



Facultad de Ciencias

El grupo especial ortogonal $SO(3)$
The special orthogonal group
 $SO(3)$

Trabajo de Fin de Grado
para acceder al
Grado en Matemáticas

Autor: Ignacio Ahumada Vélez

Director: Fernando Etayo Gordejuela

Junio 2022

Resumen

El grupo especial ortogonal de dimensión tres $SO(3)$ es un objeto matemático de gran riqueza de estructuras. En este trabajo se presenta como grupo de matrices, isomorfo al grupo de movimientos directos de $\mathbb{R}^3$, se relaciona con los grupos unitarios de matrices, con los cuaterniones, con la esfera S^3 y el espacio proyectivo tridimensional, y se presenta como grupo de Lie.

Palabras clave: grupo especial ortogonal, grupo unitario, cuaterniones, rotaciones.

Abstract

The three-dimensional special orthogonal group $SO(3)$ is a very rich mathematical object structure-wise. In this project it is presented as a matrix group, isomorphic to the direct movement group of $\mathbb{R}^3$, related to the matrix unitary group, to the quaternions, to the S^3 sphere and the three-dimensional projective space, and presented as a Lie group.

Key words: special orthogonal group, unitary group, quaternions, rotations.

Índice

1	Introducción	1
2	SO(3) como estructura algebraica	3
2.1	SO(3) como grupo	6
2.2	Determinando la rotación definida por una matriz de SO(3)	9
2.2.1	Eje	11
2.2.2	Ángulo	12
2.3	Determinando la matriz de rotación dada por un eje de rotación y un ángulo	12
2.4	Relación entre SO(3) y SU(2)	16
2.4.1	SU(2) y S^3	25
2.4.2	SO(3) y S^3	27
2.4.3	Conclusión	30
3	SO(3) y los cuaterniones	33
3.1	Introducción	33
3.2	Formas de representar los cuaterniones	33
3.2.1	Representación Vectorial	33
3.2.2	Representación como matrices reales	34
3.2.3	Representación como matrices complejas	36
3.3	Los cuaterniones y $SO(3)$	37
3.4	Determinando una rotación a partir de un cuaternión	41
3.4.1	Determinando el eje	41
3.4.2	Determinando el ángulo	41
4	SO(3) como grupo de Lie	45
4.1	El grupo de Lie SO(3)	45
4.2	El álgebra de Lie de SO(3)	47
5	Conclusión	51

1 Introducción

En este trabajo introducimos $SO(3)$ como grupo de matrices, viéndolo como un subgrupo del grupo ortogonal $O(3)$. Una vez visto esto, estudiamos cómo se relacionan dichos grupos con las isometrías del espacio real tridimensional: $O(3)$ es isomorfo al grupo formado por las isometrías del espacio tridimensional y $SO(3)$ es a su vez isomorfo al grupo de isometrías del espacio que preservan la orientación del mismo. Seguidamente pasamos a ver como una matriz de $SO(3)$ determina una rotación de $\mathbb{R}^3$ y viceversa. Primero, vemos como dada una matriz de $SO(3)$ podemos determinar la rotación que representa, hayando el eje de rotación y el ángulo que la definen. Por otro lado, vemos como obtener una matriz de $SO(3)$ a partir de una rotación, es decir, dados un eje de rotación y un ángulo. Después de esto, pasamos a ver como $SO(3)$ es homeomorfo a $SU(2)/\mathbb{Z}_2$. Para ello, introducimos primero el grupo especial unitario $SU(2)$, viendo que es subgrupo del grupo unitario $U(2)$, para ver después que $U(2)$ y $SU(2)$ son isomorfos a los grupos de isometrías de $\mathbb{C}^2$ y al grupo de las isometrías de $\mathbb{C}^2$ que preservan la orientación, respectivamente. Una vez introducido $SU(2)$, vemos que es homeomorfo a la esfera S^3 , y después que $SO(3)$ es homeomorfo al espacio proyectivo real de dimensión 3 $\mathbb{P}^3$, y por ello $SO(3)$ es homeomorfo a $S^3/\mathbb{Z}_2$. Una vez hemos visto esto, pasamos a concluir que $SO(3)$ es homeomorfo a $SU(2)/\mathbb{Z}_2$, y que el homeomorfismo nos proporciona un doble cubrimiento de $SU(2)$ en $SO(3)$.

En la siguiente sección del trabajo pasamos a estudiar los cuaterniones. Para ello vemos primero como representarlos, de forma vectorial, como matrices reales y como matrices complejas, y damos algunas propiedades básicas de este conjunto. Después pasamos a ver como un cuaternión nos determina una rotación de $SO(3)$, para después concluir como determinar la rotación que nos define un cuaternión.

En la última sección del trabajo trabajamos con $SO(3)$ como grupo de Lie, para lo cual vemos primero que $SO(3)$ es en efecto un grupo de Lie, para lo cual también vemos que es una variedad diferenciable. Después, introducimos las álgebras de Lie y vemos que el conjunto $\mathfrak{so}(3)$ formado por las matrices antisimétricas de dimensión 3 es un álgebra de Lie, para finalmente concluir que $\mathfrak{so}(3)$ es el álgebra de Lie de $SO(3)$.

Para todo esto estudiamos $SO(3)$ a través de diversos campos matemáticos que hemos estudiado durante la carrera, como la teoría de grupos, la geometría clásica, el álgebra lineal, la topología y la geometría diferencial. Además, hemos completado varias demostraciones a lo largo del trabajo.

A lo largo de este trabajo manejaremos $\mathbb{R}^3$ como espacio vectorial y las isometrías que trataremos serán lineales.

2 $SO(3)$ como estructura algebraica

En esta sección introducimos y estudiamos $SO(3)$, así como su relación con otros conjuntos, y vemos como representar una matriz de rotación. Los conceptos de esta sección han sido extraídos de [2], [3], [4] y [12].

Comencemos desde el grupo ortogonal $O(3)$. Este está compuesto por las matrices reales cuadradas de dimensión 3 tales que su inversa es igual a su traspuesta, es decir

$$O(3) = \{A \in \mathcal{M}_{3 \times 3} : A^t = A^{-1}\}$$

Visto desde un punto de vista geométrico, este grupo está formado por las transformaciones lineales del espacio $\mathbb{R}^3$ que preservan el producto escalar canónico, el cual se define como es usual, es decir, dados $a = (a_1, a_2, a_3), b = (b_1, b_2, b_3)$ respecto de la base canónica $\{e_1, e_2, e_3\}$ de $\mathbb{R}^3$, $a \cdot b = a_1b_1 + a_2b_2 + a_3b_3$. Dichas transformaciones son conocidas como isometrías lineales. Veamos esto más detenidamente.

Definición 2.1 $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ lineal es isometría si preserva el producto escalar, es decir, dados $\vec{v}, \vec{w}$ se cumple que $\vec{v} \cdot \vec{w} = f(\vec{v}) \cdot f(\vec{w})$ siendo $\cdot$ el producto escalar usual de $\mathbb{R}^3$.

Proposición 2.2 Si f es una isometría, entonces f preserva normas y ángulos.

Demostración

Sea $\vec{v}, \vec{w} \in \mathbb{R}^3$. Primero, tenemos que $\|f(\vec{v})\| = \sqrt{f(\vec{v}) \cdot f(\vec{v})} = \sqrt{\vec{v} \cdot \vec{v}} = \|\vec{v}\|$, con lo que f preserva la norma. Por otra parte, $\cos(f(\vec{v}), f(\vec{w})) = \frac{f(\vec{v}) \cdot f(\vec{w})}{\|f(\vec{v})\| \cdot \|f(\vec{w})\|} = \frac{\vec{v} \cdot \vec{w}}{\|\vec{v}\| \cdot \|\vec{w}\|} = \cos(\vec{v}, \vec{w})$ con lo que también preserva los ángulos. $\square$

Corolario 2.3 Si f es una isometría, entonces f lleva bases ortonormales en bases ortonormales, con lo que es un isomorfismo de espacios vectoriales.

Demostración

Sea e_1, e_2, e_3 una base ortonormal de $\mathbb{R}^3$. Como f preserva ángulos y normas por ser isometría, $f(e_1), f(e_2), f(e_3)$ será también una base ortonormal.

Veámos ahora como es la estructura de $O(3)$.

Proposición 2.4 $O(3)$ es un grupo respecto del producto de matrices.

Demostración

- El producto de matrices de $O(3)$ es cerrado, es decir, queremos ver que si $A, B \in O(3)$, entonces $AB \in O(3)$. Sean $A, B \in O(3)$:

Consideramos el producto de la matriz AB por su traspuesta:

$$AB(AB)^\top = ABB^\top A^\top = AIA^\top = AA^\top = I, \text{ con lo que resulta que } AB \in O(3).$$

- $O(3)$ tiene elemento neutro respecto de la multiplicación de matrices. Tomemos la matriz identidad I , que es el elemento neutro para la multiplicación de matrices. $II^T = II = I$, con lo que $I \in O(3)$.
- $O(3)$ tiene elemento inverso para la multiplicación de matrices. Sea $A \in O(3)$: Como $A^{-1} = A^T$, y $A^T A = I$, entonces se tiene que $A^{-1} \in O(3)$.
- Tenemos que ver que la operación es asociativa, es decir, para cualesquiera $A, B, C \in O(3)$, se verifica que $(AB)C = A(BC)$.
Como el producto de matrices cumple esta propiedad en general, la cumple también en $O(3)$.

Con todo esto, concluimos que $O(3)$ es un grupo respecto de la multiplicación de matrices. $\square$

Proposición 2.5 *El conjunto $Isom(\mathbb{R}^3) = \{f : \mathbb{R}^3 \rightarrow \mathbb{R}^3 : f \text{ es isometría lineal}\}$ es un grupo respecto a la composición de aplicaciones.*

Demostración

Para esta demostración, damos por sabido que el conjunto de isomorfismos lineales es un grupo, y comprobamos que la composición de dos isometrías, la identidad, y el inverso de una isometría también lo son, esto es, que también preservan el producto escalar. No es necesario comprobar la propiedad asociativa ya que esta se cumple por cumplirse para los isomorfismos lineales en general.

- Tenemos que ver que es cerrado para la composición de aplicaciones, es decir, que para cualesquiera $f, g \in O(3)$, su composición $(f \circ g) \in Isom(\mathbb{R}^3)$. Sean $f, g \in Isom(\mathbb{R}^3)$:
Queremos que ver que $(f \circ g)(\vec{v}) \cdot (f \circ g)(\vec{w}) = \vec{v} \cdot \vec{w}$ para todo $\vec{v}, \vec{w} \in \mathbb{R}^3$. Entonces, tenemos que $(f \circ g)(\vec{v}) \cdot (f \circ g)(\vec{w}) = f(g(\vec{v})) \cdot f(g(\vec{w})) = g(\vec{v}) \cdot g(\vec{w}) = \vec{v} \cdot \vec{w}$ ya que f y g son isometrías.
- El elemento neutro para la composición de aplicaciones pertenece a $Isom(\mathbb{R}^3)$, es decir, tenemos que ver que la aplicación identidad, denotada por Id pertenece a $Isom(\mathbb{R}^3)$.
Queremos ver que $Id(\vec{v}) \cdot Id(\vec{w}) = \vec{v} \cdot \vec{w}$, pero $Id(\vec{v}) = \vec{v}$, $\forall \vec{v} \in \mathbb{R}^3$, con lo que se ve trivialmente.
- $Isom(\mathbb{R}^3)$ tiene elemento inverso para toda $f \in Isom(\mathbb{R}^3)$. Sea $f \in Isom(\mathbb{R}^3)$:
Tenemos que ver que $f^{-1}(f(\vec{v})) \cdot f^{-1}(f(\vec{w})) = \vec{v} \cdot \vec{w}$ para cualesquiera $\vec{v}, \vec{w} \in \mathbb{R}^3$. Consideramos la igualdad $f(\vec{v}) \cdot f(\vec{w}) = \vec{v} \cdot \vec{w}$ que sabemos cierta por ser $f \in Isom(\mathbb{R}^3)$. Como $f^{-1}(f(v)) = v$ para cualquier v , podemos escribir la misma igualdad en términos de f^{-1} , y quedaría $\vec{v} \cdot \vec{w} = f^{-1}(f(\vec{v})) \cdot f^{-1}(f(\vec{w}))$ que es el resultado que queríamos.

Con lo que concluimos que $Isom(\mathbb{R}^3)$ tiene estructura de grupo con la composición de aplicaciones. $\square$

Y ahora que hemos visto que tanto $O(3)$ como $Isom(\mathbb{R}^3)$ tienen estructura pasamos a ver la relación entre estas.

Teorema 2.6 *La aplicación:*

$$g : O(3) \rightarrow Isom(\mathbb{R}^3)$$

$$A \mapsto f \in Isom(\mathbb{R}^3) \text{ tal que } M(f) = A$$

está bien definida y es un isomorfismo de grupos, donde $M(f)$ es la matriz de f respecto de la base canónica de $\mathbb{R}^3$.

