(v, w) = \langle \Gamma w, v \rangle = \langle v, \Gamma w \rangle.$$

2. Si $G = (g_{ij})$ es la matriz de Γ con respecto a la base (v_1, \dots, v_n) , entonces $\Gamma v_j = \sum_i g_{ij} v_i$ y por lo tanto

$$\begin{aligned} \gamma(v_k, v_j) &= \langle v_k, \Gamma v_j \rangle \\ &= \sum_i g_{ij} \langle v_k, v_i \rangle \\ &= g_{kj}. \end{aligned}$$

En consecuencia, G es también la matriz asociada a la forma bilineal γ .

Definición 1.33. El endomorfismo Γ se denomina **simétrico** o **autoadjunto** cuando la forma bilineal asociada γ es simétrica, ésto es, si para todo $v, w \in V$ se cumple:

$$\langle v, \Gamma w \rangle = \langle w, \Gamma v \rangle = \langle \Gamma v, w \rangle.$$

En términos matriciales, si G es la matriz de Γ respecto a una base ortonormal entonces $g_{ij} = g_{ji}$, o $G^t = G$.

Teorema 1.34 (Teorema espectral para operadores autoadjuntos). Sea Γ un endomorfismo autoadjunto de un espacio euclidiano V de dimensión finita. Entonces V posee una base ortonormal de vectores propios.

Si en el teorema ponemos $V = \mathbb{R}^n$, $\Gamma = G$ una matriz simétrica y $T \in O(n)$ es la transformación que manda la base ortonormal en la base canónica, entonces TGT^{-1} es una matriz diagonal. En efecto, el siguiente diagrama es conmutativo.

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{G} & \mathbb{R}^n \\ T \downarrow & & \downarrow T \\ \mathbb{R}^n & \xrightarrow{TGT^{-1}} & \mathbb{R}^n \end{array}$$

Aquí $T^t = T^{-1}$. Si $Gv_i = \lambda_i v_i$ entonces:

$$\begin{aligned} TGT^{-1}(e_i) &= TGT^{-1}T(v_i) \\ &= TG(v_i) \\ &= \lambda_i T(v_i) \\ &= \lambda_i e_i. \end{aligned}$$

Nótese que también podríamos haber escogido a $T \in SO(n)$, pues siempre podemos cambiar al primer vector de la base por su negativo, en caso de ser necesario. Como consecuencia del teorema tenemos lo siguiente.

Corolario 1.35 (Teorema de los ejes principales). Si G es una matriz real $n \times n$ simétrica, entonces existe $T \in SO(n)$, tal que TGT^{-1} es una matriz diagonal:

$$TGT^{-1}(e_i) = \lambda_i \cdot e_i.$$

La base que se tiene en el teorema es **ortonormal** para la forma bilineal $\langle \cdot, \cdot \rangle$ y al mismo tiempo **ortogonal** para γ .

Lo que sabemos de la forma bilineal $\langle \cdot, \cdot \rangle$ es que es simétrica y definida positiva, por lo que respecto a las matrices podemos decir lo siguiente:

Proposición 1.36 (Diagonalización simultánea). Si P es una matriz $n \times n$, real, simétrica y definida positiva; G una matriz real $n \times n$ simétrica arbitraria. Entonces existe $T \in GL(n, \mathbb{R})$, tal que

1. $T^t P T = I$.
2. $T^t G T$ es una matriz diagonal.

Demostración. Sea T_1 la matriz que manda la base canónica de \mathbb{R}^n en una base ortonormal para el producto interno $(x, y) \mapsto x^t P y$. Entonces por corolario 1.18 se tiene que $T_1^t P T_1 = I$. Por otro lado, por el teorema espectral existe $T_2 \in O(n)$ tal que $T_2^t (T_1^t P T_1) T_2 = (T_1 T_2)^t P (T_1 T_2)$ es diagonal. Además $(T_1 T_2)^t P (T_1 T_2) = T_2^t (T_1^t P T_1) T_2 = T_2^t T_2 = I$. Por lo tanto $T = T_1 T_2$ es la matriz buscada. \square

Lo que resta de esta sección es esencialmente la prueba del teorema espectral.

Observación 1.37. Si Γ es autoadjunta, entonces vectores propios que corresponden a distintos valores propios de Γ son ortogonales.

Demostración. Si $\Gamma(v) = \lambda v$ y $\Gamma(w) = \mu w$ donde $\lambda \neq \mu$, entonces

$$\begin{aligned} \lambda \langle v, w \rangle &= \langle \lambda v, w \rangle \\ &= \langle \Gamma(v), w \rangle \\ &= \langle v, \Gamma(w) \rangle \\ &= \langle v, \mu w \rangle \\ &= \mu \langle v, w \rangle \end{aligned}$$

en consecuencia $(\lambda - \mu) \langle v, w \rangle = 0$. Por lo tanto $\langle v, w \rangle = 0$. \square

Lema 1.38. Si $U \subset V$ es un subespacio, Γ es autoadjunta y $\Gamma U \subset U$ entonces $\Gamma(U^\perp) \subset U^\perp$.

Demostración. Si $u \in U$, $v \in U^\perp$ entonces $\langle u, \Gamma(v) \rangle = \langle \Gamma(u), v \rangle = 0$ porque $\Gamma(u) \in U$ para todo $u \in U$. Por lo tanto $\Gamma(v) \in U^\perp$. \square

Lo anterior muestra que si encontramos un valor propio real de Γ y U es el espacio propio correspondiente (con alguna base ortonormal) entonces $\Gamma U \subset U$, $\Gamma U^\perp \subset U^\perp$ y como $\dim U^\perp < \dim V$, podemos suponer inductivamente que U^\perp tiene una base ortonormal de vectores propios de $\Gamma|_{U^\perp}$. Por lo tanto, junto con la base ortonormal de U tenemos la base de V buscada. Lo que es importante mostrar es que Γ tiene valores propios reales.

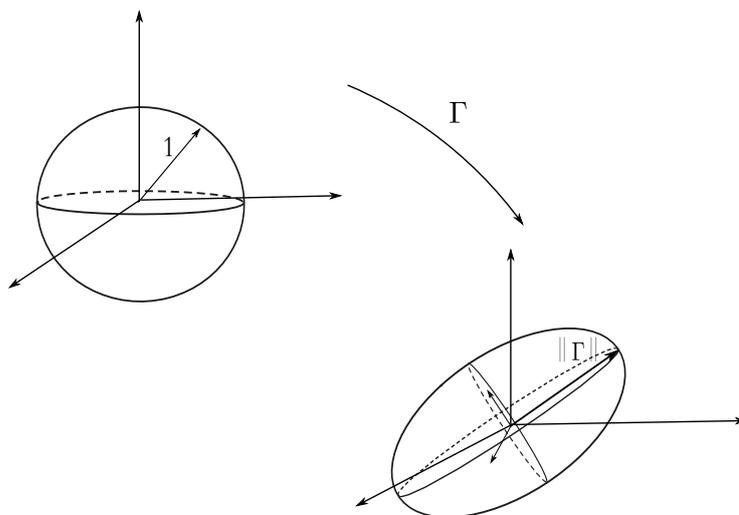
Definición 1.39. Sea Γ un endomorfismo de un espacio euclidiano de dimensión finita. La **norma** de Γ se define como:

$$\|\Gamma\| = \max \{ |\Gamma(v)| \mid v \in V, |v| = 1 \}.$$

En dimensión infinita uno debe tomar el supremo, pero en este caso es fácil mostrar que se alcanza el máximo. A saber, notemos que la esfera unitaria $\{v \in V \mid |v| = 1\}$ es compacta y por lo tanto su imagen bajo Γ también es compacta $\{\Gamma v \mid |v| = 1\}$. Finalmente, como la función norma $|\cdot| \rightarrow \mathbb{R}$ es continua, sabemos que alcanza su máximo en el conjunto anterior.

Regresando a la norma de Γ , esta se puede caracterizar también del modo siguiente: $\|\Gamma\|$ es el mínimo número real K , tal que

$$|\Gamma(v)| \leq K \cdot |v| \quad \forall v \in V.$$

Figura 1.5: Γ manda a la esfera en un elipsoide.

De esto se sigue que para cualquier otro endomorfismo $\Delta : V \rightarrow V$ se tiene:

$$\begin{aligned} \|\Gamma\Delta\| &\leq \|\Gamma\| \cdot \|\Delta\| \\ \|\Gamma + \Delta\| &\leq \|\Gamma\| + \|\Delta\|. \end{aligned}$$

En efecto:

$$|\Gamma\Delta(v)| \leq \|\Gamma\| \cdot |\Delta(v)| \leq \|\Gamma\| \cdot \|\Delta\| \cdot |v|$$

y

$$\begin{aligned} |(\Gamma + \Delta)(v)| &\leq |\Gamma(v)| + |\Delta(v)| \\ &\leq \|\Gamma\| \cdot |v| + \|\Delta\| \cdot |v| \\ &= (\|\Gamma\| + \|\Delta\|) \cdot |v|. \end{aligned}$$

Lema 1.40. Si Γ es autoadjunta entonces

$$\|\Gamma\| = \max\{|\langle \Gamma(v), v \rangle| \mid |v| = 1\} =: M.$$

Demostración. La desigualdad de Cauchy-Schwarz y el hecho de que $|v| = 1$ da como resultado:

$$|\langle \Gamma(v), v \rangle| \leq |\Gamma(v)| \cdot |v| \leq \|\Gamma\|.$$

Por lo tanto $M \leq \|\Gamma\|$. Resta mostrar que $\|\Gamma\| \leq M$ ó $|\Gamma(v)| \leq M$ cuando $|v| = 1$.

Supongamos que $\Gamma(v) \neq 0$ y $w = \frac{\Gamma(v)}{|\Gamma(v)|}$ entonces

$$\langle \Gamma(v), w \rangle = |\Gamma(v)| = \langle v, \Gamma(w) \rangle,$$

de tal forma que:

$$\langle \Gamma(v+w), v+w \rangle = \langle \Gamma(v), v \rangle + 2|\Gamma(v)| + \langle \Gamma(w), w \rangle,$$

$$\langle \Gamma(v-w), v-w \rangle = \langle \Gamma(v), v \rangle - 2|\Gamma(v)| + \langle \Gamma(w), w \rangle;$$

restando obtenemos:

$$\begin{aligned}
 4|\Gamma(v)| &= \langle \Gamma(v+w), v+w \rangle - \langle \Gamma(v-w), v-w \rangle \\
 &\leq M|v+w|^2 + M \cdot |v-w|^2 \\
 &= 2M \cdot (|v|^2 + |w|^2) \\
 &= 4M.
 \end{aligned}$$

Por lo tanto $|\Gamma(v)| \leq M$. □

A continuación concluimos la demostración del teorema espectral probando el siguiente resultado.

Teorema 1.41. *Si Γ es un endomorfismo autoadjunto de un espacio euclidiano V de dimensión finita, entonces $\|\Gamma\|$ ó $-\|\Gamma\|$ es un valor propio de Γ .*

Demostración. Elijamos $v \in V$, con $|v| = 1$ de modo que $|\langle \Gamma v, v \rangle| = \|\Gamma\|$ y pongamos $\alpha = \langle \Gamma v, v \rangle$. Entonces $\alpha^2 = \|\Gamma\|^2$ y se tiene:

$$\begin{aligned}
 0 &\leq |\Gamma v - \alpha v|^2 \\
 &= |\Gamma v|^2 - 2\alpha \langle \Gamma v, v \rangle + \alpha^2 |v|^2 \\
 &= |\Gamma v|^2 - 2\alpha^2 + \alpha^2 \\
 &= |\Gamma v|^2 - \|\Gamma\|^2 \leq 0.
 \end{aligned}$$

En consecuencia $\Gamma v = \alpha v$. □

Si ahora tenemos un espacio vectorial real de dimensión n con una forma bilineal simétrica γ donde inicialmente no existe un producto interno, entonces se puede introducir cualquier producto interno euclidiano mediante un isomorfismo arbitrario $V \cong \mathbb{R}^n$. Por lo tanto, el teorema afirma que existe una base **ortogonal** (v_1, \dots, v_n) para la forma bilineal γ .

Si ponemos $w_j = \lambda_j v_j$ se sigue que $\gamma(w_j) = \lambda_j^2 \gamma(v_j)$ y podemos determinar λ_j de tal manera que $\gamma(w_j) \in \{1, -1, 0\}$.

Teorema 1.42.

1. *Si γ es una forma bilineal simétrica sobre un espacio vectorial real de dimensión n , entonces existe una métrica euclidiana sobre V y una descomposición ortogonal:*

$$V = V_+ \oplus V_- \oplus V_0,$$

donde $\gamma(v) = |v|^2$ para $v \in V_+$, $\gamma(v) = -|v|^2$ para $v \in V_-$ y $\gamma(v) = 0$ para $v \in V_0$.

2. *Si G una matriz real simétrica, entonces existe una matriz invertible T , tal que $T^t G T$ es una matriz diagonal, con entradas 1, -1, 0 en la diagonal.*

Vemos que los valores propios de la matriz bajo la transformación $G \rightarrow T^t G T$ no se preservan.

Teorema 1.43 (De Sylvester). *Los números $\dim V_0$, $\dim V_+$, $\dim V_-$ son igual al índice de γ en el teorema (1.42). Se determinan de forma única por γ , es decir, el número de entradas 1, -1, 0 en la diagonal de T^tGT depende sólo de G , es decir, no depende de la base.*

Demostración. Sea $\Gamma : V \rightarrow V^*$ el operador autoadjunto asociado a γ entonces $V_0 = \ker(\Gamma)$ porque $\gamma(v) = 0$ para todo $v \in V_0$. El número $\dim V_+$ es la dimensión máxima de un subespacio $U \subset V_+$, por lo tanto $\gamma|_U$ es definida positiva. De hecho si $\dim(U) > \dim(V_+)$, entonces $U \cap (V_0 + V_-) \neq \{0\}$, es decir, U tendría un vector $u \neq 0$ tal que $\gamma(u) \leq 0$. \square

El índice de γ se denomina **signatura**. Lo sorprendente en la transformación de los ejes principales mediante una matriz ortogonal es que esta transformación puede verse como una forma bilineal y al mismo tiempo como la transformación de un endomorfismo: $G \mapsto T^tGT = T^{-1}GT$ donde $T \in O(n)$.

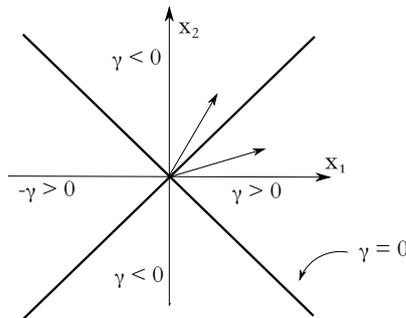


Figura 1.6: Plano hiperbólico.

Ejemplo 1.44 (El plano hiperbólico).

Sea $V = \mathbb{R}^2$ y $\gamma(x, y) = x_1y_1 - x_2y_2$. La matriz asociada a γ es:

$$G = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Una base ortogonal de γ asigna cada par de vectores $(a, b), (b, a)$ con $a \neq b$. Su forma cuadrática asociada es $\gamma(x_1, x_2) = x_1^2 - x_2^2$. La base del teorema espectral determina si los valores propios de Γ son o no todos distintos.

Ejemplo 1.45 (El tensor de inercia).

Consideremos un sólido rígido K al que asignamos un espacio euclidiano $V \cong \mathbb{R}^3$ de tal manera que los puntos en el cuerpo son los vectores en V cuya norma es constante. Cuando se mueve el cuerpo, el espacio \mathbb{R}^3 se mueve con él.

Cuando el cuerpo gira alrededor de un eje fijo que pasa por su centro de masa, la velocidad de rotación está descrita por un vector ω en la dirección del eje de rotación. De tal forma que para un punto $p \in K$ la velocidad está dada por $v = \omega \times q$, ésta se conoce como la **velocidad angular**. La energía cinética de una partícula q de masa m es: $T = \frac{1}{2}m \cdot |\omega \times q|^2$. Las partículas del cuerpo describen circunferencias centradas en el eje de rotación con una

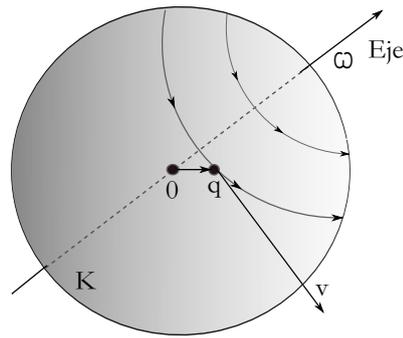


Figura 1.7: Descripción del movimiento de K .

velocidad proporcional al radio de la circunferencia que describen. La energía cinética total es la suma de las energías cinéticas de cada una de las partículas:

$$T = \frac{1}{2} \sum_j m_j |\omega \times q_j|^2.$$

Considerado como un sistema continuo, reemplazamos la suma por una integral sobre el área que ocupa el cuerpo. El paso de las fórmulas que contienen sumas sobre las partículas discretas a las fórmulas del cuerpo continuo se efectúa sustituyendo las masas de las partículas por la densidad de masa $m : K \rightarrow \mathbb{R}$ e integrando para todo el volumen del cuerpo:

$$T(\omega) = \frac{1}{2} \int_K m(q) |\omega \times q|^2 dq.$$

Esta integral como una función del vector ω es una forma cuadrática sobre V correspondiente a la forma bilineal simétrica:

$$\gamma(x, y) = \int_K m(q) \langle x \times q, y \times q \rangle dq,$$

a saber, $T(\omega) = \frac{1}{2} \gamma(\omega, \omega)$.

Ahora bien, si

$$\begin{aligned} \langle x \times q, y \times q \rangle &= \det(x, q, y \times q) \\ &= \langle x, q \times (y \times q) \rangle \end{aligned}$$

entonces podemos describirla como:

$$\begin{aligned} \gamma(x, y) &= \int_K m(q) \langle x, q \times (y \times q) \rangle dq \\ &= \langle x, \int_K m(q) q \times (y \times q) dq \rangle. \end{aligned}$$

Entonces γ es la forma bilineal del endomorfismo autoadjunto

$$\Gamma : V \rightarrow V$$

$$y \mapsto \int_K m(q) \cdot q \times (y \times q) dq.$$

Este endomorfismo se conoce como el **tensor de inercia** del sólido rígido (para el punto fijo $0 \in V$). Siempre se puede elegir un sistema de vectores ortonormales como en el teorema espectral, con los cuales el tensor de inercia tiene forma diagonal. Donde los ejes de coordenadas son los **ejes principales** del cuerpo y los tres valores propios correspondientes I_1, I_2, I_3 son los **momentos principales de inercia**.

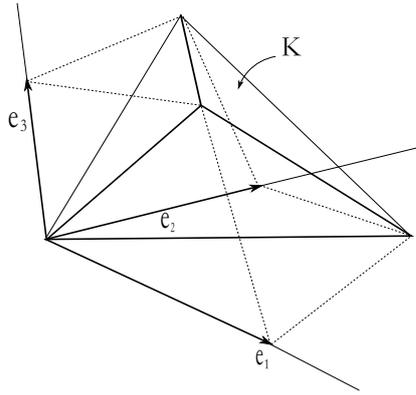


Figura 1.8: $K \subset \mathbb{R}^3$

Para estas coordenadas, se tiene: $T(\omega) = \frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2)$. Si la energía cinética se quiere constante, digamos si $T(\omega) = \frac{1}{2}$ se debe satisfacer la ecuación:

$$I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2 = 1.$$

Este es el llamado **elipsoide de inercia** del sólido rígido en el origen. El cual se obtiene a partir de la esfera unitaria $\omega_1^2 + \omega_2^2 + \omega_3^2 = 1$ multiplicando por un factor $\sqrt{I_j}$ para cada ω_j . La forma cuadrática es definida positiva, por lo tanto $I_j \geq 0$ por ello podemos tomar la raíz cuadrada. Este elipsoide permite visualizar las propiedades del tensor, es decir, determina la posición de los ejes principales, los valores propios de Γ , la función $T(\omega)$ y en consecuencia el momento del cuerpo al rotar. El elipsoide está sujeto a una restricción más: la suma de dos valores propios es mayor igual al tercero $I_1 + I_2 \geq I_3$. En otras palabras, la terna $(I_1, I_2, I_3) \in \mathbb{R}^3$ de los tensores de inercia forman el cono $K \subset \mathbb{R}^3$ de coordenadas (x, y, z) tales que $x, y, z \geq 0$ y $x + y \geq z, y + z \geq x, z + x \geq y$.

1.5. Espacios unitarios

Definición 1.46. Sea V un espacio vectorial complejo. Una **forma hermitiana** en V es una función:

$$\gamma : V \times V \rightarrow \mathbb{C}$$

que satisface:

1. $\gamma(\lambda u + \mu v, w) = \lambda \cdot \gamma(u, w) + \mu \cdot \gamma(v, w),$

$$2. \gamma(u, v) = \overline{\gamma(v, u)}$$

para todo $\lambda, \mu \in \mathbb{C}$, $u, v, w \in V$.

Por lo tanto, se tiene $\gamma(w, \lambda u + \mu v) = \bar{\lambda} \cdot \gamma(w, u) + \bar{\mu} \cdot \gamma(w, v)$, es decir, γ es lineal en la primera variable y **antilineal** en la segunda.

La forma hermitiana canónica en \mathbb{C}^n , es el producto interno usual dado por:

$$(x, y) \mapsto \langle x, y \rangle := x^t \bar{y} = \sum_{j=1}^n x_j \bar{y}_j \quad (1.5.1)$$

En este caso, (1) es claro. Para verificar (2) observemos que: $\overline{(x^t \bar{y})} = \bar{x}^t y = y^t \bar{x}$.

Observemos que $\gamma(v, v) = \overline{\gamma(v, v)}$ por lo que $\gamma(v, v)$ es un número real para todo $v \in V$.

Definición 1.47. Una forma hermitiana γ se dice **definida positiva o producto interno unitario** si $\gamma(v, v) > 0$ para todo $0 \neq v \in V$.

El producto interno usual en \mathbb{C}^n es definido positivo porque:

$$\langle x, x \rangle = x^t \bar{x} = \sum_j x_j \bar{x}_j = \sum_j |x_j|^2 > 0.$$

Definición 1.48. Un **espacio unitario** es un espacio vectorial complejo con un producto interno unitario $\langle \cdot, \cdot \rangle$.

Ejemplo 1.49. \mathbb{C}^n con el producto interno usual es un espacio unitario.

Ejemplo 1.50. El espacio $V = C^0([0, 1], \mathbb{C})$ de las funciones continuas $f : [0, 1] \rightarrow \mathbb{C}$, con el producto interno dado por

$$\langle f, g \rangle := \int_0^1 f(t) \bar{g}(t) dt,$$

es un espacio unitario.

A continuación procedemos de manera análoga al caso de los espacios euclidianos:

Definición 1.51. Sea V un espacio unitario. Si $v, w \in V$ definimos

$$1. \text{ La norma de } v: |v| = \sqrt{\langle v, v \rangle}.$$

$$2. \text{ Decimos que } v \text{ y } w \text{ son } \mathbf{ortogonales} \text{ si } \langle v, w \rangle = 0. \text{ En tal caso escribimos } v \perp w.$$

Proposición 1.52 (Desigualdad de Cauchy-Schwarz).

$$|\langle u, v \rangle| \leq |u| \cdot |v|$$

y la igualdad se cumple si y sólo si u y v son linealmente dependientes.

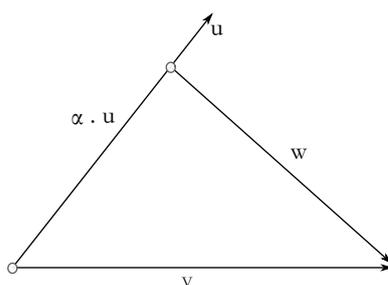


Figura 1.9: Desigualdad de Cauchy-Schwarz.

Demostración. Sean $u, v \in V$ y $u \neq 0$ entonces existe una descomposición $v = \alpha \cdot u + w$ con $\langle w, u \rangle = 0$ y $\alpha = \frac{\langle v, u \rangle}{|u|^2}$, por lo tanto $|v|^2 = |\alpha|^2|u|^2 + |w|^2 \geq |\alpha|^2|u|^2$. En consecuencia

$$|u| \cdot |v| \geq |\alpha| \cdot |u|^2 = |\langle v, u \rangle|.$$

□

Proposición 1.53 (Desigualdad del triángulo).

$$|u + v| \leq |u| + |v|.$$

Una **base ortonormal** de V es una base (v_1, \dots, v_n) tal que para todo $j, k = 1, \dots, n$ se cumple:

$$\langle v_j, v_k \rangle = \delta_{jk}.$$

Al igual que en el caso de los espacios euclidianos, se puede ver que todo espacio unitario de dimensión finita posee una base ortonormal, ya que dada una base arbitraria, siempre se puede aplicar el proceso de Gram-Schmidt. Con respecto a una tal base ortonormal, todo vector $u \in V$ tiene la siguiente expresión en coordenadas:

$$u = \sum_j \langle u, v_j \rangle \cdot v_j$$

ya que al hacer el producto interno de la combinación lineal $u = \lambda_1 v_1 + \dots + \lambda_n v_n$ con v_k se obtiene: $\langle u, v_k \rangle = \lambda_k$.

Definición 1.54. Una transformación lineal entre espacios unitarios $\alpha : V \rightarrow W$ se dice **unitaria** si

$$\langle \alpha(v), \alpha(w) \rangle_W = \langle v, w \rangle_V \quad \forall v, w \in V.$$

Notemos que α es inyectiva porque si $v \neq 0$ entonces $\langle \alpha(v), \alpha(v) \rangle = \langle v, v \rangle \neq 0$. Por lo tanto $\alpha(v) \neq 0$. Más aún, si las dimensiones de los espacios son iguales, entonces α es un isomorfismo. Por otro lado, la existencia de bases ortonormales implica que si V es un espacio unitario de dimensión n , entonces existe un isomorfismo unitario $V \cong \mathbb{C}^n$. Durante el resto de esta sección, los automorfismos unitarios serán de especial interés para nosotros.

Definición 1.55. El grupo unitario $U(n)$ es el grupo de automorfismos unitarios de \mathbb{C}^n y sus elementos son las matrices unitarias $n \times n$. El subgrupo $SU(n)$ de matrices unitarias con determinante 1 se denomina el grupo especial unitario.

Observación 1.56. Una matriz A $n \times n$ compleja se dice unitaria si y sólo si:

$$A^* := \overline{A}^t = A^{-1},$$

esto es, si las columnas de A forman una base ortonormal. Si este es el caso, entonces las matrices \overline{A}, A^t, A^* también son unitarias y por lo tanto las filas de A también son ortonormales y $|\det(A)| = 1$.

Demostración.

$$\begin{aligned} x^t \overline{y} = \langle x, y \rangle &= \langle Ax, Ay \rangle \\ &= (Ax)^t (\overline{Ay}) \\ &= x^t A^t \overline{Ay} \\ \iff A^t \overline{A} &= I \text{ (poniendo } x = e_j, y = e_k) \\ \iff I &= \overline{I} = A^* A, \end{aligned}$$

lo que muestra la primera afirmación. Ahora, si $\overline{A}^t A = I$, entonces al transponer tenemos $A^t \overline{A} = I$, de donde se sigue que $(\overline{A})^* \overline{A} = I$. Del mismo modo se tiene que $I = \overline{A} A^t = (A^t)^* A^t$. En consecuencia \overline{A} y A^t son también unitarias. La ecuación $A^t \overline{A} = I$ indica que la columnas forman una familia ortonormal. Finalmente

$$\begin{aligned} |\det(A)|^2 &= \overline{\det(A)} \cdot \det(A) \\ &= \overline{\det(A^t)} \cdot \det(A) \\ &= \det(A^*) \cdot \det(A) \\ &= \det(A^* A) \\ &= \det(I) \\ &= 1. \end{aligned}$$

□

Ejemplo 1.57. En dimensiones bajas tenemos:

1. $U(1) = \{a \in \mathbb{C}^n \mid |a| = 1\} = S^1$.
2. $SU(1) = \{1\}$.
3. $SU(2) = \left\{ \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix} \mid |a|^2 + |b|^2 = 1 \right\}$.

Una matriz $A \in SU(2)$ está entonces determinada unívocamente por su primera columna, la que a su vez es un vector arbitrario de norma 1 en \mathbb{C}^2 . Por lo tanto, al pensar a la 3-esfera S^3 como un subconjunto de $\mathbb{C}^2 = \mathbb{R}^4$, se tiene una biyección:

$$S^3 \rightarrow SU(2)$$

$$\begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix}.$$

El grupo $SU(2)$ también se conoce como **el grupo de los cuaterniones** o como **Spin(3)**. El conjunto de todas las matrices complejas 2×2 de la forma

$$\begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix}$$

constituyen el espacio \mathbb{H} de los cuaterniones. Así, tenemos:

$$\mathbb{H} = \{r \cdot A \mid r \in \mathbb{R}, A \in SU(2)\} =: \mathbb{R} \cdot SU(2).$$

Como espacio vectorial $\mathbb{H} \cong \mathbb{C}^2$ ó $\mathbb{H} \cong \mathbb{R}^4$, sin embargo \mathbb{H} es también un álgebra sobre \mathbb{R} con la suma y multiplicación de matrices y todo elemento $B \in \mathbb{H}$, con $B \neq 0$, tiene inverso multiplicativo. A saber, si $B = r \cdot A$, con $r \in \mathbb{R}$ y $A \in SU(2)$, entonces

$$B^{-1} = r^{-1}A^* = r^{-2}B^*.$$

La multiplicación en \mathbb{H} no es conmutativa, sin embargo se cumplen todos los demás axiomas de campo. Por lo tanto, los cuaterniones forman un **anillo de división**. El álgebra lineal (salvo la teoría de determinantes) se puede desarrollar también para anillos de división, pero se debe distinguir bien entre multiplicar por la derecha y por la izquierda. En el capítulo IV hablaremos más al respecto.

Definición 1.58. Sea V un espacio unitario, un endomorfismo $\Gamma : V \rightarrow V$ se dice **hermitano** o **autodjunto**, si la función:

$$\gamma : V \times V \rightarrow \mathbb{C}$$

$$(v, w) \mapsto \langle v, \Gamma w \rangle.$$

es una forma hermitiana, es decir, sí y sólo si $\langle v, \Gamma w \rangle = \langle \Gamma v, w \rangle$ para todo $v, w \in V$.

Si (v_1, \dots, v_n) es una base ortonormal de V y $G = (g_{jk})$ es la matriz de Γ , entonces $\Gamma v_k = \sum_j g_{jk} v_j$. Por lo tanto:

$$\langle v_s, \Gamma v_k \rangle = \sum_j \bar{g}_{jk} \langle v_s, v_j \rangle = \bar{g}_{sk}$$

$$\langle \Gamma v_s, v_k \rangle = \sum_j g_{js} \langle v_j, v_k \rangle = g_{ks}.$$

En consecuencia $G^* = G$, es decir G es una matriz **hermitiana**.

Teorema 1.59 (Teorema espectral).

1. Sea $\Gamma : V \rightarrow V$ un endomorfismo autoadjunto de un espacio unitario V de dimensión finita. Entonces los valores propios de Γ son reales y V posee una base ortonormal de vectores propios.

2. Si G una matriz $n \times n$ compleja hermitiana, entonces existe una matriz unitaria $T \in U(n)$ tal que $T^{-1}GT$ es una matriz diagonal real.