Demostración

Veamos primero que está bien definida. Para ello tenemos que ver que $g(A) = f$ es única y que existe dicha f en $Isom(\mathbb{R}^3)$ tal que $g(A) = f$. Es claro que dicha f es única. Queda ver que tenemos tal $f \in Isom(\mathbb{R}^3)$. Sean ahora $v = \sum v_i e_i$ y $w = \sum w_i e_i$. Entonces $f(v) \cdot f(w) = (Av)^\top Aw = v^\top A^\top Aw = v^\top w = v \cdot w$ y efectivamente una matriz $A \in O(3)$ tiene como imagen f que será isomorfismo lineal porque $\det(M(f)) = \det(A) \neq 0$, y por ello $f \in Isom(\mathbb{R}^3)$. Con lo que g está bien definida.

Veamos ahora que $g : O(3) \rightarrow Isom(\mathbb{R}^3)$ es inyectiva. Para ello, tenemos que ver que a dos elementos distintos de $O(3)$ nunca les corresponde la misma isometría $f \in Isom(\mathbb{R}^3)$. Razonamos por reducción al absurdo. Supongamos que la aplicación no es inyectiva. Entonces, para algunas matrices $A, B \in O(3)$, $A \neq B$, se tendría que $g(A) = g(B)$. En ese caso tendríamos $f \in Isom(\mathbb{R}^3)$ de forma que $g(A) = g(B) = f$, es decir, tendríamos que $M(f) = A = B$. Pero esto es absurdo, puesto que hemos tomado matrices distintas. Como hemos supuesto que la aplicación no era inyectiva y hemos llegado a un absurdo, concluimos que la aplicación es inyectiva.

Veamos ahora que la aplicación es sobreyectiva. Tenemos que ver que dada para todo elemento de $Isom(\mathbb{R}^3)$ tenemos otro elemento en $O(3)$ que sea su contraimagen. Sea $A = M(f)$ con $f \in Isom(\mathbb{R}^3)$. Como f es isometría, conserva el producto escalar, es decir, $\vec{v} \cdot \vec{w} = f(\vec{v}) \cdot f(\vec{w})$ para todo $v, w \in \mathbb{R}^3$. Matricialmente, el producto escalar se puede expresar de la siguiente forma:

$$f(\vec{v}) \cdot f(\vec{w}) = (A\vec{v})^\top A\vec{w} = \vec{v}A^\top A\vec{w} = \vec{v} \cdot \vec{w}$$

Estudiemos más detenidamente la última igualdad. Sea $\vec{v}$ un vector cuya entrada i -ésima es igual a 1, y el resto son 0, y $\vec{w}$ un vector un vector cuya entrada j -ésima es igual a 1, y el resto son 0. Entonces, el elemento b_{ij} de la matriz $A^\top A$ será:

$$b_{ij} = \vec{v}^\top (A^\top A) \vec{w} = (0, \dots, 1, \dots, 0) (A^\top A) \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \vec{v}^\top \vec{w} = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

de donde concluimos que $A^\top A = I$ y por lo tanto $A \in O(3)$ será su contraimagen.

Como hemos visto que $g : O(3) \rightarrow Isom(\mathbb{R}^3)$ es inyectiva y sobreyectiva, se tiene que g es una biyección.

Por último, para poder concluir que es un isomorfismo de grupos, tenemos que ver que la aplicación g y su inversa g^{-1} son homomorfismos, es decir, que las aplicaciones g y g^{-1} respetan las operaciones al pasar de un conjunto a otro. Para ello tenemos que comprobar que dadas $A, B \in O(3)$, $g(AB) = g(A) \circ g(B)$. Sean $f, h \in Isom(\mathbb{R}^3)$ tales que $g(A) = f, g(B) = h$, es decir $M(f) = A$ y $M(h) = B$. Ahora bien, la composición $(f \circ h)(x) = f(h(x))$ se expresa matricialmente como $M(f)M(h)x = ABx$ para $x \in \mathbb{R}^3$, con lo que efectivamente, la imagen de AB a través de g será la composición $f \circ h$ con $M(f) = A$ y $M(h) = B$. Por otra parte, también tenemos que ver que para cualesquiera $f, h \in Isom(\mathbb{R}^3)$, se tiene que $g^{-1}(f \circ h) = g^{-1}(f)g^{-1}(h)$. La aplicación inversa hará corresponder a cada $f \in Isom(\mathbb{R}^3)$ la matriz A tal que $M(f) = A$. Sean $f, h \in Isom(\mathbb{R}^3)$, y consideramos su composición $f \circ h$. Como ya hemos visto, la composición $(f \circ h)(x) = f(h(x))$ para $x \in \mathbb{R}^3$ se expresa matricialmente como $M(f)M(h)x$. Sean ahora A, B tales que $M(f) = A$ y $M(h) = B$. Entonces la matriz de la composición $f \circ h$ será AB y tendremos que $g^{-1}(f \circ h) = AB = g^{-1}(f)g^{-1}(h)$ que es lo que queríamos ver. Con esto, g es un homomorfismo de grupos.

Como hemos probado que g es un homomorfismo biyectivo de grupos, concluimos que g es un isomorfismo de grupos. $\square$

Al probar este resultado hemos llegado a que podemos identificar cualquier matriz del grupo ortogonal con una isometría de $\mathbb{R}^3$.

2.1 $SO(3)$ como grupo

Estudiemos ahora la igualdad $AA^\top = I$ que cumplen por definición todas las matrices de $O(3)$. Si consideramos el determinante a ambos lados de la igualdad llegamos a que $1 = \det(A)\det(A^\top) = \det(A)^2$, y por lo tanto $\det(A) = \pm 1$. Las matrices de $O(3)$ que cumplen $\det(A) = 1$ forman un subgrupo, conocido como grupo especial ortogonal y se denota por $SO(3)$. Estas isometrías preservan la orientación del espacio, mientras que las que tienen determinante -1 la invierten.

Definición 2.7 *El conjunto definido de la siguiente forma:*

$$\{A \in \mathcal{M}_{3 \times 3}(\mathbb{R}) : AA^\top = I, \det(A) = 1\}$$

donde I es la matriz identidad de dimensión 3 se denomina grupo especial ortogonal o $SO(3)$.

Como su nombre indica, $O(3)$ tiene estructura de grupo con el producto de matrices puesto que el producto de dos matrices que preservan el producto escalar lo seguirá preservando.

Veamos ahora cómo se relacionan las estructuras de $O(3)$ y $SO(3)$.

Proposición 2.8 *$SO(3)$ es subgrupo de $O(3)$.*

Demostración

- $SO(3) \subseteq O(3)$

Se ve trivialmente por definición.

- El producto de matrices de $SO(3)$ es cerrado, es decir, queremos ver que si $C, D \in SO(3)$, entonces $CD \in SO(3)$. Sean $C, D \in SO(3)$:

Consideramos el determinante del producto de matrices CD :

$\det(CD) = \det(C)\det(D) = 1$ ya que $CD \in SO(3)$, luego $CD \in SO(3)$ y es cerrado para la multiplicación de matrices.

- $SO(3)$ tiene al elemento neutro para la multiplicación de matrices. Tomamos la matriz I , que es el elemento neutro de $O(3)$:

Como $\det(I) = 1$, entonces $I \in SO(3)$ y $SO(3)$ tiene elemento neutro para la multiplicación.

- Tiene elemento inverso para la multiplicación de matrices. Sea $C \in SO(3)$:

Como $\det(C^{-1}) = \frac{1}{\det(C)} = 1$, entonces la inversa de C pertenece también a $SO(3)$.

Con todo esto hemos demostrado que $SO(3)$ es un subgrupo de $O(3)$. $\square$

Además, es un grupo no conmutativo. Por ejemplo:

$$\text{Sean } A = \begin{pmatrix} \frac{4}{5} & \frac{3}{5} & 0 \\ \frac{-3}{5} & \frac{4}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix}, B = \begin{pmatrix} \frac{-2}{3} & \frac{-2}{3} & \frac{1}{3} \\ \frac{-2}{3} & \frac{1}{3} & \frac{-2}{3} \\ \frac{1}{3} & \frac{-2}{3} & \frac{-2}{3} \end{pmatrix}. A, B \in SO(3).$$

Consideremos ahora los siguientes productos de matrices:

$$AB = \begin{pmatrix} \frac{-14}{15} & \frac{-1}{3} & \frac{-2}{15} \\ \frac{-2}{15} & \frac{2}{3} & \frac{-11}{15} \\ \frac{1}{15} & \frac{-2}{3} & \frac{-2}{3} \end{pmatrix}, BA = \begin{pmatrix} \frac{-2}{15} & \frac{-14}{15} & \frac{1}{3} \\ \frac{-11}{15} & \frac{-2}{15} & \frac{-2}{3} \\ \frac{2}{15} & \frac{-1}{3} & \frac{-2}{3} \end{pmatrix}$$

Como se puede apreciar, $AB \neq BA$. Esto tiene sentido geoméricamente ya que dados dos ejes de rotación se obtienen resultados distintos según cual elijamos para rotar primero.

Hasta ahora hemos introducido el conjunto $O(3)$, hemos visto que es un grupo y su relación con $Isom(\mathbb{R}^3)$ por medio de un isomorfismo de grupos. Por otra parte hemos introducido $SO(3)$ y hemos visto que es un subgrupo de $O(3)$. Entonces, es razonable pensar que $SO(3)$ estará relacionado con un subgrupo de $Isom(\mathbb{R}^3)$. Efectivamente esto es así, veámoslo con detenimiento:

Definición 2.9 f preserva la orientación si $\det(M(f)) > 0$.

Proposición 2.10 El conjunto

$$Isom^+(\mathbb{R}^3) = \{f : \mathbb{R}^3 \rightarrow \mathbb{R}^3 : f \text{ es isometría lineal y preserva la orientación}\}$$

es subgrupo de $Isom(\mathbb{R}^3)$.

Demostración

- $Isom^+(\mathbb{R}^3) \subseteq Isom(\mathbb{R}^3)$

Se ve trivialmente por definición.

- La composición de aplicaciones es cerrada, es decir, queremos ver que dadas $f, g \in Isom^+(\mathbb{R}^3)$, entonces $(f \circ g) \in Isom^+(\mathbb{R}^3)$. Sean $f, g \in Isom^+(\mathbb{R}^3)$:

Sean $A = M(f)$ y $B = M(g)$. Considero la composición de aplicaciones $f \circ g$. Matricialmente, se expresaría como $(f \circ g)(v) = f(g(v)) = f(Bv) = ABv$ para $v \in \mathbb{R}^3$. Resulta que la matriz de $f \circ g$ es AB . Pero entonces $\det(AB) = \det(A)\det(B) > 0$ por ser producto de dos números mayores que 0. Con lo que efectivamente, $f \circ g$ preserva la orientación.

- El elemento neutro de la composición de aplicaciones pertenece a $Isom^+(\mathbb{R}^3)$. Sea id la aplicación identidad, que es el neutro para la composición de aplicaciones:

$M(Id) = I_3$ la matriz identidad. $\det(I_3) = 1 > 0$, con lo que el elemento neutro pertenece a $Isom^+(\mathbb{R}^3)$.

- Para toda $f \in Isom^+(\mathbb{R}^3)$, se tiene que el elemento inverso de f pertenece a $Isom^+(\mathbb{R}^3)$. Sea $f \in Isom^+(\mathbb{R}^3)$

Si $M(f) = A$, entonces $M(f^{-1}) = A^{-1}$. Entonces, $\det(A^{-1}) = \frac{1}{\det(A)} > 0$ ya que $\det(A) > 0$ por ser $f \in Isom^+(\mathbb{R}^3)$, con lo que $f^{-1} \in Isom^+(\mathbb{R}^3)$.

Con todo esto, concluimos que $Isom^+(\mathbb{R}^3)$ es subgrupo de $Isom(\mathbb{R}^3)$ respecto a la composición de aplicaciones. $\square$

Y ahora pasamos a ver la relación entre $SO(3)$ y $Isom^+(\mathbb{R}^3)$:

Teorema 2.11 *La aplicación:*

$$g|_{SO(3)} : SO(3) \rightarrow Isom^+(\mathbb{R}^3)$$

$$A \mapsto f \in Isom^+(\mathbb{R}^3) \text{ tal que } M(f) = A$$

donde g es la aplicación definida en el teorema 2.6 es un isomorfismo de grupos.

Demostración

Podemos ver que $g|_{SO(3)}$ es un homomorfismo de la misma forma que en el teorema 2.6, y lo mismo ocurre con la inyectividad.

En cuanto a la sobrectividad, tenemos que ver que para todo elemento de $Isom^+(\mathbb{R}^3)$ tenemos un elemento de $SO(3)$ que sea su contraimagen. Sea $f \in Isom^+(\mathbb{R}^3)$, con $M(f) = A$. Como $f \in Isom^+(\mathbb{R}^3)$, $\det(A) > 0$. Sean $\vec{v}, \vec{w} \in \mathbb{R}^3$. Expresamos matricialmente el producto escalar como hicimos en el teorema 2.6:

$$g(\vec{v}) \cdot g(\vec{w}) = \vec{v}A^\top A\vec{w} = \vec{v}\vec{w}$$

Y estudiamos una vez más la última igualdad de la misma forma que hicimos previamente, y llegamos a que $A^\top A = I$. Nos faltaría ver que $\det(A) = 1$ para poder concluir que $A \in SO(3)$. Viendo el determinante a ambos lados de la igualdad obtenemos que $1 = \det(I) = \det(A^\top A) = \det(A^\top)\det(A)$. Tenemos que $\det(A^\top) = \det(A)$, luego $\det(A) = \pm 1$. Pero sabemos que $\det(A) > 0$ ya que $A = M(f)$ y $f \in Isom^+(\mathbb{R}^3)$. Con lo que $A \in SO(3)$ y la aplicación es sobreyectiva.

Como hemos visto que es un homomorfismo, que es inyectiva y biyectiva concluimos que es un isomorfismo de grupos. $\square$

Con todo lo que hemos visto en esta sección, nos quedaría el siguiente esquema:

$$\begin{array}{ccc} O(3) & \xleftarrow{g} & Isom(\mathbb{R}^3) \\ \cup & & \cup \\ SO(3) & \xleftarrow{g} & Isom^+(\mathbb{R}^3) \end{array}$$

Con el cual vemos no solo que $SO(3)$ e $Isom^+(\mathbb{R}^3)$ son subgrupos de $O(3)$ e $Isom(\mathbb{R}^3)$ respectivamente, si no también como las relaciones entre estos se mantienen a través del isomorfismo g al pasar de un lado a otro.

Visto desde un punto de vista geométrico, los elementos de $SO(3)$ son las rotaciones del espacio $\mathbb{R}^3$, es decir, transformaciones lineales que son un giro respecto de una recta vectorial de $\mathbb{R}^3$ que queda fija a través de la rotación. Veamos esto con más detalle:

2.2 Determinando la rotación definida por una matriz de $SO(3)$

Como ya se ha mencionado previamente, las matrices de $SO(3)$ definen rotaciones vectoriales en $\mathbb{R}^3$. Toda rotación distinta de la identidad deja fijo un subespacio de dimensión 1 al que conocemos como eje de la rotación. Una vez conocemos el eje de rotación podemos ver la rotación en tres dimensiones como una rotación en dos dimensiones en el plano ortogonal al eje en cada punto de este.

Además, para determinar la rotación necesitamos un ángulo que nos diga cuánto girar. También hay que determinar una forma de saber en que sentido rotamos para un cierto eje generado por $t \in \mathbb{R}^3$, pero podemos solucionar esto asumiendo un sentido de rotación y si queremos girar en el sentido opuesto simplemente tomamos $-t$ como generador del eje.

Ahora sea $A \in SO(3)$. En esta sección estudiaremos cómo encontrar la rotación cuya matriz es A .

Proposición 2.12 *La matriz A siempre posee $\lambda = 1$ como autovalor. El resto de autovalores dependerán de cómo sea la rotación que define A . Si tenemos la rotación trivial aparecerá el autovalor $\lambda = 1$ otras 2 veces. Si tenemos una rotación de ángulo π , aparecerá el autovalor $\lambda = -1$ dos veces. En el caso general tendremos dos autovalores complejos conjugados.*

Demostración

Sea $t \in \mathbb{R}^3$ tal que $|t| = 1$ y sean $u, v \in \mathbb{R}^3$ ortonormales a t . Estos vectores forman una base ortonormal, luego podemos expresar una matriz de rotación en función de

ellos. Supongamos que t genera el eje de la rotación. Como ya hemos mencionado, las columnas de la matriz de rotación serán la imagen de los vectores de la base canónica a través de la rotación. Como t es el eje de esta, quedará invariante, luego la primera columna será $(1, 0, 0)$. Por otra parte, como los otros dos vectores de la base son ortonormales a t , la rotación actuará sobre ellos como si fuese una rotación en el plano. Veámoslo detenidamente:

Sea $f \in Isom(\mathbb{R}^3)$ la isometría tal que $M(f) = A$ y sea Π el plano generado por u, v . Queremos ver que Π queda fijo a través de f , es decir, $f(\Pi) = \Pi$. Tenemos la base $\{u, v\}$ de Π por lo que $\Pi \perp t$ ya que $u \perp t$ y $v \perp t$. Vemos ahora que pasa con $f(u)$ y $f(v)$. Como f es una isometría, preserva los ángulos por lo que $f(u) \perp f(t)$ y $f(v) \perp f(t)$. Pero $f(t) = t$ por ser t el generador del eje de rotación, con lo que $f(u) \perp t$ y $f(v) \perp t$. Como consecuencia de esto, tendremos que $f(u), f(v) \in \Pi$, y por ello y por ser f isomorfismo, tendremos que $f(\Pi) = \Pi$ que es lo que queríamos.

Con todo esto, una rotación de eje t y ángulo θ tendrá la siguiente forma en la base descrita:

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$$

Como el polinomio característico de la matriz es independiente de la base escogida, puedo usar esta matriz para calcular los autovalores de una rotación en general, puesto que dada una rotación cualquiera siempre la puedo expresar en función de una base de la que se ha usado. Hagámoslo:

$$\det(M - \lambda I) = \det \begin{pmatrix} 1 - \lambda & 0 & 0 \\ 0 & \cos \theta - \lambda & -\sin \theta \\ 0 & \sin \theta & \cos \theta - \lambda \end{pmatrix} = (1 - \lambda)((\cos \theta - \lambda)^2 + \sin^2 \theta)$$

Obtenemos del primer paréntesis la solución $\lambda = 1$. Desarrollamos el segundo y llegamos a la ecuación de segundo grado $\lambda^2 - 2\lambda \cos \theta + \cos^2 \theta + \sin^2 \theta = \lambda^2 - 2\lambda \cos \theta + 1$. Resolvemos y llegamos a $\lambda = \cos \theta \pm \sqrt{\cos^2 \theta - 1}$. Como $\cos^2 \theta \in [0, 1]$, tendremos una raíz de un número negativo siempre que $\theta \notin \{0, \pi\}$, y tendremos en ese caso dos autovalores complejos $\lambda_1 = e^{i\theta}$ y $\lambda_2 = e^{-i\theta}$, siendo uno el conjugado del otro. Si $\theta = 0$ o $\theta = \pi$, la raíz se anulará y tendremos repetidos los autovalores $\lambda = 1$ y $\lambda = -1$ respectivamente.

El autovalor 1 siempre está presente en la primera raíz del polinomio puesto que es el que se corresponde con el eje de la rotación, que queda fijo por la misma. En cuanto a las otras dos raíces, si $\theta = 0$ solo tenemos el 1 repetido otras dos veces como autovalor, puesto que si rotamos el espacio un ángulo $\theta = 0$ el espacio entero se queda fijo. Por otra parte, si $\theta = \pi$ nos encontramos el autovalor -1 repetido dos veces ya que girar el espacio π rad es equivalente a transformar cada vector en su opuesto. En el resto de casos nos encontraremos con dos autovalores complejos conjugados ya que el polinomio coaracterístico es real. $\square$

2.2.1 Eje

Proposición 2.13 *El eje de la rotación dada por la siguiente matriz*

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

está generado por $t = (a_{32} - a_{23}, a_{13} - a_{31}, a_{21} - a_{12})$.

Demostración

Sabemos que el eje de la rotación es invariante por la rotación definida por A , con lo que si llamamos t a un generador del eje, se cumplirá:

$$At = t, \text{ que es equivalente a } At = It$$

Lo cual nos lleva a que $(A - I)t = 0$, es decir t pertenece al núcleo de $A - I$, o visto de otra forma, $\lambda = 1$ es valor propio de A y t es un vector propio asociado. Como ya hemos visto, las matrices de rotación siempre poseen este valor propio, mientras que los otros dos serán complejos en el caso general. De aquí deducimos que una matriz de rotación tiene un único vector propio real salvo multiplicación por constante.

Ahora, para llegar al eje de rotación, consideramos la siguiente cadena de igualdades:

$$0 = A^T 0 + 0 = A^T(A - I)t + (A - I)t = (A^T A - A^T + A - I)t = (A - A^T)t$$

Ahora estudiamos cómo es la matriz $A - A^T$. Para $1 \leq i, j \leq 3$, siendo a_{ij} el elemento de la i -ésima fila y de la j -ésima columna de la matriz A tendremos que $(A - A^T)_{ij} = a_{ij} - a_{ji}$. Concretamente, si $i = j$, $(A - A^T)_{ij} = 0$. Si consideramos su traspuesta, vemos que $(A - A^T)_{ij}^T = a_{ji} - a_{ij} = -(A - A^T)_{ij}$. Con esto, concluimos que la matriz $A - A^T$ es antisimétrica.

Ahora que sabemos que la matriz $A - A^T$ es antisimétrica, podemos tomar un t de forma que:

$$[t]_{\times} = (A - A^T)$$

donde

$$t = (t_1, t_2, t_3), \text{ y } [t]_{\times} = \begin{pmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{pmatrix}$$

Con esto, consideramos el siguiente producto:

$$[t]_{\times} t = \begin{pmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} = \begin{pmatrix} t_2 t_3 - t_3 t_2 \\ t_3 t_1 - t_1 t_3 \\ t_1 t_2 - t_2 t_1 \end{pmatrix} = t \times t = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Con esto, el t que hemos escogido nos garantiza que el resultado siempre sea 0. Entonces:

Si

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \text{ entonces } t = (a_{32} - a_{23}, a_{13} - a_{31}, a_{21} - a_{12}).$$

Y queda probado lo que queríamos ver. $\square$

Si la matriz A fuese simétrica este proceso no sería válido. En este caso, tendríamos que $A - A^T = 0$. Al seguir el proceso descrito, tomaríamos $[t]_{\times}$ y por ello llegaríamos a que el vector $t = (0, 0, 0)$ es el eje de la rotación, lo cual no tiene sentido. En este caso una forma alternativa de calcular el eje de la rotación sería diagonalizar la matriz y calcular el vector propio correspondiente al valor propio 1.

2.2.2 Ángulo

Proposición 2.14 *El ángulo de la rotación definida por una matriz $A \in SO(3)$ es $\alpha = \arccos \left(\frac{(Av)v}{|Av||v|} \right)$.*

Demostración

Sea una matriz $A \in SO(3)$ y sea $v \in \mathbb{R}^3$ ortogonal al eje de rotación dado por A . Entonces, v estará en el plano ortogonal al eje de rotación, y también lo estará su transformado Av . La forma más sencilla de hallar el ángulo de rotación es calculando el ángulo entre v y Av , es decir, calcularíamos $\alpha = \arccos \left(\frac{(Av)v}{|Av||v|} \right)$. $\square$

Otra forma de llegar a este resultado es considerar la traza de la matriz A . La traza de la matriz es igual a la suma de los valores propios, y como ya conocemos estos, $tr(A) = 1 + 2\cos\theta$. De aquí deducimos que $\theta = \arccos \left(\frac{tr(A) - 1}{2} \right)$.

Podemos llegar a este resultado por otro camino, ya que como la traza de la matriz es un invariante algebraico, $tr(A)$ será igual a la traza de la matriz que expresa la misma rotación en función de la base ortonormal formada por el eje de la rotación con norma igual a 1, y dos vectores ortonormales a este. Como ya se ha visto previamente en la sección 2.2 de este trabajo, nos quedará una matriz con la forma de la M descrita en dicha sección. La traza de esa matriz es la misma que a la que hemos llegado con el método anterior, y concluiríamos de la misma forma que $\theta = \arccos \left(\frac{tr(A) - 1}{2} \right)$.

2.3 Determinando la matriz de rotación dada por un eje de rotación y un ángulo

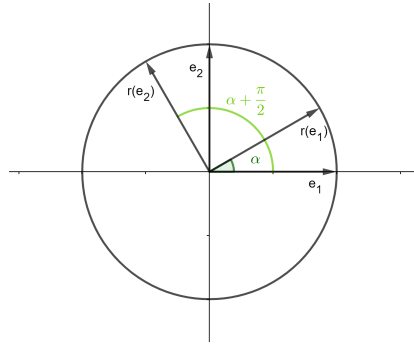
Cuando aplicamos una rotación al espacio, lo que estamos haciendo realmente es asignarle una nueva base ortonormal. Si consideramos la base canónica de $\mathbb{R}^3$, $\{e_1, e_2, e_3\}$, las columnas de una matriz $A \in SO(3)$ serán Ae_1, Ae_2 y Ae_3 . Como sabemos que esta matriz representa una isometría, sabemos que preservará la norma, por lo que los vectores Ae_1, Ae_2, Ae_3 también serán ortogonales entre si y seguirán teniendo norma 1, luego formarán una nueva base ortonormal.