Procederemos como en el caso real.

Observación 1.60. Los valores propios de Γ son reales.

Demostración. En efecto, si $\Gamma v = \lambda v$ con $v \neq 0$, entonces

$$\begin{aligned} \lambda \cdot \langle v, v \rangle &= \langle \lambda v, v \rangle \\ &= \langle \Gamma v, v \rangle \\ &= \langle v, \Gamma v \rangle \\ &= \langle v, \lambda v \rangle \\ &= \bar{\lambda} \langle v, v \rangle. \end{aligned}$$

Por lo tanto $\lambda = \bar{\lambda} \in \mathbb{R}$. □

Observación 1.61. Si $U \subset V$ es un subespacio y $\Gamma(U) \subset U$, entonces

$$\Gamma(U^\perp) \subset U^\perp.$$

Demostración. Si $v \in U^\perp$, entonces $\langle v, u \rangle = 0$ para todo $u \in U$. Por lo tanto

$$\langle \Gamma v, u \rangle = \langle v, \Gamma u \rangle = 0 \quad \forall u \in U,$$

en consecuencia $\Gamma v \in U^\perp$. □

Para concluir la demostración del teorema espectral solo resta mostrar que Γ posee al menos un valor propio, es decir, que el polinomio característico de Γ posee al menos una raíz. Pero esto se sigue del teorema fundamental del álgebra. Al igual que en el caso real:

Corolario 1.62. Si Γ es un operador autoadjunto en un espacio unitario de dimensión finita, entonces:

1. $\|\Gamma\| = \max\{|\Gamma(v)| \mid |v| = 1\}$
2. $\|\Gamma\|$ ó $-\|\Gamma\|$ es un valor propio de Γ .

Demostración. Por el teorema espectral sabemos que existe una base ortonormal de vectores propios, (v_1, \dots, v_n) con respecto a la cual Γ tiene una representación como matriz diagonal:

$$\Gamma = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Afirmamos que $\|\Gamma\| = \max\{|\lambda_j| \mid j = 1, \dots, n\}$. En efecto, si $v = (z_1, \dots, z_n)$ y $|v| = 1$, entonces $\|\Gamma(v)\| = \|\lambda_1|z_1|^2 + \dots + \lambda_n|z_n|^2\| \leq \max|\lambda_j|(|z_1|^2 + \dots + |z_n|^2) = \max|\lambda_j|$. Luego $\|\Gamma(v)\| \leq \max\{|\lambda_j|\}$. Por otro lado, notemos que para todo $j = 1, \dots, n$:

$$|\lambda_j| = |\Gamma(v_j)| \leq \|\Gamma\|$$

de donde se sigue que $\max|\lambda_j| \leq \|\Gamma\|$. En consecuencia la afirmación implica inciso (2). □

Proposición 1.63. *Las formas hermitianas sobre \mathbb{C}^n satisfacen:*

$$(x, y) = x^t G \bar{y}, \quad G = G^*$$

y la forma es definida positiva si y sólo si $G = T^ T$ para alguna $T \in GL(n, \mathbb{C})$.*

Demostración. Si T es la transformación de una base ortonormal tras la ortogonalización de Gram - Schmidt, entonces

$$x^t G y = (Tx)^t (\overline{Ty}) = x^t (T^t \overline{T}) \bar{y},$$

por lo tanto $G = T^t \overline{T} = \overline{T^*} \overline{T}$ para toda $\overline{T} \in GL(n, \mathbb{C})$ □

Se puede deducir fácilmente el teorema espectral real a partir del complejo: si G es una matriz real simétrica, entonces G es una matriz hermitiana cuyos coeficientes son todos reales. Se sigue del teorema espectral complejo que el polinomio característico de G tiene raíces reales puras. Dicho de otra manera: Si V es un espacio unitario, entonces V es ciertamente un espacio vectorial real por la inclusión $\mathbb{R} \subset \mathbb{C}$ y también por que la métrica

$$\langle v, w \rangle_{\mathbb{R}} := \operatorname{Re} \langle v, w \rangle_{\mathbb{C}}, \tag{1.5.2}$$

por la derecha es unitaria y por la izquierda es un producto interno euclidiano. Es claro que es una métrica real, bilineal y definida positiva. Entonces de acuerdo al teorema espectral real, existe una base ortonormal de vectores propios reales correspondientes a valores propios de G . Sin embargo como los espacios propios son subespacios complejos y son espacios propios con distintos valores propios ortogonales, es fácil encontrar la base ortonormal compleja buscada de vectores propios de G .

Capítulo 2

Superficies cuádricas

2.1. Superficies de segundo orden

Consideremos un campo K de característica distinta de dos, es decir $1 + 1 \neq 0$ en K y $f : K^n \rightarrow K$ una función cuadrática:

$$f(x) = \sum_{i,j} g_{ij}x_i x_j + 2 \sum_i a_i x_i + b =: x^t G x + 2a^t x + b. \quad (2.1.1)$$

Sin pérdida de generalidad podemos suponer que G es una matriz **simétrica**.

Definición 2.1. El conjunto $M = \{x \in K^n \mid f(x) = 0\}$ determinado por f se denomina **cuádrica afín** o **superficie de segundo orden**.

Definición 2.2. Un punto $x \in M$ para el cual $x^t G + a^t \neq 0$ se denomina **punto regular** o **suave**, a los puntos que no sean regulares les llamaremos **puntos singulares**.

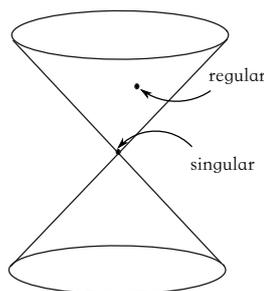


Figura 2.1: Puntos en una superficie.

Ejemplo 2.3.

1. Si $a = 0$ y $b \neq 0$ entonces sólo existen puntos regulares. En particular la esfera $S^{n-1} \subset \mathbb{R}^n$ tiene ecuación $x^t I x - 1 = 0$ y no tiene puntos singulares.
2. El cono $|x|^2 - z^2 = 0$ en \mathbb{R}^3 , tiene al punto $x = 0$ como punto singular.

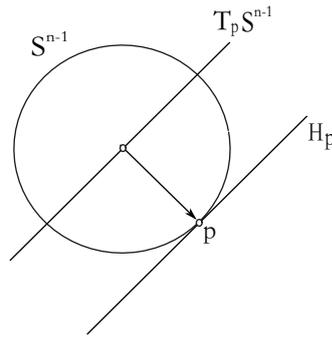


Figura 2.2: Hiperplano afín y espacio tangente en $p \in S^{n-1}$.

Definición 2.4. Sea $p \in M$ un punto regular. Se define el **espacio tangente** a M en el punto p como el hiperplano afín $H_p = \{x \mid p^t Gx + a^t p + a^t x + b = 0\}$ y el espacio vectorial asociado $T_p M = \{x \mid p^t Gx + a^t x = 0\}$ se conoce como el **espacio vectorial tangente** de M en el punto p .

Ejemplo 2.5. Si $M = S^{n-1} = \{x \in \mathbb{R}^n \mid \langle x, x \rangle - 1 = 0\}$ entonces:

$$H_p = \{x \mid \langle p, x \rangle - 1 = 0\} \quad y \quad T_p S^{n-1} = \{x \mid \langle p, x \rangle = 0\}.$$

La definición de H_p y $T_p M$ se justifica como sigue: sean $p, p+h \in M$. Entonces $p^t Gp + 2a^t p + b = 0$ y $(p+h)^t G(p+h) + 2a^t(p+h) + b = 0$, restando ambas ecuaciones se tiene: $p^t Gh + h^t Gp + h^t Gh + 2a^t h = 0$.

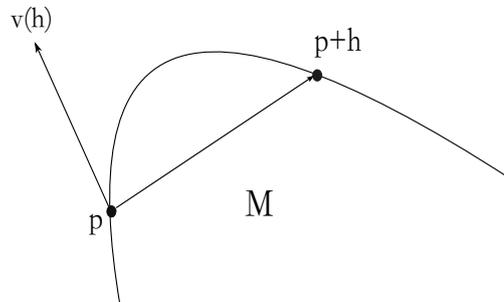


Figura 2.3

Pongamos $v(h) = 2(p^t G + a^t) + h^t G$ de modo que $v(h) \cdot h = 0$. Pensando a $v(h)$ como vector columna tenemos $\langle v(h), h \rangle = 0$ es decir, h está en el complemento ortogonal del vector $v(h)$. Tomando el límite cuando $h \rightarrow 0$, obtenemos un vector ortogonal al espacio tangente afín H_p

$$v(0) = 2(p^t G + a^t).$$

Por lo tanto el espacio vectorial tangente $T_p M$ consiste de todos los vectores ortogonales a $v(0)$. Finalmente la ecuación para H_p se obtiene observando que $H_p = p + T_p M$, porque $p \in H_p$. La ecuación de H_p es fácil de recordar, sustituímos en el término cuadrático $x^t Gx$ de la ecuación (2.1.1) el primer factor x por p y en el término lineal $a^t x + a^t x$, en un sumando sustituímos x por p .

Ahora nos preguntamos, ¿Cómo se ve una superficie de segundo orden?. Asumiremos que $K = \mathbb{R}$, sin embargo los cálculos son similares para cualquier campo de característica distinta de 2. Una pregunta natural es, ¿cómo se ve $M \subset \mathbb{R}^n$ bajo un movimiento rígido?. Esto es, una función de la forma:

$$\begin{aligned}\mathbb{R}^n &\rightarrow \mathbb{R}^n \\ x &\mapsto Ax + v\end{aligned}$$

donde $A \in O(n)$ y $v \in \mathbb{R}^n$. Si componemos una transformación ortogonal $T_1 : x \mapsto T_1x$ que diagonalice a x^tGx , obtenemos:

$$fT_1(x) = \sum_{i=1}^k \lambda_i x_i^2 + 2\langle a, x \rangle + b$$

con $\lambda_i \neq 0$ para $i = 1, \dots, k$. (Obviamente con nuevas constantes a, b). Si ponemos ahora $T_2x = x + d$ se obtiene:

$$fT_1T_2(x) = \sum_{i=1}^k \lambda_i x_i^2 + 2 \sum_{i=1}^k (a_i + \lambda_i d_i) x_i + 2 \sum_{i=k+1}^n a_i x_i + b,$$

(de nuevo con una nueva constante b) y podemos escoger d de tal modo que $a_i + \lambda_i d_i = 0$ para $i = 1, \dots, k$, puesto que $\lambda_i \neq 0$. Así tenemos:

$$fT_1T_2(x) = \sum_{i=1}^k \lambda_i x_i^2 + 2 \sum_{i=k+1}^n a_i x_i + b.$$

Caso 1: Si en la última ecuación se tiene $a_{k+1} = \dots = a_n = 0$ podemos multiplicar por una constante distinta de cero para obtener una ecuación de la forma (con nuevas λ_i):

$$\sum_{i=1}^k \lambda_i x_i^2 = 0 \quad \text{ó} \quad \sum_{i=1}^k \lambda_i x_i^2 = 1.$$

Caso 2: Consideremos $L(e_{k+1}, \dots, e_n)$ el espacio generado por la $n - k$ componente de un múltiplo del primer vector $(0, \dots, 0, a_{k+1}, \dots, a_n)^t$ como una base ortonormal. Entonces la expresión anterior se transforma en:

$$\sum_{i=1}^k \lambda_i x_i^2 + 2cx_{k+1} + d = 0.$$

Si sustituímos x_{k+1} por $x_{k+1} - \frac{d}{2c}$ y dividimos por $-c$:

$$\sum_{i=1}^k \lambda_i x_i^2 = 2x_{k+1}.$$

Resumiendo, tenemos el siguiente teorema.

Teorema 2.6. *Bajo un movimiento rígido de \mathbb{R}^n , la ecuación (2.1.1) multiplicada por un factor distinto de cero se transforma en alguna de las expresiones:*

$$\sum_{i=1}^k \lambda_i x_i^2 = 0, \quad \sum_{i=1}^k \lambda_i x_i^2 = 1, \quad \sum_{i=1}^k \lambda_i x_i^2 = 2x_{k+1}$$

con $k \geq 0$. En particular bajo un movimiento rígido, la cuádrica $M = \{x \mid f(x) = 0\}$ se convierte en cualquiera de las expresiones anteriores.

Si ponemos $x_i \mapsto \sqrt{|\lambda_i|}x_i$ las ecuaciones anteriores se transforman en:

$$\begin{aligned} |x|^2 - |y|^2 &= 0, \\ |x|^2 - |y|^2 &= 1, \\ |x|^2 - |y|^2 &= 2z. \end{aligned} \tag{2.1.2}$$

con $x \in \mathbb{R}^r$, $y \in \mathbb{R}^s$, $z \in \mathbb{R}$. Se tiene la componente λ_i positiva para el vector x y negativa para el vector y . La función no depende de las variables en (2.1.2), si M es el conjunto de ceros y conocemos $M \cap \mathbb{R}^{r+s}$ entonces

$$M = (M \cap \mathbb{R}^{r+s}) \times \mathbb{R}^{n-r-s}.$$

Consideremos los resultados que figuran en el plano \mathbb{R}^2 , los casos degenerados:

1. $0 = 1$ ó $-|y|^2 = 1$, el conjunto vacío.
2. $|x|^2 = 0$, el punto.
3. $z = 0$ ó $x^2 = 0$, la recta.
4. $0 = 0$, el plano

De acuerdo a este modelo, se pueden obtener todos los espacios afín de dimensión mayor. La ecuación $x^2 = 1$ define **dos rectas paralelas** $x = \pm 1$.

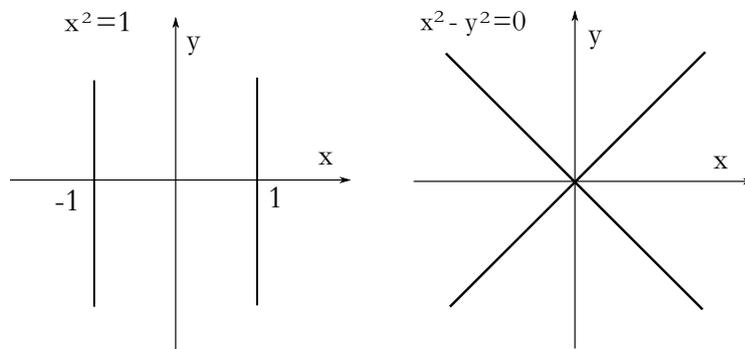


Figura 2.4: Rectas.

La ecuación $x^2 - y^2 = (x + y)(x - y) = 0$ define **dos rectas que se cruzan** $y = \pm x$.

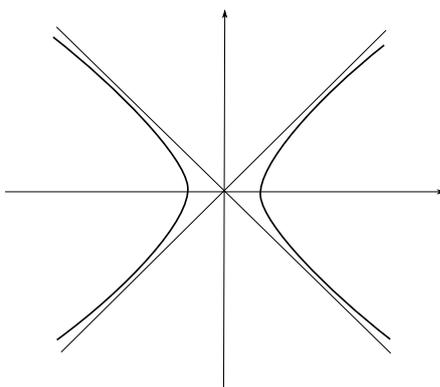


Figura 2.5: Hipérbola.

Conocemos la gráfica de la hipérbola $y = \frac{1}{x}$ con ejes $y = 0, x = 0$, por lo tanto si sustituimos $(x + y)$ para y y $(x - y)$ para x , obtenemos la **hipérbola** $(x + y)(x - y) = x^2 - y^2 = 1$ con asíntota $y = \pm x$.

La ecuación $|x|^2 = 1, x \in \mathbb{R}^2$ describe un **círculo**. Si aplicamos la transformación

$$x_1 \mapsto \lambda x_1$$

$$x_2 \mapsto \mu x_2$$

obtenemos la **elipse** $\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$. En el siguiente dibujo los ejes coordenados son los ejes principales y $2a, 2b$ son las longitudes de los ejes mayor y menor.

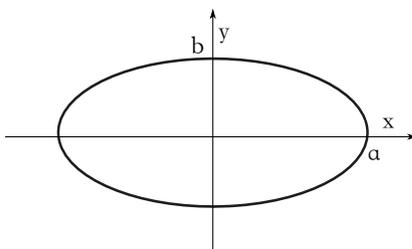


Figura 2.6: Elipse.

La parábola tiene como ecuación: $2z = x^2$.

En dimensiones más altas, en las ecuaciones (2.1.2), consideremos la norma de x , de y ó ambas según corresponda. Esto significa que el conjunto M se forma a partir de una de las curvas planas por la rotación del espacio de coordenadas (ponemos $|x|$ en lugar de x). Por ejemplo, la ecuación $|x|^2 = 1$ en \mathbb{R}^3 con coordenadas x_1, x_2, y , corresponde a la ecuación $x^2 = 1$. El conjunto correspondiente surge de las dos rectas por rotación del eje y : la ecuación $|x|^2 = 1$ corresponde al **cilindro** $S^1 \times \mathbb{R}$.

Del mismo modo, a partir de las dos rectas $x^2 - y^2 = 0$ obtenemos el **cono**, con ecuación $|x|^2 - y^2 = 0$.

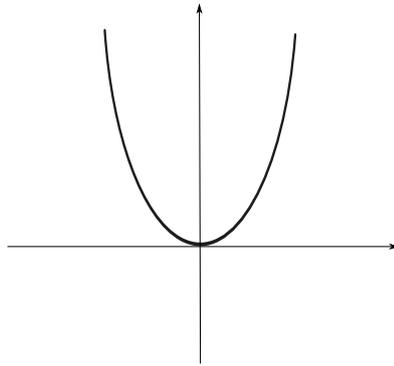
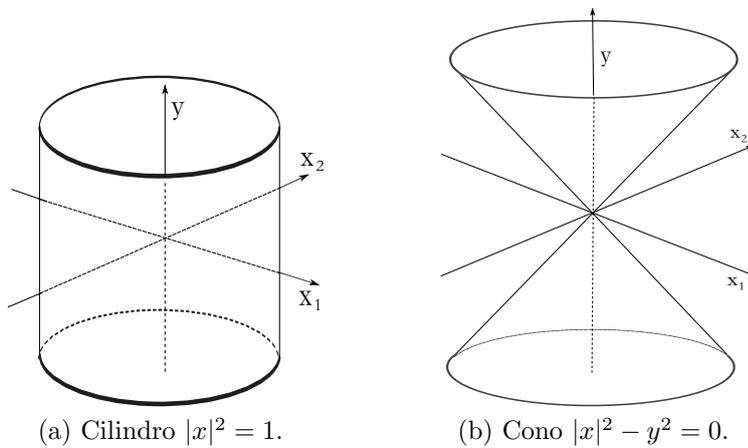


Figura 2.7: Gráfica de la parábola $z = x^2$.



(a) Cilindro $|x|^2 = 1$.

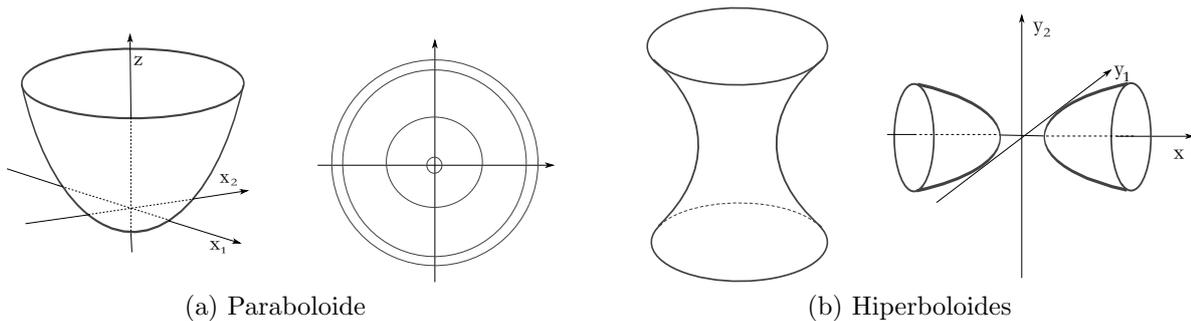
(b) Cono $|x|^2 - y^2 = 0$.

Figura 2.8: Rotación de las rectas.

A partir de la ecuación de la parábola $z = x^2$, se obtiene la ecuación $z = |x|^2$ y se denomina **paraboloide de revolución**. Las curvas de nivel de la función $z = |x|^2$ están dadas por la ecuación $z = c$. Tales curvas forman un sistema concéntrico de círculos.

En el caso de la ecuación de la hipérbola $x^2 - y^2 = 1$, hay dos tipos de superficies en \mathbb{R}^3 , dependiendo de si se utiliza la norma del vector x ó y : $|x|^2 - y^2 = 1$, es el **hiperboloide de una hoja** como resultado de la rotación del eje y . Y $x^2 - |y|^2 = 1$ es el **hiperboloide de dos hojas** resultado de la rotación del eje x .

La ecuación $|x|^2 = 1 \in \mathbb{R}^3$ describe una **esfera** y como antes, al aplicar la transformación



(a) Paraboloide

(b) Hiperboloides

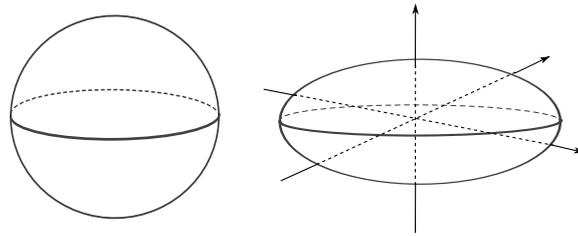


Figura 2.9: Esfera y elipsoide.

$x_j \mapsto \frac{x_j}{a}$, obtenemos la familia de **elipsoides**.

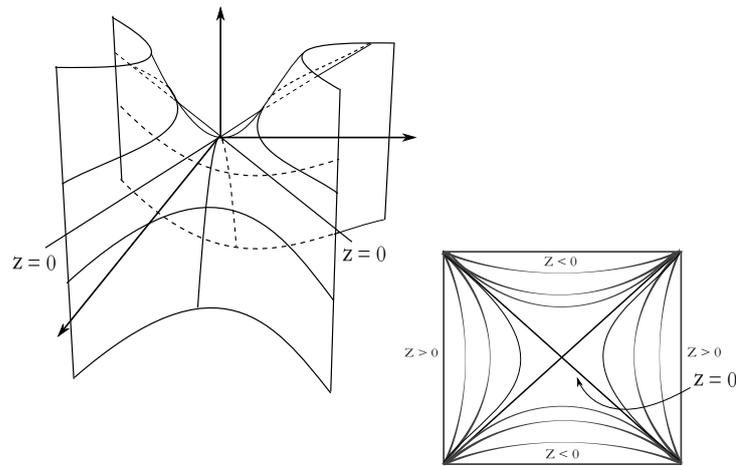


Figura 2.10: Superficie silla de montar y sus curvas de nivel.

Existe también en \mathbb{R}^3 una nueva ecuación: $z = x^2 - y^2$ es la superficie denominada **silla de montar**.

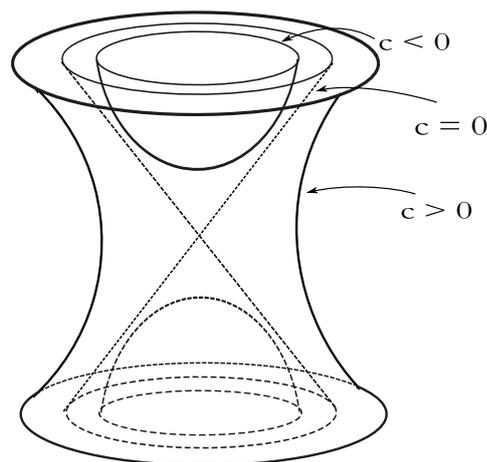


Figura 2.11: Familia de superficies.

Por último tenemos las superficies menos notables: dos planos $x^2 = 1$, el cilindro sobre otras curvas de subespacios afines y el conjunto vacío. Finalmente la figura (2.11) muestra la familia de las superficies $|x|^2 - y^2 = c$, con $c \in \mathbb{R}$: el hiperboloide de una hoja $c > 0$, sobre el cono $c = 0$ y el hiperboloide de dos hojas $c < 0$.

2.2. Secciones cónicas y superficies regladas.

Consideremos el campo de los números reales \mathbb{R} . A las curvas de nivel de segundo orden también se les conoce como secciones cónicas, como veremos a continuación.

Definición 2.7. Se denomina **parábola** al lugar geométrico de los puntos del plano que distan de un punto fijo F llamado **foco** y de una recta fija L llamada **directriz**.

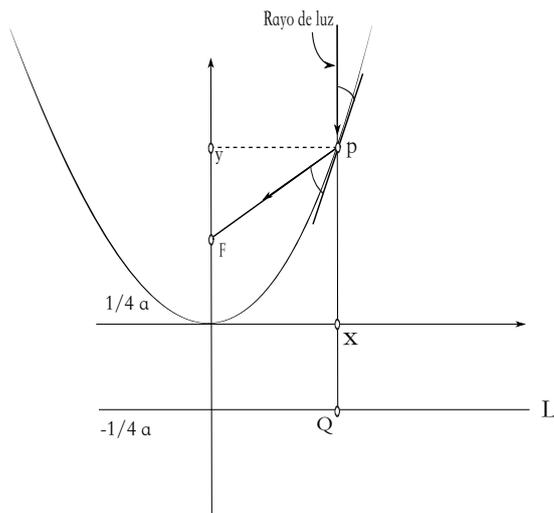


Figura 2.12: Rayo de luz en una parábola.

Proposición 2.8. La forma normal de la parábola está dada por la ecuación $y = ax^2$. Con foco $F = \begin{pmatrix} 0 \\ \frac{1}{4a} \end{pmatrix}$ y la directriz es el polar de los focos $y = -\frac{1}{4a}$.

En efecto, si $p = \begin{pmatrix} x \\ y \end{pmatrix}$ entonces la distancia al punto focal es $(x^2 + (y - \frac{1}{4a})^2)^{\frac{1}{2}}$ y la distancia a la directriz es $y + \frac{1}{4a}$. Haciendo $y = ax^2$, obtenemos $(ax^2 + \frac{1}{4a})^2 = x^2 + (ax^2 - \frac{1}{4a})^2$. El vector $v = (p - F) + (p - Q) = 2p + F - Q$ está en dirección del ángulo bisector FpQ y

$$\begin{aligned} v &= 2 \begin{pmatrix} x \\ ax^2 \end{pmatrix} - \begin{pmatrix} 0 \\ \frac{1}{4a} \end{pmatrix} - \begin{pmatrix} x \\ -\frac{1}{4a} \end{pmatrix} \\ &= \begin{pmatrix} x \\ 2ax^2 \end{pmatrix} = x \begin{pmatrix} 1 \\ 2ax \end{pmatrix}. \end{aligned}$$

Pero la parábola tiene pendiente $2ax$ en p ó en nuestro lenguaje: $\begin{pmatrix} 1 \\ 2ax \end{pmatrix}$ se encuentra en el espacio tangente en p a la parábola, dicho bisector es tangente a la parábola en p . La figura nos muestra que los rayos de luz que chocan paralelos al eje de la parábola se reflejan en el punto focal.

Definición 2.9. Se denomina **elipse** al lugar geométrico de los puntos del plano cuya suma de distancia a otros dos fijos llamados focos es constante.

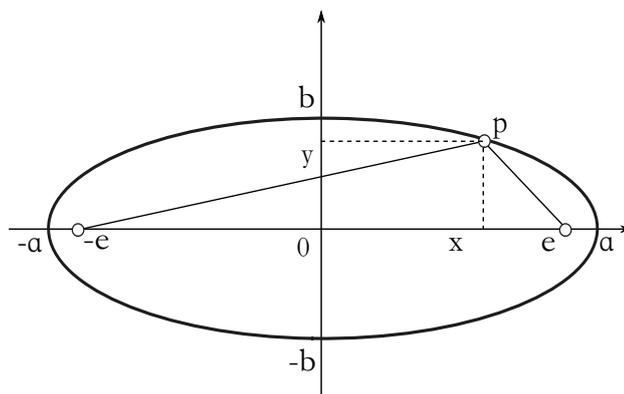


Figura 2.13: Elipse.

Proposición 2.10. La forma normal de la elipse está dada por la ecuación $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ con $a \geq b$. $2a$ es la longitud del **eje mayor**, $2e = 2\sqrt{a^2 - b^2}$ es la **excentricidad** y los puntos

$$F_1 = \begin{pmatrix} -e \\ 0 \end{pmatrix}, \quad F_2 = \begin{pmatrix} e \\ 0 \end{pmatrix}$$

son los **focos**.

La tangente en el punto $p = (p_1, p_2)^t$ satisface la ecuación

$$\frac{p_1 x}{a^2} + \frac{p_2 y}{b^2} = 1.$$

Los vectores tangentes son ortogonales al vector $\begin{pmatrix} p_1/a^2 \\ p_2/b^2 \end{pmatrix}$. De acuerdo a la definición, la elipse se calcula como la suma de las distancias de un punto $p = \begin{pmatrix} x \\ y \end{pmatrix}$ a los focos cuya constante es $2a$, por lo tanto

$$((x + e)^2 + y^2)^{\frac{1}{2}} = 2a - ((x - e)^2 + y^2)^{\frac{1}{2}}$$

en consecuencia

$$a((x - e)^2 + y^2)^{\frac{1}{2}} = a^2 - ex$$

y dividiendo por $a^2(a^2 - e^2)$ obtenemos:

$$\frac{x^2}{a^2} + \frac{y^2}{a^2 - e^2} = 1$$

donde $b^2 = a^2 - e^2$.

De la figura (2.13), observamos que un rayo que pasa sobre la elipse en un punto p al foco F_1 se refleja hasta el otro foco F_2 en otras palabras, la bisectriz del ángulo $F_1 p F_2$ es perpendicular a la tangente de la elipse en el punto p .

Definición 2.11. Se denomina **hipérbola** al lugar geométrico de los puntos del plano cuya diferencia de las distancias a otros dos fijos llamados focos, es constante.

Proposición 2.12. La ecuación de la hipérbola en su forma normal es $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$, la **excentricidad** es $2e = 2\sqrt{a^2 + b^2}$, los **focos** están en $F_1 = \begin{pmatrix} -e \\ 0 \end{pmatrix}$, $F_2 = \begin{pmatrix} e \\ 0 \end{pmatrix}$.

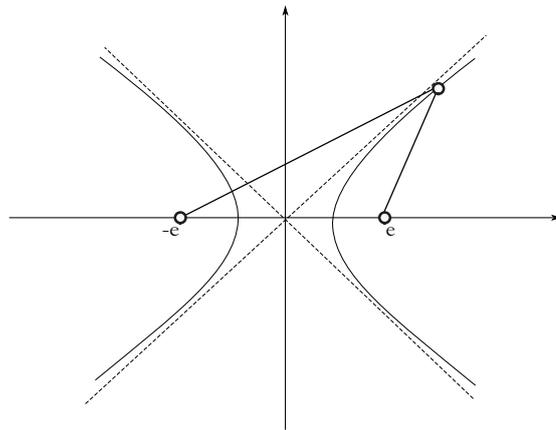


Figura 2.14: Hipérbola.

De manera análoga a la elipse, se muestra que la hipérbola es el lugar geométrico cuya diferencia de las distancias a los focos es la constante $\pm 2a$.

Definición 2.13. Una superficie reglada es una superficie S que para cada punto p en la superficie existe una línea contenida en S .