Para determinar una matriz de rotación con un eje cualquiera, primero tenemos que conocer las matrices de rotación de ejes x,y o z. Consideremos primero una matriz de

rotación en dos dimensiones. Como esta matriz es una transformación lineal, sus columnas serán la imagen de los vectores de la base canónica a través de esta transformación.

Como los vectores de la base canónica e_1, e_2 tienen norma 1, los estaremos desplazando a través de la circunferencia S^1 tanto como nos diga el ángulo θ de rotación. Por tanto, si giramos $e_1 = (1, 0)$ su imagen será $(\cos\theta, \sin\theta)$. Por otra parte, el transformado de $e_2 = (0, 1)$ estará girado $\frac{\pi}{2}$ más que el transformado de e_1 ya que estos son ortogonales entre si. Por tanto podemos ver la imagen de e_2 como la imagen de rotar e_1 el ángulo $\theta + \frac{\pi}{2}$, es decir $(\cos(\theta + \frac{\pi}{2}), \sin(\theta + \frac{\pi}{2}))$.

Por un lado, se tiene que $\sin(\theta + \frac{\pi}{2}) = \cos\theta$. Por otro, usando la misma igualdad, $\cos(\theta + \frac{\pi}{2}) = \sin(\theta + \pi)$ que es a su vez $-\sin\theta$, con lo que el transformado de e_2 será $(-\sin\theta, \cos\theta)$.



En esta imagen vemos lo descrito en el párrafo anterior, donde $r(v)$ denota el transformado de un vector v a través de una rotación de ángulo α .

Con todo esto, llegamos a que una matriz de rotación en dos dimensiones de ángulo θ tiene la siguiente forma:

$$M = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

Si ahora consideramos esa misma rotación en un espacio de tres dimensiones, es decir, una rotación del plano xy, podemos expresarla extendiendo la matriz que ya teníamos y (coloquialmente hablando) "rellenándola" con componentes de la matriz identidad. Esto resultará en una matriz que no solo gira el plano xy, si no todo el espacio alrededor del eje z:

$$M_z = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Si ahora en vez de rellenar la última columna con la identidad, "rellenamos" la primera o la segunda columna obtendremos matrices que giran todo el espacio alrededor de los ejes x e y respectivamente:

$$M_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}, M_y = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix}$$

Ahora que conocemos las matrices de rotación elementales, sea $v = (v_1, v_2, v_3) \in \mathbb{R}^3$ un eje de rotación cualquiera y θ un ángulo. Para llegar a la matriz que describe la rotación de eje v y ángulo θ el proceso que seguiremos será el de transformarla en una rotación alrededor del eje z mediante rotaciones cuya forma ya conocemos.

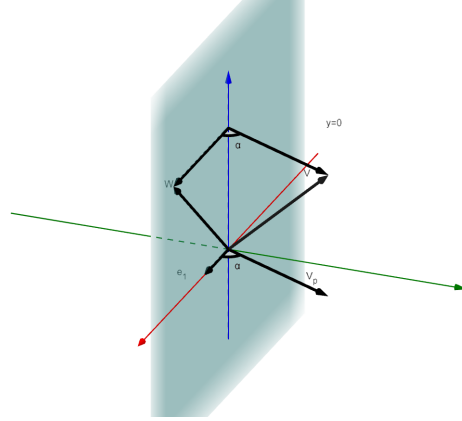
1. Lo primero que haremos será girar nuestro eje alrededor del eje z hasta que esté en el plano xz . Como ya conocemos las matrices de rotación alrededor del eje z , lo único que necesitamos conocer es el ángulo que debemos rotar.

Para ello consideramos los vectores $v_p = (v_1, v_2, 0)$ y $e_1 = (1, 0, 0)$. Si calculo el ángulo entre estos vectores tendré el ángulo para construir la rotación que busco, ya que el vector e_1 está en el plano xz y en el plano xy , y el vector v_p es la proyección de v sobre el plano xy . Con esto, calculamos $\cos(\alpha) = \frac{v_p \cdot e_1}{|v_p||e_1|} = \frac{v_1}{\sqrt{v_1^2 + v_2^2}}$. Como

$\sin^2 \alpha + \cos^2 \alpha = 1$, $\sin \alpha = \pm \frac{v_2}{\sqrt{v_1^2 + v_2^2}}$. El signo que satisface que se anule la segunda coordenada y por ello nos proporciona el resultado que queremos es el menos, así que con todo esto la matriz de rotación de este primer paso es:

$$P_1 = \begin{pmatrix} \frac{v_1}{\sqrt{v_1^2 + v_2^2}} & \frac{v_2}{\sqrt{v_1^2 + v_2^2}} & 0 \\ -\frac{v_2}{\sqrt{v_1^2 + v_2^2}} & \frac{v_1}{\sqrt{v_1^2 + v_2^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

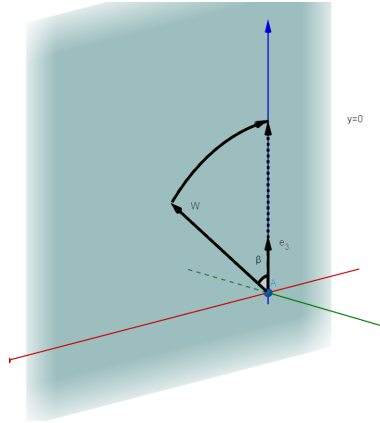
Al vector que obtenemos como transformado de v le llamaremos $w = (\sqrt{v_1^2 + v_2^2}, 0, v_3)^\top$.



2. Después, tendremos que transformar w de forma que se superponga con el eje z . Para ello tendremos que rotar alrededor del eje y y el ángulo formado entre w y $e_3 = (0, 0, 1)$. Calculamos $\cos(\beta) = \frac{w \cdot e_3}{|w||e_3|} = \frac{v_3}{\sqrt{v_1^2 + v_2^2 + v_3^2}}$, y siguiendo el mismo proceso que en el paso anterior, $\sin \beta = \pm \frac{\sqrt{v_1^2 + v_2^2}}{\sqrt{v_1^2 + v_2^2 + v_3^2}}$. Esta vez el

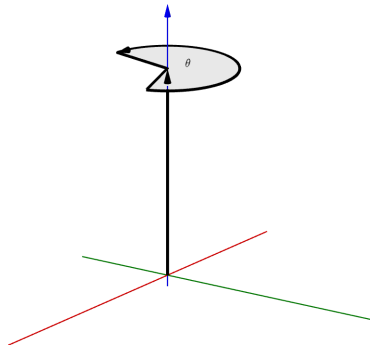
signo que hace que se anule el valor de la primera coordenada es el más. Con esto, la matriz de rotación de este paso es:

$$P_2 = \begin{pmatrix} \frac{v_3}{\sqrt{v_1^2 + v_2^2 + v_3^2}} & 0 & -\frac{\sqrt{v_1^2 + v_2^2}}{\sqrt{v_1^2 + v_2^2 + v_3^2}} \\ 0 & 1 & 0 \\ \frac{\sqrt{v_1^2 + v_2^2}}{\sqrt{v_1^2 + v_2^2 + v_3^2}} & 0 & \frac{v_3}{\sqrt{v_1^2 + v_2^2 + v_3^2}} \end{pmatrix}$$



3. Ahora que ya hemos trasladado nuestro eje original de forma que se superpone con el eje z, tenemos que rotar respecto de este el ángulo de nuestra rotación original θ . Entonces, la matriz de rotación será:

$$P_3 = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



4. Una vez ya hemos hecho la rotación original en el eje z, tenemos que "deshacer el camino" que hemos seguido en los dos primeros pasos. Para ello, tenemos que

hacer los movimientos opuestos a los que hemos hecho. Como las matrices de rotación que hemos usado pertenecen a $SO(3)$, cumplen que su inversa es su traspuesta, por lo que no tenemos que calcular nuevas matrices, con trasponer las que ya hemos usado obtendremos lo que necesitamos.

Para deshacer la rotación del paso 2, usaremos la siguiente matriz:

$$P_4 = P_2^\top = \begin{pmatrix} \frac{v_3}{\sqrt{v_1^2 + v_2^2 + v_3^2}} & 0 & \frac{\sqrt{v_1^2 + v_2^2}}{\sqrt{v_1^2 + v_2^2 + v_3^2}} \\ 0 & 1 & 0 \\ -\frac{\sqrt{v_1^2 + v_2^2}}{\sqrt{v_1^2 + v_2^2 + v_3^2}} & 0 & \frac{v_3}{\sqrt{v_1^2 + v_2^2 + v_3^2}} \end{pmatrix}$$

5. Igual que en el paso anterior, para deshacer el paso 1, uso la matriz traspuesta:

$$P_5 = P_1^\top = \begin{pmatrix} \frac{v_1}{\sqrt{v_1^2 + v_2^2}} & -\frac{v_2}{\sqrt{v_1^2 + v_2^2}} & 0 \\ \frac{v_2}{\sqrt{v_1^2 + v_2^2}} & \frac{v_1}{\sqrt{v_1^2 + v_2^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Con todo esto, hemos expresado una rotación de eje cualquiera en función de rotaciones de ejes x, y o z. Como $SO(3)$ es un grupo con respecto a la multiplicación de matrices, el producto de todas las matrices que hemos usado será la matriz de rotación de eje v y ángulo θ que queríamos obtener:

$$R_{(v,\theta)} = P_5 P_4 P_3 P_2 P_1$$

2.4 Relación entre $SO(3)$ y $SU(2)$

El grupo unitario $U(2)$ está formado por las matrices cuadradas complejas de dimensión 2 unitarias, es decir, cuya inversa es igual a su traspuesta conjugada. Estas matrices representan las transformaciones lineales de $\mathbb{C}^2$ que conservan el producto escalar de $\mathbb{C}^2$, es decir, las isometrías lineales de $\mathbb{C}^2$. Pero para considerar las isometrías de $\mathbb{C}^2$, tenemos que conocer primero su producto escalar.

Si tratásemos de usar el mismo producto que en $\mathbb{R}^2$, nos encontraríamos con un problema. Consideremos la siguiente aplicación $((z_1, z_2), (w_1, w_2)) \mapsto z_1 w_1 + z_2 w_2$ como producto escalar en $\mathbb{C}^2$. Si fuese un producto escalar, se verificaría que dado $(z_1, z_2) \in \mathbb{C}^2$, $z_1 = a + bi$, $z_2 = c + di$ tales que $(z_1, z_2) \neq (0, 0)$, entonces $(z_1, z_2) \cdot (z_1, z_2) > 0$. Pero esto no se cumple. Por ejemplo, si tomamos $z_1 = bi$ y $z_2 = di$ tendríamos $(z_1, z_2) \cdot (z_1, z_2) = z_1 z_1 + z_2 z_2 = a^2 - b^2 + 2abi + c^2 + d^2 + 2cdi = -b^2 - d^2 < 0$ y la propiedad no se verificaría. De hecho, en un caso general no tendría ni por que pertenecer a $\mathbb{R}$. Es por esto que tenemos que conocer primero cómo son los productos escalares en espacios vectoriales complejos, y por tanto, cómo definimos un espacio vectorial complejo:

Definición 2.15 Sea V un espacio vectorial sobre $\mathbb{C}$. Una aplicación $\varphi : V \times V \rightarrow \mathbb{C}$ es una forma sesquilineal si cumple las siguientes propiedades:

1. $\varphi(x + y, z) = \varphi(x, z) + \varphi(y, z)$
2. $\varphi(x, y + z) = \varphi(x, y) + \varphi(x, z)$
3. $\varphi(\lambda x, z) = \lambda \varphi(x, z)$
4. $\varphi(x, \lambda z) = \bar{\lambda} \varphi(x, z)$

para todo $x, y, z \in V$, $\lambda \in \mathbb{C}$ y $\bar{z}$ denota el conjugado de un número complejo.

Definición 2.16 Sea $\varphi : V \times V \rightarrow \mathbb{C}$ una forma sesquilineal. Si $\varphi(x, y) = \overline{\varphi(y, x)}$, entonces es una forma hermítica.

Observación 2.17 Si queremos ver si una aplicación es una forma hermítica, bastaría con comprobar que se cumplen las condiciones 1 y 3 de la definición 2.15 y que $\varphi(x, y) = \overline{\varphi(y, x)}$, ya que con todas estas se pueden deducir las otras dos condiciones para que sea sesquilineal.

En particular, si φ es hermítica, se tiene que $\varphi(x, x) = \overline{\varphi(x, x)}$, con lo que $\varphi(x, x) \in \mathbb{R}$, y la siguiente definición tiene sentido:

Definición 2.18 Una forma hermítica $\varphi : V \times V \rightarrow \mathbb{C}$ se dice definida positiva si $\varphi(x, x) > 0 \forall x \neq 0$.