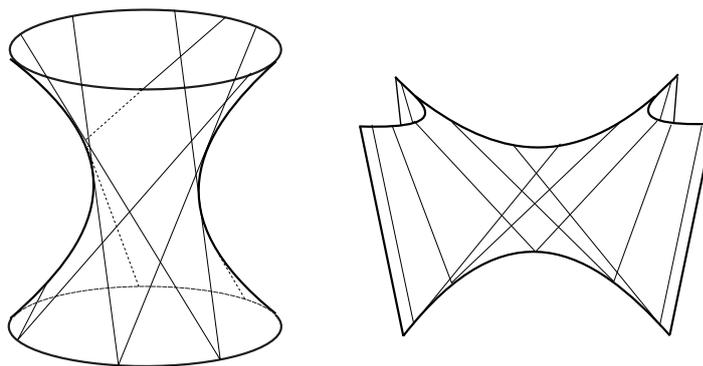


Figura 2.15: Superficies regladas.

Ejemplo 2.14.

1. El cilindro.
2. El cono como unión de líneas que pasan por el origen.
3. El hiperboloide y la silla de montar.

Son ejemplos de superficies regladas.

Nótese que el hiperboloide es único, dada la ecuación $|x|^2 - y^2 = 1$ entonces $x_1^2 - y^2 = 1 - x_2^2$ de tal manera que:

$$(x_1 - y)(x_1 + y) = (1 - x_2)(1 + x_2)$$

y contiene a las rectas

$$x_1 - y = 1 - x_2, \quad x_1 + y = 1 + x_2$$

y

$$x_1 - y = 1 + x_2, \quad x_1 + y = 1 - x_2$$

Cada una de estas rectas se obtiene por rotación del eje y .

En el caso de la silla de montar tenemos: $z = x^2 - y^2 = (x + y)(x - y)$, con las ecuaciones de las rectas

$$x + y = a, \quad a(x - y) = z$$

y

$$x - y = a, \quad a(x + y) = z$$

Se entenderá mejor en la siguiente sección.

2.3. El espacio proyectivo

La geometría proyectiva surge como una solución a los problemas de perspectiva que aparecen al tratar de representar en un plano objetos del espacio. Esta geometría estudia las propiedades de las figuras que se conservan por proyección. Como veremos más adelante, en geometría proyectiva no sólo no hay noción de distancia como en la geometría euclidiana, sino que tampoco hay noción de paralelismo como en la geometría afín. Sólo la colinealidad y la incidencia son típicamente proyectivas. Es la geometría asociada al modo en que el ojo humano percibe el mundo.

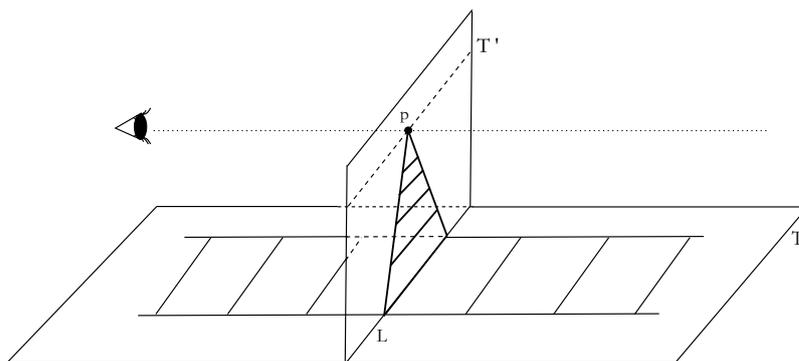


Figura 2.16: Modelo geométrico de como el ojo humano percibe el mundo.

Supongamos que el ojo es el origen en \mathbb{R}^3 , de modo que los rayos de luz que llegan a él son rectas que pasan por el origen. Supongamos que estamos mirando un paisaje plano T , las vías de un tren por ejemplo. Los objetos que viven en T quedan representados en el plano T' por medio del punto de corte de la recta correspondiente con T' . Dos rectas perpendiculares a $L = T \cap T'$ parecen cortarse en cierto punto p , el cual no corresponde a ningún punto de T . En teoría de perspectiva a este punto le denominan *punto evanescente*

que es algo como un **punto en el infinito** que nuestro ojo añade a T por su forma de percibir la realidad. Obviamente todos los puntos de la recta paralela a L que pasan por p son de este tipo. Por otra parte, las traviesas de la vía se ven en T' como segmentos cada vez más pequeños, de tal manera que la proyección a T' no conserva ni el paralelismo ni las distancias.

Definición 2.15. Sea V un espacio vectorial sobre K . El **espacio proyectivo** $P(V)$ es el conjunto de rectas $p \subset V$ que pasan por el origen. Sus elementos se denominan **puntos**. Si $\dim V = n$ entonces $\dim P(V) := n - 1$. Si $U \subset V$ es un subespacio entonces $P(U) \subset P(V)$ es un **subespacio proyectivo**.

Si $p \in P(V)$ y $v \in p, v \neq 0$ entonces la recta p consiste de todos los múltiplos del vector v , lo cual denotaremos por $p = [v]$. En consecuencia $[v] = [\lambda \cdot v]$ para toda $0 \neq \lambda \in K$ y el espacio proyectivo se puede describir como:

$$P(V) = (V \setminus \{0\})/K^* \quad (2.3.1)$$

Los elementos de $P(V)$ son clases de equivalencia de vectores $v \neq 0$ en V bajo la relación de equivalencia $v \sim \lambda v$ para $\lambda \in K^*$. Si $V = K^{n+1}$ y $x = (x_0, \dots, x_n)$, $x \neq 0$ visto como vector fila, escribimos

$$[x] = [x_0, \dots, x_n] \in P(K^{n+1}) = KP^n$$

donde las x_j se conocen como las **coordenadas homogéneas**. Nótese que están determinadas salvo un factor escalar. También se suele escribir $[x] = [x_0 : x_1 : \dots : x_n]$. Ahora se puede interpretar ciertos conceptos de la teoría de espacios vectoriales con los conceptos de geometría proyectiva.

Una **recta proyectiva** es un espacio proyectivo de dimensión uno, un **hiperplano proyectivo** $H \subset P(V)$ es un subespacio proyectivo de codimensión 1 es decir, de dimensión $\dim P(V) - 1$.

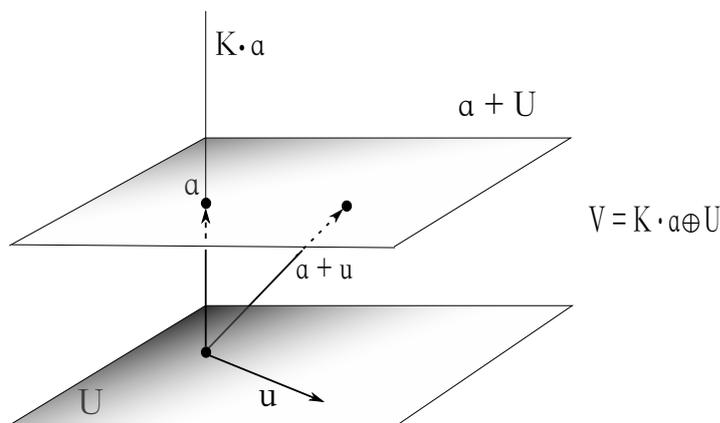


Figura 2.17: El espacio afín depende de la elección de un punto.

Fijemos ahora un hiperplano $U \subset V$ y con ello un hiperplano proyectivo $H = P(U) \subset P(V)$, que llamaremos el **hiperplano al infinito**. Sea $A = P(V) \setminus H$ y sea $a \in V \setminus U$

fijo. Entonces $V = K \cdot a \oplus U$ y A consiste de todos los puntos $[\kappa a, u]$ con $\kappa \neq 0$. Como $[\kappa a, u] = [a, \frac{1}{\kappa}u]$ para $\kappa \neq 0$, obtenemos las biyecciones:

$$U \rightarrow a + U \rightarrow A$$

$$u \mapsto a + u \mapsto [a + u]$$

En otras palabras, A se ha identificado con el espacio afín asociado al espacio vectorial U . Luego el espacio proyectivo $P(V)$ puede verse como la unión disjunta del espacio afín A y el hiperplano al infinito H .

Por otro lado, si partimos de un espacio vectorial U y quisieramos realizarlo como el espacio afín de un espacio proyectivo, ponemos $V = K \times U$ y tomemos $H = P(U)$ como el hiperplano al infinito y $a = (1, 0)$. Entonces tenemos la biyección:

$$U \rightarrow A$$

$$u \mapsto [1, u].$$

La descomposición de $P(V)$ como unión del espacio afín A y el hiperplano al infinito se sigue de la descomposición:

$$K \times U \xrightarrow{\cong} V$$

$$(\kappa, u) \mapsto \kappa + u.$$

Queremos mantener esta descomposición y comparar la geometría en U ó mejor aún, la de $1 \times U$ con $P(V) = P(K \times U)$. Escribimos U para $0 \times U \subset K \times U$. Además, se puede probar que hay una biyección

$$\{\text{subespacios afín de dim} = k \text{ de } 1 \times U\} \leftrightarrow \{\text{subespacios proyectivos de dim} = k \text{ de } P(K \times U) \text{ que no están en } P(U)\}$$

de la siguiente manera: si $P(W) \subset P(V)$ y que no está contenido en $P(U)$ entonces $W \cap (1 \times U)$ es el subespacio afín asociado de $1 \times U$.

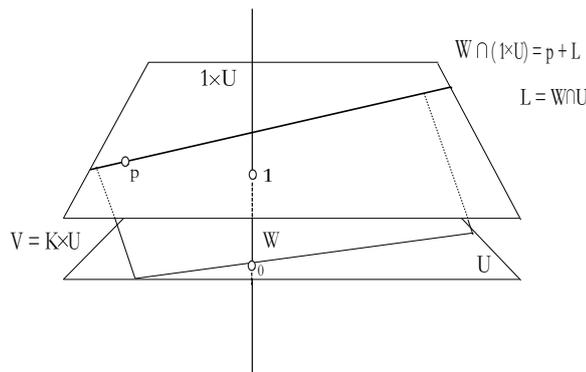


Figura 2.18: Subespacio afín asociado $W \cap (1 \times U)$.

En efecto, si $W \not\subset U = 0 \times U$ entonces W contiene un punto p con la primer componente 1, por lo tanto $p \in 1 \times U$ y entonces $1 \times U = p + U$, por lo tanto $W \cap (1 \times U) = W \cap (p + U) = p + (W \cap U)$ y $\dim(W \cap U) = \dim(W) = 1$ porque $W + U = V$.

Recíprocamente, si $p + L$ es un subespacio afín de dimensión k en $1 \times U$ con un espacio

vectorial asociado L , entonces $W = K \cdot p + L$ es el subespacio de dimensión $\dim(L) + 1$ de V , porque $L \subset U$ y $p \notin U$. Esto nos proporciona el mapeo $p + L \rightarrow P(W)$.

Para contestar preguntas sobre el espacio proyectivo, a veces es mejor hacer cálculos en el espacio afín. Por ejemplo: tres puntos $p, q, x \in U$ son colineales si y sólo si los tres puntos $(1, p), (1, q), (1, x)$ están en un subespacio afín de $1 \times U$ de dimensión 1 si y sólo si $[1, p], [1, q], [1, x] \in P(K \times U)$ están en un subespacio proyectivo de dimensión 1. Esto sucede si y sólo si los correspondientes vectores $(1, p), (1, q), (1, x)$ son linealmente dependientes.

Observación 2.16. *La ecuación de la recta que pasa por dos puntos distintos $p, q \in K^2$ es*

$$\det \begin{pmatrix} 1 & x_1 & x_2 \\ 1 & p_1 & p_2 \\ 1 & q_1 & q_2 \end{pmatrix} = 0.$$

Definición 2.17. *Una k -tupla de puntos $[v_1], \dots, [v_k]$ de $P(V)$ se dice **independiente** o en **posición general** cuando los vectores v_1, \dots, v_k en V son linealmente independientes.*

Lo anterior nos dice que los puntos no están en un subespacio proyectivo de dimensión menor que $k - 1$. En consecuencia, determinan a un único subespacio $P(L(v_1, \dots, v_k)) \subset P(V)$ en el que están contenidos y se dice que generan este subespacio. Por ejemplo, puntos en posición general generan un plano. Una tupla de puntos que están sobre una recta se dicen **colineales**. Ahora ilustremos en dimensiones bajas el espacio proyectivo real y complejo.

Notemos que se tiene una proyección canónica:

$$\begin{aligned} \pi : S^n &\rightarrow \mathbb{R}P^n \\ v &\mapsto [v], \end{aligned}$$

y que todo punto en $\mathbb{R}P^n$ tiene dos pre imágenes $v, -v \in S^n$, ya que toda recta que pasa por el origen en \mathbb{R}^{n+1} intersecta a S^n en dos puntos antipodales. De esta manera se puede visualizar a $\mathbb{R}P^n$ como el conjunto de pares no ordenados de puntos antipodales $\{v, -v\}$ de la esfera S^n , es decir: $\mathbb{R}P^n$ se obtiene a partir de S^n por identificación de puntos antipodales.

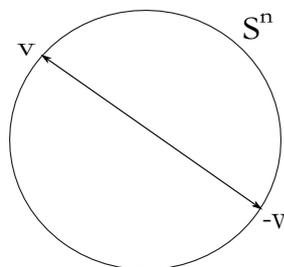


Figura 2.19: Puntos antipodales de la esfera S^n .

Esta descripción de $\mathbb{R}P^n$ es útil sobre todo en dimensiones altas. Por ejemplo KP^0 el espacio de rectas en K consta de un solo punto y no hay nada que describir. En el espacio

KP^1 el hiperplano al infinito es un solo punto ∞ y por tanto $KP^1 = K \cup \{\infty\}$. Vamos a ilustrarlo para $K = \mathbb{R}$ y \mathbb{C} :

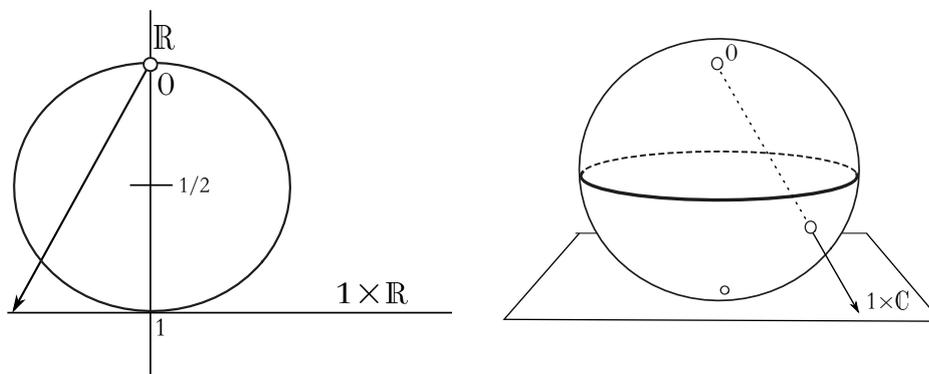


Figura 2.20: Proyección estereográfica.

Consideremos una esfera de dimensión 1 ó 2 en $\mathbb{R} \times K$ de radio $1/2$ y centro $(1/2, 0)$. Toda recta que pasa por el origen en \mathbb{R}^2 intersecta al círculo en exactamente un punto. Los puntos distintos de 0 del círculo corresponden biyectivamente a puntos afines en $1 \times \mathbb{R}$, el punto ∞ corresponde al polo norte.

Además, una recta compleja $[z_1, z_2]$ con $z_1 \neq 0$ contiene los puntos $\lambda(|z_1|^2, \bar{z}_1 z_2)$ con $\lambda \in \mathbb{R}$ cuya primer componente es real, así las rectas complejas que pasa por el origen en \mathbb{C}^2 se corresponden biyectivamente a las rectas reales que pasan por el origen en $\mathbb{R} \times \mathbb{C}$, siempre que toda primer componente del vector que define a cada recta no desaparezca. Ahora justo como la recta real antes mencionada, se corresponde biyectivamente los puntos de intersección de la esfera, con la excepción del polo norte 0 por lo cual se añade el punto al infinito, por lo tanto se tiene

$$\mathbb{R}P^1 = S^1, \quad \mathbb{C}P^1 = S^2. \tag{2.3.2}$$

A $\mathbb{C}P^1$ se le conoce como la **esfera de Riemann**.

El espacio $\mathbb{R}P^2$ se obtiene por la identificación de puntos antipodales en S^2 . Los puntos en el hemisferio norte $\{x \mid x_3 > 0\}$ tienen un asociado en el hemisferio sur $\{x \in S^2 \mid x_3 \leq 0\}$, los puntos antipodales del ecuador $\{(x_1, x_2, 0) \mid |x| = 1\}$ se identifican.

Gracias a tal unión surge la banda C que es una banda de Möbius, en el borde de un círculo que tiene circunferencia $A \cup B$, localmente se ve como \mathbb{R}^2 .

Similarmente $\mathbb{C}P^1 = S^2$ se obtiene de la esfera $S^3 = \{z \in \mathbb{C}^2 \mid |z| = 1\}$. Cada punto $[z] \in \mathbb{C}P^1$ tiene por lo menos un representante $z \in S^3 \subset \mathbb{C}^2$. La función:

$$\begin{aligned} \pi : S^3 &\rightarrow S^2 = \mathbb{C}P^1, \\ z &\mapsto [z]. \end{aligned}$$

se denomina **fibración de Hopf**, tal función es sobreyectiva y la imagen inversa de un punto $p = [z] \in \mathbb{C}P^1$ es el conjunto $\{\lambda \cdot z \mid \lambda \in \mathbb{C}, |\lambda| = 1\}$. El cual es un círculo, si la

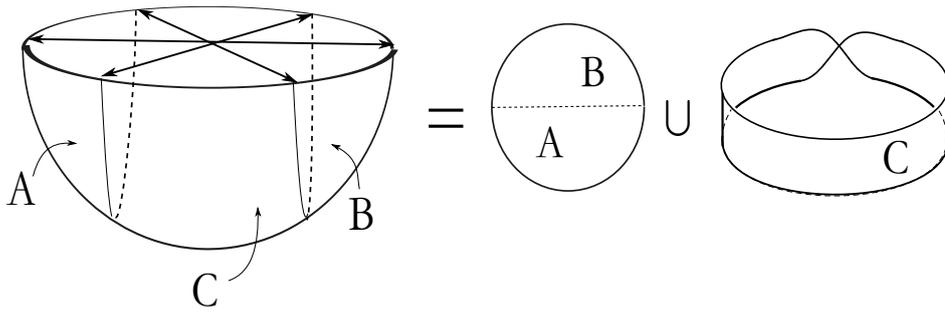


Figura 2.21: $\mathbb{R}P^2 = \text{Disco} \cup \text{banda de Möbius}$.

imagen inversa de p corresponde a un punto z , existe una función biyectiva y continua:

$$S^1 \rightarrow \pi^{-1}\{p\}$$

$$\lambda \mapsto \lambda \cdot z$$

Podríamos creer que existe una biyección continua entre $S^1 \times S^2 \rightarrow S^3$ pero esto es falso. La elección de $z \in [z]$ no puede hacerse de manera continua. La esfera S^n no puede ser producto de dos espacios $A \times B$ a menos que uno de ellos sea un punto y el otro una esfera.

2.4. Ecuaciones homogéneas.

Una ecuación de la forma $f(x) = 1 + x_0 + 2x_1^2 + \dots = 0$ tiene sentido en K^{n+1} sin embargo no tiene sentido en el espacio proyectivo KP^n como una ecuación en coordenadas homogéneas de un punto $[x_0, \dots, x_n]$, porque para $\lambda \neq 1$, $f(\lambda x) = 0$ es una ecuación muy diferente de $f(x) = 0$ y las coordenadas homogéneas están definidas salvo un factor en común.

Definición 2.18. Una **cuádrica proyectiva** se define como el conjunto

$$N = \{[v] \mid \gamma(v) = 0, v \neq 0\} \subset P(V),$$

y el **tangente proyectivo** es el conjunto $\{[x] \mid \gamma(v, x) = 0, x \neq 0\}$ para $[v] \in N$. La **dimensión de la cuádrica** es $\dim(P(V) - 1)$.

Aquí γ es una forma bilineal simétrica, no degenerada.

Definición 2.19. Una transformación $f : V \rightarrow W$ entre espacios vectoriales sobre un campo K se dice **homogénea de grado k** si

$$f(\lambda \cdot v) = \lambda^k \cdot f(v)$$

para todo $\lambda \in K$.

Por ejemplo, $x \mapsto x_0^{\alpha_0} \cdot \dots \cdot x_n^{\alpha_n}$ es homogénea de grado $\alpha_0 + \dots + \alpha_n$. Una transformación homogénea (de cualquier grado) $f : V \rightarrow W$ tiene como conjunto de ceros

$$M = \{x \in V \mid f(x) = 0\}$$

un **cono**, que es la unión de rectas que pasan por el origen ya que $f(x) = 0$ si y sólo si $f(\lambda x) = \lambda^k f(x) = 0$ para toda $\lambda \neq 0$. El cono en V con vértice en el origen, se puede escribir como una unión (posiblemente infinita) de rectas que pasan por el origen, así que es un subconjunto de $P(V)$. Las ecuaciones homogéneas $f : V \rightarrow W$ definen entonces el conjunto de ceros:

$$M(f) = \{[x] \in P(V) \mid f(x) = 0\} \subset P(V)$$

De tal manera que la propiedad de que f sea cero o no, depende solamente de la clase de equivalencia del punto. Ahora bien, si tenemos ecuaciones no homogéneas

$$A \cdot x - b = 0, \quad x^t G x + 2a^t x + b = 0 \quad (2.4.1)$$

para puntos x en un espacio vectorial U . Éstas ecuaciones se pueden llevar a ecuaciones homogéneas de la siguiente manera: realizamos el espacio $K \times U$ cuyos puntos se indican mediante las coordenadas (y, x) y sustituimos en la ecuación (2.4.1) para llevarlas a su forma correspondiente homogénea o ecuación proyectiva asociada:

$$A \cdot x - b \cdot y = 0, \quad x^t G x + 2a^t x \cdot y + b \cdot y^2 = 0. \quad (2.4.2)$$

El método es análogo para grados mayores. En la parte afín

$$1 \times U \subset P(V)$$

$$(1, u) \mapsto [1, u]$$

las ecuaciones vuelven a las originales (2.4.1) de tal forma que, si $f = 0$ es una ecuación de la forma (2.4.2) entonces $M(f) \cap (1 \times U)$ es el espacio de ceros de la ecuación (2.4.1) correspondiente.

Como sabemos, la ecuación no homogénea $Ax - b = 0$ en U es un conjunto solución de un espacio afín para el espacio vectorial $\ker(A)$, o es vacía y así la ecuación proyectiva $Ax - by = 0$ tiene como conjunto solución un subespacio proyectivo cuyo espacio vectorial asociado es el núcleo de la transformación

$$\begin{pmatrix} y \\ x \end{pmatrix} \mapsto (-b \mid A) \cdot \begin{pmatrix} y \\ x \end{pmatrix}.$$

Si la ecuación no homogénea tiene solución única v entonces la ecuación homogénea tiene la solución $(0, \ker(A)) \oplus K(1, v)$. Si no tiene una solución, entonces la solución es $(0, \ker(A))$ de la ecuación homogénea en el hiperplano al infinito.

En geometría proyectiva, las superficies de segundo orden se describen por una forma cuadrática y su clasificación se reduce a clasificar las formas cuadráticas de transformaciones lineales salvo un factor. Recordando el teorema 1.43 (de Sylvester) donde tratamos el índice o signatura de la forma cuadrática; para $K = \mathbb{C}$ el rango de la forma bilineal asociada es invariante y para $K = \mathbb{R}$ el rango y el índice son invariantes, es decir, en estos casos las cuádricas son equivalentes. En consecuencia, sólo existe la ecuación:

$$|x|^2 - |y|^2 = 0.$$

Por ejemplo en \mathbb{R}^3 tenemos rango 3 y sólo existe un tipo de cuádrica no degenerada

$$|x|^2 = y^2, \quad x \in \mathbb{R}^2,$$

que es el cono. Las imágenes afines se forman mediante la intersección de la transformación lineal del cono con el hiperplano $1 \times \mathbb{R}^2$.

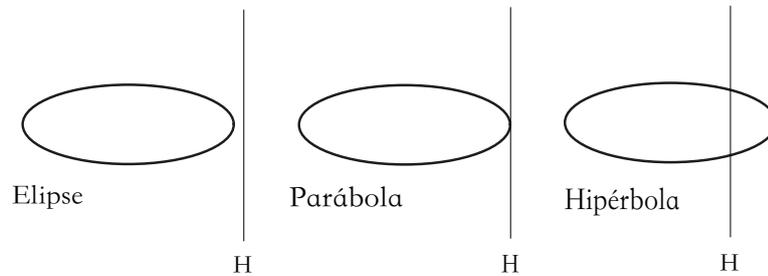


Figura 2.22: Secciones cónicas.

De ahí el nombre de sección cónica. La curva proyectiva completa $|x|^2 - y^2 = 0$ es una elipse y dependiendo de como ésta elipse está a cierta distancia de la recta H (la toca o no) surge la parte afín $\mathbb{R}^2 = \mathbb{R}P^2 \setminus H$: la elipse, parábola e hipérbola. Este no es necesario calcular, sólo se tiene que escoger uno de ellos y familiarizarse con curvas afín de segundo orden, de modo que hay que cuestionar su forma.

Inclusive las rectas de la superficie silla de montar y el hiperboloide de una hoja pueden ser más claras en geometría proyectiva. La **inmersión de Segre** es utilizada en geometría proyectiva para considerar el producto cartesiano de dos espacios proyectivos como una variedad proyectiva. En particular tenemos:

$$s : KP^1 \times KP^1 \rightarrow KP^3$$

$$([x], [y]) \mapsto [x_0y_0, x_0y_1, x_1y_0, x_1y_1] = [s_0, s_1, s_2, s_3].$$

Teorema 2.20. *La inmersión de Segre en este caso es inyectiva y su imagen es la cuádrica proyectiva M con ecuación $s_0s_3 - s_1s_2 = 0$.*

Demostración. Si $[x] \neq [x']$ y $y_j \neq 0$, entonces (x_0y_j, x_1y_j) no es múltiplo de $(x'_0y'_j, x'_1y'_j)$ por lo tanto $s([x], [y]) \neq s([x'], [y'])$. Análogamente para $[y] \neq [y']$. Esto muestra que s es inyectiva. Para mostrar que es cierto que los puntos en la imagen cumplen que $s_0s_3 = x_0y_0x_1y_1 = s_1s_2$. Si todo punto de la cuádrica M está en la imagen de s , distinguimos los siguientes casos:

1. $s_0 = s_1 = 0$. Hacemos $x = (0, 1), y = (s_2, s_3)$.
2. $s_0 = s_2 = 0$. Hacemos $x = (s_1, s_3), y = (0, 1)$.
3. $s_0 \neq 0$, sin pérdida de generalidad sea $s_0 = 1$. Hacemos $x = (1, s_2), y = (1, s_1)$.

□

En el producto $KP^1 \times KP^1$ podemos ver dos rectas que se intersectan: pasan por el punto (p, q) a través de la recta $p \times KP^1$ y $KP^1 \times q$ y también las imágenes de estas rectas bajo s son rectas.

En el caso de las superficies regladas afines M es como sigue: si sustituímos $s_0 = u - v$, $s_3 = u + v$, $s_1 = w$ y consideramos la parte afín de $s_2 = 1$ entonces tenemos la superficie **silla de montar**:

$$u^2 - v^2 = w.$$

Si sustituímos $s_0 = u - v$, $s_3 = u + v$, $s_1 = w - t$, $s_2 = w + t$ y consideramos la parte afín $t = 1$, entonces obtenemos el **hiperboloide**:

$$u^2 + w^2 - v^2 = 1.$$

Ambas superficies regladas son la parte afín de la misma superficie reglada proyectiva $M = KP^1 \times KP^1$. Tenemos una buena imagen de $KP^1 = K \cup \{\infty\}$ en el caso del campo real y en el caso complejo: $\mathbb{R}P^1 = S^1$ y $\mathbb{C}P^1 = S^2$. En consecuencia $\mathbb{R}P^1 \times \mathbb{R}P^1 = S^1 \times S^1$ es el **toro** y $\mathbb{C}P^1 \times \mathbb{C}P^1 = S^2 \times S^2$.

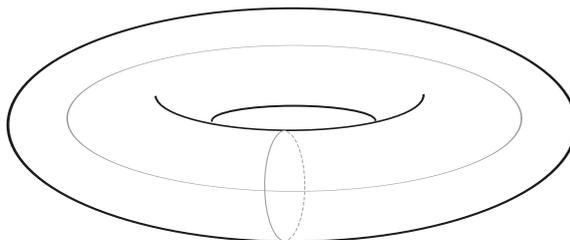


Figura 2.23: $\mathbb{R}P^1 \times \mathbb{R}P^1$.

2.5. Tipo topológico de las cuádricas

Nuestra intuición inmediata no es suficiente para conocer a las cuádricas en dimensiones superiores. Sin embargo, se puede entender a las cuádricas a partir de espacios más simples. Asumiremos que se conoce lo que es un espacio topológico y lo que es un homeomorfismo $f : X \rightarrow Y$ entre espacios topológicos: una función continua $f : X \rightarrow Y$ con inversa $f^{-1} : Y \rightarrow X$ continua. Si un homeomorfismo existe, decimos que los espacios son homeomorfos y escribimos $X \approx Y$. Nuestra intención es describir cuádricas salvo homeomorfismos. Los homeomorfismos presentan fórmulas bastante explícitas, ellos son diferenciales, analíticos. Por lo que no se necesita gran cantidad de conceptos para nuestro propósito. Empezemos con las cuádricas afines reales.

Denotemos por (x, y) un punto en $\mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$ y a (x, y, z) un punto en $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} = \mathbb{R}^{n+m+1}$. En la primera sección de este capítulo vimos que para las transformaciones afines existen tres tipos de ecuaciones cuádricas: supondremos que la forma cuadrática tiene rango máximo:

$$|x|^2 - |y|^2 = z$$

$$\begin{aligned} |x|^2 - |y|^2 &= 1 \\ |x|^2 - |y|^2 &= 0 \end{aligned}$$

Sean $Q(n, m, z)$, $Q(n, m, 1)$, $Q(n, m, 0)$ las cuádricas asociadas, de tal manera que los conjuntos solución de tales ecuaciones están en \mathbb{R}^{n+m+1} ó \mathbb{R}^{n+m} . Entonces tenemos:

$$Q(n, m, z) \approx \mathbb{R}^{n+m} \quad (2.5.1)$$

En efecto, $Q(n, m, z)$ es la gráfica de la función $z(x, y)$. El homeomorfismo está dado por: $(x, y, z) \mapsto (x, y)$ e inversa $(x, y) \mapsto (x, y, |x|^2 - |y|^2)$.

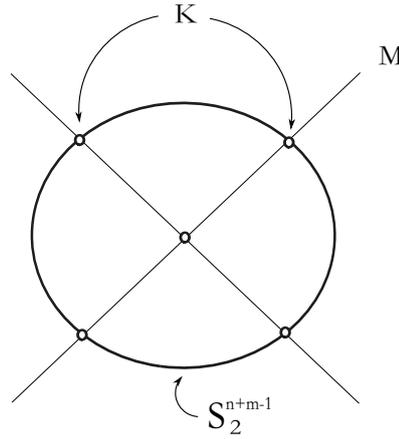


Figura 2.24: $K(n, m) = Q(n, m, 0) \cap S_2^{n+m-1}$.

$$Q(n, m, 1) \approx S^{n-1} \times \mathbb{R}^m. \quad (2.5.2)$$

En efecto, si $r(y) := \sqrt{1 + |y|^2}$, entonces $S^{n-1} \times \mathbb{R}^m \rightarrow Q(n, m, 1)$, $(\xi, y) \mapsto (r(y) \cdot \xi, y)$ es un homeomorfismo cuya inversa es $(x, y) \mapsto (r(y)^{-1} \cdot x, y)$.

La tercera ecuación $|x|^2 = |y|^2$ es homogénea, la cuádrica es por lo tanto un cono en \mathbb{R}^{m+n} , mas precisamente:

Si S_2^{n+m-1} es la esfera de radio 2 en \mathbb{R}^{n+m} con centro en el origen y

$$K(n, m) = Q(n, m, 0) \cap S_2^{n+m-1}.$$

Entonces $Q(n, m, 0)$ es la unión de los rayos que pasan por el origen a través de $K(n, m)$, entonces decimos que: $Q(n, m, 0)$ es el **cono** sobre $K(n, m)$. Por lo tanto es suficiente describir a $K(n, m)$.

$$K(n, m) = Q(n, m, 0) \cap S_2^{n+m-1} = S^{n-1} \times S^{m-1}. \quad (2.5.3)$$

El producto de esferas está en $\mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$.

En efecto, $K(n, m)$ es el conjunto de puntos que satisfacen las ecuaciones: $|x|^2 = |y|^2$, $|x|^2 + |y|^2 = 2$. Es decir $|x|^2 = 1$, $|y|^2 = 1$ y definen la $S^{n-1} \times S^{m-1}$ esfera. En el caso complejo, las ecuaciones (2.1.2) corresponden a los siguientes 3 tipos:

$$z_1^2 + \dots + z_n^2 = w,$$

$$\begin{aligned} z_1^2 + \dots + z_n^2 &= 1, \\ z_1^2 + \dots + z_n^2 &= 0 \end{aligned} \tag{2.5.4}$$

Las cuádricas complejas afines las denotaremos por $Q(n, w), Q(n, 1), Q(n, 0)$. Como antes:

$$Q(n, w) \approx \mathbb{C}^n. \tag{2.5.5}$$

Para el cálculo de $Q(n, 1)$ hacemos $z = u + iv$, $u, v \in \mathbb{R}^n$ parte real e imaginaria de z . Entonces $u_1^2 - v_1^2 + \dots + u_n^2 - v_n^2 = 1$, $2i(u_1v_1 + \dots + u_nv_n) = 0$, por lo tanto

$$|u|^2 - |v|^2 = 1, \quad \langle u, v \rangle = 0,$$

Ahora bien, $\{x \in \mathbb{R}^n \mid |x|^2 = 1\}$ define la esfera $S^{n-1} \in \mathbb{R}^n$ y $\{y \in \mathbb{R}^n \mid \langle x, y \rangle = 0\}$ es el espacio vectorial tangente en el punto $x: T_x S^{n-1}$ (ver ejemplo 2.5). El espacio

$$\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mid x \in S^{n-1}, y \in T_x S^{n-1}\} =: TS^{n-1}$$

se denomina **fibrado tangente** de S^{n-1} . Este consiste en la unión disjunta de los espacios tangentes de todos los puntos de S^{n-1} :

$$TS^{n-1} = \bigcup_{x \in S^{n-1}} \{x\} \times T_x S^{n-1}.$$

Sin embargo, en general TS^{n-1} no es homeomorfo al producto $S^{n-1} \times \mathbb{R}^{n-1}$, es decir, no

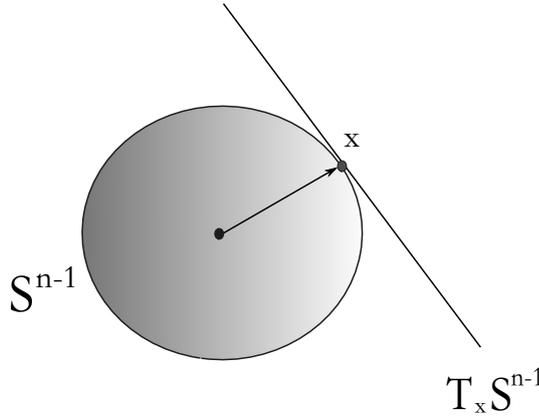


Figura 2.25: Espacio tangente en $\{x\}$ a S^{n-1}

es trivial. Por ejemplo, para $n = 2$ no hay una función continua que asigne a cada $x \in S^n$ un vector tangente $v(x) \neq 0$ de S^n en x . Pero en todo caso, TS^{n-1} es el conjunto de puntos $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ con

$$|x|^2 = 1, \quad \langle x, y \rangle = 0,$$

y análogamente como en (2.5.2) se obtiene el homeomorfismo

$$Q(n, 1) \approx TS^{n-1} \tag{2.5.6}$$

a saber, si $r(v) = \sqrt{1 + |v|^2}$ entonces tenemos la función

$$TS^{n-1} \rightarrow Q(n, 1)$$

$$(x, y) \mapsto (r(v) \cdot x, v)$$

cuya inversa es $(u, v) \mapsto (r(v)^{-1}u, v)$.

Las consideraciones para describir a $Q(n, 0)$ son similares: sea $K(n) := Q(n, 0) \cap S_2^{2n-1}$, la sección del cono $Q(n, 0)$ con la esfera de radio 2 con centro en el origen. Para K se tienen las ecuaciones

$$\begin{aligned} z_1^2 + \dots + z_n^2 &= 0, \\ |z_1|^2 + \dots + |z_n|^2 &= 2. \end{aligned}$$

La descomposición en parte real e imaginaria de $z = u + iv$ da como resultado:

$$|u|^2 = |v|^2, \quad \langle u, v \rangle = 0, \quad |u|^2 + |v|^2 = 2$$

es decir,

$$|u|^2 = 1, \quad |v|^2 = 1, \quad \langle u, v \rangle = 0. \quad (2.5.7)$$

En general, una k -tupla de vectores (v_1, \dots, v_k) en \mathbb{R}^n con $|v_j|^2 = 1$, $\langle v_i, v_j \rangle = 0$, se denomina un **k-marco ortonormal** y al espacio de todos los k -marcos ortonormales en \mathbb{R}^n se denomina una **variedad de Stiefel** $V_{k,n}$. La ecuación (2.5.7) significa:

$$Q(n, 0) \cap S_2^{2n-1} = K(n) \approx V_{2,n}. \quad (2.5.8)$$

Dicho de otro modo, $V_{2,n} \subset TS^{n-1}$ es el espacio de todos los vectores tangentes de longitud 1.

Ejemplo 2.21. *Un 2-marco ortonormal $(v, w) \in \mathbb{R}^3$ se puede completar de manera única a un 3-marco ortonormal $(v, w, v \times w)$ orientado positivamente, por lo tanto a una matrix en $SO(3)$, es decir,*

$$K(3) \approx V_{2,3} \approx SO(3).$$

Nuevamente, $Q(n, 0)$ es el cono de todos los rayos que emanan desde el origen (reales) en $K(n)$.

Nos preguntamos ahora ¿cuáles son las cuádricas proyectivas?. En el caso real tenemos la ecuación homogénea $|x|^2 = |y|^2$ en \mathbb{R}^{n+m} sobre S_2^{n+m-1} definido como el producto de $S^{n-1} \times S^{m-1}$. En el caso proyectivo, el conjunto $\mathbb{R}P^{n+m-1}$ definido por la identificación de (x, y) con $(-x, -y)$ resulta ser:

Observación 2.22. *La cuádrica proyectiva de ecuación $|x|^2 - |y|^2 = 0$, es homeomorfa a*

$$S^{n-1} \times S^{m-1} / (x, y) \sim -(x, y).$$

En general esta descripción debe ser suficiente para nosotros, sin embargo si $m = 2$ y $n = 2k$ entonces $S^{m-1} = S^1 \subset \mathbb{C}$ y $S^{n-1} \subset \mathbb{C}^k$ y tenemos el homeomorfismo

$$\kappa : S^{2k-1} \times S^1 \rightarrow S^{2k-1} \times S^1$$

$$(\omega, \xi) \mapsto (\xi \cdot \omega, \xi).$$

Y el siguiente diagrama es conmutativo:

$$\begin{array}{ccccc}
 (\omega, \xi) & S^{2k-1} \times S^1 & \xrightarrow{\kappa} & S^{2k-1} \times S^1 & (\omega, \xi) \\
 \downarrow & \downarrow & & \downarrow & \downarrow \\
 (-\omega, -\xi) & S^{2k-1} \times S^1 & \xrightarrow{\kappa} & S^{2k-1} \times S^1 & (\omega, -\xi)
 \end{array}$$

El espacio $(S^{2k-1} \times S^1)/(x, y) \sim -(x, y)$ es por lo tanto homeomorfo a $(S^{2k-1} \times S^1)/(x, y) \sim (x, -y)$, de este modo $S^{2k-1} \times \mathbb{R}P^1 \approx S^{2k-1} \times S^1$. Análogamente podemos ver: \mathbb{R}^n actuando por el grupo $S^0 = \{1, -1\}$, $(\mathbb{C}^2)^k$ actuando por el grupo S^3 y S^7 de los números de Cayley actuando por $(\mathbb{R}^8)^k$.

Teorema 2.23. *Si $m = 1$ ó $n=2k$ y $m=2$, ó $n =4k$ y $m=4$, ó $n =8k$ y $m=8$, entonces la cuádrica proyectiva real de ecuación $|x|^2 - |y|^2 = 0$ en $P(\mathbb{R}^n \times \mathbb{R}^m)$ es homeomorfa a $S^{n-1} \times \mathbb{R}P^{m-1}$. A la vez que $\mathbb{R}P^1 \approx S^1$.*

La cuádrica proyectiva compleja con ecuación $z_1^2 + \dots + z_n^2 = 0$ surge de la intersección de la cuádrica afín compleja $K(n)$ con la esfera identificando los puntos $z = u + iv$, convirtiéndose así bajo la acción de S^1 en los puntos de la cuádrica. Un 2-marco ortonormal (u, v) es una base ortonormal del plano generado por ellos $L(u, v) \subset \mathbb{R}^n$ y mediante las transformaciones ortogonales que preservan orientación, elementos de $S^1 = SO(2)$ van de dicha base a todos las bases ortonormales positivas del plano $L(u, v) \subset \mathbb{R}^n$, de modo que los puntos de la cuádrica proyectiva de ecuación $z_1^2 + \dots + z_n^2 = 0$ se corresponden exactamente con los planos orientados en \mathbb{R}^n .

Ejemplo 2.24. *La cuádrica proyectiva compleja no degenerada $z_0^2 + z_1^2 + z_2^2 = 0$ es homeomorfa a S^2 .*

Demostración. El homeomorfismo está dado por $[z] = [u + iv] \mapsto (u \times v)/|u \times v|$. Se debe tener en cuenta que un plano orientado $L(u, v)$ en \mathbb{R}^3 con base ortonormal positiva (u, v) (no depende de la elección de la base, sólo de la orientación) se determina de manera única por su vector normal $u \times v \in S^2$. □

Para este homeomorfismo hay una explicación puramente algebraica.

Observación 2.25. *La transformación*

$$KP^1 \rightarrow KP^1, \quad [x_0, x_1] \mapsto [x_0^2, x_0x_1, x_1^2],$$

es inyectiva y su imagen es la cuádrica proyectiva $y_0y_2 - y_1^2 = 0$.

Demostración. En un punto de la cuádrica si $y_0 \neq 0$ ó $y_2 \neq 0$ entonces (por simetría) sin pérdida de generalidad $y_0 = 1$. Entonces el punto tiene identificada la pre-imagen $[1, y_1]$ □

Para $K = \mathbb{C}$ esta es la única cuádrica no degenerada en $\mathbb{C}P^2$, que de este modo se puede transformar en la cuádrica $z_0^2 + z_1^2 + z_2^2 = 0$. Para $K = \mathbb{R}$ es la cuádrica $y_0^2 + y_1^2 - y_2^2 = 0$. La observación muestra que esta cuádrica en ambos casos es homeomorfa a KP^1 . Por lo tanto se tiene que para $K = \mathbb{C}$: $\mathbb{C}P^1 = S^2$; para $K = \mathbb{R}$: $\mathbb{R}P^1 = S^1$, como vimos en la sección 2.3.

La función inversa de una cuádrlica $M \subset KP^2$ no degenerada para KP^1 se puede describir de la siguiente manera:

Sea M la cuádrlica no degenerada para la forma cuadrática $\gamma : K^3 \rightarrow K$. Como γ tiene rango 3, γ no se anula en un subespacio de dimensión 2, por lo tanto M no contiene ninguna recta:

$$rg \left(\begin{array}{cc|c} 0 & 0 & ? \\ 0 & 0 & ? \\ ? & ? & ? \end{array} \right) \leq 2$$

Supongamos ahora que M contiene al menos un punto p . Entonces elegimos cualquier recta proyectiva $KP^1 \cong L \subset KP^2$ tal que $p \notin L$ y proyectamos M de p a L , donde p se proyecta en sí mismo por la tangente en p . Por lo tanto: definimos una función $\varphi : M \rightarrow L$ de la siguiente manera:

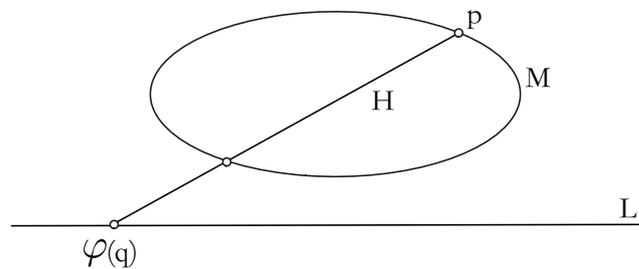


Figura 2.26: Proyección de un punto en una cuádrlica a una recta proyectiva.

Para cada recta H que pasa por p , la restricción de γ es una cuádrlica no trivial que contiene al punto p y por lo tanto existe otro punto q (ó H es la tangente en p) y $\varphi(q) \in L$ es el punto de intersección de H y L , es decir: un polinomio cuadrático que tiene una raíz p , o que tiene una raíz q , o la raíz p es doble. Entonces hemos definido una función biyectiva (continua) de la cuádrlica en $L \cong KP^1$. Si elegimos la cuádrlica como en la observación anterior y $p = [0, 0, 1]$, $L = \{[y_0, y_1, y_2] \mid y_2 = 0\}$, entonces esta función es la inversa de la función de (2.24), porque $[x, y] \mapsto [x^2, xy, y^2] \mapsto [x^2, xy] = [x, y]$, para $x \neq 0$. Concluimos esta sección con una aplicación en el área de los determinantes.

Ejemplo 2.26. En el espacio $End_k(K^2) \cong K^4$ se tiene la cuadrática de ecuación $\det(A) = 0$. Por lo tanto para $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\det(A) = ad - bc = 0$.

Esta ecuación es linealmente equivalente a la ecuación $x_1^2 + x_2^2 = y_1^2 + y_2^2$ cuando $2 \neq 0$ en K . En consecuencia: el determinante de la superficie $\{\det(A) = 0\} \subset End_{\mathbb{R}}(\mathbb{R}^2)$ es homeomorfo al cono sobre el toro $S^1 \times S^1$. El vértice del cono corresponde a la matriz cero. El determinante de la superficie $\{\det(A) = 0\} \subset End_{\mathbb{C}}(\mathbb{C}^2)$ es homeomorfo al cono sobre la variedad de Stiefel $V_{2,4}$.

Capítulo 3

Grupos lineales y álgebras de Lie

3.1. Descomposición polar y de Iwasawa

En este capítulo consideremos $K = \mathbb{R}$ ó \mathbb{C} y los siguientes subgrupos de matrices de $GL(n, K)$:

Definición 3.1.

$$B^+(n, K) = \{T \in GL(n, K) \mid t_{ik} = 0 \text{ para } i > k\} \quad (3.1.1)$$

es el grupo de **matrices triangulares superiores** (con coeficientes en la diagonal distintos de cero).

$$N^+(n, K) = \{T \in B^+(n, K) \mid t_{ii} = 1 \text{ para } i = 1, \dots, n\} \quad (3.1.2)$$

es el grupo de **matrices unipotentes triangulares superiores**.

$$D(n, K) = \{T \in GL(n, K) \mid t_{ij} = 0 \text{ para } i \neq j\} \quad (3.1.3)$$

es el subgrupo de **matrices diagonales** con componentes en la diagonal distintos cero.

$$B^-(n, K) = \{T^t \mid T \in B^+(n, K)\} \quad (3.1.4)$$

$$N^-(n, K) = \{T^t \mid T \in N^+(n, K)\}$$

son los grupos de matrices **triangulares inferiores** y matrices **unipotentes triangulares inferiores** respectivamente.

Denotemos por $D(n)$ al grupo de matrices diagonales con coeficientes reales positivos en la diagonal. Entonces

$$D(n) \cdot N^+(n, K) = N^+(n, K) \cdot D(n) \subset B^+(n, K)$$

es el grupo de **matrices triangulares superiores con diagonal positiva**. La anterior es una igualdad de conjuntos (ó grupos), los elementos de $D(n)$ no necesariamente conmutan con los de $N^+(n, K)$. Tenemos una sucesión exacta corta de grupos:

$$1 \rightarrow N^+(n, K) \rightarrow D(n) \cdot N^+(n, K) \rightarrow D(n) \rightarrow 1$$

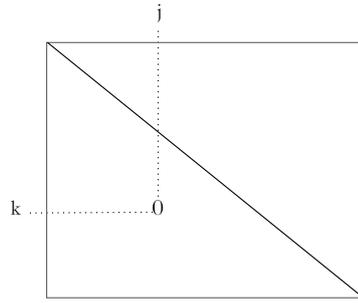


Figura 3.1

Ahora bien, si multiplicamos una matriz $A \in GL(n, K)$ por la derecha con $T \in D(n) \cdot N^+(n, K)$ entonces mediante operaciones elementales por columna, a saber, aquellas en las que la j -ésima columna a_j se sustituye por una combinación lineal $\sum_k t_{kj} a_k$ con $t_{kj} = 0$ para $k > j$ y $t_{kk} > 0$. Estas son precisamente las transformaciones que se utilizan en el método de ortogonalización de Gram-Schmidt: la ortogonalización del j -ésimo vector corresponde a multiplicar con $N \in N^+(n, K)$ y la normalización corresponde con $D \in D(n)$. El método por lo tanto proporciona para cada matriz $A \in GL(n, K)$ una matriz $D \in D(n)$, una $N \in N^+(n, K)$ y una matriz $U \in O(n)$ ó $U(n)$, tal que $ADN = U$, de manera que $A = U(DN)^{-1}$. En otras palabras.

Teorema 3.2 (Descomposición de Iwasawa). *El producto de matrices induce las biyecciones:*

$$O(n) \times D(n) \times N^+(n, \mathbb{R}) \rightarrow GL(n, \mathbb{R}) \quad (3.1.5)$$

$$U(n) \times D(n) \times N^+(n, \mathbb{C}) \rightarrow GL(n, \mathbb{C})$$

Por lo tanto, toda matriz $A \in GL(n, \mathbb{R})$ se puede escribir de manera única como producto de una matriz ortogonal, una matriz diagonal positiva y una matriz unipotente triangular superior. Un resultado similar se vale para $GL(n, \mathbb{C})$.

Demostración. (Unicidad) Si $A = UDN = U'D'N'$, entonces $U'^{-1}U = D'N'N^{-1}D^{-1}$. El lado izquierdo es una matriz unitaria, el lado derecho es una matriz triangular superior con diagonal positiva $D'D^{-1}$. Inductivamente columna por columna se concluye que el lado derecho es la matriz identidad, entonces $U' = U, D' = D$. Por lo tanto $N' = N$. \square

Ahora bien, al aplicar transpuesta, inversa (ó ambas) se obtiene las descomposiciones correspondientes:

$$N^-(n, \mathbb{R}) \times D(n) \times O(n) \longleftrightarrow GL(n, \mathbb{R}) \quad (3.1.6)$$

$$N^-(n, \mathbb{C}) \times D(n) \times U(n) \longleftrightarrow GL(n, \mathbb{C})$$

en donde el lado izquierdo también se puede escribir en orden inverso.

Observemos que las biyecciones (3.1.5) y (3.1.6) **no** son homomorfismos de grupos.

Comenzando ahora con las biyecciones:

$$\mathbb{R}^{n(n-1)/2} \rightarrow N^+(n, \mathbb{R})$$

$$\mathbb{C}^{n(n-1)/2} \rightarrow N^+(n, \mathbb{C})$$

$$\mathbb{R}^n \xrightarrow{\text{exp}} \mathbb{R}_+^n \rightarrow D(n)$$

obtenemos las biyecciones:

$$GL(n, \mathbb{R}) \rightarrow O(n) \times \mathbb{R}^k$$

$$GL(n, \mathbb{C}) \rightarrow U(n) \times \mathbb{R}^s$$

con $k = n(n+1)/2$, $s = n^2$. La geometría de \mathbb{R}^n y \mathbb{C}^n está contenida en la parte compacta de $O(n)$ y $U(n)$, respectivamente.

Si G es una matriz $n \times n$, real, simétrica y definida positiva. Entonces existe una matriz $A \in GL(n, \mathbb{R})$ tal que $x^t G y = (Ax)^t (Ay)$. Esta matriz A describe la transformación ortogonal: $(\mathbb{R}^n, \text{producto interno } x^t G y) \rightarrow (\mathbb{R}^n, \langle \cdot, \cdot \rangle)$. Además $G = A^t A$, como sabíamos. Si descomponemos A según el teorema como en (3.1.5) entonces

$$G = A^t A = (UDN)^t (UDN) = N^t D^2 N$$

Observación 3.3. Sean $H^+(n, \mathbb{R})$ el conjunto de matrices reales, simétricas, definidas positivas y $H^+(n, \mathbb{C})$ el conjunto de matrices hermitianas complejas, respectivamente. Entonces tenemos las biyecciones:

$$N^+(n, \mathbb{R}) \times D(n) \rightarrow H^+(n, \mathbb{R}),$$

$$N^+(n, \mathbb{C}) \times D(n) \rightarrow H^+(n, \mathbb{C}),$$

$$(N, D) \mapsto N^t D N.$$

$$(N, D) \mapsto N^* D N.$$

Teorema 3.4. Toda matriz definida positiva $H \in H^+(n, K)$ posee una única raíz \sqrt{H} en $H^+(n, K)$ cuyo cuadrado es igual a H .

Demostración. Recordemos que toda matriz simétrica es semejante a una matriz diagonal. Entonces $THT^{-1} = D \in D(n)$ con $T \in O(n)$ ó $U(n)$. Por lo tanto, es suficiente probar el teorema para D . Como $(\sqrt{D})^2 = D$ entonces

$$\begin{aligned} THT^{-1} &= (\sqrt{D})^2 \\ H &= T^{-1} \sqrt{D} \sqrt{D} T \\ &= (T^{-1} \sqrt{D} T)(T^{-1} \sqrt{D} T) \end{aligned}$$

Sea $B^2 = H$ entonces $THT^{-1} = TB^2T^{-1} = D$. Por lo tanto $D = (TBT^{-1})(TBT^{-1})$. Así $TBT^{-1} = \sqrt{D}$, en consecuencia $B = T^{-1} \sqrt{D} T$. \square

Teorema 3.5 (Descomposición polar). *El producto de matrices induce las biyecciones:*

$$H^+(n, \mathbb{R}) \times O(n) \rightarrow GL(n, \mathbb{R})$$

$$H^+(n, \mathbb{C}) \times U(n) \rightarrow GL(n, \mathbb{C})$$

Demostración. Mostraremos la biyección $H^+(n, \mathbb{C}) \times U(n) \rightarrow GL(n, \mathbb{C})$.

Injectividad: si $H \in H^+(n, \mathbb{C})$, $U \in U(n)$ tales que $H \cdot U = A$. Entonces $A^* = (HU)^* = U^* H^* = U^{-1} H$. Por lo tanto $H^2 = HUU^{-1}H = AA^*$ entonces $H = \sqrt{AA^*}$, de manera

que H está determinada por A y en consecuencia $U = H^{-1}A$.

Sobreyectividad: sea $H = \sqrt{AA^*}$ y $U = H^{-1}A$. Debemos mostrar que $U \in U(n)$,

$$\begin{aligned} U^*U &= (H^{-1}A)^*H^{-1}A \\ &= A^*(H^{-1})^*H^{-1}A \\ &= A^*H^{-2}A \\ &= A^*(AA^*)^{-1}A \\ &= I. \end{aligned}$$

En consecuencia $U \in U(n)$. □

Por la descomposición de Iwasawa hay exactamente una matriz triangular superior con diagonal positiva. Aquí le denominamos A , entonces el teorema muestra como la matriz unitaria $U = H^{-1}A$ es definida positiva y $H = \sqrt{AA^*}$ de tal manera que $HU = A$.

3.2. Funciones exponencial y logaritmo

Las versiones matriciales de las funciones exponencial $\exp(x)$ y logaritmo $\log(x)$ (reales y complejas), son fundamentales para el estudio de los grupos de matrices, en particular de los grupos de Lie. La función exponencial proporciona el vínculo entre un grupo de matrices y su álgebra de Lie, como se verá en la sección 3.3.

La manera más natural para definir la exponencial de una matriz es usar la serie de potencias de la función $f(x) = e^x$. Recordemos que si $A \in \text{End}(\mathbb{C}^n)$, entonces la **norma** de A se define como:

$$\|A\| = \max\{|Ax| \mid |x| = 1\} \quad (3.2.1)$$

y es fácil verificar que: $\|A \cdot B\| \leq \|A\| \cdot \|B\|$ y $\|A + B\| \leq \|A\| + \|B\|$. Por otro lado, dada una serie de potencias

$$f(z) = \sum_{j=0}^{\infty} c_j z^j$$

con radio de convergencia R , ésta converge absoluta y uniformemente sobre cada bola $\{z \mid |z| < r\}$ con $r < R$. Por lo tanto la serie de matrices

$$f(A) := \sum_{j=0}^{\infty} c_j A^j \quad (3.2.2)$$

también converge uniformemente cuando $\|A\| < r$, puesto que:

$$\left\| \sum_{j=N}^{N+k} c_j A^j \right\| \leq \sum_{j=N}^{N+k} |c_j| \|A\|^j \leq \sum_{j=N}^{\infty} |c_j| r^j$$

el último término tiende a cero cuando $N \rightarrow \infty$.

Proposición 3.6. Si $T \in GL(n, \mathbb{C})$ entonces

$$Tf(A)T^{-1} = f(TAT^{-1}). \quad (3.2.3)$$

Demostración. Notemos que para k y N fijos se tiene:

$$\left\| \sum_{j=N}^{N+k} c_j (TAT^{-1})^j \right\| = \left\| T \left(\sum_{j=N}^{N+k} c_j A^j \right) T^{-1} \right\| \leq \|T\| \cdot \|T^{-1}\| \cdot \left\| \sum_{j=N}^{N+k} c_j A^j \right\|$$

y el último término tiende a cero cuando $N \rightarrow \infty$. Luego, la serie $f(TAT^{-1}) = \sum_{j=0}^{\infty} c_j (TAT^{-1})^j$ converge. Ahora la afirmación se obtiene conjugando término a término en las sumas parciales. \square

Las observaciones anteriores nos permiten definir la exponencial de matrices complejas $n \times n$ por medio de la serie:

$$\exp(A) = \sum_{j=0}^{\infty} \frac{1}{j!} A^j$$

La proposición anterior hace que el problema de calcular la exponencial de una matriz diagonalizable sea muy sencillo de resolver, puesto que si TAT^{-1} es diagonal:

$$T \cdot \exp(A) T^{-1} = \exp(TAT^{-1}),$$

de donde

$$\exp(A) = T^{-1} \cdot \exp(TAT^{-1}) \cdot T.$$

Sin embargo existen matrices que no son diagonalizables y el caso general se debe tratar por medio de la forma canónica de Jordan.

Proposición 3.7. *Si las matrices A y B conmutan, entonces*

$$\exp(A + B) = \exp(A) \cdot \exp(B)$$

Demostración.

$$\begin{aligned} \exp(A) \cdot \exp(B) &= \left(\sum_{k=0}^{\infty} \frac{A^k}{k!} \right) \left(\sum_{k=0}^{\infty} \frac{B^k}{k!} \right) \\ &= \sum_{k=0}^{\infty} \sum_{r=0}^k \frac{A^r B^{k-r}}{r!(k-r)!} \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{r=0}^k \binom{k}{r} A^r B^{k-r} \\ &= \sum_{k=0}^{\infty} \frac{(A+B)^k}{k!} \end{aligned}$$

\square

Observación 3.8. *Notemos que $\exp(A)$ siempre es una matriz invertible.*

Demostración. Esto se debe a que A y $-A$ conmutan y por lo tanto

$$I = \exp(0) = \exp(A - A) = \exp(A) \cdot \exp(-A)$$

Luego $\exp(-A) = \exp(A)^{-1}$. □

A continuación enunciamos sin demostración una de las propiedades más importantes de la exponencial.