Con todo esto, ahora ya podemos pasar a ver como es la versión de los productos escalares en espacios vectoriales complejos. Consideramos la aplicación dada de la siguiente forma:

$$\begin{aligned} \Phi : \mathbb{C}^2 \times \mathbb{C}^2 &\rightarrow \mathbb{C} \\ ((z_1, z_2), (w_1, w_2)) &\mapsto z_1 \bar{w}_1 + z_2 \bar{w}_2 \end{aligned}$$

con $z_i, w_i \in \mathbb{C}$.

Proposición 2.19 La aplicación Φ define una forma hermítica definida positiva en $\mathbb{C}^2$.

Demostración

Sean $z = (z_1, z_2), w = (w_1, w_2), x = (x_1, x_2) \in \mathbb{C}^2$ y $a, b \in \mathbb{C}$:

- Φ verifica la hermiticidad:

$$\overline{\Phi(w, z)} = \overline{(w_1 \bar{z}_1 + w_2 \bar{z}_2)} = z_1 \bar{w}_1 + z_2 \bar{w}_2 = \Phi(z, w), \text{ con lo que efectivamente } \Phi(z, w) = \overline{\Phi(w, z)}.$$

- Φ es lineal en la primera coordenada:

$$\Phi(az + bw, x) = (az_1 + bw_1) \bar{x}_1 + (az_2 + bw_2) \bar{x}_2 = az_1 \bar{x}_1 + z_2 \bar{x}_2 + bw_1 \bar{x}_1 + bw_2 \bar{x}_2 = a\Phi(z, x) + b\Phi(w, x), \text{ luego se verifica que } \Phi(az + bw, x) = a\Phi(z, x) + b\Phi(w, x).$$

- Φ es definida positiva:

Sea $z_1 = c + di$ y $z_2 = e + fi$. Entonces $\Phi(z, z) = z_1 \overline{z_1} + z_2 \overline{z_2} = c^2 + d^2 + e^2 + f^2 > 0$, con lo que $\Phi(z, z) \geq 0$. Además, de aquí deducimos que si $\Phi(z, z) = 0$ entonces $c = d = e = f = 0$ y por ello $\Phi(z, z) = 0 \iff z = (0, 0)$.

Con lo que efectivamente Φ es una forma hermítica en $\mathbb{C}^2$. $\square$

Definición 2.20 Si φ es una forma hermítica definida positiva, entonces φ define una norma con $\sqrt{\varphi(x, x)}$ a la que denominamos norma inducida por φ .

A partir de ahora, cuando hablemos del producto escalar de $\mathbb{C}^2$, nos estaremos refiriendo a Φ ya descrita previamente, es decir, para $z, w \in \mathbb{C}^2$ con $z = (z_1, z_2)$ y $w = (w_1, w_2)$ tendremos que $z \cdot w = \Phi(z, w)$ y cuando hablemos de la norma en $\mathbb{C}^2$, estaremos hablando de la norma inducida por la aplicación Φ , es decir, tendremos que para $z = (z_1, z_2) \in \mathbb{C}^2$, su norma será $\|z\| = \sqrt{z_1 \overline{z_1} + z_2 \overline{z_2}}$.

Además, ahora ya podemos hablar de isometrías en $\mathbb{C}^2$, que se definen igual que en $\mathbb{R}^3$:

Definición 2.21 $f : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ lineal es una isometría si preserva el producto escalar, es decir, para cualesquiera $z, w \in \mathbb{C}^2$, se verifica que $z \cdot w = f(z) \cdot f(w)$.

Ahora que ya hemos definido un producto escalar, y con el la norma y las isometrías en $\mathbb{C}^2$, pasamos a ver como es la estructura de $U(2)$ y su relación con las isometrías.

Proposición 2.22 $U(2) = \{A \in \mathcal{M}_{2 \times 2}(\mathbb{C}) : A^* = A^{-1}\}$ donde A^* denota la traspuesta conjugada de la matriz A es un grupo respecto del producto de matrices.

Demostración

- El producto de matrices en $U(2)$ es cerrado, es decir, para cualesquiera $A, B \in U(2)$ se tiene que $AB \in U(2)$. Para verlo, sean $A, B \in U(2)$:

Consideramos el producto de AB por su traspuesta conjugada:

$(AB)^* AB = B^* A^* AB = B^* IB = B^* B = I$ ya que $A, B \in U(2)$, con lo que resulta que $AB \in U(2)$ que es lo que queríamos ver.

- $U(2)$ tiene al neutro respecto de la multiplicación de matrices. Tomamos la matriz identidad I , que es el neutro de la multiplicación de matrices:

Se tiene que $I^* = I$, con lo que $I^* I = I$ y resulta que $I \in U(2)$.

- $U(2)$ tiene elemento inverso para toda $A \in U(2)$. Sea $A \in U(2)$:

Tenemos que ver que $A^{-1} \in U(2)$. Resulta que $A^{-1} = A^*$, y con ello $AA^* = I$ y por tanto $A^{-1} \in U(2)$.

- Tenemos que ver que la operación es asociativa, es decir, para cualesquiera $A, B, C \in U(2)$, se verifica que $(AB)C = A(BC)$.

Como el producto de matrices cumple esta propiedad en general, la cumple también en $U(2)$.

Con todo esto concluimos que $U(2)$ es grupo respecto del producto de matrices. $\square$

Y por otro lado, vemos también que las isometrías de $\mathbb{C}^2$ también son un grupo:

Proposición 2.23 *El conjunto $Isom(\mathbb{C}^2) = \{f : \mathbb{C}^2 \rightarrow \mathbb{C}^2 : f \text{ es isometría lineal}\}$ es un grupo respecto de la composición de aplicaciones.*

Demostración

Para esta demostración, damos por sabido que el conjunto de isomorfismos lineales es un grupo, y comprobamos que la composición de dos isometrías, la identidad, y el inverso de una isometría también lo son, esto es, que también preservan el producto escalar. No es necesario comprobar la propiedad asociativa ya que esta se cumple por cumplirse para los isomorfismos lineales en general.

- Queremos ver que es cerrado para la composición de aplicaciones, es decir, que para cualesquiera $f, g \in Isom(\mathbb{C}^2)$ se cumple que $(f \circ g) \in Isom(\mathbb{C}^2)$. Sean $f, g \in Isom(\mathbb{C}^2)$:

Sean $z, w \in \mathbb{C}^2$. Consideramos las siguientes igualdades:

$$(f \circ g)(z) \cdot (f \circ g)(w) = f(g(z)) \cdot f(g(w)) = g(z) \cdot g(w) = z \cdot w \text{ luego efectivamente } (f \circ g) \text{ es isometría de } \mathbb{C}^2 \text{ y con ello } (f \circ g) \in Isom(\mathbb{C}^2).$$

- Tenemos que ver que el elemento neutro para la composición de aplicaciones pertenece a $Isom(\mathbb{C}^2)$. Tomamos la aplicación identidad, que denotamos por Id , y es el neutro para la composición de aplicaciones:

Como $Id(z) = z, \forall z \in \mathbb{C}^2$, entonces se tiene que $Id(z) \cdot Id(w) = z \cdot w$ para cualesquiera $z, w \in \mathbb{C}^2$, con lo que $Id \in Isom(\mathbb{C}^2)$

- Tenemos que ver que $Isom(\mathbb{C}^2)$ tiene elemento inverso para cualquier $f \in Isom(\mathbb{C}^2)$. Sea $f \in Isom(\mathbb{C}^2)$:

Queremos ver que $f^{-1}(z) \cdot f^{-1}(w) = z \cdot w$ para cualesquiera $z, w \in \mathbb{C}^2$. Sabemos que $f(z) \cdot f(w) = z \cdot w$. Podemos escribir esta misma igualdad en términos de f^{-1} ya que sabemos que $f^{-1}(f(z)) = z$ para cualquier z . Obtenemos así que $f^{-1}(z) \cdot f^{-1}(w) = z \cdot w$ y por ello $f^{-1} \in Isom(\mathbb{C}^2)$.

Con esto hemos visto que $Isom(\mathbb{C}^2)$ es un grupo respecto de la composición de aplicaciones.

Y ahora que hemos visto que $Isom(\mathbb{C}^2)$ y $U(2)$ son grupos, vemos como están relacionados:

Teorema 2.24 *La aplicación*

$$h : U(2) \rightarrow Isom(\mathbb{C}^2)$$

$$A \mapsto f \in Isom(\mathbb{C}^2) \text{ tal que } M(f) = A$$

está bien definida y es un isomorfismo de grupos, donde $M(f)$ es la matriz de f respecto de la base canónica de $\mathbb{R}^3$.

Demostración

Primero vemos que está bien definida, es decir, vemos que $h(A) = f$ es única y que existe dicha f que verifica $h(A) = f$. Es claro que esa f es única. Tenemos que ver que existe esa $f \in \text{Isom}(\mathbb{C}^2)$. Sea $A \in U(2)$ y sean $z, w \in \mathbb{C}^2$. Entonces, podemos expresar matricialmente el producto escalar complejo como $(Az)^*Aw = z^*A^*Aw = z^*w = z \cdot w$ y tenemos la f dada por $A = M(f)$.

Veamos ahora que h es inyectiva por reducción al absurdo. Supongamos que existe algún par de matrices $A, B \in U(2)$ distintas tales que $h(A) = h(B) = f$. Entonces resultaría que $M(f) = A = B$, pero como hemos tomado matrices distintas hemos llegado a un absurdo, con lo que h es inyectiva.

Para ver la sobrectividad, tenemos que ver que para todo elemento de $\text{Isom}(\mathbb{C}^2)$ hay un elemento de $U(2)$ que es su contraimagen. Sea $f \in \text{Isom}(\mathbb{C}^2)$ con $A = M(f)$. Como f es isometría, tenemos que $f(z) \cdot f(w) = z \cdot w$ para cualesquiera $z, w \in \mathbb{C}^2$. Expresamos matricialmente el producto escalar complejo de la siguiente forma:

$$f(z) \cdot f(w) = (Az)^*Aw = z^*A^*Aw = z \cdot w$$

Veamos más detalladamente la última igualdad. Sea $\vec{v}$ un vector cuya entrada i -ésima es 1, y el resto son 0, y $\vec{w}$ un vector un vector cuya entrada j -ésima es 1, y el resto son 0. Entonces, el elemento b_{ij} de la matriz A^*A será:

$$b_{ij} = \vec{v}^\top (A^\top A) \vec{w} = (0, \dots, 1, \dots, 0)(A^*A) \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \vec{v}^\top \vec{w} = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Con lo que llegamos a que $A^*A = I$ y por ello $A \in U(2)$ será la contraimagen de f .

Para concluir que g es un isomorfismo, queda ver que es un homomorfismo, o lo que es lo mismo, que g respeta las operaciones al pasar de un conjunto a otro. Tenemos que comprobar que dadas $A, B \in U(2)$, se verifica que $h(AB) = h(A) \circ h(B)$. Sean entonces $A, B \in U(2)$ con $A = M(f)$ y $B = M(g)$. Tenemos por un lado que $h(AB) = f \circ g$. Por otro, tenemos que la expresión matricial de $(f \circ g)(z)$ para $z \in \mathbb{C}^2$ es $M(f)M(g)z = ABz$, con lo que g respeta las operaciones internas al pasar de un conjunto a otro y por ello es un homomorfismo.

Como hemos probado que es un homomorfismo biyectivo, concluimos que es un isomorfismo. $\square$

Como consecuencia de lo que hemos visto, como las matrices de $U(2)$ representan isometrías, estas preservan el producto escalar, con lo que sus columnas formarán una base ortonormal de $\mathbb{C}^2$ ya que serán las imágenes de los vectores de la base canónica de $\mathbb{C}^2$.

Con esto, el grupo unitario especial $SU(2)$ está definido de la siguiente forma:

$$SU(2) = \{A \in \mathcal{M}_{2 \times 2} : AA^* = I, \det(A) = 1\}$$

Donde I es la matriz identidad y A^* representa la traspuesta conjugada de la matriz.

Proposición 2.25 $SU(2)$ es subgrupo de $U(2)$.

Demostración

- Hay que ver que $SU(2) \subseteq U(2)$

Se ve trivialmente por definición

- Tenemos que ver que el producto de matrices es cerrado en $SU(2)$, es decir, que para cualesquiera $A, B \in SU(2)$, se verifica que $AB \in SU(2)$. Para ello, sean $A, B \in SU(2)$.

Consideramos el determinante de AB : $\det(AB) = \det(A)\det(B) = 1$ ya que $A, B \in SU(2)$, con lo que $AB \in SU(2)$.

- Tenemos que ver que el elemento neutro para la multiplicación de matrices pertenece a $SU(2)$. Tomamos la matriz identidad I , que es el neutro para la multiplicación de matrices:

Como $\det(I) = 1$, entonces $I \in SU(2)$ y el elemento neutro pertenece a $SU(2)$.

- Tiene elemento inverso para la multiplicación de matrices. Sea $A \in SU(2)$:

Consideramos el determinante de la inversa de A : $\det(A^{-1}) = \frac{1}{\det(A)} = 1$ ya que $\det(A) = 1$ porque $A \in SU(2)$, luego $A^{-1} \in SU(2)$.

Con lo que concluimos que $SU(2)$ es subgrupo de $U(2)$.

Siguiendo el mismo razonamiento que hicimos son $SO(3)$ e $Isom(\mathbb{R}^3)$, podemos esperar que a $SU(2)$ le corresponda un subgrupo de $Isom(\mathbb{C}^2)$. Veámoslo a continuación:

Proposición 2.26 *El conjunto*

$$Isom^+(\mathbb{C}^2) = \{f : \mathbb{C}^2 \rightarrow \mathbb{C}^2 : f \text{ es isometría lineal y preserva la orientación}\}$$

es subgrupo de $Isom(\mathbb{C}^2)$.

Demostración

- $Isom^+(\mathbb{C}^2) \subseteq Isom(\mathbb{C}^2)$

Trivial por definición.

- La composición de aplicaciones es cerrada, es decir, tenemos que ver que para cualesquiera $f, g \in Isom^+(\mathbb{C}^2)$ se tiene que $(f \circ g) \in Isom^+(\mathbb{C}^2)$. Sean $f, g \in Isom^+(\mathbb{C}^2)$:

Sean $A = M(f)$ y $B = M(g)$. Considero la composición de aplicaciones $f \circ g$. Matricialmente, esta se expresaría como $(f \circ g)(v) = f(g(v)) = f(Bv) = ABv$ para $v \in \mathbb{C}^2$. Entonces, tenemos que la matriz de $f \circ g$ es AB . Pero entonces $\det(AB) = \det(A)\det(B) > 0$ por ser producto de dos números mayores que ya que $f, g \in Isom^+(\mathbb{C}^2)$. Con lo que $f \circ g$ preserva la orientación por ello $f \circ g \in Isom^+(\mathbb{C}^2)$.

- $Isom^+(\mathbb{C}^2)$ tiene al elemento neutro para la composición de aplicaciones. Tomamos la aplicación identidad Id , que es el neutro para la composición:

$$M(Id) = I \text{ y } \det(I) = 1 > 0, \text{ con lo que } Id \in Isom^+(\mathbb{C}^2).$$

- Queda ver que para cualquier $f \in Isom^+(\mathbb{C}^2)$, su elemento inverso correspondiente pertenece a $Isom^+(\mathbb{C}^2)$. Sea $f \in Isom^+(\mathbb{C}^2)$:

Sea $A = M(f)$. Entonces $M(f^{-1}) = A$. Entonces tenemos que $\det(A^{-1}) = \frac{1}{\det(A)} > 0$ ya que $\det(A) > 0$ por ser $f \in Isom^+(\mathbb{C}^2)$. Con lo que $f^{-1} \in Isom^+(\mathbb{C}^2)$.

Y con ello concluimos que $Isom^+(\mathbb{C}^2)$ es subgrupo de $Isom(\mathbb{C}^2)$. $\square$

Teorema 2.27 *La aplicación*

$$h|_{SU(2)} : SU(2) \rightarrow Isom^+(\mathbb{C}^2)$$

$$A \mapsto f \in Isom^+(\mathbb{C}^2) \text{ tal que } M(f) = A$$

donde h es la aplicación definida en el teorema 2.24 es un isomorfismo de grupos.

Vemos que es un homomorfismo y que es inyectivo de la misma forma que en el teorema 2.24.

Para ver que es sobreyectiva, comenzamos el razonamiento de la misma forma, tomando $f \in Isom^+(\mathbb{C}^2)$ con $M(f) = A$. Y expresando la igualdad $f(z) \cdot f(w) = v \cdot w$ que f verifica de forma matricial:

$$f(z) \cdot f(w) = (Az)^*Aw = z^*A^*Aw = z \cdot w$$

Con el mismo razonamiento que seguimos previamente, llegamos a que $A^*A = I$. Queda ver que $\det(A) = 1$. Tomando el determinante a ambos lados de la igualdad $A^*A = I$, llegamos a que $1 = \det(I) = \det(A^*A) = \det(A^*)\det(A)$. Como $\det(A^*) = \overline{\det(A)}$ y sabemos que $\mathbb{R} \ni \det(A) > 0$, llegamos a que $\det(A^*) = \det(A)$. Con esto, tenemos que $\det(A)^2 = \pm 1$, pero como sabemos que $\det(A) > 0$ concluimos que $\det(A) = 1$, con lo que $A \in SU(2)$ será la contraimagen de f y la aplicación será sobreyectiva.

Con todo esto, concluimos que $h|_{SU(2)}$ es un isomorfismo de grupos. $\square$

Como consecuencia de todo lo que hemos visto, tenemos que $U(2)$ e $Isom(\mathbb{C}^2)$ siguen un esquema similar al que vimos para $O(3)$ e $Isom(\mathbb{R}^3)$:

$$\begin{array}{ccc} U(2) & \xleftarrow{h} & Isom(\mathbb{C}^2) \\ \cup & & \cup \\ SU(2) & \xleftarrow{h} & Isom^+(\mathbb{C}^2) \end{array}$$

Definición 2.28 *Una sucesión exacta de grupos es una secuencia*

$$G_0 \xrightarrow{f_1} G_1 \xrightarrow{f_2} G_2 \xrightarrow{f_3} \dots \xrightarrow{f_n} G_n$$

donde G_i son grupos y f_i son homomorfismos tales que $\text{im}(f_i) = \ker(f_{i+1}) \forall 1 \leq i < n$ con $n \in \mathbb{N}$.

Si la sucesión tiene la forma

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

se denomina sucesión exacta de grupos corta.

Observación 2.29 Realmente podemos dar una definición de sucesión exacta para otros tipos de estructuras que no sean grupos. Por ejemplo, si los conjuntos fuesen espacios vectoriales y las aplicaciones fuesen lineales, tendríamos secuencias exactas de espacios vectoriales.

Vamos a establecer una sucesión exacta de grupos corta que tiene a $SU(2)$ y $U(2)$ como conjuntos, pero para hacerlo vamos a ver un resultado previo que necesitaremos. Es por ello que estudiamos el determinante de las matrices de $U(2)$.

Sea $A \in U(2)$, tal que $A = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$ con $x, y, z, t \in \mathbb{C}$. Entonces, su determinante será $\det(A) = xt - yz \in \mathbb{C}$.

Proposición 2.30 Si $A \in U(2)$, entonces $|\det(A)| = 1$, y por ello $\det(A) \in S^1$.

Demostración

Sea $A \in U(2)$. Sabemos por definición que $A^*A = I$. Considerando el determinante a ambos lados de la igualdad, obtenemos que $\det(I) = \det(A^*A)$, que a su vez podemos expresar como $\det(I) = \det(A^*)\det(A)$. Como $\det(A^*) = \overline{\det(A)}$, tenemos que $1 = \det(I) = \overline{\det(A)}\det(A) = |\det(A)|^2$. Como la norma es siempre mayor que cero, podemos descartar la solución negativa del cuadrado y concluimos que $|\det(A)| = 1$, que es el resultado que queríamos. $\square$

Por otra parte, se puede ver que $S^1 = U(1)$, ya que $U(1) = \{A \in \mathcal{M}_{1 \times 1}(\mathbb{C}) : A^* = A^{-1}\}$. Las matrices complejas de dimensión 1×1 claramente son elementos de $\mathbb{C}$, y al estar tratando con estos elementos la el trasponer la matriz no supone ningún cambio, con lo que podemos reescribir el conjunto como $U(1) = \{a + bi = z \in \mathbb{C} : z\bar{z} = |z|^2 = 1\}$. Usando el mismo razonamiento que en la anterior demostración, $|z|^2 = 1$ nos da que $|z| = 1$ y tendríamos $U(1) = \{a + bi = z \in \mathbb{C} : |z| = a^2 + b^2 = 1\}$ que es la definición de S^1 .

Ahora que hemos visto todo esto, podemos ver el siguiente resultado:

Proposición 2.31 La sucesión

$$0 \longrightarrow SU(2) \xrightarrow{i} U(2) \xrightarrow{\det} U(1) \longrightarrow 0$$

donde i denota la inyección es una sucesión exacta corta de grupos.

Demostración

Para ver esto tenemos que demostrar tres cosas: que cada conjunto es un grupo, que las aplicaciones son homomorfismos, y que la imagen de una aplicación coincide con el núcleo de la siguiente. Veámoslo:

El conjunto 0 es el grupo trivial, y ya hemos probado en las proposiciones 2.22 y 2.25 que $U(2)$ y $SU(2)$ son grupos, respectivamente. Además, es sabido que la circunferencia unidad es un grupo respecto de la multiplicación vista como el conjunto de números complejos de módulo uno.

Por otra parte, es trivial que la primera y la última aplicación son homomorfismos. La inyección i también lo es, puesto que es la identidad restringida a $SU(2)$. Con lo que nos queda el determinante. Sean $A, B \in U(2)$, tenemos que ver que $\det(AB) = \det(A)\det(B)$, pero esto es una propiedad que cumple el determinante con lo que también es un homomorfismo.

Por último, tenemos que ver que la imagen de cada aplicación es igual al núcleo de la siguiente:

- En el primer caso, la imagen de la aplicación $0 \rightarrow A$ es el neutro, ya que por ser homomorfismo el neutro, que es el único elemento del primero conjunto, tiene como imagen el neutro del segundo, que en nuestro caso es la matriz identidad I . Tenemos que considerar ahora el núcleo de la inyección i . Pero como es la inyección, es claro ver que $\ker(i) = I$ y se verifica la propiedad en este primer caso.
- Ahora tenemos que ver que $\text{im}(i) = \ker(\det)$. Es claro que la imagen de i será $SU(2) \subset U(2)$. Por otra parte, para ver el núcleo de $\det$ tendríamos que ver el conjunto $\{A \in U(2) : \det(A) = 1\}$ o lo que es lo mismo, $\{A \in M_{2 \times 2} : AA^* = I, \det(A) = 1\}$ que es la definición de $SU(2)$, con lo que tenemos que $\text{im}(i) = \ker(\det) = SU(2)$ que es lo que queríamos probar.
- Por último, queremos demostrar que la imagen de $\det$ es igual al núcleo de la última aplicación. El núcleo de la aplicación cero es todo $U(1)$. Sabemos por la proposición 2.30 que la imagen está contenida en $U(1)$, con lo que solo quedaría ver que para cada $z \in S^1$ hay un elemento en $U(2)$ que lo tiene como imagen, equivalente a probar que es sobreyectiva. Pero esto es sencillo, ya que podemos tomar una matriz diagonal que tenga z en uno de sus elementos distintos de cero, y que el otro sea 1, es decir, construimos una matriz A de la siguiente forma:

$$A = \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix}, \text{ que verifica } A^*A = 1$$

y con esto tendríamos la matriz tal que $\det(A) = z$ para cualquier $z \in S^1$. Por esto, $\text{im}(\det) = U(1)$ y se cumple la última propiedad que queríamos demostrar

Como hemos visto que los conjuntos son grupos, y que las aplicaciones son homomorfismos que satisfacen que la imagen de cada una de ellas es el núcleo de la siguiente, concluimos que efectivamente tenemos una sucesión exacta corta de grupos. $\square$

Ahora que ya conocemos $SU(2)$, vamos a ver la relación entre este y $SO(3)$. Para ello, vamos a ver primero la relación entre cada uno de estos con la esfera S^3 .

2.4.1 $SU(2)$ y S^3

Ahora vamos a ver que $SU(2)$ es isomorfo a $S^3 = \{(a, b, c, d) \in \mathbb{R}^4 : a^2 + b^2 + c^2 + d^2 = 1\}$. Podemos identificar los puntos de esta esfera con elementos de $\mathbb{C}^2$, de forma que tendríamos $x, y \in \mathbb{C}^2$, $x = a + bi$, $y = c + di$ tales que $|x|^2 + |y|^2 = a^2 + b^2 + c^2 + d^2 = 1$. Para ver que son isomorfos consideramos la siguiente aplicación:

$$\begin{aligned} \varphi : S^3 &\rightarrow SU(2) \\ (x, y) &\mapsto \begin{pmatrix} x & y \\ -\bar{y} & \bar{x} \end{pmatrix} \text{ con } x, y \in \mathbb{C} \end{aligned}$$

Teorema 2.32 φ es un homeomorfismo.

Demostración

Tenemos que demostrar que φ es un homeomorfismo, es decir, tenemos que ver que φ es biyectiva, continua, y que su inversa es continua.