Teorema 3.9. *Si A es una matriz compleja $n \times n$ entonces*

$$\exp(\operatorname{tr}(A)) = \det(\exp(A))$$

Demostración. ¹ □

A primera vista la ecuación del teorema anterior pareciera incorrecta porque el lado izquierdo sólo depende de los elementos de la diagonal de A y esto no es inmediatamente claro para el lado derecho. El punto aquí es que \det y \exp son invariantes bajo conjugación, al igual que la traza, por lo tanto si T es invertible

$$\det(\exp(TAT^{-1})) = \det(T \exp(A) T^{-1}) = \det(\exp(A)).$$

Finalmente notemos que

$$\exp(A^t) = \exp(A)^t, \quad \exp(\bar{A}) = \overline{\exp(A)}. \quad (3.2.4)$$

Pasemos ahora al caso de la función logaritmo. Así como $\exp(x)$ está definida para todo $x \in \mathbb{R}$ y $\log(x)$ está definido sólo para $x > 0$, el logaritmo de una matriz estará definido sólo para matrices cercanas a la identidad. De manera más precisa, si $\|A - I\| < 1$ definimos:

$$\log(A) = \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j} (A - I)^j$$

Proposición 3.10. *Las funciones \exp y \log tienen las siguientes propiedades:*

1. Si $\|A - I\| < 1$ entonces $\exp(\log(A)) = A$.
2. Si $\|\exp(B) - I\| < 1$ entonces $\log(\exp(B)) = B$.

Demostración. Como $\exp(B) - I = B + \frac{B^2}{2!} + \frac{B^3}{3!} + \frac{B^4}{4!} + \dots$, entonces

$$\begin{aligned} \log(\exp(B)) &= \left(B + \frac{B^2}{2!} + \frac{B^3}{3!} + \frac{B^4}{4!} + \dots \right) - \frac{1}{2} \left(B + \frac{B^2}{2!} + \frac{B^3}{3!} + \frac{B^4}{4!} + \dots \right)^2 \\ &\quad + \frac{1}{3} \left(B + \frac{B^2}{2!} + \frac{B^3}{3!} + \frac{B^4}{4!} + \dots \right)^3 - \frac{1}{4} \left(B + \frac{B^2}{2!} + \frac{B^3}{3!} + \frac{B^4}{4!} + \dots \right)^4 \\ &\quad + \dots \\ &= B + \left(\frac{B^2}{2!} - \frac{B^2}{2} \right) + \left(\frac{B^3}{6} - \frac{B^3}{2} + \frac{B^3}{3} \right) + \dots \\ &= B. \end{aligned}$$

¹ver [4]

Por otro lado, $\log(A) = (A - I) - \frac{(A - I)^2}{2} + \frac{(A - I)^3}{3} - \frac{(A - I)^4}{4} + \dots$, entonces

$$\begin{aligned} \exp(\log(A)) &= \left((A - I) - \frac{(A - I)^2}{2!} + \frac{(A - I)^3}{3} - \frac{(A - I)^4}{4} + \dots \right) \\ &+ \frac{1}{2!} \left((A - I) - \frac{(A - I)^2}{2!} + \frac{(A - I)^3}{3} - \frac{(A - I)^4}{4} + \dots \right)^2 \\ &+ \frac{1}{3!} \left((A - I) - \frac{(A - I)^2}{2} + \frac{(A - I)^3}{3} - \frac{(A - I)^4}{4} + \dots \right)^3 + \dots \\ &= A - \frac{(A - I)^2}{2} + \frac{(A - I)^2}{2} + \left(\frac{(A - I)^3}{3} - \frac{(A - I)^3}{2} + \frac{(A - I)^3}{6} \right) \\ &+ \dots \\ &= A. \end{aligned}$$

□

Observación 3.11. *Se puede probar que la función $\exp : \text{End}(\mathbb{C}^n) \rightarrow \text{GL}(n, \mathbb{C}^n)$ es sobreyectiva.*

Notemos que también podemos definir raíces, aunque no de manera única:

$$A^\alpha := \exp(\alpha \log(A)). \quad (3.2.5)$$

Sea U el conjunto abierto $U = \{A \in \text{End}(\mathbb{C}^n) \mid \|A\| < 1\}$ y pongamos $1 + U = \{I + A \mid A \in U\}$. Observemos que $\text{GL}(n, \mathbb{C})$ también es un conjunto abierto en $\text{End}(\mathbb{C}^n) \cong \mathbb{C}^{n \cdot n}$. En efecto, notemos que

$$\begin{aligned} \text{GL}(n, \mathbb{C}) &= \{A \in M_n(\mathbb{C}) \mid \det A \neq 0\} \\ &= \{A \in M_n(\mathbb{C}) \mid \det A \in \mathbb{C} \setminus \{0\}\} \\ &= \det^{-1}(\mathbb{C} \setminus \{0\}) \end{aligned}$$

y $\det : \text{End}(\mathbb{C}^n) \rightarrow \mathbb{C}$ es una función continua.

Observación 3.12. *El conjunto $1 + U$ es abierto en $\text{GL}(n, \mathbb{C})$ y se tiene una función logaritmo bien definida y continua*

$$\log : 1 + U \rightarrow \text{End}(\mathbb{C}^n)$$

tal que $\exp(\log(I + A)) = I + A$.

Demostración. Recordemos que la expansión en serie $\frac{1}{1+x} = 1 - x + x^2 - \dots$ tiene radio de convergencia 1, luego para $\|A\| < 1$ se tiene que $(I + A)^{-1} = I - A + A^2 - \dots$. Por lo tanto $I + A \in \text{GL}(n, \mathbb{C})$. Además la serie $\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots$ tiene radio de convergencia 1. En consecuencia el logaritmo está bien definido en $1 + U$. □

La función log definida en $1 + U$ resulta ser inyectiva, ya que exp es una inversa izquierda.

Lema 3.13. $\|\exp(A) - I\| \leq \exp(\|A\|) - 1$.

Demostración.

$$\begin{aligned} \|\exp(A) - I\| &= \left\| A + \frac{A^2}{2} + \frac{A^3}{3!} + \frac{A^4}{4!} \dots \right\| \\ &\leq \|A\| + \left\| \frac{A^2}{2} \right\| + \left\| \frac{A^3}{3!} \right\| + \left\| \frac{A^4}{4!} \right\| + \dots \\ &\leq \|A\| + \frac{\|A\|^2}{2!} + \frac{\|A\|^3}{3!} + \frac{\|A\|^4}{4!} + \dots \\ &= \exp(\|A\|) - 1 \end{aligned}$$

□

De lo anterior se deduce que la imagen de la función logaritmo en (3.12) contiene a la vecindad abierta del cero $W = \frac{1}{2}U$. En efecto, si $\|A\| < \frac{1}{2}$ entonces $\|\exp(A) - I\| < \sqrt{e} - 1 < 1$. Por lo tanto, $\exp(A) = 1 + (\exp(A) - 1) \in 1 + U$ y $\log(\exp(A)) = A$ para $\|A\| < \frac{1}{2}$ por (3.10).

Tenemos las vecindades $W = \{A \mid \|A\| < \frac{1}{2}\}$ del 0 en $End(\mathbb{C}^n)$, $V = \log^{-1}(W) = \exp(W)$ del 1 en $GL(n, \mathbb{C})$; y difeomorfismos inversos entre si:

$$1 \in V \xleftarrow[\log]{\exp} W \ni 0$$

Para los fines de este trabajo realmente no es necesario describir a estas vecindades con tanta precisión. De cualquier modo, las observaciones anteriores se pueden resumir en el siguiente resultado:

Teorema 3.14. *Existen vecindades V , de la identidad en $GL(n, \mathbb{C})$, W , del 0 en $End(\mathbb{C}^n)$ junto con difeomorfismos*

$$V \xrightarrow{\log} W \xrightarrow{\exp} V.$$

inversos uno del otro. Más aún, W puede elegirse de modo que $\forall A \in W$ se cumple: $|\operatorname{tr}(A)| < 2\pi$ y además $-A, A^t, \bar{A} \in W$.

Demostración. Mostremos solo la última parte. Claramente, la condición $|\operatorname{tr}(A)| < 2\pi$ se satisface en alguna vecindad $W_1 \subset W$ del 0, gracias a la continuidad de la traza. A continuación reemplazamos a W_1 por $W_1 \cap (-W_1) \cap (W_1^t) \cap \overline{W_1}$ con $W_1^t = \{A^t \mid A \in W_1\}$. □

3.3. Álgebras de Lie

Definición 3.15. *El álgebra de Lie $L = \mathfrak{gl}(n, K)$ es el espacio vectorial $End(K^n)$ de las matrices $n \times n$ con coeficientes en K junto con la operación dada por el **conmutador de matrices** (o corchete de Lie):*

$$[A, B] = AB - BA. \quad (3.3.1)$$

Esta estructura algebraica tiene las siguientes propiedades:

1. El álgebra de Lie L es un espacio vectorial sobre K y el corchete de Lie

$$L \times L \rightarrow L$$

$$(X, Y) = [X, Y]$$

es *bilineal*.

2. $[X, X] = 0$ y se cumple que $[X, Y] = -[Y, X] \quad \forall X, Y \in L$.

3. $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0 \quad \forall X, Y, Z \in L$, llamada la **identidad de Jacobi**.

Observación 3.16. Estas propiedades se verifican fácilmente en $\mathfrak{gl}(n, K)$.

En general, un espacio vectorial L equipado con una operación $[\cdot, \cdot]$ que satisface las propiedades anteriores, se conoce como álgebra de Lie.

Definición 3.17. Un homomorfismo de álgebras de Lie $\alpha : L \rightarrow L'$ es una transformación lineal tal que

$$[\alpha(X), \alpha(Y)] = \alpha([X, Y])$$

para todo $X, Y \in L$.

Nos interesan principalmente las siguientes álgebras de Lie reales y complejas, contenidas en $\mathfrak{gl}(n, \mathbb{C})$ con el corchete de Lie definido anteriormente.

Definición 3.18. Definimos a $\mathfrak{sl}(n, \mathbb{R})$ y $\mathfrak{sl}(n, \mathbb{C})$ como las álgebras de Lie de matrices $n \times n$ de traza cero, con coeficientes en \mathbb{R} y \mathbb{C} respectivamente. Definimos también a $\mathfrak{so}(n)$ como el álgebra de Lie de las matrices $n \times n$ antisimétricas.

Una matriz A $n \times n$ se dice **antisimétrica** cuando $A^t = -A$. Toda matriz $n \times n$ se puede escribir de manera única como la suma de una matriz simétrica y una antisimétrica:

$$A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t) \quad (3.3.2)$$

Definición 3.19. Definimos a $\mathfrak{u}(n)$ como el álgebra de Lie de las matrices complejas antihermitianas y a $\mathfrak{su}(n) = \mathfrak{u}(n) \cap \mathfrak{sl}(n, \mathbb{C})$ como la subálgebra de Lie de las matrices antihermitianas de traza cero.

Una matriz compleja A $n \times n$ se dice **antihermitiana** si $A^t = -\bar{A}$, en otras palabras $A^* = -A$. Similarmente, toda matriz $n \times n$ se puede expresar de manera única como suma de una matriz hermitiana y una antihermitiana:

$$A = \frac{1}{2}(A + A^*) + \frac{1}{2}(A - A^*). \quad (3.3.3)$$

Nótese que los conjuntos de matrices hermitianas y antihermitianas son subespacios vectoriales *reales* de $End_{\mathbb{C}}(\mathbb{C}^n)$ pero no un subespacios vectoriales sobre \mathbb{C} . Por ejemplo, A es hermitiana si y sólo si iA es antihermitiana, por lo tanto multiplicación por i intercambia los sumandos de la descomposición (3.3.2).

Definición 3.20. *Definimos a $\mathfrak{d}(n)$ como el álgebra de Lie de las matrices diagonales reales (con el corchete de Lie trivial). Definimos también a $\mathfrak{n}^+(n, \mathbb{R})$, $\mathfrak{n}^+(n, \mathbb{C})$ como las álgebras de Lie de matrices triangulares superiores, nilpotentes reales ó complejas respectivamente (esto es, triangulares superiores con diagonal cero).*

Es fácil ver que todos estos subespacios vectoriales son álgebras de Lie con el corchete dado por (3.3.1). Notemos además que la notación para todos ellos coincide con la de los grupos de matrices, con la única diferencia de que los grupos se denotan con letras mayúsculas y las álgebras de Lie con letras góticas. Por convención, si G es uno de tales grupos de matrices, LG denotará a la correspondiente álgebra de Lie. Por lo tanto:

$$\begin{aligned}
 \mathfrak{sl}(n, \mathbb{R}) &= L SL(n, \mathbb{R}) \\
 \mathfrak{sl}(n, \mathbb{C}) &= L SL(n, \mathbb{C}) \\
 \mathfrak{so}(n) &= L SO(n) \\
 \mathfrak{u}(n) &= L U(n) \\
 \mathfrak{su}(n) &= L SU(n) \\
 \mathfrak{d}(n) &= L D(n) \\
 \mathfrak{n}^+(n, \mathbb{R}) &= L N^+(n, \mathbb{R}) \\
 \mathfrak{n}^+(n, \mathbb{C}) &= L N^+(n, \mathbb{C}).
 \end{aligned}
 \tag{3.3.4}$$

Teorema 3.21. *Sea W una vecindad del 0 en $\mathfrak{gl}(n, \mathbb{C})$ que satisface las condiciones del teorema (3.14). Entonces para todos los grupos G y su correspondiente álgebra de Lie LG en (3.3.4), la función exponencial induce un homeomorfismo*

$$\exp : W \cap LG \rightarrow V \cap G$$

de una vecindad del 0 en LG en una vecindad del I en G .

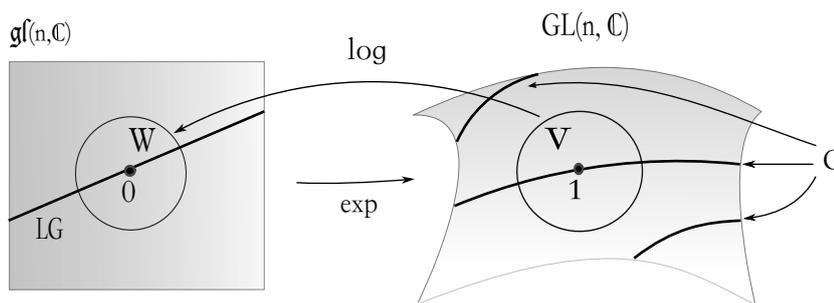


Figura 3.2: Teorema 3.21

Demostración. En cada caso, la función $\exp : W \rightarrow V$ es un homeomorfismo que tiene como inversa a la función \log . Si $\text{tr}(A) = 0$, entonces $\det(\exp(A)) = e^{\text{tr}(A)} = e^0 = 1$ por (3.9). Por lo tanto $A \in \mathfrak{sl}(n, \mathbb{C})$ implica que $\exp(A) \in SL(n, \mathbb{C})$ y obviamente $A \in \mathfrak{sl}(n, \mathbb{R})$ implica que $\exp(A) \in SL(n, \mathbb{R})$

Si $A + A^t = 0$, entonces

$$\begin{aligned} \exp(A) \cdot (\exp(A))^t &= \exp(A) \cdot \exp(A^t) \\ &= \exp(A + A^t) \\ &= \exp(0) \\ &= I. \end{aligned}$$

Si además $A \in \mathfrak{so}(n)$, entonces $\det(\exp(A)) = 1$ (como el párrafo anterior). Luego, si $A \in \mathfrak{so}(n)$ entonces $\exp(A) \in SO(n)$. Similarmente, si $A + A^* = 0$, entonces

$$\begin{aligned} \exp(A) \cdot (\exp(A))^* &= \exp(A) \cdot \exp(A^*) \\ &= \exp(A + A^*) \\ &= \exp(0) \\ &= I \end{aligned}$$

y si además $\operatorname{tr}(A) = 0$, entonces $\det(\exp(A)) = 1$. Por lo tanto, si $A \in \mathfrak{u}(n)$ entonces $\exp(A) \in U(n)$ y si $A \in \mathfrak{su}(n)$ entonces $\exp(A) \in SU(n)$. Algo similar se cumple trivialmente para los grupos restantes. En consecuencia

$$\exp(LG) \subset G.$$

Recíprocamente, si $A \in W$ y $\exp(A) \in SL(n, \mathbb{C})$, entonces $e^{\operatorname{tr}(A)} = \det(\exp(A)) = 1$, de donde se sigue que $\operatorname{tr}(A) = 0$ ya que $|\operatorname{tr}(A)| < 2\pi$. Por lo tanto $A \in \mathfrak{sl}(n, \mathbb{C})$. Si además $\exp(A)$ es una matriz real, entonces $\exp(A) = \overline{\exp(A)} = \exp(\overline{A})$, de donde $A = \overline{A}$ (pues la función exponencial es inyectiva en W), es decir, A es real y por lo tanto $A \in \mathfrak{sl}(n, \mathbb{R})$. Más aún (también en el caso real), si $\exp(A) \cdot \exp(A)^t = I$, entonces $\exp(A^t) = \exp(-A)$ y como $-A \in W$ se sigue que $A^t = -A$, es decir $A \in \mathfrak{so}(n)$. Similarmente se prueba que si $A \in W$ y $\exp(A) \in U(n)$ entonces $A \in \mathfrak{u}(n)$ y en consecuencia, si $\exp(A) \in SU(n)$ entonces $A \in \mathfrak{su}(n)$.

Finalmente $\exp(A)$ es una matriz diagonal, entonces también lo es $A = \log(\exp(A))$ y si $\exp A = 1 + N$ es una matriz triangular superior unipotente entonces $A = \log(\exp(A)) = N - N^2/2 + \dots$ es una matriz triangular superior nilpotente. \square

Este teorema nos habla acerca de la estructura geométrica de los grupos. La función \exp es un difeomorfismo de una vecindad del 0 en $\mathfrak{gl}(n, \mathbb{C})$ en una vecindad del I en $GL(n, \mathbb{C})$. Veremos a continuación que localmente los subgrupos que hemos considerado $G \subset GL(n, \mathbb{C})$ van bajo \log a subespacios de $\mathfrak{gl}(n, \mathbb{C}) \cong \mathbb{C}^{n \cdot n}$. Excepto por el difeomorfismo G que localmente se ve como un subconjunto de $GL(n, \mathbb{C})$ del 1, al igual que LG se ve localmente como un subconjunto de $\mathfrak{gl}(n, \mathbb{C}) \cong \mathbb{C}^{n \cdot n}$ del 0. Para otro punto $A \in G$ la situación no es distinta: tenemos el difeomorfismo

$$(W, 0) \xrightarrow{\exp} (V, 1) \xrightarrow{A \cdot} (A \cdot V, A)$$

con $A \cdot V = \{A \cdot B \mid B \in V\}$ y es la vecindad $W \cap LG$ del 0 en LG en la vecindad $A \cdot V \cap G$ de A en G . Con los conceptos de topología diferencial podemos resumir con el siguiente teorema:

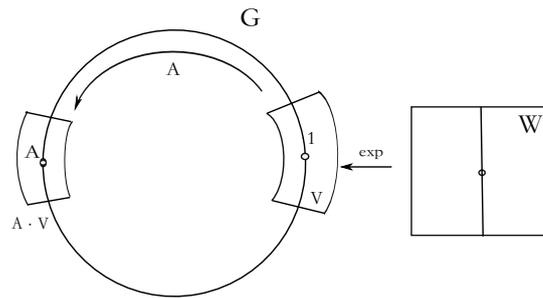


Figura 3.3

Teorema 3.22. *Los subgrupos $GL(n, \mathbb{R})$, $SL(n, \mathbb{C})$, $SL(n, \mathbb{R})$, $SO(n)$, $O(n)$, $U(n)$, $SU(n)$, $D(n)$, $N^+(n, \mathbb{C})$ y $N^+(n, \mathbb{R})$ son subvariedades suaves de $GL(n, \mathbb{C})$. El producto $G \times G \rightarrow G$ es una función suave y la función $G \rightarrow G$, $g \mapsto g^{-1}$ es suave.*

Demostración. Mostraremos la última parte, para ello sólo se necesita verificar la afirmación para $GL(n, \mathbb{C})$. Notemos que la multiplicación de matrices es diferenciable y la función $g \mapsto g^{-1} = \det(g)^{-1} \text{Adj}(g)$, por la regla de Cramer también es diferenciable. \square

Definición 3.23. *Un grupo de Lie G es un grupo topológico con la estructura de variedad suave, tal que las operaciones $\cdot : G \times G \rightarrow G$, $^{-1} : G \rightarrow G$ son suaves.*

Por definición, la dimensión del grupo de Lie es la dimensión del álgebra de Lie asociada.

Observación 3.24.

$$\begin{array}{ll}
 \dim GL(n, \mathbb{C}) &= 2n^2, & \dim O(n) &= n(n-1)/2 \\
 \dim GL(n, \mathbb{R}) &= n^2, & \dim SO(n) &= n(n-1)/2, \\
 \dim SL(n, \mathbb{C}) &= 2n^2 - 2, & \dim D(n) &= n, \\
 \dim SL(n, \mathbb{R}) &= n^2 - 1, & \dim N^+(n, \mathbb{C}) &= n(n-1), \\
 \dim U(n) &= n^2, & \dim N^+(n, \mathbb{R}) &= n(n-1)/2. \\
 \dim SU(n) &= n^2 - 1, & &
 \end{array}$$

3.4. La representación adjunta

Hasta el momento hemos especificado las álgebras de Lie LG de los grupos G considerados, aunque uno también puede hallarlos de la siguiente manera. Consideremos curvas suaves $\varphi(t)$ con $t \in (-\epsilon, \epsilon)$ tales que $\varphi(0) = I$ y $\varphi(t) \in G$ para toda t . A tales curvas les corresponde un vector tangente (o vector velocidad)

$$\dot{\varphi}(t) = \left. \frac{d}{dt} \right|_{t=0} \varphi(t) \in \mathfrak{gl}(n, \mathbb{C})$$

y se cumple lo siguiente:

Observación 3.25. *LG es el conjunto de todos los vectores de la forma $\dot{\varphi}(0)$, donde φ corre sobre todas las curvas suaves $\varphi : (-\epsilon, \epsilon) \rightarrow G$ con $\varphi(0) = I$. Esto es, LG es el espacio tangente de G en I .*

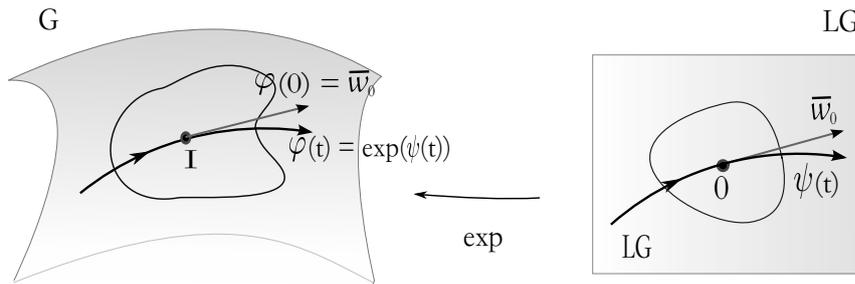


Figura 3.4: Observación 3.25

Demostración. Recordemos que la función exponencial es localmente invertible y su diferencial es 0 en la identidad. Nótese que al componer las curvas φ con \log , obtenemos curvas de la función $\psi : (-\epsilon, \epsilon) \rightarrow LG$, cuyos vectores tangentes son todos los vectores de LG . Luego, si $\bar{w}_0 \in LG$ existe ψ tal que $\psi(0) = 0$ y $\dot{\psi}(0) = \bar{w}_0$. Por lo tanto, si definimos $\varphi(t) = \exp(\psi(t))$ tenemos: $\dot{\varphi}(t) = \exp'(\psi(t)) \cdot \dot{\psi}(t)$, de donde $\dot{\varphi}(0) = \exp'(0) \cdot \dot{\psi}(0) = \bar{w}_0$. \square

De manera más explícita tenemos lo siguiente:

Observación 3.26. Si $A \in \text{End}_{\mathbb{C}}(\mathbb{C}^n)$ es una matriz $n \times n$ fija, entonces la función $\varphi_A : \mathbb{R} \rightarrow GL(n, \mathbb{C})$ dada por $\varphi_A(t) = \exp(tA)$ es un homomorfismo de grupos y satisface que $\dot{\varphi}_A(0) = A$.

Demostración. La primera afirmación es equivalente a $\exp(t+s)A = \exp(tA) \cdot \exp(sA)$, pero esta igualdad se sigue del hecho de que tA y sA conmutan. La segunda afirmación se sigue de la serie $\exp(tA) = 1 + tA + \frac{(tA)^2}{2!} + \dots$ \square

Es importante enfatizar que la función exponencial no es un homomorfismo de grupos ya que en general la ecuación $\exp(A+B) = \exp(A) \cdot \exp(B)$ no se cumple. Un homomorfismo $\varphi : \mathbb{R} \rightarrow G$ se conoce como un **subgrupo uniparamétrico** de G . Así, LG es el espacio de los vectores velocidad $\dot{\varphi}(0)$ de todos los subgrupos uniparamétricos de G .

Estas observaciones nos serán de utilidad para establecer la relación del corchete de Lie con la estructura del grupo y demostrar por qué el espacio LG (que bajo \exp se mapea en G de manera localmente sobreyectiva), debe tener estructura de álgebra de Lie.

Notemos que la conjugación con $T \in GL(n, \mathbb{C})$ induce las transformaciones:

$$\begin{aligned} ad_T : \mathfrak{gl}(n, \mathbb{C}) &\rightarrow \mathfrak{gl}(n, \mathbb{C}), & X &\mapsto TXT^{-1} \\ c_T : GL(n, \mathbb{C}) &\rightarrow GL(n, \mathbb{C}), & A &\mapsto TAT^{-1} \end{aligned}$$

y la ecuación (3.2.3) muestra que $\exp \circ ad_T = c_T \circ \exp$, es decir, el siguiente diagrama es conmutativo:

$$\begin{array}{ccc} \mathfrak{gl}(n, \mathbb{C}) & \xrightarrow{ad_T} & \mathfrak{gl}(n, \mathbb{C}) \\ \exp \downarrow & & \downarrow \exp \\ GL(n, \mathbb{C}) & \xrightarrow{c_T} & GL(n, \mathbb{C}) \end{array} \tag{3.4.1}$$

Si tomamos ahora $T \in G$ y $Y \in LG$, entonces $c_T \exp(sY) \in G$ para todo $s \in \mathbb{R}$. Por lo tanto $\exp ad_T(sY) \in G$ y la derivada en 0 de este subgrupo uniparamétrico es

$$\left. \frac{d}{ds} \right|_0 \exp ad_T(sY) = ad_T Y = TYT^{-1} \in LG,$$

es decir, el conjugado de un elemento de LG por un elemento de G es de nuevo un elemento de LG .

Finalmente, supongamos que $\varphi : \mathbb{R} \rightarrow G$ es un subgrupo uniparamétrico con $\dot{\varphi}(0) = X \in LG$ y sea $Y \in LG$. Entonces $\varphi(t)Y\varphi(-t) \in LG$ y por lo tanto

$$\left. \frac{d}{dt} \right|_0 \varphi(t)Y\varphi(-t) \in LG.$$

Desarrollando en serie de Taylor tenemos que:

$$\begin{aligned} \varphi(t) &= I + tX + \dots \\ \varphi(t)Y\varphi(-t) &= (I + tX + \dots)Y(I - tX - \dots) \\ &= Y + t(XY - YX) + \dots, \end{aligned}$$

En particular, al tomar $\varphi(t) = \exp(tX)$ tenemos:

$$[X, Y] = \left. \frac{d}{dt} \right|_0 \exp(tX) \cdot Y \cdot \exp(-tX). \quad (3.4.2)$$

Teorema 3.27. *Si $T \in G$, entonces T induce el automorfismo de álgebras de Lie $ad_T : LG \rightarrow LG$ dado por $ad_T(X) = TXT^{-1}$ y en consecuencia G actúa sobre LG por automorfismos de álgebras de Lie. En otras palabras, se tiene el homomorfismo de grupos*

$$G \rightarrow \text{Aut}(LG), \quad T \mapsto ad_T.$$

Tal homomorfismo se denomina la **representación adjunta** del grupo G .

Demostración. Se debe mostrar que $[ad_T(X), ad_T(Y)] = ad_T([X, Y])$. En efecto:

$$\begin{aligned} [ad_T(X), ad_T(Y)] &= [T(X)T^{-1}, T(Y)T^{-1}] \\ &= (TXT^{-1})(TYT^{-1}) - (TYT^{-1})(TXT^{-1}) \\ &= (TXYT^{-1}) - (TYXT^{-1}) \\ &= T[X, Y]T^{-1} \\ &= ad_T([X, Y]). \end{aligned}$$

Resta mostrar que $ad_{TS}(X) = ad_T \circ ad_S$:

$$\begin{aligned} ad_{TS}(X) &= TSX(TS)^{-1} \\ &= TSXS^{-1}T^{-1} \\ &= ad_T(SXS^{-1}) \\ &= ad_T \circ ad_S. \end{aligned}$$

□

En vez de continuar con la teoría de Lie general, nos concentraremos en los grupos $SU(n)$ y $SO(m)$. Para ésto, dotemos primero a $\mathfrak{su}(n)$ del producto escalar real euclidiano dado por:

$$\langle A, B \rangle := -\frac{1}{2}tr(A \cdot B). \quad (3.4.3)$$

Es claro que $\langle \cdot, \cdot \rangle$ es bilineal, además es real ya que:

$$\begin{aligned} \overline{tr(AB)} &= tr(\overline{A} \ \overline{B}) \\ &= tr(A^t B^t) \\ &= tr(BA)^t \\ &= tr(BA) \\ &= tr(AB). \end{aligned}$$

Este producto es definido positivo debido a que:

$$\begin{aligned} 2\langle A, A \rangle &= -tr(A \cdot A) \\ &= tr(A \cdot A^*) \\ &= \sum_{i,j} a_{ij} \cdot \bar{a}_{ij} \\ &= \sum_{i,j} |a_{ij}|^2. \end{aligned}$$

Por lo tanto, $\mathfrak{su}(n)$ es un espacio euclidiano de dimensión $n^2 - 1$ (sobre los reales).

Teorema 3.28. *Si $T \in U(n)$ entonces*

$$ad_T : \mathfrak{su}(n) \rightarrow \mathfrak{su}(n), \quad X \mapsto TXT^{-1}$$

es una transformación ortogonal. En consecuencia ad_T define un homomorfismo

$$ad : U(n) \rightarrow SO(\mathfrak{su}(n)) \cong SO(n^2 - 1)$$

$$T \mapsto ad_T.$$

(Si V es un espacio euclidiano entonces $SO(V)$ es el grupo de automorfismos ortogonales de V con determinante 1).

Demostración. La transformación $ad_T : \mathfrak{su}(n) \rightarrow \mathfrak{su}(n)$ es ortogonal porque:

$$\begin{aligned} \langle ad_T(X), ad_T(Y) \rangle &= -\frac{1}{2}tr(TXT^{-1})(TYT^{-1}) \\ &= -\frac{1}{2}tr(TXYT^{-1}) \\ &= -\frac{1}{2}tr(XY) \\ &= \langle X, Y \rangle. \end{aligned}$$

□

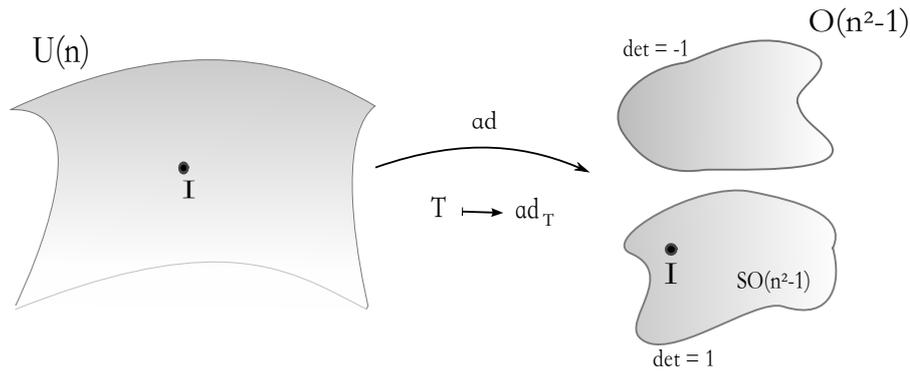


Figura 3.5: Teorema 3.28

y que el determinante de ad_T es positivo se sigue del siguiente lema:

Lema 3.29. $U(n)$ es conexo.

Demostración. Si $T \in U(n)$ está en la forma canónica de Jordan, entonces es una matriz diagonal de la forma

$$T = \begin{pmatrix} e^{i\varphi_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\varphi_n} \end{pmatrix}.$$

Luego, es posible dar una trayectoria $\gamma : [0, 1] \rightarrow U(n)$, a saber

$$\gamma(t) = \begin{pmatrix} e^{i\varphi_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\varphi_n} \end{pmatrix} \in U(n),$$

tal que $\gamma(0) = I$ y $\gamma(1) = T$.

Ahora bien, si $B = STS^{-1} \in U(n)$ con T como antes, consideremos la trayectoria

$$\rho(t) = S \cdot \gamma(t) \cdot S^{-1} \in U(n),$$

la cual comienza en I y termina en $STS^{-1} = B$. Por lo tanto, la función

$$U(n) \xrightarrow{ad} O(n^2 - 1) \xrightarrow{\det} \mathbb{R} \setminus \{0\}$$

debe tener signo constante y para $T = I$ se tiene que $\det(ad_T) = 1$. \square

Un resultado análogo se tiene ahora para el grupo $SO(n)$ ya que $SO(n)$ es el subgrupo de matrices reales en $SU(n)$ y del mismo modo, $\mathfrak{so}(n)$ es el subespacio de matrices reales en $\mathfrak{su}(n)$. Se sigue entonces del teorema anterior:

Teorema 3.30. Si $T \in SO(n)$ entonces

$$ad_T : \mathfrak{so}(n) \rightarrow \mathfrak{so}(n), \quad X \mapsto TXT^{-1}$$

es una transformación ortogonal. Por lo tanto, ad_T define un homomorfismo

$$ad : SO(n) \rightarrow SO(\mathfrak{so}(n)) \cong SO\left(\frac{n(n-1)}{2}\right)$$

$$T \mapsto ad_T.$$

Capítulo 4

Cuaterniones y grupos ortogonales

Los grupos ortogonales y unitarios en dimensiones bajas tienen sus propias características y son de gran importancia en física teórica. La estructura de estos grupos la describiremos a través del estudio del álgebra de cuaterniones lo que nos llevará a estudiar otros grupos lineales fuera de los que hemos visto en los capítulos anteriores. Por último, hablaremos acerca de la estructura del grupo de Lorentz.

4.1. El grupo $SO(3)$ y su álgebra de Lie

La representación adjunta de $SU(2)$ en (3.28) es un homomorfismo

$$ad : SU(2) \rightarrow SO(3),$$

el cual establece una estrecha relación entre estos grupos. Por otro lado, la representación adjunta de $SO(3)$ es un isomorfismo $SO(3) \rightarrow SO(3)$. El corchete de Lie en $\mathfrak{so}(3)$ corresponde al producto cruz en \mathbb{R}^3 y este hecho explica la aparición del producto cruz en la mecánica clásica. La afirmación exacta es la siguiente.

Teorema 4.1. *Existe un isomorfismo de espacios vectoriales $\kappa : \mathfrak{so}(3) \rightarrow \mathbb{R}^3$ con las siguientes propiedades:*

1. $|\kappa(X)|^2 = -\frac{1}{2}tr(X^2)$, lo que implica que κ es ortogonal.
2. $\kappa(TXT^{-1}) = T(\kappa(X))$ para $T \in SO(3)$, es decir, el siguiente diagrama es conmutativo:

$$\begin{array}{ccc} \mathfrak{so}(3) & \xrightarrow{ad_T} & \mathfrak{so}(3) \\ \kappa \downarrow & & \downarrow \kappa \\ \mathbb{R}^3 & \xrightarrow{T} & \mathbb{R}^3 \end{array}$$

3. $\kappa([X, Y]) = \kappa(X) \times \kappa(Y)$, esto es, el siguiente diagrama conmuta:

$$\begin{array}{ccc} \mathfrak{so}(3) \times \mathfrak{so}(3) & \xrightarrow{[\cdot, \cdot]} & \mathfrak{so}(3) \\ \kappa \times \kappa \downarrow & & \downarrow \kappa \\ \mathbb{R}^3 \times \mathbb{R}^3 & \xrightarrow{\times} & \mathbb{R}^3 \end{array}$$

4. Si $X \in \mathfrak{so}(3)$, entonces $\exp(X) \in SO(3)$ es una rotación en \mathbb{R}^3 con eje de rotación $\kappa(X)$.

5. El isomorfismo inverso $\kappa^{-1} : \mathbb{R}^3 \rightarrow \mathfrak{so}(3)$ manda al vector v en la transformación lineal antisimétrica:

$$\begin{aligned} \kappa^{-1}(v) : \mathbb{R}^3 &\rightarrow \mathbb{R}^3 \\ w &\mapsto v \times w \end{aligned}$$

6. La forma explícita para κ es la siguiente:

$$\begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Demostración. Definimos $\kappa^{-1} : \mathbb{R}^3 \rightarrow \text{End}_{\mathbb{R}}(\mathbb{R}^3)$ como en el inciso (5) y mostremos que $\kappa^{-1}(v)$ es antisimétrica: sean $v, u, w \in \mathbb{R}^3$

$$\begin{aligned} \langle \kappa^{-1}(v)u, w \rangle &= \langle v \times u, w \rangle \\ &= \det(v, u, w) \\ &= -\det(v, w, u) \\ &= -\langle v \times w, u \rangle \\ &= -\langle u, v \times w \rangle \\ &= -\langle u, \kappa^{-1}(v)w \rangle. \end{aligned}$$

Al usar coordenadas, vemos que $\kappa^{-1}(xe_1 + ye_2 + ze_3)$ es la transformación dada por:

$$e_1 \mapsto \begin{pmatrix} 0 \\ z \\ -y \end{pmatrix}, \quad e_2 \mapsto \begin{pmatrix} -z \\ 0 \\ x \end{pmatrix}, \quad e_3 \mapsto \begin{pmatrix} y \\ -x \\ 0 \end{pmatrix},$$

de donde la matriz de $\kappa^{-1}(xe_1 + ye_2 + ze_3)$ con respecto a la base e_1, e_2, e_3 está dada por:

$$\begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix}.$$

Esto muestra que κ^{-1} es un isomorfismo, cuyo inverso está dado por el inciso (6).

Para ver el inciso (4) notemos que: de (5) se tiene

$$\kappa^{-1}(v) : v \mapsto 0$$

por lo tanto, si $\kappa(X) = v$ entonces $X \cdot \kappa(X) = \kappa^{-1}(v)v = 0$, así

$$\exp(X) \cdot \kappa(X) = \left(1 + X + \frac{X^2}{2!} + \dots\right) \cdot \kappa(X) = \kappa(X)$$

por lo que $\kappa(X)$ permanece fijo bajo $\exp(X)$.

Para probar (1): se tiene que $-\frac{1}{2}\text{tr}(X^2) = \frac{1}{2}\text{tr}(X^t \cdot X) = \sum_{i,j} x_{ij}^2$ y la afirmación se sigue de (6).

Probemos (2): para todo $v, w \in \mathbb{R}^3$ y $T \in SO(3)$ se cumple:

$$T(v) \times T(w) = T(v \times w), \quad (4.1.1)$$

ésto es, $\kappa^{-1}(T(v))T(w) = T(v) \times T(w) = T\kappa^{-1}(v)w$. Como la ecuación anterior es cierta para todo w , se tiene que

$$\kappa^{-1}(T(v)) \cdot T = T \cdot \kappa^{-1}(v)$$

o lo que es lo mismo

$$\kappa^{-1}(T(v)) = T \cdot \kappa^{-1} \cdot T^{-1}.$$

Poniendo $X = \kappa^{-1}(v)$ tenemos:

$$\kappa^{-1}(T\kappa(X)) = T \cdot X \cdot T^{-1}.$$

de donde $T(\kappa(X)) = \kappa(T \cdot X \cdot T^{-1})$.

Finalmente veamos (3): si $X = \kappa^{-1}(v), Y = \kappa^{-1}(w)$, entonces debemos probar que:

$$\kappa^{-1}(v \times w) = [\kappa^{-1}(v), \kappa^{-1}(w)].$$

Aplicando ambos lados a u , obtenemos:

$$(v \times w) \times u = v \times (w \times u) - w \times (v \times u)$$

lo cual es equivalente a la identidad de Jacobi para el producto cruz. \square

Supongamos que un cuerpo rígido B gira alrededor de un punto fijo $0 \in B$ y elijamos un sistema de coordenadas euclidiano sobre B , centrado en 0 y al mismo tiempo escogemos un sistema de coordenadas (en reposo) con origen 0 en el espacio ambiente \mathbb{R}^3 , en donde B se mueve. Entonces el movimiento de B está descrito por una curva $t \mapsto A(t) \in SO(3)$, a saber: $A(t)$ es la matriz que manda el sistema de coordenadas en reposo en el sistema de coordenadas móvil (sobre B), en el instante t .

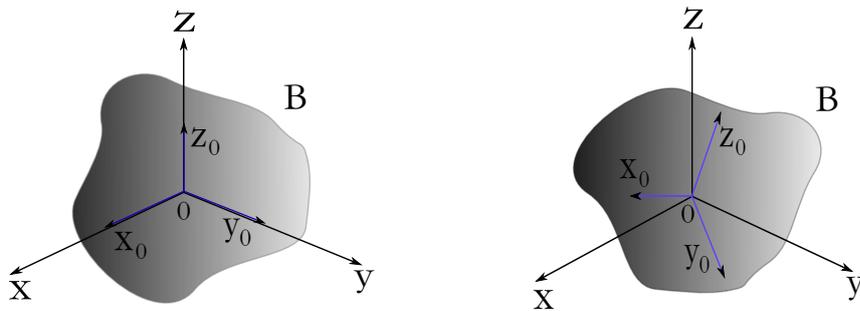


Figura 4.1: Cuerpo rígido B que gira alrededor de un punto fijo.

Ahora bien, la elección de los sistemas de coordenadas fué hecha de tal manera que $A(0) = id$, luego como se ha visto anteriormente

$$\left. \frac{d}{dt} \right|_{t=0} A(t) = X \in \mathfrak{so}(3) \cong \mathbb{R}^3.$$

Por lo tanto, la velocidad de rotación está descrita por un vector $\omega = \kappa(X)$ en la dirección del eje de rotación instantáneo, como se afirma en inciso (4). La velocidad de un punto $q \in B$ está dada entonces por:

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=0} A(t) \cdot q &= \dot{A}(0) \cdot q \\ &= X \cdot q \\ &= \kappa^{-1}(\omega) \cdot q \\ &= \omega \times q, \end{aligned}$$

por inciso (5).

Queremos estudiar con más precisión a la rotación e^{tX} con velocidad angular $X \in \mathfrak{so}(3)$ en el tiempo $t = 0$. Sabemos que $\exp(TXT^{-1}) = T \exp(X)T^{-1}$ y que $\kappa(TXT^{-1}) = T \cdot \kappa(X)$. Por lo tanto, dado X , podemos escoger T de tal manera que $\kappa(TXT^{-1}) = t \cdot e_3$ para algún $t \in \mathbb{R}$. Así tenemos:

$$\begin{aligned} \exp(t\kappa^{-1}(e_3)) &= T \cdot \exp(X) \cdot T^{-1} \\ &= \exp(TXT^{-1}). \end{aligned}$$

ó

$$T^{-1} \cdot \exp(t\kappa^{-1}(e_3)) \cdot T = \exp(X),$$

es decir, el siguiente diagrama es conmutativo:

$$\begin{array}{ccc} \mathbb{R}^3 & \xrightarrow{\exp(X)} & \mathbb{R}^3 \\ T \downarrow & & \downarrow T \\ \mathbb{R}^3 & \xrightarrow{\exp(t\kappa^{-1}(e_3))} & \mathbb{R}^3 \end{array}$$

Por lo tanto, es importante estudiar el grupo uniparamétrico de rotaciones:

$$\{\exp(t\kappa^{-1}(e_3)) \mid t \in \mathbb{R}\}.$$

Como tales rotaciones fijan a e_3 , se trata esencialmente de rotaciones del plano $L(e_1, e_2)$ y por el teorema (4.1) inciso (6) son de la forma:

$$\exp \left(t \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right)$$

En notación compleja, la primera matriz corresponde simplemente multiplicar por i :

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} : \mathbb{R}^2 \cong \mathbb{C} \xrightarrow{i} \mathbb{C} \cong \mathbb{R}^2$$

$$1 \mapsto i$$

$$i \mapsto -1$$

Por lo tanto:

$$\exp(ti) = \cos(t) + i \sin(t) = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}.$$

Si restringimos t al intervalo $I = [-\pi, \pi]$, entonces para $t_1, t_2 \in I$ se tiene que $\exp(t_1 i) = \exp(t_2 i)$ sólo si $t_1 = t_2$ ó $t_1 = -t_2 = \pm\pi$.

Consideremos ahora la bola cerrada de centro 0 y radio π en \mathbb{R}^3 :

$$D_\pi^3 = \{v \in \mathbb{R}^3 \mid |v| \leq \pi\}.$$

Teorema 4.2. *La función*

$$\delta : D_\pi^3 \rightarrow SO(3)$$

$$v \mapsto \exp(\kappa^{-1}(v))$$

es sobreyectiva y cumple que: $\delta(v) = \delta(w)$ si y sólo si $v = w$ ó $v = -w$ y $|w| = \pi$.

Demostración. La función δ es sobreyectiva porque a partir de $\exp(t\kappa^{-1}(e_3))$ se obtienen todas las rotaciones alrededor del eje e_3 con $t \in [-\pi, \pi]$ y si $v = T(e_3)$ entonces a partir de $\exp(t\kappa^{-1}(v)) = \exp(t\kappa^{-1}(T(e_3))) = T \exp(t\kappa^{-1}e_3)T^{-1}$ se obtienen todas las rotaciones en el eje $T(e_3)$.

Ahora, si $\delta(v) = \delta(w)$ entonces v, w deben ser vectores linealmente dependientes. De lo contrario $\exp(v) = \exp(w)$ sería una rotación que deja fijos a dos vectores linealmente independientes, luego $\exp(v) = \exp(w) = id$ y se tendría que $v = w = 0$. Por lo tanto si $\delta(v) = \delta(w)$ entonces $\delta(v)$ y $\delta(w)$ tienen el mismo eje de rotación, entonces $v = w$ ó $v = -w$ y $|v| = |w| = \pi$ pues ambos tienen longitud menor o igual a π . \square

Este teorema nos da una fórmula explícita para producir una rotación a partir de un vector dado v , de modo que el eje de rotación está dado por la dirección de v y el ángulo de rotación por $|v|$ (donde las rotaciones por π y $-\pi$ sobre el mismo eje producen el mismo efecto). El teorema nos da por lo tanto una descripción geométrica del espacio $SO(3)$.

Corolario 4.3. *La función δ induce un homeomorfismo de $\mathbb{R}P^3$ en $SO(3)$.*

Demostración. Al identificar puntos antipodales en la frontera de D_π^3 , δ induce un homeomorfismo entre $D_\pi^3 / \sim = \mathbb{R}P^3$ y $SO(3)$.

$$\begin{array}{ccc} D_\pi^3 & \xrightarrow{\delta} & SO(3) \\ p \downarrow & \nearrow \approx & \\ D_\pi^3 / \sim & & \end{array}$$

\square

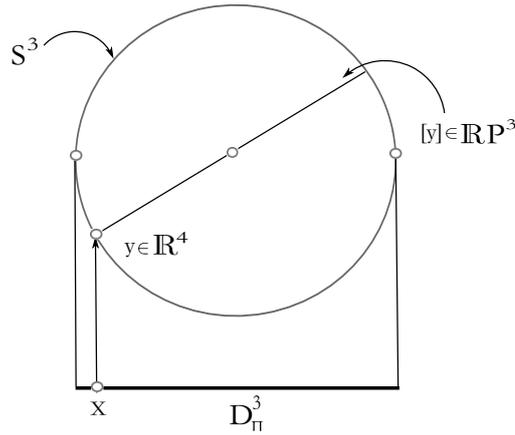


Figura 4.2: Corolario 4.3

4.2. Cuaterniones

El álgebra (real) \mathbb{H} de los **cuaterniones** es el álgebra de las matrices complejas 2×2 de la forma

$$A = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}. \quad (4.2.1)$$

Análogamente a la conjugación en \mathbb{C} , \mathbb{H} posee el antiautomorfismo:

$$\mathbb{H} \rightarrow \mathbb{H}, \quad A \mapsto A^* = \bar{A}^t$$

el cual satisface:

$$\begin{aligned} (A + B)^* &= A^* + B^*, \\ (AB)^* &= B^* \cdot A^*. \end{aligned} \quad (4.2.2)$$

Definición 4.4. La norma del cuaternio A dado en (4.2.1) se define como:

$$N(A) = |a|^2 + |b|^2 = \det(A). \quad (4.2.3)$$

Por lo tanto $N(A \cdot B) = N(A) \cdot N(B)$. Definimos la magnitud de A por $|A| = \sqrt{N(A)}$ y se tiene entonces $|A \cdot B| = |A| \cdot |B|$. El cuaternio conjugado se denotará por A^* y no debe confundirse con la matriz conjugada \bar{A} . Es fácil ver que $A \cdot A^* = N(A) \cdot Id$, por lo tanto todo cuaternio $A \in \mathbb{H}$ distinto de cero posee un inverso multiplicativo (como en los complejos):

$$A^{-1} = N(A)^{-1} \cdot A^*. \quad (4.2.4)$$

El álgebra de los cuaterniones es un espacio vectorial de dimensión cuatro sobre \mathbb{R} y la forma cuadrática N hace de \mathbb{H} un espacio euclidiano. También tenemos una inclusión canónica $\mathbb{C} \hookrightarrow \mathbb{H}$

$$c \mapsto \begin{pmatrix} c & 0 \\ 0 & \bar{c} \end{pmatrix}, \quad (4.2.5)$$

mediante la cual podemos identificar a \mathbb{C} como un subcampo de \mathbb{H} . En $\mathbb{C} \subset \mathbb{H}$, la conjugación de \mathbb{H} coincide con la conjugación usual en \mathbb{C} . Se debe tener cuidado de distinguir entre coeficientes complejos de las matrices y los elementos de $\mathbb{C} \subset \mathbb{H}$. Mediante la inclusión

(4.2.5) también podemos ver a \mathbb{H} como espacio vectorial de dimensión 2 sobre \mathbb{C} , donde el subcampo \mathbb{C} actúa sobre \mathbb{H} por multiplicación izquierda. También debemos distinguir entre la multiplicación izquierda y multiplicación derecha ya que \mathbb{H} no es conmutativo. Como espacio vectorial sobre \mathbb{C} , \mathbb{H} tiene la siguiente **base canónica**:

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (4.2.6)$$

y se cumplen las siguientes reglas de multiplicación:

$$j^2 = -1, \quad zj = j\bar{z} = \begin{pmatrix} 0 & z \\ -\bar{z} & 0 \end{pmatrix} \quad \forall z \in \mathbb{C}. \quad (4.2.7)$$

Entonces podemos escribir

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} = a \cdot 1 + b \cdot j = a + bj, \quad a, b \in \mathbb{C}.$$

El espacio vectorial complejo \mathbb{H} viene equipado con un producto interno unitario:

$$\langle a + bj, c + dj \rangle = a\bar{c} + b\bar{d} \quad (4.2.8)$$

respecto al cual la base (4.2.6) es ortonormal.

La inclusión (4.2.5) induce también una inclusión:

$$\mathbb{R} \hookrightarrow \mathbb{H}, \quad r \mapsto r \cdot 1.$$

Como espacio vectorial real \mathbb{H} tiene la **base canónica**

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad i = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad k = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad (4.2.9)$$

Observación 4.5. La regla de multiplicación de \mathbb{H} como álgebra sobre \mathbb{R} está dada por:

$$i^2 = j^2 = k^2 = -1, \quad ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j$$

La base anterior es ortonormal con el producto interno euclidiano dado por la norma N vista como forma cuadrática sobre \mathbb{H} . A los cuaterniones de la forma $ai + bj + ck$, con $a, b, c \in \mathbb{R}$ les llamaremos **cuaterniones imaginarios puros**. Estos forman un espacio euclidiano de dimensión tres $Im(\mathbb{H})$, que identificaremos con \mathbb{R}^3 mediante la base i, j, k . En consecuencia, tenemos la descomposición ortogonal canónica de \mathbb{H} en el subcampo \mathbb{R} y el espacio euclidiano $Im(\mathbb{H}) = \mathbb{R}^3$ de cuaterniones imaginarios puros:

$$\mathbb{H} = \mathbb{R} \oplus \mathbb{R}^3, \quad \mathbb{R}^3 = Im(\mathbb{H}) \quad (4.2.10)$$

de tal manera que

$$A = Re(A) + Im(A) \in \mathbb{H}.$$

La regla de multiplicación en (4.5) nos recuerda a las reglas del producto cruz (producto vectorial) y muestra que éste se puede expresar en términos de la multiplicación de cuaterniones:

Observación 4.6. Para $A, B \in \mathbb{R}^3 = \text{Im}(\mathbb{H})$

$$A \times B = \text{Im}(A \cdot B).$$

La descomposición (4.2.10) está determinada de manera única por la estructura de anillo de \mathbb{H} , como demostraremos. Recordemos que el **centro** de un anillo R se define como:

$$Z(R) = \{z \in R \mid zx = xz \text{ para toda } x \in R\}.$$

Teorema 4.7. El centro del anillo \mathbb{H} es el subcampo \mathbb{R} .

Este resultado se deduce inmediatamente del siguiente lema.

Lema 4.8. Sean $A \in \mathbb{H} \setminus \mathbb{R}$ y $X \in \mathbb{H}$ tales que $XA = AX$ entonces $X \in \mathbb{R} + \mathbb{R} \cdot A$.

Demostración. Notemos que $XA = AX$ si y sólo si $X\text{Im}(A) = \text{Im}(A)X$. Por lo tanto podemos suponer que $A \in \text{Im}(\mathbb{H})$ con $A \neq 0$. Para un cuaternio puro $A = ai + bj + ck$ se cumple:

$$(ai + bj + ck)^2 = -(a^2 + b^2 + c^2) = -N(A) \quad (4.2.11)$$

Sin pérdida de generalidad podemos suponer que $X = \text{Im}(X)$. Claramente, la conmutatividad de A y X no se altera al multiplicar por factores reales. Luego podemos suponer (después de normalizar en caso de ser necesario) que $X^2 = A^2 = -1$. Luego $(X - A)(X + A) = X^2 + XA - AX - A^2 = 0$ y por lo tanto $X = \pm A$. \square

En consecuencia, el subcampo \mathbb{R} es el centro del anillo \mathbb{H} . De (4.2.11) tenemos que si $A \in \text{Im}(\mathbb{H})$ entonces $A^2 \in \mathbb{R}^- \cup \{0\}$. Esta propiedad caracteriza a $\text{Im}(\mathbb{H})$. En efecto, si $A = r + Q$ con $r \in \mathbb{R}$ y $Q \in \text{Im}(\mathbb{H})$, entonces $A^2 = r^2 + 2rQ + Q^2$ si y sólo si $A^2 \in \mathbb{R}$ cuando $rQ = 0$, por lo tanto $r = 0$ ó $Q = 0$ y no es positiva si y sólo si $r = 0$. Por lo tanto,

$$A \in \text{Im}(\mathbb{H}) \iff A^2 \in \mathbb{R}^- \cup \{0\}. \quad (4.2.12)$$

La conjugación $A \mapsto A^*$ también queda determinada por la estructura de anillo de \mathbb{H} ya que por (4.2.9):

$$(\text{Re}(A) + \text{Im}(A))^* = \text{Re}(A) - \text{Im}(A) \quad (4.2.13)$$

Por lo tanto, la estructura de \mathbb{H} como espacio euclidiano queda determinada por su estructura de anillo. Recordemos que el único automorfismo de anillos $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ es la identidad. Para \mathbb{H} tenemos el siguiente resultado:

Teorema 4.9. Todo automorfismo de anillos $\varphi : \mathbb{H} \rightarrow \mathbb{H}$ satisface:

1. $\varphi|_{\mathbb{R}} = \text{id}_{\mathbb{R}}$.
2. $\varphi(A^*) = (\varphi(A))^*$.
3. $|\varphi(A)| = |A|$.

Demostración. Como $\mathbb{R} = Z(\mathbb{H})$ entonces $\varphi(\mathbb{R}) = \mathbb{R}$ y como el único automorfismo de anillos en \mathbb{R} es la identidad, entonces $\varphi|_{\mathbb{R}} = id_{\mathbb{R}}$. De aquí se sigue que $\varphi(Im(\mathbb{H})) = Im(\mathbb{H})$. Ahora, si $A = B + C$ es la descomposición en parte real e imaginaria de A , entonces $\varphi(A) = \varphi(B) + \varphi(C) = B + \varphi(C)$, por lo tanto $\varphi(A)^* = B - \varphi(C) = \varphi(B - C) = \varphi(A^*)$. Finalmente, del hecho de que $A \cdot A^* = N(A) \in \mathbb{R}$ se sigue que:

$$\begin{aligned} N(\varphi(A)) &= \varphi(A) \cdot \varphi(A)^* \\ &= \varphi(A) \cdot \varphi(A^*) \\ &= \varphi(A \cdot A^*) \\ &= \varphi(N(A)) \\ &= N(A). \end{aligned}$$

Por lo tanto $N(\varphi(A)) = N(A)$, en consecuencia $|\varphi(A)| = |A|$. \square

A continuación presentamos algunas reglas para el cálculo con cuaterniones las cuales utilizaremos más adelante. Por definición de norma:

$$X \cdot X^* = X^* \cdot X = |X|^2.$$

Si se sustituye X por $X + Y$ entonces

$$\langle X, Y \rangle = \frac{1}{2}(X \cdot Y^* + Y \cdot X^*), \quad (4.2.14)$$

en donde $\langle \cdot, \cdot \rangle$ representa el producto interno real en \mathbb{R}^4 .

Observación 4.10. $X \cdot Y^* = -Y \cdot X^*$ si y sólo si $\langle X, Y \rangle = 0$ si $X, Y \in Im(\mathbb{H})$ esto es equivalente a que: $XY = -YX$

Corolario 4.11. Sean $I, J \in Im(\mathbb{H})$ ortonormales, si $K = IJ$ entonces (I, J, K) es una base ortonormal positiva de $Im(\mathbb{H})$ y se cumple que:

$$I^2 = J^2 = K^2 = -1, \quad IJ = -JI = K, \quad JK = -KJ = I, \quad KI = -IK = J$$

Demostración. $K^2 = IJIJ = -I^2J^2 = -1$, por lo tanto $K \in Im(\mathbb{H})$ por (4.2.12). La primera regla de multiplicación se sigue también de (4.2.12) y de (4.2.14). De la observación anterior, vemos que K es ortogonal a I y J . También tenemos que $|K|^2 = K \cdot K^* = -K^2 = 1$. Finalmente la base es positiva porque

$$K = Im(K) = Im(IJ) = I \times J.$$

\square

En consecuencia, la inclusión $\mathbb{C} \hookrightarrow \mathbb{H}$ no está determinada por la estructura de anillo de \mathbb{H} , ya que toda transformación ortogonal positiva $Im(\mathbb{H}) \rightarrow Im(\mathbb{H})$ define un automorfismo de \mathbb{H} . Por último, resulta ser útil el siguiente resultado:

Proposición 4.12 (Identidad triple). Para $X, Y \in \mathbb{H}$ tenemos:

$$YXY = 2\langle X^*, Y \rangle Y - \langle Y, Y \rangle X^*.$$

Demostración. Por (4.2.14) tenemos: $2\langle X^*, Y \rangle = X^*Y^* + YX$ multiplicando a la derecha por Y y usando el hecho de que $Y^* \cdot Y = \langle Y, Y \rangle$ se sigue el resultado. \square

Hasta aquí en cuanto al cálculo con cuaterniones. La notación \mathbb{H} se usa en honor a Hamilton quien fué su descubridor. Históricamente los cuaterniones no han sido tan relevantes como los complejos, sin embargo son de gran importancia para el estudio de las formas cuadráticas y los principales grupos de matrices.

El conjunto de los cuaterniones de norma 1 es el grupo multiplicativo:

$$Sp(1) = \{A \in \mathbb{H} \mid N(A) = 1\}. \quad (4.2.15)$$

Como espacio $Sp(1) = S^3 \subset \mathbb{R}^4 = \mathbb{H}$ aunque como grupo es mejor conocido como:

$$Sp(1) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \mid |a|^2 + |b|^2 = 1 \right\} \cong SU(2),$$

Este grupo es tan importante que recibe un tercer nombre: **Spin(3)** y lo estudiaremos con más detalle en las siguientes secciones pues tiene una estrecha relación con $SO(3)$.

Concluimos esta sección con una pequeña aplicación a la teoría de números. A partir de la fórmula $N(A) \cdot N(B) = N(A \cdot B)$ y de la expresión $N(a_1 + a_2i + a_3j + a_4k) = a_1^2 + a_2^2 + a_3^2 + a_4^2$ obtenemos una fórmula que expresa el producto de dos sumas de 4 cuadrados nuevamente como suma de 4 cuadrados:

Identidad de los cuatro cuadrados:

$$\begin{aligned} (a_1^2 + \dots + a_4^2)(b_1^2 + \dots + b_4^2) &= N((a_1 + a_2i + a_3j + a_4k)(b_1 + b_2i + b_3j + b_4k)) \\ &= (a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4)^2 \\ &+ (a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3)^2 \\ &+ (a_1b_3 + a_3b_1 + a_4b_2 - a_2b_4)^2 \\ &+ (a_1b_4 + a_4b_1 + a_2b_3 - a_3b_2)^2. \end{aligned}$$

Esta identidad es particularmente interesante cuando $a_1, \dots, a_4, b_1, \dots, b_4 \in \mathbb{Z}$.

4.3. Los grupos $SU(2)$, $SO(3)$ y $SO(4)$

Consideremos nuevamente la base canónica de \mathbb{H} sobre los reales:

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad i = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad k = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad (4.3.1)$$

y observemos que (i, j, k) es una base ortonormal para el espacio $\mathfrak{su}(2)$ de matrices anti-hermitianas de traza cero, por lo tanto

$$\mathfrak{su}(2) = \text{Im}(\mathbb{H}). \quad (4.3.2)$$

Esto resulta bastante natural pues $SU(2) = S^3 \subset \mathbb{H}$ es la esfera unitaria y $\mathfrak{su}(2)$, el espacio tangente a $SU(2)$ en la identidad, es el complemento ortogonal de $\mathbb{R} \subset \mathbb{H}$.

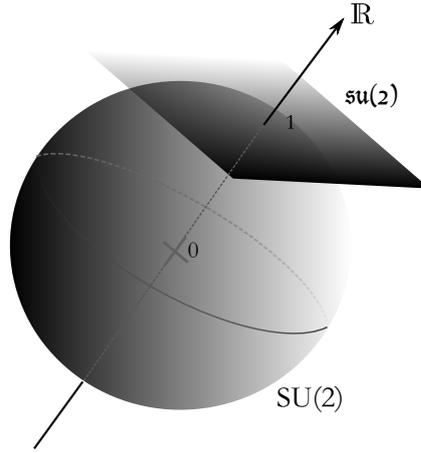


Figura 4.3: Representación del grupo $SU(2)$ y su espacio tangente en I

En la sección anterior vimos como multiplicar cuaterniones, esto nos permite obtener la estructura de $\mathfrak{su}(2)$ como álgebra de Lie:

$$[i, j] = ij - ji = 2ij = 2k, \quad [j, k] = 2i, \quad [k, i] = 2j.$$

El factor 2 se puede eliminar fácilmente. En efecto, si ponemos $e_1 = i/2, e_2 = j/2, e_3 = k/2$ obtenemos:

$$[e_1, e_2] = e_3, \quad [e_2, e_3] = e_1, \quad [e_3, e_1] = e_2 \quad (4.3.3)$$

y podemos concluir lo siguiente; ver (4.1):

Teorema 4.13. *Existe un isomorfismo de álgebras de Lie:*

$$\mathfrak{su}(2) \xrightarrow{\cong} \mathfrak{so}(3).$$

Entenderemos mejor este isomorfismo si estudiamos la relación entre los grupos correspondientes. Recordemos primero un resultado importante sobre los grupos ortogonales.