Primero veamos que es inyectiva. Tenemos que ver que dos elementos distintos tienen siempre una imagen distinta. Razonamos por reducción al absurdo. Para ello, sean $(x_1, y_1) \in S^3$ con $x_1, y_1 \in \mathbb{C}$ y $(x_2, y_2) \in S^3$ con $x_2, y_2 \in \mathbb{C}$ tales que $(x_1, y_1) \neq (x_2, y_2)$. Suponemos que $\varphi(x_1, y_1) = \varphi(x_2, y_2)$. Esto implicaría la siguiente igualdad entre matrices:

$$\begin{pmatrix} x_1 & y_1 \\ -\bar{y}_1 & \bar{x}_1 \end{pmatrix} = \begin{pmatrix} x_2 & y_2 \\ -\bar{y}_2 & \bar{x}_2 \end{pmatrix}$$

Para que se verifique la igualdad, las matrices tendrían que ser iguales coordenada a coordenada. Entonces, tendríamos que $x_1 = x_2$ y que $y_1 = y_2$, con lo que resultaría que $(x_1, y_1) = (x_2, y_2)$. Pero esto es absurdo, puesto que hemos tomado (x_1, y_1) y (x_2, y_2) distintos, con lo que concluimos que φ es inyectiva.

Veamos ahora que es sobreyectiva. Para demostrar esto, tenemos que ver que para cualquier matriz $A \in SU(2)$ hay un elemento en S^3 que es la contraimagen de A . Para ello, estudiamos las condiciones que satisfacen todas las matrices que pertenecen a $SU(2)$.

Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SU(2)$, con $a, b, c, d \in \mathbb{C}$. Sabemos que A satisface $AA^* = 1$, es decir:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{d} & \bar{b} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \text{ De aquí deducimos que } \begin{cases} a\bar{a} + b\bar{b} = |a|^2 + |b|^2 = 1 \\ a\bar{c} + b\bar{d} = 0 \\ c\bar{a} + d\bar{b} = 0 \\ c\bar{c} + d\bar{d} = |c|^2 + |d|^2 = 1 \end{cases}$$

Por otra parte, también se cumplirá que $A^*A = 1$:

$$\begin{pmatrix} \bar{a} & \bar{c} \\ \bar{d} & \bar{d} \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \text{ De aqu\u00ed deducimos que } \begin{cases} a\bar{a} + c\bar{c} = |a|^2 + |c|^2 = 1 \\ a\bar{b} + c\bar{d} = 0 \\ b\bar{a} + d\bar{c} = 0 \\ b\bar{b} + d\bar{d} = |b|^2 + |d|^2 = 1 \end{cases}$$

De $a\bar{a} + b\bar{b} = 1$ y $a\bar{a} + c\bar{c} = 1$ deducimos que $b\bar{b} = c\bar{c}$ lo cual implica que $c = \frac{b\bar{b}}{\bar{c}}$. De forma similar, usando $a\bar{a} + b\bar{b} = 1$ y $b\bar{b} + d\bar{d} = 1$ llegamos a $d\bar{d} = \bar{a}a$ y con ello $d = \frac{\bar{a}a}{\bar{d}}$. Habiendo visto estas dos cosas, podemos reescribir nuestra matriz como

$$A = \begin{pmatrix} a & b \\ \frac{b\bar{b}}{\bar{c}} & \frac{a\bar{a}}{\bar{d}} \end{pmatrix} \in SU(2)$$

Ahora, como $a\bar{c} + b\bar{d} = 0$ con lo que $b\bar{d} = -a\bar{c}$ y de aqu\u00ed $\frac{-a}{\bar{d}} = \frac{b}{\bar{c}}$, y de nuevo, reescribimos nuestra matriz:

$$A = \begin{pmatrix} a & b \\ -\frac{a}{\bar{d}}\bar{b} & \frac{a}{\bar{d}}\bar{a} \end{pmatrix} \in SU(2)$$

Veamos que podemos averiguar sobre $\frac{a}{\bar{d}}$. Tenemos de los sistemas de ecuaciones previos que $c\bar{c} + d\bar{d} = |c|^2 + |d|^2 = 1$ y $a\bar{a} + c\bar{c} = |a|^2 + |c|^2 = 1$, de donde llegamos a que $|a|^2 = |d|^2$. Como $|z|^2 = |\bar{z}|^2$ para cualquier $z \in \mathbb{C}$, entonces tenemos que $|a|^2 = |\bar{d}|^2$ y con ello $\frac{|a|^2}{|\bar{d}|^2} = 1$. Como las normas siempre son mayores que 0, podemos eliminar el -1 al despejar el cuadrado y quedarnos con $\frac{|a|}{|\bar{d}|} = 1$. Esta igualdad nos dice que a y $\bar{d}$ tienen la misma norma, y por ello se verifica que $\left| \frac{a}{\bar{d}} \right| = 1$. Con esto hemos llegado a que $\frac{a}{\bar{d}}$ es un n\u00famero complejo con norma igual a 1. Entonces, para alg\u00fan \u00e1ngulo α , podremos expresarlo como $\frac{a}{\bar{d}} = \cos(\alpha) + i\sin(\alpha)$ que a su vez podemos escribir como $e^{i\alpha}$. Con todo esto, volvemos a reescribir nuestra matriz:

$$A = \begin{pmatrix} a & b \\ -e^{i\alpha}\bar{b} & e^{i\alpha}\bar{a} \end{pmatrix} \in SU(2)$$

Finalmente, como las matrices de $SU(2)$ cumplen que su determinante vale 1, evaluamos el determinante de la matriz a la que hemos llegado, que ser\u00e1 $\det(A) = e^{i\alpha}a\bar{a} + e^{i\alpha}b\bar{b} = e^{i\alpha}(|a|^2 + |b|^2)$. De las igualdades que deducimos de las matrices sabemos que $|a|^2 + |b|^2 = 1$ y llegamos finalmente a que $\det(A) = e^{i\alpha} = 1$, con lo que concluimos que nuestra matriz tiene la nueva forma

$$A = \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} \in SU(2)$$

y sabemos que verifica que $|a|^2 + |b|^2 = 1$ luego el par (a, b) pertenece a S^3 y será la contraimagen de la matriz A .

Por último, para poder hablar de continuidad, tenemos que hablar previamente de topología. Como $S^3 \subset \mathbb{R}^4$ luego no hay problema con la topología usual de $\mathbb{R}^4$. Por otro lado, las matrices de $SU(2)$ quedan determinadas por [whoswho](#). Como las funciones de cada componente de la matriz de la imagen a la que hemos llegado no presentan discontinuidades la aplicación será continua. Definimos la inversa como $\varphi^{-1}\left(\begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix}\right) = (x, y)$ y claramente es continua.

Con todo esto concluimos que φ es un homeomorfismo y con ello $S^3 \cong SU(2)$. $\square$

2.4.2 $SO(3)$ y S^3

Ahora vamos a estudiar la relación entre $SO(3)$ y S^3 , para luego relacionar $SO(3)$ y $SU(2)$ a través de S^3 . Vamos a ver que $SO(3)$ es homeomorfo al espacio proyectivo $\mathbb{P}^3$ de dos formas, y estableceremos de esta forma la relación que buscamos ya que $\mathbb{P}^3 \cong S^3/\mathbb{Z}_2$:

Para ello construyamos primero el espacio proyectivo $\mathbb{P}^3$. Seguimos el mismo proceso que con $\mathbb{P}^2$. Para este último, considerábamos todos los vectores del espacio tridimensional $\mathbb{R}^3$ y dada la relación $x \sim y$ si y solo si $y = \lambda x$, con $x, y \in \mathbb{R}^3$ y $\lambda \in \mathbb{R}$, construíamos el cociente $\mathbb{R}^3/\sim$. Podemos ver esto como que cada elemento de $\mathbb{P}^2$ es una recta vectorial de $\mathbb{R}^3$, o como el resultado de tomar la esfera S^2 e identificar cada punto con su antipodal.

Ahora, seguimos el mismo proceso para $\mathbb{R}^4$. Establecemos la relación $x \sim y$ si y solo si $y = \lambda x$, con $x, y \in \mathbb{R}^4$ y $\lambda \in \mathbb{R}$. Al igual que en $\mathbb{R}^3$, podemos ver cada elemento como una recta vectorial o como un elemento de la esfera identificando puntos antipodales. Centrémonos en esto último. Identificar puntos antipodales es relacionar $x \sim y$ si y solo si $y \in \{x, -x\}$, con lo que tendríamos que $\mathbb{P}^3 \cong S^3/\sim$, o lo que es lo mismo $\mathbb{P}^3 \cong S^3/\mathbb{Z}_2$.

Para la primera demostración, sabemos que necesitamos una forma de identificar los elementos de $\mathbb{P}^3$ con los de $SO(3)$. Como ya hemos visto, cada elemento de $SO(3)$ se puede identificar de manera unívoca dado un eje de rotación y un ángulo.

Consideramos la bola $\overline{B}_\pi^3(0) = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq \pi^2\}$ y la siguiente aplicación:

$$f : \overline{B}_\pi^3(0) \rightarrow SO(3)$$

$$p \mapsto R_{(\vec{0}p, ||\vec{0}p||)}$$

donde la matriz $R_{(\vec{v}, \theta)}$ denota la matriz de rotación cuyo eje está generado por $\vec{v}$ y cuyo ángulo es θ .

Pero esta aplicación daría lugar a un problema, que un punto de $\overline{B}_\pi^3(0)$ y su antipodal darían lugar a la misma rotación. Es por ello que consideramos la relación $\sim$, donde

$p \sim q$ si y solo si $q \in \{p, -p\}$ que identificaría los puntos que dan lugar a la misma matriz de rotación de $SO(3)$.

Con ello consideramos la misma aplicación f ya descrita, pero que relaciona puntos de $\overline{B}_\pi^3(0)/\sim$ con matrices de $SO(3)$, en vez de puntos de $\overline{B}_\pi^3(0)$:

$$f^* : \overline{B}_\pi^3(0)/\sim \cong \mathbb{P}^3 \rightarrow SO(3)$$

$$p \mapsto R_{(\vec{0}p, \|\vec{0}p\|)}$$

Proposición 2.33 *La aplicación f^* es un homeomorfismo.*

Demostración

Para poder considerar un homeomorfismo tenemos que hablar previamente de las topologías de ambos espacios. Primero, en cuanto a $\mathbb{P}^3$, este vive en $\mathbb{R}^4$, luego lo vemos como subconjunto de este y podemos usar su topología usual. Por otra parte, $SO(3)$ es una matriz cuadrada de dimensión 3, por lo que podemos verlo un elemento de $SO(3)$ como uno de $\mathbb{R}^9$ el cual obtendríamos poniendo sus componentes en una 9-tupla. Por ello, podemos usar para $SO(3)$ la topología usual de $\mathbb{R}^9$.

Esta aplicación es continua por serlo las funciones de las componentes de la matriz $R_{(\vec{0}p, \|\vec{0}p\|)}$. Además, f^* es inyectiva ya que nos hemos deshecho de los puntos de $\overline{B}_\pi^3(0)$ que determinaban rotaciones repetidas, y es sobreyectiva porque podemos determinar cualquier rotación de $SO(3)$ mediante un generador del eje y un ángulo dados por un elemento $p \in \overline{B}_\pi^3(0)/\sim$. Obtenemos estos con $\vec{0}p$ y $\|\vec{0}p\|$ respectivamente.

La inversa de esta aplicación relacionaría cada matriz de $SO(3)$ de eje t y ángulo $\theta \in [0, \pi)$ con un punto de $\overline{B}_\pi^3(0)/\sim$ que obtendríamos con $\theta t/\|t\|$. La matriz identidad estaría relacionada con el centro de $\overline{B}_\pi^3(0)/\sim$. Con esto, la aplicación inversa sería continua ya que $\|t\|$ nunca se anulará.

Con todo esto la aplicación f es un homeomorfismo y concluimos que $\mathbb{P}^3 \cong \overline{B}_\pi^3(0)/\sim \cong SO(3)$. $\square$

En cuanto a la segunda forma de ver este resultado, vamos a trabajar con cuaterniones. Los cuaterniones son tratados de forma más detallada en una sección posterior de este trabajo.