Definición 4.14. *Sea V un espacio euclidiano y $v \in V$ tal que $|v| = 1$. Entonces la reflexión asociada a v está dada por:*

$$R_v : V \rightarrow V \quad (4.3.4)$$

$$x \mapsto x - 2\langle x, v \rangle \cdot v.$$

Esta transformación es la identidad en el hiperplano ortogonal a v , $H_v = \{x \mid \langle x, v \rangle = 0\}$ y $R_v(v) = -v$, por lo tanto R_v es la reflexión en el hiperplano H_v . En particular $\det(R_v) = -1$ si V es de dimensión finita. Si $T \in O(V)$ entonces claramente

$$T \circ R_v \circ T^{-1} = R_{T(v)} \quad (4.3.5)$$

ya que ambos lados mandan el vector $T(v)$ en $-T(v)$ y dejan fijo al hiperplano $H_{T(v)}$. Además se tiene $s_v^2 = id$.

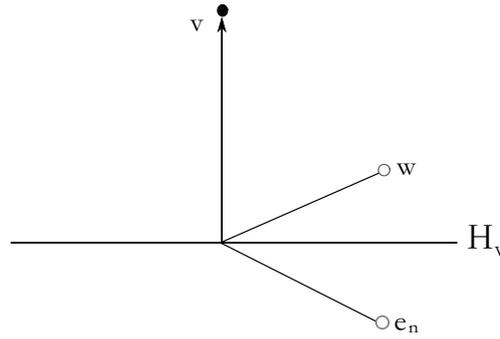


Figura 4.4: Teorema 4.15

Teorema 4.15 (Cartan–Dieudonné). *Si V es un espacio euclidiano de dimensión n , entonces todo automorfismo ortogonal de V es composición de k reflexiones con $k \leq n$. El número k de reflexiones es par si y sólo si el automorfismo está en $SO(V)$.*

Demostración. Sea α tal automorfismo y (e_1, \dots, e_n) una base ortonormal de V . Procederemos por inducción sobre n . Si $\alpha(e_n) = e_n$, entonces α es por hipótesis producto de a lo más $n - 1$ reflexiones y si $\alpha(e_n) = w \neq e_n$, entonces pongamos $v = \frac{(w - e_n)}{|w - e_n|}$, por lo tanto $R_v(w) = e_n$ y $R_v \circ \alpha$ es producto de a lo más $n - 1$ reflexiones. Por lo tanto $\alpha = R_v \circ R_v \circ \alpha$ es producto de a lo más n reflexiones. La última afirmación es clara ya que $\det(R_v) = -1$. \square

Ahora identifiquemos \mathbb{H} con \mathbb{R}^4 mediante la base (4.2.9) y estudiemos el grupo ortogonal $O(\mathbb{H}) = O(4)$.

Si $A, B \in SU(2) \subset \mathbb{H}$, entonces las transformaciones

$$\mathbb{H} \rightarrow \mathbb{H}$$

$$X \mapsto AXB \text{ y } X \mapsto AX^*B$$

son ortogonales porque $|AXB| = |A| \cdot |X| \cdot |B| = |X| = |AX^*B|$. Veremos que todas las transformaciones ortogonales de \mathbb{H} son de este tipo.

Consideremos las transformaciones de la forma:

$$p_A : \mathbb{H} \rightarrow \mathbb{H}, \quad X \mapsto AXA$$

con $A \in SU(2)$. Entonces se cumple:

Teorema 4.16. *Toda transformación ortogonal positiva $\mathbb{H} \rightarrow \mathbb{H}$ es composición de a lo más cuatro transformaciones de la forma p_A con $A \in SU(2)$. El grupo $O(\mathbb{H})$ está generado por el conjunto de todas las p_A con $A \in SU(2)$ y la conjugación $X \mapsto X^*$.*

Demostración. Aplicando la fórmula (4.2.14) para $Y = 1$ se tiene que $\langle 1, X \rangle = \frac{1}{2}(X \cdot 1 + 1 \cdot X^*)$ por lo tanto

$$R_1(X) = -X^*. \quad (4.3.6)$$

Por otra parte, se cumple

$$p_A = R_A \circ R_1, \quad (4.3.7)$$

ya que al aplicar la identidad triple (4.12) tenemos:

$$\begin{aligned} R_A \circ R_1(X) &= -R_A(X^*) = -X^* + 2\langle A, X^* \rangle A \\ &= 2\langle X^*, A \rangle A - \langle A, A \rangle X^* \\ &= AXA. \end{aligned}$$

Finalmente nótese que

$$R_A \circ R_B = p_A \circ p_{-B^*} \quad (4.3.8)$$

En efecto,

$$\begin{aligned} R_A \circ R_B &= R_A \circ (R_1 \circ R_1) \circ R_B \\ &= (R_A \circ R_1) \circ (R_1 \circ R_B) \\ &= (R_A \circ R_1) \circ (R_1 \circ R_B \circ R_1^{-1}) \circ R_1 \\ &= p_A \circ p_{-B^*}. \end{aligned}$$

por (4.3.6) y (4.3.5).

La primera afirmación del teorema se sigue ahora de (4.3.8) y del teorema (4.15).

Si $\alpha \in O(\mathbb{H}) \setminus SO(\mathbb{H})$ y $\kappa : \mathbb{H} \rightarrow \mathbb{H}$ es la conjugación, entonces $\alpha = (\alpha \circ \kappa) \circ \kappa$ con $\alpha \circ \kappa \in SO(\mathbb{H})$ porque la primera afirmación es composición de a lo más cuatro transformaciones p_A con $A \in SU(2)$. \square

El resultado importante que queremos exponer se hará poco a poco. Notemos que el grupo $SU(2) \times SU(2)$ actúa ortogonalmente sobre el espacio euclidiano \mathbb{H} , a saber (A, B) actúa sobre $X \in \mathbb{H}$ por:

$$(A, B) : X \mapsto AXB^*.$$

Esto da lugar a un homomorfismo

$$\varphi : SU(2) \times SU(2) \rightarrow SO(\mathbb{H}) \cong SO(4) \quad (4.3.9)$$

que manda a (A, B) en la transformación ortogonal

$$\mathbb{H} \rightarrow \mathbb{H}, \quad X \mapsto AXB^*.$$

Por otro lado tenemos la representación adjunta

$$\alpha : SU(2) \rightarrow SO(\text{Im}(\mathbb{H})) \cong SO(3) \quad (4.3.10)$$

que manda A sobre la transformación ortogonal $\text{Im}(\mathbb{H}) \rightarrow \text{Im}(\mathbb{H})$, $X \mapsto AXA^*$. El hecho de que la imagen de φ ó α consista de transformaciones que preserven orientación se sigue de la continuidad del determinante y de que $SU(2)$ es conexo.

Teorema 4.17.

1. El homomorfismo φ es sobreyectivo y $\ker(\varphi) = \{1, -1\}$.
2. La representación adjunta α es sobreyectiva y $\ker(\alpha) = \{1, -1\}$

El primer inciso fué descubierto por Hamiton, el segundo por Cayley. El elemento identidad en $SU(2) \times SU(2)$ es obviamente (I, I) .

Demostración. Al tomar $B = A^*$ vemos que los generadores p_A están en la imagen de φ . Por lo tanto, φ es sobreyectiva. Por otro lado, si $f \in SO(Im(\mathbb{H}))$ entonces f se puede ver como una transformación ortogonal en $SO(\mathbb{H})$ poniendo $f|_{\mathbb{R}} = id_{\mathbb{R}}$. Por lo anterior, existen $A, B \in SU(2)$ tales que $f(X) = AXB$. Como $f(1) = 1$, entonces $AB = 1$ y por lo tanto $B = A^*$, luego $f(X) = AXA^* = AXA^{-1}$. En consecuencia la representación adjunta α es sobreyectiva.

Sea ahora $A \in \ker(\alpha)$ entonces $AXA^* = X$ para todo $X \in Im(\mathbb{H})$, por lo tanto para todo $X \in \mathbb{H}$ se tiene que $AX = XA$, es decir, A está en el centro de \mathbb{H} que es \mathbb{R} , por lo tanto $A \in \mathbb{R} \cap SU(2) = \{1, -1\}$. Finalmente si $(A, B) \in \ker(\varphi)$, entonces $AXB^* = X$ para toda $X \in \mathbb{H}$. Haciendo $X = 1$, se sigue que $A = B$ de manera que $A \in \ker(\alpha)$. En consecuencia $(A, B) = (A, A) \in \{(1, 1), -(1, 1)\}$. \square

Corolario 4.18.

1. Si $f \in O(\mathbb{H}) \setminus SO(\mathbb{H})$, existen $A, B \in SU(2)$ tales que $f(X) = AX^*B$ para toda $X \in \mathbb{H}$.
2. Si $f \in O(Im(\mathbb{H})) \setminus SO(Im(\mathbb{H}))$ existe $A \in SU(2)$ tal que $f(X) = -AXA^*$.

Demostración. Sea $\kappa : X \mapsto X^*$ la conjugación y notemos que $\kappa(X) = -X$ para $X \in Im(\mathbb{H})$. Entonces para el primer inciso: $f \circ \kappa \in SO(\mathbb{H})$ por lo tanto existen $A, B \in SU(2)$ tales que $f(X) = f \circ \kappa(X^*) = AX^*B$. Para el segundo inciso: $f \circ \kappa \in SO(Im(\mathbb{H}))$ por lo tanto existe $A \in SU(2)$ tal que $f(X) = f \circ \kappa(-X) = A(-X)A^* = -AX^*A$. \square

Teorema 4.19. *Todo automorfismo de anillos $\varphi : \mathbb{H} \rightarrow \mathbb{H}$ es de la forma $\varphi(X) = AXA^{-1}$ para algún $A \in SU(2)$.*

Demostración. Por teorema el (4.9) φ es ortogonal. Como $\varphi(1) = 1$ entonces $\varphi(X) = AXA^{-1}$ ó $\varphi(X) = AX^*A^{-1}$ para alguna $A \in SU(2)$. Pero el segundo caso es imposible porque

$$\begin{aligned} \varphi(XY) &= A(XY)^*A^{-1} \\ &= AY^*X^*A^{-1} \\ &= (AY^*A^{-1})(AX^*A^{-1}) \\ &= \varphi(Y)\varphi(X). \end{aligned}$$

\square

Observemos que el teorema (4.17) exhibe una parametrización de los grupos ortogonales $SO(3)$ y $SO(4)$ por los grupos geoméricamente más simples $SU(2) = S^3$ y $SU(2) \times SU(2) = S^3 \times S^3$. En particular, observemos que:

Teorema 4.20. *Como espacios topológicos: $SO(3) \approx \mathbb{R}P^3$ y $SO(4) \approx S^3 \times \mathbb{R}P^3$.*

Demostración. $SO(3) \cong S^3/\{1, -1\} = S^3/(x \sim -x) = \mathbb{R}P^3$. Por otro lado,

$$\begin{aligned} SO(4) &\cong (S^3 \times S^3)/\{(1, 1), -(1, 1)\} \\ &= (S^3 \times S^3)/((x, y) \sim -(x, y)) \\ &\approx S^3 \times \mathbb{R}P^3 \end{aligned}$$

□

De hecho, $SO(4)$ es una cuádrica proyectiva real, a saber: en $\mathbb{H} \times \mathbb{H} = \mathbb{R}^8$ se tiene la forma cuadrática

$$\gamma : (X, Y) \mapsto N(X) - N(Y)$$

La cuádrica proyectiva asociada en $P(\mathbb{H} \times \mathbb{H}) \cong \mathbb{R}P^7$ es

$$\begin{aligned} Q &= \{[(X, Y)] \mid N(X) - N(Y) = 0\} \\ &= \{[(X, Y)] \mid |X|^2 - |Y|^2 = 0\} \\ &= (S^3 \times S^3)/((X, Y) \sim -(X, Y)). \end{aligned}$$

y ésta es $SO(4)$.

El subgrupo normal $SU(2) \times \{1\}$ de $SU(2) \times SU(2)$ es mapeado por φ de manera inyectiva sobre un subgrupo normal $G \trianglelefteq SO(4)$, isomorfo a $SU(2)$, porque $\ker(\varphi) \cap (SU(2) \times 1) = \{1\}$ y φ es sobreyectiva. Por otra parte podemos obtener un isomorfismo de $SO(3)$ en un subgrupo (que no es normal) $H \leq SO(4)$ del siguiente modo: consideremos el monomorfismo diagonal $\delta : SU(2) \rightarrow SU(2) \times SU(2)$, dado por $A \mapsto (A, A)$. Tenemos entonces un diagrama conmutativo de homomorfismos de grupos

$$\begin{array}{ccc} SU(2) & \xrightarrow{\delta} & SU(2) \times SU(2) \\ \alpha \downarrow & & \downarrow \varphi \\ SO(3) & \xrightarrow{d} & SO(4) \end{array}$$

donde $d : SO(3) \rightarrow SO(4)$ está inducido por δ ya que $\ker(\alpha) = \{1, -1\} = \ker(\varphi \circ \delta)$. Esta ecuación también muestra que d es inyectiva: si $d(\alpha(X)) = 1$, entonces $\varphi(\delta(X)) = 1$ de donde $\alpha(X) = 1$.

Teorema 4.21. *Existe un subgrupo normal $SU(2) \cong G \trianglelefteq SO(4)$ y un subgrupo $SO(3) \cong H \leq SO(4)$ tales que $G \cap H = 1$, $G \cdot H = SO(4)$. Por lo tanto $SO(4)$ es producto semidirecto de $SU(2)$ y $SO(3)$.*

Demostración. Hagamos $\varphi(A, 1) = \varphi(B, B) \in G \cap H$, entonces $\varphi(A^{-1}B, B) = 1$ por lo tanto $B = \pm 1$ y $\varphi(B, B) = 1$. Así mismo $\varphi(A, B) = \varphi(AB^{-1}, 1) \cdot \varphi(B, B) \in G \cdot H$. □

Todos estos homomorfismos son por definición continuos y por lo tanto, esta descomposición afirma que como espacios $SO(4) \approx S^3 \times \mathbb{R}P^3$. En particular, el teorema afirma que $SO(4)/G = GH/G \cong H \cong SO(3)$, por lo tanto: $SO(4)$ actúa en \mathbb{R}^3 mediante el epimorfismo $SO(4) \rightarrow SO(3)$.

La representación adjunta $\alpha : S^3 = SU(2) \rightarrow SO(3)$ parametriza las rotaciones de manera única salvo signo, por puntos en S^3 , es decir, por medio de los parámetros $(\kappa, \lambda, \mu, \nu)$ con $\kappa^2 + \lambda^2 + \mu^2 + \nu^2 = 1$. Podemos calcular explícitamente la matriz asociada $A \in SO(3)$:

Observación 4.22 (Parametrización de Euler de $SO(3)$).

$$A = \begin{pmatrix} \kappa^2 + \lambda^2 - \mu^2 - \nu^2 & -2\kappa\nu + 2\lambda\mu & 2\kappa\mu + 2\lambda\nu \\ 2\kappa\nu + 2\lambda\mu & \kappa^2 - \lambda^2 - \nu^2 & -2\kappa\lambda + 2\mu\nu \\ -2\kappa\mu + 2\lambda\nu & 2\kappa\lambda + 2\mu\nu & \kappa^2 - \lambda^2 - \mu^2 + \nu^2 \end{pmatrix}$$

Así a todo punto $(\kappa, \lambda, \mu, \nu) \in S^3$ le corresponde una matriz $A \in SO(3)$ y a toda matriz $A \in SO(3)$ le corresponde una única 4-tupla (salvo signo) en S^3 .

En $SU(2)$ podemos conectar $1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ con $-1 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ mediante la trayectoria $t \mapsto A(t) = \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}$, $t \in [0, \pi]$. Como elemento de \mathbb{H} :

$$A(t) = \cos t \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin t \cdot \begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix} = \cos t + (\sin t) \cdot i$$

y se puede verificar fácilmente que la rotación

$$\varphi(A(t)) : X \mapsto (\cos t + i \sin t)X(\cos t - i \sin t)$$

correspondiente a $A(t)$ manda el plano $L(j, k) \subset \text{Im}(\mathbb{H})$ en sí mismo y con respecto a la base (j, k) está dada por la matriz

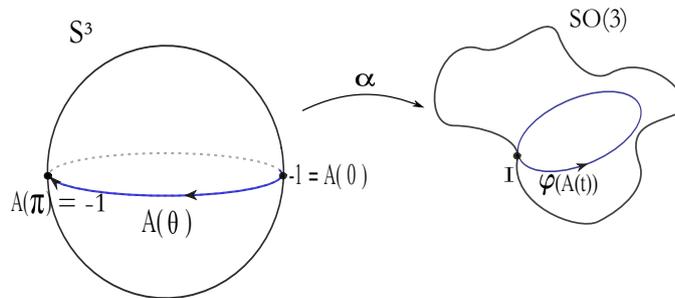


Figura 4.5: $\varphi(A(t)) \in SO(3)$ describe una rotación completa, mientras $A(\theta)$ no.

$$\begin{pmatrix} \cos 2t & \sin 2t \\ -\sin 2t & \cos 2t \end{pmatrix}, \quad 0 \leq t \leq \pi.$$

Por lo tanto: cuando $t \in [0, \pi]$, $\varphi(A(t)) \in SO(3)$ describe una rotación completa alrededor de un eje, mientras que $A(\pi) = -A(0)$.

Notación física: Para nosotros, el álgebra de Lie de $S^1 = U(1)$ es $i\mathbb{R}$ y la función exponencial es

$$\exp : it \mapsto e^{it}.$$

Los físicos prefieren trabajar con números reales que con números imaginarios y prefieren matrices hermitianas a las antihermitianas. En vez de las álgebras de Lie LG que hemos considerado, prefieren considerar el álgebra de Lie iLG y pasar de LG a iLG vía multiplicación por $(-i)$. Siempre que tengamos una matriz $A \in LG$, para los físicos es $-iA$ y el correspondiente corchete de Lie para los físicos es:

$$[A', B'] = i(A'B' - B'A').$$

Explícitamente:

$$\begin{array}{ccc} (A, B) & \xrightarrow{\text{Matematicos}} & AB - BA \\ \begin{array}{c} \downarrow (-i) \\ (-iA, -iB) = (A', B') \end{array} & \xrightarrow{\text{Fisicos}} & i(A'B' - B'A') = -i(AB - BA) \\ & & \downarrow (-i) \end{array}$$

Del mismo modo la exponencial para los físicos está dada por:

$$A \mapsto \exp(iA).$$

Esta convención la observamos también en $\mathfrak{su}(2)$ y en lugar de la base

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

del espacio de matrices antihermitianas 2×2 de traza 0, se multiplica por $(-i)$ para obtener una base de las matrices hermitianas 2×2 de traza cero:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Estas matrices tienen un nombre especial, se denominan las **matrices de Pauli** y aparecen en mecánica cuántica al estudiar la interacción del espín de una partícula con un campo electromagnético externo.

4.4. El grupo simpléctico

Los resultados básicos de álgebra lineal, a saber, el teorema de completación de bases y sus consecuencias también son válidos para anillos de división (no necesariamente conmutativos). Entonces tenemos espacios vectoriales sobre \mathbb{H} , donde la multiplicación por escalares de \mathbb{H} es siempre a la izquierda. Un espacio vectorial finitamente generado sobre \mathbb{H} tiene una dimensión bien definida n y por lo tanto es isomorfo a \mathbb{H}^n . Una transformación \mathbb{H} -lineal $\alpha : \mathbb{H}^n \rightarrow \mathbb{H}^n$ se describe por una matriz $n \times n$ $A = (a_{\lambda\mu})$ con coeficientes en \mathbb{H} de la siguiente manera: si $e_\nu \in \mathbb{H}^n$ es el ν -ésimo vector canónico, entonces el escalar $a_{\lambda\nu}$ está determinado por

$$\alpha(e_\nu) = \sum_{\lambda} a_{\lambda\nu} e_\lambda.$$

Si $h = (h_1, \dots, h_n)^t \in \mathbb{H}^n$ entonces

$$\alpha(h) = \alpha\left(\sum_{\nu} h_{\nu} e_{\nu}\right) = \sum_{\nu} h_{\nu} \alpha(e_{\nu}) = \sum_{\nu, \lambda} h_{\nu} a_{\lambda\nu} e_{\lambda}$$

por lo tanto la componente a lo largo de e_{λ} es:

$$\alpha(h)_{\lambda} = \sum_{\nu} h_{\nu} a_{\lambda\nu}. \quad (4.4.1)$$

Esta es la fórmula usual, sin embargo notemos que los $a_{\lambda\nu}$ aparecen a la derecha. Podemos entonces identificar el grupo de \mathbb{H} -automorfismos de \mathbb{H}^n con el grupo de matrices invertibles $n \times n$ con coeficientes en \mathbb{H} :

$$GL(n, \mathbb{H}) := Aut_{\mathbb{H}}(\mathbb{H}^n).$$

Las matrices aquí se multiplican de la siguiente manera:

$$(a_{\lambda\mu}) \cdot (b_{\mu\nu}) = (c_{\lambda\nu}) \quad (4.4.2)$$

$$c_{\lambda\nu} = \sum_{\mu} b_{\mu\nu} \cdot a_{\lambda\mu}.$$

El intercambio del orden de los factores en la última igualdad es consecuencia de que las transformaciones \mathbb{H} -lineales del \mathbb{H} -espacio vectorial \mathbb{H} en si mismo, son de la forma:

$$[A] : \mathbb{H} \rightarrow \mathbb{H}, \quad X \mapsto X \cdot A$$

con $A \in \mathbb{H}$. Es decir, están dadas por multiplicación a la derecha por un elemento fijo A (la imagen del 1), entonces $[A] \circ [B](X) = [A](XB) = XBA = [BA](X)$. Por lo tanto tenemos el anti-isomorfismo:

$$\mathbb{H} \rightarrow End_{\mathbb{H}}(\mathbb{H}), \quad A \mapsto [A]$$

$$[A + B] = [A] + [B], \quad [AB] = [B] \cdot [A].$$

Las componentes de las matrices son ahora elementos en $End_{\mathbb{H}}(\mathbb{H})$ y como tales se deben multiplicar en el orden correcto. Así mismo, tenemos el isomorfismo:

$$\mathbb{H} \rightarrow End_{\mathbb{H}}(\mathbb{H}), \quad A \mapsto [A^*].$$

Recordemos que la teoría de determinantes para campos se puede generalizar al caso de anillos conmutativos, pero no para anillos de división no conmutativos. Aunque esto no es tan malo, ya que una transformación \mathbb{H} -lineal es en particular una transformación \mathbb{R} -lineal de espacios vectoriales reales y es invertible (un isomorfismo) si es invertible como transformación \mathbb{R} -lineal, es decir, si es biyectiva. Esto se cumple si y sólo si el determinante de dicha transformación real (con respecto a una base real) es distinta de cero y ésta condición define un subconjunto abierto de $End_{\mathbb{H}}(\mathbb{H}^n)$. Por lo tanto, $GL(n, \mathbb{H})$ es un subconjunto abierto del \mathbb{H} -espacio vectorial $End_{\mathbb{H}}(\mathbb{H}) \cong \mathbb{H}^{2n} \cong \mathbb{R}^{4n^2}$. En consecuencia

$GL(n, \mathbb{H})$ es un grupo de Lie.

Como espacio vectorial complejo, identificamos a \mathbb{H} con \mathbb{C}^2 mediante el isomorfismo $\mathbb{H} = \mathbb{C} + \mathbb{C} \cdot j \cong \mathbb{C}^2$, ver (4.2.6). Esto produce el isomorfismo canónico de espacios vectoriales complejos

$$\mathbb{H}^n = \mathbb{C}^n + \mathbb{C}^n \cdot j = \mathbb{C}^n \oplus \mathbb{C}^n \cong \mathbb{C}^{2n}. \quad (4.4.3)$$

Por lo tanto, podemos ver a un endomorfismo \mathbb{H} -lineal α de \mathbb{H}^n como un tipo muy especial de endomorfismo \mathbb{C} -lineal de \mathbb{C}^{2n} :

$$\mathbb{C}^{2n} = \mathbb{C}^n \oplus \mathbb{C}^n \cong \mathbb{C}^n + \mathbb{C}^n j \xrightarrow{\alpha} \mathbb{C}^n + \mathbb{C}^n j \cong \mathbb{C}^{2n},$$

a saber, como uno que conmute con la transformación \mathbb{R} -lineal (pero no \mathbb{C} -lineal) dado por multiplicación izquierda por j , es decir, con la transformación:

$$J : \mathbb{C}^n \oplus \mathbb{C}^n \rightarrow \mathbb{C}^n \oplus \mathbb{C}^n$$

$$(u, v) = u + vj \mapsto j(u + vj) = -\bar{v} + \bar{u}j = (-\bar{v}, \bar{u}).$$

El hecho de que un endomorfismo complejo de $\mathbb{C}^{2n} = \mathbb{C}^n \oplus \mathbb{C}^n$ conmute con la transformación J significa que la matriz del endomorfismo tiene la forma:

$$\begin{pmatrix} A & -\bar{B} \\ B & \bar{A} \end{pmatrix} \quad A, B \in \text{End}(\mathbb{C}^n), \quad (4.4.4)$$

como se puede observar del diagrama:

$$\begin{array}{ccc} (u, 0) & \xrightarrow{j} & (0, \bar{u}) \\ \left(\begin{array}{cc} A & ? \\ B & ? \end{array} \right) \downarrow & & \downarrow \left(\begin{array}{cc} ? & -\bar{B} \\ ? & \bar{A} \end{array} \right) \\ (Au, Bu) & \xrightarrow{j} & (-\bar{B}\bar{u}, \bar{A}\bar{u}) \end{array}$$

Por lo tanto podemos identificar $GL(n, \mathbb{H})$ con el grupo de matrices complejas $2n \times 2n$ de la forma (4.4.4). Como matriz $n \times n$ con coeficientes en \mathbb{H} , (4.4.4) es la matriz $A + Bj$, la cual satisface $(A + Bj)u = Au + Buj$ para $u \in \mathbb{C}^n \subset \mathbb{H}^n$, por (4.4.1).

En \mathbb{H}^n se tiene el **producto interno simpléctico usual**:

$$\langle h, k \rangle = \sum_{\nu=1}^n h_{\nu} k_{\nu}^* \quad (4.4.5)$$

para $h = (h_1, \dots, h_n)^t, k = (k_1, \dots, k_n)^t$ en \mathbb{H}^n . La **norma** asociada es $N(h) = \langle h, h \rangle \geq 0$

Definición 4.23. *El grupo simpléctico $Sp(n)$ es el grupo de los automorfismos de \mathbb{H}^n que preservan la norma:*

$$Sp(n) = \{ \alpha \in GL(n, \mathbb{H}) \mid N(\alpha(h)) = N(h) \quad \forall h \in \mathbb{H}^n \}.$$

Al igual que en el caso complejo, se puede ver que los \mathbb{H} -endomorfismos que preservan norma, el endomorfismo \mathbb{H} -lineal también preserva el producto interno simpléctico. Si identificamos \mathbb{H}^n con \mathbb{C}^{2n} como antes, entonces la norma simpléctica es la norma $|h|^2$ unitaria usual en \mathbb{C}^{2n} , por lo tanto por (4.4.4) $Sp(n)$ es el subgrupo de $U(2n)$ de matrices:

$$\begin{pmatrix} A & -\bar{B} \\ B & \bar{A} \end{pmatrix} \in U(2n), \quad A, B \in \text{End}(\mathbb{C}^n). \quad (4.4.6)$$

Los elementos de este grupo de matrices complejas se denominan **matrices simplécticas**. De manera natural $Sp(1)$ es el grupo de cuaterniones de norma 1, entonces

$$Sp(1) \cong SU(2) \cong S^3. \quad (4.4.7)$$

Notemos que la transformación $j : \mathbb{C}^{2n} = \mathbb{H}^n \rightarrow \mathbb{H}^n = \mathbb{C}^{2n}$, que manda $(u, v) = u + vj$ en $j(u + vj) = (-\bar{v}, \bar{u})$ no es \mathbb{C} -lineal. Sin embargo es la composición de la transformación \mathbb{C} -lineal (incluso \mathbb{H} -lineal) dada por multiplicación derecha por j y la conjugación compleja $c : \mathbb{C}^{2n} \rightarrow \mathbb{C}^{2n}$, $w \mapsto \bar{w}$. La multiplicación derecha con j induce la transformación:

$$J : \mathbb{C}^{2n} = \mathbb{H}^n \xrightarrow{j} \mathbb{H}^n = \mathbb{C}^{2n} \quad (u, v) \mapsto (-v, u)$$

$$J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \quad (4.4.8)$$

donde I es la matriz identidad en $GL(n, \mathbb{C})$. En consecuencia, una matriz unitaria $S \in U(2n)$ es simpléctica si y sólo si $ScJ = cJS$. Como $Sc = c\bar{S}$, esta condición es equivalente $c\bar{S}J = cJS$, es decir, $\bar{S}J = JS$. Como S es unitaria, $\bar{S} = (S^t)^{-1}$ en consecuencia S es simpléctica si y sólo si $S \in U(2n)$ y

$$S^t JS = J. \quad (4.4.9)$$

Esta igualdad implica que la transformación lineal S deja invariante a la forma bilineal con la matriz asociada J , es decir, a la forma bilineal $(u, v) \rightarrow u^t Jv$. Esta es una forma bilineal **antisimétrica**, por lo tanto $u^t Jv = -v^t Ju$ porque $J^t = -J$.

Dejando a un lado la condición de que S sea unitaria, obtenemos el **grupo simpléctico complejo**:

$$Sp(n, \mathbb{C}) = \{S \in GL(2n, \mathbb{C}) \mid S^t JS = J\}.$$

El álgebra de Lie $\mathfrak{sp}(n)$ de $Sp(n)$ es el álgebra de las matrices antihermitianas en $\text{End}_{\mathbb{H}}(\mathbb{H}^n) \subset \text{End}(\mathbb{C}^{2n})$ y tales matrices son las matrices complejas $2n \times 2n$ de la forma

$$\begin{pmatrix} A & -\bar{B} \\ B & \bar{A} \end{pmatrix}, \quad A^* = -A, \quad B^t = B. \quad (4.4.10)$$

La función exponencial manda los $\text{End}_{\mathbb{H}}(\mathbb{H}^n)$ en los $\text{Aut}(\mathbb{H})$ y las matrices antihermitianas en las matrices unitarias, en consecuencia la exponencial manda el álgebra $\mathfrak{sp}(n)$ en el grupo $Sp(n)$, es decir, la función exponencial manda una vecindad del 0 de $\mathfrak{sp}(n)$ en una vecindad de la I de $Sp(n)$. La dimensión de $Sp(n)$ es la dimensión del espacio tangente en la identidad del grupo, por lo tanto por (4.4.10) tenemos:

$$\dim \mathfrak{sp}(n) = 2n^2 + n. \quad (4.4.11)$$

Nos preguntamos ahora, ¿Cuál es la forma canónica de Jordan de una matriz simpléctica A ? Sea A una matriz simpléctica y $w \in \mathbb{H}^n$ un vector propio, consideremos a A como la transformación \mathbb{C} -lineal entonces $Aw = \lambda w$ donde $\lambda \in \mathbb{C}$. Por lo tanto $Ajw = jAw = j\lambda w = \bar{\lambda}jw$, así jw es un vector propio con valor propio $\bar{\lambda}$. En consecuencia los vectores w generan \mathbb{H} -subespacios de \mathbb{H}^n los cuales son invariantes bajo una base compleja de vectores propios de la forma w, jw . El complemento ortogonal de $L_{\mathbb{H}}(w)$ es invariante bajo A y por lo tanto de manera inductiva, existe una base ortonormal (w_1, \dots, w_n) de \mathbb{H}^n tal que $Aw_\nu = \lambda_\nu w_\nu$ entonces $A(jw_\nu) = \bar{\lambda}_\nu jw_\nu$ con $\lambda_\nu \in \mathbb{C}$. En otras palabras:

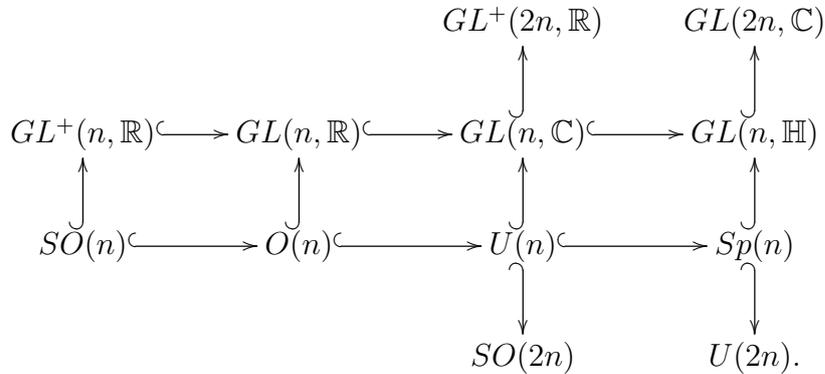
Teorema 4.24. *Para toda matriz $A \in Sp(n)$ existe una transformación $T \in Sp(n)$ tal que TAT^{-1} es una matriz diagonal compleja, con diagonal $(\lambda_1, \dots, \lambda_n, \bar{\lambda}_1, \dots, \bar{\lambda}_n)$.*

Observemos que, aunque los w_ν son vectores propios de la transformación A que es \mathbb{H} -lineal, no podemos concluir que el espacio generado $L_{\mathbb{H}}(w_\nu)$ de la transformaciones \mathbb{H} -lineales consta de los vectores propios (y el 0).

En conclusión, tenemos tres situaciones análogas para \mathbb{R}, \mathbb{C} y \mathbb{H} con su producto interno correspondiente:

$$\begin{aligned} Sp(n) &\subset GL(n, \mathbb{H}) \text{ con el producto interno simpléctico,} \\ U(n) &\subset GL(n, \mathbb{C}) \text{ con el producto interno hermitiano,} \\ O(n) &\subset GL(n, \mathbb{R}) \text{ con el producto interno euclidiano.} \end{aligned}$$

donde $GL(n, \mathbb{R}) \subset GL(n, \mathbb{C}) \subset GL(n, \mathbb{H})$ por las inclusiones $\mathbb{H}^n = \mathbb{C}^{2n}$, $\mathbb{C}^n = \mathbb{R}^{2n}$ y $GL(n, \mathbb{H}) \subset GL(2n, \mathbb{C})$, $GL(n, \mathbb{C}) \subset GL(2n, \mathbb{R})$. El siguiente diagrama muestra algunas inclusiones que hemos mencionado:



4.5. El grupo de Lorentz

Sea V un espacio vectorial real de dimensión finita y $\gamma : V \times V \rightarrow \mathbb{R}$ una forma bilineal. Entonces se tiene el **grupo ortogonal** correspondiente

$$O(\gamma) = \{A \in Aut(V) \mid \gamma(Av, Aw) = \gamma(v, w) \text{ para todo } v, w \in V\}.$$

Si G es la matriz asociada a γ , entonces $(Av)^t G(Aw) = v^t Gw$, por lo tanto $A^t G A = G$. De esto se sigue que $\det(A)^2 \cdot \det(G) = \det(G)$. Por lo tanto si γ es no degenerada entonces

$$\det(A) = \pm 1.$$

Podemos considerar apropiadamente cualquier campo y una forma bilineal más general pero aquí nos interesa el caso en donde γ es una forma bilineal sobre \mathbb{R} , simétrica y no degenerada. Dada una base podemos elegir un isomorfismo $V \cong \mathbb{R}^n = \mathbb{R}^p \oplus \mathbb{R}^q$ de tal manera que la forma cuadrática asociada a $\mathbb{R}^p \oplus \mathbb{R}^q$ este dada por

$$\gamma : \mathbb{R}^p \oplus \mathbb{R}^q \rightarrow \mathbb{R}, \quad (x, y) \mapsto |x|^2 - |y|^2. \quad (4.5.1)$$

Si γ tiene índice q . Esta forma bilineal especial le denotaremos aquí con $\langle v, w \rangle$, a la forma cuadrática asociada con $\langle v, v \rangle$ y al grupo ortogonal correspondiente con $O(p, q)$. Particularmente en física se tiene especial interés en el **grupo de Lorentz** $L = O(1, 3)$. El espacio cuadrático $\mathbb{R} \oplus \mathbb{R}^3$ con la métrica (4.5.1) se denomina el **espacio de Minkowski** con la **métrica de Minkowski** definida por:

$$\langle v, w \rangle = v_0 w_0 - (v_1 w_1 + v_2 w_2 + v_3 w_3). \quad (4.5.2)$$

En teoría de la relatividad con la velocidad de la luz normalizada, la primera coordenada se denomina *temporal* y las tres restantes son las *coordenadas espaciales*. Los elementos de $O(1, 3)$, es decir, las transformaciones del espacio 4-dimensional que dejan invariante la métrica de Minkowski se les denomina **transformaciones de Lorentz**.

Para fijar ideas, consideremos el **plano hiperbólico** $\mathbb{R} \oplus \mathbb{R}$ con la forma cuadrática $(x, y) \mapsto x^2 - y^2$ entonces tenemos el correspondiente grupo hiperbólico de isometrías $O(1, 1)$. Donde $SO^\uparrow(1, 1)$ es el **grupo de rotaciones hiperbólicas**, el cual es la componente conexa en la identidad de $O(1, 1)$. Nos preguntamos, ¿cómo es la matriz de una rotación hiperbólica?:

Si $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ es una matriz de rotación hiperbólica, entonces $a^2 - b^2 = 1$ por lo tanto $a^2 \geq 1$ de tal manera que $a \geq 1$ ó $a \leq -1$. Pero un conjunto abierto en $O(1, 1)$ contiene a la identidad, entonces $a \geq 1$, en particular para las rotaciones hiperbólicas (el signo del coeficiente se indica mediante la flecha \uparrow). Por lo tanto:

$$a \geq 1, \quad a^2 - b^2 = 1, \quad ad - bc = 1, \quad ac - bd = 0.$$

Una matriz de rotación hiperbólica tiene entonces la forma:

$$D(R) = \begin{pmatrix} \alpha(R) & R \\ R & \alpha(R) \end{pmatrix}, \quad \alpha(R) = \sqrt{1 + R^2}, \quad R \in \mathbb{R} \quad (4.5.3)$$

y el conjunto de tales matrices es, en efecto la componente conexa en la identidad del grupo $O(1, 1)$. Si sustituímos $R = \sinh(x) = \frac{1}{2}(e^x - e^{-x})$ y $\alpha(R) = \cosh(x) = \frac{1}{2}(e^x + e^{-x})$ que podemos escribir como $\cosh^2(x) - \sinh^2(x) = 1$, por lo tanto

$$D(R) = A(x) = \begin{pmatrix} \cosh(x) & \sinh(x) \\ \sinh(x) & \cosh(x) \end{pmatrix} \quad (4.5.4)$$

y por la identidad de la suma de funciones trigonométricas hiperbólicas

$$A(x + y) = A(x) \cdot A(y). \quad (4.5.5)$$

En consecuencia, tenemos el siguiente resultado:

Observación 4.25. *La transformación $x \mapsto A(x)$ dada por (4.5.4) induce un isomorfismo entre \mathbb{R} y el grupo de rotaciones hiperbólicas $SO^\uparrow(1, 1)$.*

En general, es mejor hacer cálculos con la matriz $D(R)$ dada en (4.5.3) porque los coeficientes están parametrizados por funciones algebraicas. Los vectores $(\alpha(R), R)$ se mueven sobre la rama hiperbólica $\alpha^2 - R^2 = 1, \alpha \geq 1$. Los múltiplos positivos o negativos de los vectores, son justamente los que tiene la forma cuadrática positiva.

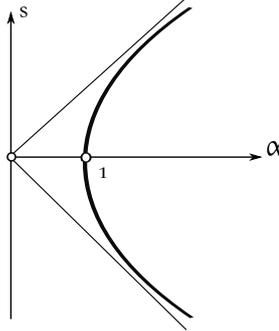


Figura 4.6: Rama hiperbólica.

El vector (x, y) con $|y| < x, x > 0$ está representado por la rotación hiperbólica $D(R)$ de un múltiplo de e_1 , pues si $(\alpha, -R) = r \cdot (x, y)$, con $r \in \mathbb{R}^+$ ya que $D(R)^{-1} = D(-R)$.

Ahora bien, para el grupo de Lorentz sobre el espacio de Minkowski $V = \mathbb{R}^4$ tenemos la base ortogonal (e_0, e_1, e_2, e_3) con $\langle e_0, e_0 \rangle = -1, \langle e_j, e_j \rangle = 1$ para $j = 1, 2, 3$. Por lo tanto en V tenemos: tres planos hiperbólicos $E_j = L(e_0, e_j)$ para $j = 1, 2, 3$ y Δ_j el grupo de transformaciones de Lorentz. Donde las rotaciones hiperbólicas mandan E_j en sí mismo y éste corresponde al complemento ortogonal $L(e_i, e_k)$ con $\{0, j\} \cap \{i, k\} = \emptyset$. En consecuencia, por (4.25) tenemos que $\Delta_j \cong \mathbb{R}$. La matriz (4.5.4) que corresponde a la rotación hiperbólica en Δ_j le denotaremos por $D_j(R)$. Por otro lado, tenemos la inclusión $O(1) \times O(3) \subset O(1, 3) = L$ y se cumple:

Teorema 4.26 (Descomposición de Iwasawa del grupo de Lorentz). *La multiplicación induce la biyección:*

$$\Delta_1 \times \Delta_2 \times \Delta_3 \times O(1) \times O(3) \rightarrow O(1, 3) = L.$$

En particular $L \approx \mathbb{R}^3 \times O(1) \times O(3)$ como variedad suave.

Demostración. Sea $A = (a_{ij}) \in L$. Como $a_{00}^2 - (a_{10}^2 + a_{20}^2 + a_{30}^2) = \langle Ae_0, Ae_0 \rangle = 1$, si $a_{00}^2 \geq a_{j0}^2 + 1$, entonces tenemos las rotaciones hiperbólicas $D_j \in \Delta_j$ tales que $D_3 D_2 D_1 A : e_0 \mapsto \epsilon \cdot e_0$; podemos asumir que $a_{00} > 0$ y tenemos que $D_j = D_j(R_j)$ de tal manera que $(\alpha(R_j), -R_j) = r(a_{00}, a_{j0})$. Pero si $B \in L$ y $Be_0 = \epsilon e_0$ entonces $\epsilon = \pm 1$ y B manda el complemento ortogonal $L(e_1, e_2, e_3)$ en sí mismo, por lo tanto $B \in O(1) \times O(3)$. \square

Corolario 4.27. *Si $A = (a_{ij}) \in L$, entonces el grupo de Lorentz se divide en cuatro*

componentes conexas, las cuales están definidas por:

$$\begin{aligned} L_+^\uparrow &: \det(A) = 1, \quad a_{00} \geq 1 \\ L_+^\downarrow &: \det(A) = 1, \quad a_{00} \leq -1 \\ L_-^\uparrow &: \det(A) = -1, \quad a_{00} \geq 1 \\ L_-^\downarrow &: \det(A) = -1, \quad a_{00} \leq -1 \end{aligned}$$

Las transformaciones de Lorentz en $L^\uparrow = L_+^\uparrow \cup L_-^\uparrow$ preservan la **orientación temporal** y $L_+^\uparrow \cup L_-^\downarrow$ se denomina el **espacio de orientación**.

Las rotaciones hiperbólicas, en particular las que están contenidas en Δ_1 son las transformaciones de Lorentz que usualmente encontramos en los libros de física. Estas transformaciones no son todas las transformaciones de Lorentz y para hablar más sobre esto, se debe estudiar con mayor detalle la estructura global del grupo de Lorentz. Para ello, basta estudiar a L_+^\uparrow el **grupo de Lorentz restringido** del espacio-tiempo para conocer las propiedades de todo L .

En la descomposición de Iwasawa (4.26), la componente conexa de la identidad es una variedad:

$$L_+^\uparrow = \Delta_1 \cdot \Delta_2 \cdot \Delta_3 \cdot SO(3) \approx \mathbb{R}^3 \times SO(3) \approx \mathbb{R}^3 \times \mathbb{R}P^3. \quad (4.5.6)$$

En consecuencia $\dim L_+^\uparrow = 6$.

Por otra parte, el álgebra de Lie $L(L_+^\uparrow)$ se calcula como el espacio de las velocidades iniciales de las curvas $s \mapsto \exp(sA)$ en L_+^\uparrow , A tiene la siguiente condición:

$$(1 + sA + \dots)\eta(1 + sA + \dots) = \eta$$

donde η es una matriz diagonal, con diagonal $(1, -1, -1, -1)$. Lo que significa que $\eta + s(A^t\eta + \eta A) + \dots = \eta$, por lo tanto $A^t\eta = -\eta A$.

Observación 4.28. *El álgebra de Lie de L es el espacio de las matrices A reales 4×4 con $A^t\eta = -\eta A$. Tales matrices son de la forma:*

$$\left(\begin{array}{c|c} 0 & b^t \\ \hline b & C \end{array} \right), \quad C \in \text{End}(\mathbb{R}^3), \quad C^t = -C.$$

Por otro lado, existe una relación entre los grupos $SL(2, \mathbb{C})$ y L_+^\uparrow similar a la de $SU(2)$ y $SO(3)$, la explicaremos a continuación:

Sea $H(2, \mathbb{C})$ el espacio vectorial real de las matrices hermitianas 2×2 con coeficientes en los complejos. Entonces $\dim H(2, \mathbb{C}) = 4$ y tal espacio tiene como base:

$$\tau_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \tau_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (4.5.7)$$

Recordemos que τ_1, τ_2, τ_3 son las matrices de Pauli y que en la notación de los físicos, $H(2, \mathbb{C})$ se puede ver como el álgebra de Lie de $LU(2)$.

En $H(2, \mathbb{C})$ definimos una forma cuadrática real, dada por el determinante:

$$\det : H(2, \mathbb{C}) \rightarrow \mathbb{R}, \quad X \mapsto \det(X) =: \langle X, X \rangle \quad (4.5.8)$$

$$\begin{pmatrix} x_{11} & x_{12} \\ \bar{x}_{12} & x_{22} \end{pmatrix} \mapsto x_{11}x_{22} - |x_{12}|^2$$

con la forma bilineal asociada:

$$\langle X, Y \rangle = \frac{1}{2}(x_{11}y_{22} + x_{22}y_{11} - \operatorname{Re}(x_{12}y_{21} + x_{21}y_{12})).$$

En consecuencia (4.5.7) es una base ortogonal para esta forma bilineal con el producto interno:

$$\begin{aligned} \langle \tau_0, \tau_0 \rangle &= 1, & \langle \tau_j, \tau_j \rangle &= -1 & j &= 1, 2, 3 \\ \langle \tau_j, \tau_k \rangle &= 0 & & & j &\neq k. \end{aligned}$$

En otras palabras:

Observación 4.29. *La base (4.5.7) define un isomorfismo ortogonal entre el espacio de Minkowski $\mathbb{R}^4 = \mathbb{R} \oplus \mathbb{R}^3$ y el espacio $H(2, \mathbb{C})$ de matrices hermitianas 2×2 , con el determinante como forma cuadrática.*

Nos interesa ver a $H(2, \mathbb{C})$ como el espacio de Minkowski y al grupo de Lorentz (especial) como el grupo de las transformaciones de $H(2, \mathbb{C})$.

Consideremos:

$$H(2, \mathbb{C}) \rightarrow H(2, \mathbb{C}), \quad X \mapsto \tilde{X}$$

que a cada matriz la manda a su adjunta y la base cumple que:

$$\tilde{\tau}_0 = \tau_0, \quad \tilde{\tau}_j = -\tau_j \quad \text{para } j = 1, 2, 3.$$

Observación 4.30. *La transformación en la adjunta induce una transformación ortogonal $H(2, \mathbb{C}) \rightarrow H(2, \mathbb{C})$, $X \mapsto \tilde{X}$ de orden 2 en L_-^\uparrow y $\det(X) = \det(\tilde{X})$, $X \cdot \tilde{X} = \tilde{X} \cdot X = \langle X, X \rangle \cdot id$.*

El grupo $SL(2, \mathbb{C})$ actúa linealmente sobre $H(2, \mathbb{C})$ y $A \in SL(2, \mathbb{C})$ por

$$\Lambda(A) : H(2, \mathbb{C}) \rightarrow H(2, \mathbb{C})$$

$$X \mapsto AXA^*.$$

A saber, $\Lambda(A)$ es ortogonal respecto a la forma cuadrática (4.5.8), por lo tanto es un elemento del grupo de Lorentz porque

$$\det(\Lambda(A)X) = \det(AXA^*) = \det(A) \cdot \det(X) \cdot \det(A^*) = \det(X).$$

De tal manera que Λ define un homomorfismo $SL(2, \mathbb{C}) \rightarrow L$ y como $SL(2, \mathbb{C})$ es conexo y Λ manda el uno en el uno, la imagen de Λ está contenida en la componente de la identidad de L_+^\uparrow .

Teorema 4.31. *Mediante $A \mapsto \Lambda(A)$, $\Lambda(A) : X \mapsto AXA^*$ se define un homomorfismo sobreyectivo*

$$\Lambda : SL(2, \mathbb{C}) \rightarrow L_+^\uparrow$$

con $\ker(\Lambda) = \{1, -1\}$. La función Λ se denomina **representación espinorial** del grupo de Lorentz restringido.

Demostración. Mostremos primero que $\ker(\Lambda) = \{1, -1\}$: si $A \in \ker(\Lambda)$ entonces $AXA^* = X$ para toda $X \in H(2, \mathbb{C})$ entonces si $X = \tau_0 = id$ en particular se tiene que $AA^* = 1$, por lo tanto $A^* = A^{-1}$. De tal manera que la condición $AX = XA$ se cumple para toda X hermitiana entonces se cumple para toda matriz antihermitiana porque éstas son de la forma iX con $X \in H(2, \mathbb{C})$. En consecuencia A conmuta con toda matriz X 2×2 , ya que toda matriz se puede escribir como la suma de una matriz hermitiana y una antihermitiana. Por lo tanto si $A = \zeta \cdot id$ y como $\det(A) = \zeta^2 = 1$ entonces $A = \pm 1$.

Ahora, el teorema se sigue de lo siguiente: ambos grupos tienen dimensión 6, Λ es diferenciable, Λ es un homomorfismo porque si $A, B, X \in SL(2, \mathbb{C})$ entonces $\Lambda(AB)(X) = ABX(AB)^* = A(BXB^*)A^* = \Lambda(A)(BXB^*) = \Lambda(A)\Lambda(B)(X)$. Del diagrama:

$$\begin{array}{ccc} SL(2, \mathbb{C}) & \xrightarrow{\Lambda} & L_+^\uparrow \\ A \cdot \downarrow \cong & & \cong \downarrow \Lambda(A) \\ SL(2, \mathbb{C}) & \xrightarrow{\Lambda} & L_+^\uparrow \end{array}$$

Como Λ es inyectiva en una vecindad del 1, entonces el rango (tras el teorema del rango de cálculo diferencial) es 6. Por lo tanto Λ es localmente sobreyectiva alrededor del 1 y la imagen de Λ contiene una vecindad U de 1 en L_+^\uparrow . Sin pérdida de generalidad sea $U = U^{-1}$ un abierto entonces es fácil ver que U es un sistema de generadores de L_+^\uparrow , porque L_+^\uparrow es conexo y por lo tanto Λ es sobreyectiva. En efecto,

$$\bigcup_{i=1}^{\infty} U^n = \{u_1, \dots, u_n \mid u_j \in U, n \in \mathbb{N}\}$$

es un subgrupo abierto de L_+^\uparrow , las clases laterales son abiertas y sólo posee una porque L_+^\uparrow es conexo. □

La descomposición de Iwasawa o la ortogonalización de Gram-Schmidt proporciona el homeomorfismo:

$$SL(2, \mathbb{C}) \approx SU(2) \times \mathbb{R}^3 \approx S^3 \times \mathbb{R}^3. \tag{4.5.9}$$

Para ver de manera explícita que Λ es sobreyectiva: en primer lugar si $SU(2) \subset SL(2, \mathbb{C})$ y la acción de $SU(2)$ por Λ sobre $H(SL(2, \mathbb{C}))$, deja fijo τ_0 , porque $A \cdot A^* = 1$ para $A \in SU(2)$. Y el complemento ortogonal es el generado de las matrices de Pauli, el espacio de matrices hermitianas con traza 0, por lo tanto $i\mathfrak{su}(2)$. Entonces $SU(2)$ actúa sobre Λ como su representación adjunta como α en (4.3.9). Más explícitamente tenemos la parametrización de Euler y las observaciones posteriores. Por otro lado se tiene en $A \in SL(2, \mathbb{C})$ a los elementos

$$A = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}, \quad a > 0$$

y $\Lambda(A)$ es la transformación de Lorentz

$$\left. \begin{array}{l} \tau_0 \mapsto \alpha\tau_0 + s\tau_1 \\ \tau_1 \mapsto s\tau_0 + \alpha\tau_1 \end{array} \right\} \alpha = 1/2(a^2 + a^{-2}), \quad s = 1/2(a^2 - a^{-2})$$

$$\tau_\nu \mapsto \tau_\nu \text{ para } \nu > 1.$$

Estas son las rotaciones hiperbólicas y si $a \in (0, \infty)$, entonces $s \in (-\infty, \infty)$ de tal manera que todas las rotaciones hiperbólicas del plano $L(\tau_0, \tau_1)$ se obtienen por conjugación con una rotación que manda τ_1 en τ_i . Entonces tenemos junto con (4.5.6) el grupo especial de Lorentz.

La inclusión $SL(2, \mathbb{C}) \rightarrow GL(2, \mathbb{C})$ indica que el subgrupo $\{1, -1\}$ está en el subgrupo \mathbb{C}^* de múltiplos de la matriz identidad e induce el homomorfismo:

$$L_+^\uparrow \cong SL(2, \mathbb{C})/\{1, -1\} \rightarrow GL(2, \mathbb{C})/\mathbb{C}^* = PGL(2, \mathbb{C}),$$

Este es un isomorfismo, a saber es inyectivo porque $\mathbb{C}^* \cap SL(2, \mathbb{C}) = \{1, -1\}$ y es sobreyectivo porque para $A \in GL(2, \mathbb{C})$, $\det(A) = \lambda^2$, se cumple que $A = \lambda(\lambda^{-1}A)$, $\lambda\mathbb{C}^*$, $\lambda^{-1}A \in SL(2, \mathbb{C})$. Por lo tanto: el grupo de Lorentz restringido L_+^\uparrow es el grupo de las transformaciones lineales de la esfera de Riemann:

$$L_+^\uparrow \cong PGL(2, \mathbb{C}). \quad (4.5.10)$$

Recordemos que una proyectividad de $\mathbb{C}P^1$ determina la imagen de tres puntos distintos $0, 1, \infty$, es decir, dado un punto en el espacio $X = \{(x, y, z) \in S^2 \times S^2 \times S^2 \mid x \neq y \neq z \neq x\}$. Obtenemos el lugar geométrico

$$L_+^\uparrow \approx \mathbb{R}^3 \times \mathbb{R}P^3 \approx X.$$

Conclusiones

El problema de investigación que nos planteamos en este trabajo consistió en usar herramientas elementales del álgebra lineal para estudiar las formas bilineales, las superficies cuádricas, los grupos lineales y sus álgebras de Lie; así como los cuaterniones y los grupos ortogonales. Para encontrar la solución, formulamos ciertos objetivos los cuales nos dieron los siguientes resultados:

El primer objetivo marcado fué el estudio de las formas bilineales y cuadráticas, con ello nos detuvimos en un tipo muy concreto que son los productos internos, lo cual nos llevó al estudio de las propiedades de los espacios euclidianos. Esto nos permite concluir que las transformaciones lineales ortogonales de \mathbb{R}^n en \mathbb{R}^n con el producto interno usual forman el grupo ortogonal $O(n)$. Análogamente, abordar el estudio de los espacios unitarios para el caso del espacio vectorial complejo nos permite concluir que, los automorfismos unitarios de \mathbb{C}^n forman un grupo, el grupo unitario $U(n)$.

En cuanto al segundo objetivo, observamos que la clasificación de las superficies cuádricas en \mathbb{R}^n está íntimamente ligada a la teoría de formas cuadráticas y por ende, al álgebra de matrices simétricas. La clasificación de las cuádricas en \mathbb{R}^n es de los siguientes tres tipos: $|x|^2 - |y|^2 = 0$, $|x|^2 - |y|^2 = 1$ y $|x|^2 - |y|^2 = 2z$ con $x \in \mathbb{R}^r$, $y \in \mathbb{R}^s$, $z \in \mathbb{R}$. Por otro lado en cuanto a la topología, notamos que las cuádricas asociadas a tales ecuaciones son:

$$Q(n, m, z) \approx \mathbb{R}^{n+m}, \quad Q(n, m, 1) \approx S^{n-1} \times \mathbb{R}^m, \quad Q(n, m, 0) \cap S_2^{n+m-1} = S^{n-1} \times S^{m-1},$$

concluimos entonces que las cuádricas pueden verse como un producto de esferas. En el caso complejo, obtuvimos:

$$Q(n, w) \approx \mathbb{C}^n, \quad Q(n, 1) \approx TS^{n-1}, \quad Q(n, 0) \cap S_2^{2n-1} \approx V_{2,n}.$$

En el caso de los espacios proyectivos, la clasificación de las cuádricas se reduce a la sola ecuación $|x|^2 - |y|^2 = 0$, de tal manera que las superficies cuádricas serán homeomorfas a $S^{n-1} \times S^{m-1}/(x, y) \sim -(x, y)$ para algún $n, m \in \mathbb{Z}$.

Referente al tercer objetivo, estudiamos que canónicamente a cada subgrupo de $GL(n, K)$ ($K = \mathbb{R}$ o \mathbb{C}) se le puede asociar un espacio tangente en la identidad del grupo, el cual se denomina álgebra de Lie, de tal forma que las álgebras de Lie asociadas a los grupos de

Lie correspondientes son:

$$\begin{aligned}
\mathfrak{sl}(n, \mathbb{R}) &= L SL(n, \mathbb{R}) \\
\mathfrak{sl}(n, \mathbb{C}) &= L SL(n, \mathbb{C}) \\
\mathfrak{so}(n) &= L SO(n) \\
\mathfrak{u}(n) &= L U(n) \\
\mathfrak{su}(n) &= L SU(n) \\
\mathfrak{d}(n) &= L D(n) \\
\mathfrak{n}^+(n, \mathbb{R}) &= L N^+(n, \mathbb{R}) \\
\mathfrak{n}^+(n, \mathbb{C}) &= L N^+(n, \mathbb{C}).
\end{aligned} \tag{4.5.11}$$

Asimismo, la representación adjunta de $U(n)$ induce el homomorfismo:

$$ad : U(n) \rightarrow SO(\mathfrak{su}(n)) \cong SO(n^2 - 1), \quad T \mapsto ad_T,$$

y en particular

$$ad : SO(n) \rightarrow SO(\mathfrak{so}(n)) \cong SO\left(\frac{n(n-1)}{2}\right), \quad T \mapsto ad_T.$$

Finalmente, nuestro último objetivo era el estudio de los cuaterniones y los grupos ortogonales con lo cual podemos concluir que, la representación adjunta induce el homomorfismo $ad : SU(2) \rightarrow SO(3)$ y un isomorfismo entre las correspondientes álgebras de Lie. Asimismo, concluimos que existe un isomorfismo entre $\mathfrak{so}(3)$ con el corchete de Lie y \mathbb{R}^3 con el producto cruz. El estudio de los cuaterniones, nos permitió identificar \mathbb{H} con \mathbb{R}^4 como espacio euclidiano mediante la norma como forma cuadrática, la cual es una herramienta que da lugar a lo siguiente: existe un homomorfismo de grupos

$$\varphi : SU(2) \times SU(2) \rightarrow SO(\mathbb{H}) \cong SO(4).$$

y la representación adjunta:

$$\alpha : SU(2) \rightarrow SO(\text{Im}(\mathbb{H})) \cong SO(3).$$

En consecuencia, como espacios topológicos: $SO(3) \approx \mathbb{R}P^3$ y $SO(4) \approx S^3 \times \mathbb{R}P^3$.

Las propiedades de los cuaterniones también nos permiten concluir que los automorfismos de \mathbb{H}^n que preservan la norma (simpléctica) forman el grupo simpléctico $Sp(n)$. De tal manera que, tenemos tres situaciones análogas para \mathbb{R}, \mathbb{C} y \mathbb{H} : $Sp(n) \subset GL(n, \mathbb{H})$, $U(n) \subset GL(n, \mathbb{C})$ y $O(n) \subset GL(n, \mathbb{R})$ con el producto interno correspondiente.

Por último dada una forma bilineal especial (real, simétrica y definida positiva), podemos hacer corresponder un grupo ortogonal generalizado $O(p, q)$. En nuestro caso, estudiamos a $L = O(1, 3)$ el grupo de Lorentz. Podemos concluir que: topológicamente $L \approx \mathbb{R}^3 \times O(1) \times O(3)$ es una variedad suave, la componente conexa $L_+^\uparrow \approx \mathbb{R}^3 \times \mathbb{R}P^3$. Asimismo, concluimos que existe una relación entre los grupos $SL(2, \mathbb{C})$ y L_+^\uparrow similar a la de $SU(2)$ con $SO(3)$, de tal manera que

$$L_+^\uparrow \cong PGL(2, \mathbb{C}) \quad \text{y} \quad L_+^\uparrow \approx \mathbb{R}^3 \times \mathbb{R}P^3.$$

Bibliografía

- [1] BAKER, A. *Matrix Groups an introduction to lie group theory*. Springer Undergraduate Mathematics. Springer, 2007.
- [2] BERGER, M. *Geometry I, II*, vol. 1 y 2. Springer, 1987.
- [3] BRÖCKER, T. *Lineare Algebra und Analytische Geometrie: Ein Lehrbuch Für Physiker Und Mathematiker*. Grundstudium Mathematik. Springer Basel Ag, 2004.
- [4] CURTIS, M. *Matrix Groups*, 2 ed. Universitext. Springer, 1984.
- [5] DAVIDOVICH, L., L. E. *Mecánica*, vol. 1. Reverté, 1978.
- [6] MNEIMNÉ, R. *Introduction a la théorie des groupes de lie classiques*. Hermann, 1986.
- [7] NABER, G. *The geometry of Minkowski spacetime*, 2 ed., vol. 92. Springer, 2012.
- [8] PIERRE, S. *Géométrie projective*, 1 ed. Presses Universitaires de France, 1986.
- [9] POGORELOV, A. *Geometry*. MIR, 1987.
- [10] ROE, J. *Elementary Geometry*. Oxford Science Publications, 1993.
- [11] SATTINGER, D. *Lie groups and algebras with applications to physics, geometry and mechanics*, vol. 61. Springer-Verlag, 1986.
- [12] STILLWELL, J. *Naive lie theory*. Springer, 2008.
- [13] TAPP, K. *Matrix Groups for Undergraduates*, vol. 29. AMS, 2005.