Consideramos la base $\{1, i, j, k\}$ de $\mathbb{H}$ y el producto escalar φ definido en la sección 3.2.3 de este trabajo. Podemos identificar la esfera S^3 con los cuaterniones de norma 1. Llamemos $\mathbb{H}_p$ al subespacio vectorial de $\mathbb{H}$ generado por $\{i, j, k\}$ y sea $x \in S^3$. Definimos la siguiente aplicación:

$$F_x : \mathbb{H}_p \rightarrow \mathbb{H}_p$$

$$y \mapsto xyx^{-1}$$

Proposición 2.34 *La aplicación F_x está bien definida, es lineal y es una isometría.*

Demostración

1. Bien definida:

Sean $y = bi + cj + dk \in \mathbb{H}_p$ y $x = e1 + fi + gj + hk \in S^3$, esto es $e^2 + f^2 + g^2 + h^2 = 1$. Identificando $\mathbb{H}$ con $\mathbb{R}^4$, tenemos lo siguiente:

$$\begin{aligned}
& xyx^{-1} = \\
& = \\
& [(-fb-gc-hd)\mathbf{1}+(eb+gd-hc)\mathbf{i}+(ec-fd+hb)\mathbf{j}+(ed+fc-gb)\mathbf{k}](e\mathbf{1}-f\mathbf{i}-g\mathbf{j}-h\mathbf{k}) = \\
& = [(-fbe-gce-hde+ebf+gbf-hcf+ecg-fdg+hbg+edh+fch-gbh)\mathbf{1}+ \\
& \quad (f^2b+gcf+hdfe^b+gde-hce-ech+fdh-h^2b+edg+fcg-g^2b)\mathbf{i}+ \\
& \quad (fbg+g^c+hdg+ebh+gdh-h^2c+e^2c-fde+hbe-edf-f^2c+gbf)\mathbf{j}+ \\
& \quad (fbh+gch+h^2d-ebg-g^2d+hcg+ecf-f^2d+hbh+e^d+fce-gbe)\mathbf{k}] = \\
& = [(b(f^2+e^2-h^2-g^2)+2(gcf+hdg+gde-hce))\mathbf{i}+(c(g^2-h^2+e^2-f^2)+2(fbg+ \\
& \quad hdg+ebh-fde))\mathbf{j}+(d(h^2-g^2-f^2+e^2)+2(fbh+gch+ecf-gbe))\mathbf{k}] \in \mathbb{H}_p
\end{aligned}$$

Ya que la parte real se anula.

2. Lineal:

Sea $\lambda \in \mathbb{R}$. Entonces $F_x(\lambda y) = x\lambda yx^{-1} = \lambda xyx^{-1} = \lambda F_x(y)$

3. Conserva la el producto escalar, y por ello es isometría:

Como $\|uv\| = \|u\| \cdot \|v\|$ para $u, v \in \mathbb{H}$, entonces $\|xyx^{-1}\| = \|x\| \cdot \|y\| \cdot \|x^{-1}\|$. Usando que $\|x\| = \|x^{-1}\| = 1$ ya que $x \in S^3$, llegamos a que $\|F_x(y)\| = \|xyx^{-1}\| = \|x\| \cdot \|y\| \cdot \|x^{-1}\| = \|y\|$.

Con lo que F_x está bien definida y es una isometría lineal. $\square$

Ahora consideramos la aplicación F , la cual asocia a cada elemento de S^3 con su matriz de la aplicación F_x , la cual denotamos A_x respecto de la base $\{i, j, k\}$:

$$F : S^3 \rightarrow SO(3)$$

$$x \mapsto A_x = \begin{pmatrix} a^2 + b^2 - c^2 - d^2 & 2(bc - ad) & 2(ac + bd) \\ 2(bc + ad) & a^2 - b^2 + c^2 - d^2 & 2(cd - ab) \\ 2(bd - ac) & 2(ab + cd) & a^2 - b^2 - c^2 + d^2 \end{pmatrix} \text{ matriz de } F_x$$

donde $x = a + bi + cj + dk$. La matriz que obtenemos con F es la matriz de rotación asociada al cuaternión x .

Nos encontramos con el mismo problema de antes, que es que tenemos elementos de S^3 que definen la misma rotación. Por ello consideramos el núcleo de la aplicación F .

Proposición 2.35 *El núcleo de F es $\{1, -1\}$.*

Demostración

Sean $y = bi + cj + dk \in \mathbb{H}_p$ e $x = e\mathbf{1} + f\mathbf{i} + g\mathbf{j} + h\mathbf{k} \in S^3$. Consideramos la siguiente cadena de implicaciones:

$$x \in \ker F \iff F_x = id \iff F_x(y) = y, \forall y \in \mathbb{H}_p \iff xyx^{-1}, \forall y \in \mathbb{H}_p \iff xy = yx, \forall y \in \mathbb{H}_p$$

Entonces, $x \in \ker F \iff xy = yx \forall y \in \mathbb{H}_p$. Con esto, desarrollando la expresion $xy - yx = 0$:

$$\begin{aligned} xy - yx &= [(-fb - gc - hd)\mathbf{1} + (eb + gd - hc)\mathbf{i} + (ec - fd + hb)\mathbf{j} + (ed + fc - gb)\mathbf{k}] - \\ &\quad [(-bf - cg - dh)\mathbf{1} + (be + ch - dg)\mathbf{i} + (ce - bh + df)\mathbf{j} + (de + bg - cf)\mathbf{k}] = \\ &\quad 2[(gd - ch)\mathbf{i} + (hb - fd)\mathbf{j} + (fc - gb)\mathbf{k}] = 0 \end{aligned}$$

con lo que $f, g, h = 0$ y nos quedaría $x = e\mathbf{1} \in S^3$, es decir, $x \in \{1, -1\}$ que es lo que queríamos ver. $\square$

Corolario 2.36 $F(x) = F(y) \iff x \in \{y, -y\}$

Demostración Tenemos que $F(x) = F(y)$, y queremos llegar a que $x \in \{y, -y\}$. Consideramos las siguientes implicaciones:

$$\begin{aligned} F(x) = F(y) &\iff F_x(z) = F_y(z) \forall z \in \mathbb{H}_p \iff xzx^{-1} = yzy^{-1} \forall z \in \mathbb{H}_p \iff \\ &\iff y^{-1}xzx^{-1} = zy^{-1} \forall z \in \mathbb{H}_p \iff (y^{-1}x)z(y^{-1}x)^{-1} = z \forall z \in \mathbb{H}_p \iff \\ &\iff F_{y^{-1}x} = id \iff y^{-1}x \in \ker F \iff y^{-1}x = \pm 1 \iff x = \pm y \iff x \in \{y, -y\} \end{aligned}$$

Que es lo que queríamos probar. Como todas las implicaciones son del tipo si y solo si, podemos seguir el razonamiento para demostrar ambos sentidos de implicación, luego queda probado que $F(x) = F(y) \iff x \in \{y, -y\}$. $\square$

Proposición 2.37 $SO(3)$ es homeomorfo a $\mathbb{P}^3$.

Demostración

Sabemos que para una aplicación lineal $f : V \rightarrow W$, se tiene que $Im(f) \cong V/\ker F$. Como un cuaternión determina de forma unívoca una rotación, esta aplicación es sobreyectiva y por ello su imagen es todo $SO(3)$. Con esto, podemos aplicar la propiedad mencionada y concluimos que $SO(3) \cong S^3/\{1, -1\} \cong \mathbb{P}^3$. $\square$

2.4.3 Conclusión

Una vez visto todo esto, hemos llegado por un lado a que $SO(3) \cong \mathbb{P}^3 \cong S^3/\mathbb{Z}_2$ y por otro a que $SU(2) \cong S^3$, por lo que podemos concluir que $SO(3) \cong SU(2)/\mathbb{Z}_2$.

Es más, podemos construir un doble cubrimiento de $SU(2)$ en $SO(3)$. Veámoslo.

Ya hemos visto que las matrices de $SU(2)$ tienen la siguiente forma:

$$\begin{pmatrix} x & y \\ -\bar{y} & \bar{x} \end{pmatrix} \text{ con } |x|^2 + |y|^2 = 1 \text{ y } x, y \in \mathbb{C}$$

y en la sección de este trabajo que trata los cuaterniones vemos que también podemos identificar este tipo de matrices con los cuaterniones. Concretamente, si tenemos una matriz como la anterior, y $x = a + bi, y = c + di$, entonces estaría asociada al cuaternion $q = a + bi + cj + dk$ y tenemos la siguiente aplicación que identificaría elementos de $SU(2)$ con cuaterniones:

$$\begin{pmatrix} x & y \\ -\bar{y} & \bar{x} \end{pmatrix} \mapsto a + bi + cj + dk \in \mathbb{H}$$

con $|x|^2 + |y|^2 = 1$ y $x = a + bi, y = c + di \in \mathbb{C}$. La imagen de esta aplicación será entonces el conjunto de cuaterniones cuya norma es 1. Sea $q \in SU(2) \subset \mathbb{H}$. Consideremos ahora la aplicación $F_q : x \mapsto qxq^{-1}$.

En la sección dedicada a los cuaterniones se demuestra que este producto deja fija la parte real de v , y envía $\mathbb{H}_p$ a sí mismo. Además, como $\|u \cdot v\| = \|u\| \cdot \|v\|$ para cualesquiera cuaterniones u, v se tiene que f_q también envía $S^1 \subset \mathbb{H}$ a S^1 y fija q .

De forma similar a como hemos hecho previamente, definimos la aplicación F :

$$F : SO(3) \supset SU(2) \rightarrow SO(3)$$

$$q \mapsto A_x = \begin{pmatrix} a^2 + b^2 - c^2 - d^2 & 2(bc - ad) & 2(ac + bd) \\ 2(bc + ad) & a^2 - b^2 + c^2 - d^2 & 2(cd - ab) \\ 2(bd - ac) & 2(ab + cd) & a^2 - b^2 - c^2 + d^2 \end{pmatrix}$$

donde $q = a + bi + cj + dk$. La matriz que obtenemos con F es la matriz de rotación asociada al cuaternión q .

Teorema 2.38 *La aplicación F nos da un doble cubrimiento de $SU(2)$ en $SO(3)$.*

Demostración

Para ver que tenemos un doble cubrimiento, basta ver que q y $-q$ tienen la misma imagen. Se tiene que $(-q)x(-q)^{-1} = (-q)x(-\bar{q})/\|q\| = qx\bar{q}/\|q\| = qxq^{-1}$, con lo que efectivamente, la matriz que define su rotación será la misma y por ello tenemos un doble cubrimiento. $\square$

3 SO(3) y los cuaterniones

3.1 Introducción

A lo largo de esta sección vamos a introducir el conjunto de los cuaterniones y algunas de sus propiedades, y veremos como se relaciona con $SO(3)$. Los resultados de esta sección han sido extraídos de [6], [7] y [10].

Comencemos considerando la circunferencia unidad S^1 de $\mathbb{C}$. Si relacionamos los elementos de $\mathbb{C}$ con los de $\mathbb{R}^2$ de forma que $(x + iy) \in \mathbb{C}$ se identifica con $(x, y) \in \mathbb{R}^2$, podemos obtener rotaciones vectoriales en el plano: la rotación de ángulo θ respecto del origen se obtiene multiplicando por $e^{i\theta}$, es decir, obtendríamos la rotación de $z = (x + iy) \in \mathbb{C}$ con $ze^{i\theta}$. Por ejemplo, si tomo $z = (1 + i)/\sqrt{2}$ y quiero girarlo $\theta = \pi/2$ me quedaría:

$$ze^{i\theta} = \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i \right) e^{i\pi/2} = \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i \right) i = -\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i$$

Así, la rotación tendrá un vector fijo o invariante (el vector $(0, 0)$), y todo el resto del plano gira a su alrededor.

Si ahora consideramos las rotaciones en $\mathbb{R}^3$, en lugar de un punto invariante tenemos el eje de la rotación en su lugar, y los vectores rotan alrededor de este a través de su plano ortogonal. Una forma de expresar estas rotaciones es mediante los cuaterniones.

3.2 Formas de representar los cuaterniones

3.2.1 Representación Vectorial

Definición 3.1 *El conjunto:*

$$\mathbb{H} = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}, \quad a, b, c, d \in \mathbb{R}$$

en el se siguen las siguientes relaciones:

$$\begin{aligned} \mathbf{i}^2 &= \mathbf{j}^2 = \mathbf{k}^2 = -\mathbf{1} \\ \mathbf{ij} &= -\mathbf{ji} = \mathbf{k} \\ \mathbf{jk} &= -\mathbf{kj} = \mathbf{i} \\ \mathbf{ki} &= -\mathbf{ik} = \mathbf{j} \end{aligned}$$

se conoce como los cuaterniones o números cuaterniónicos.

El cuaternión nulo se denominará como $\mathbf{0}$, y los cuaterniones con $a = 0$ se conocen como cuaterniones puros, y se denotan con $\mathbb{H}_p$. Gracias a estas relaciones podemos considerar el producto entre cuaterniones, que dados $X = a_1\mathbf{1} + b_1\mathbf{i} + c_1\mathbf{j} + d_1\mathbf{k}$ e $Y = a_2\mathbf{1} + b_2\mathbf{i} + c_2\mathbf{j} + d_2\mathbf{k}$ será:

$$\begin{aligned} XY &= (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2)\mathbf{1} + (a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2)\mathbf{i} + \\ &\quad (a_1c_2 + c_1a_2 + b_1d_2 - d_1b_2)\mathbf{j} + (a_1d_2 + d_1a_2 + b_1c_2 - c_1b_2)\mathbf{k} \end{aligned}$$

Una vez conocemos este producto vemos que los cuaterniones forman un anillo, y un espacio vectorial de dimensión 4. Además, por ser un espacio vectorial con un producto interno, deducimos que también forman un álgebra.

También podemos ver los números reales $\mathbb{R}$ como un subanillo de $\mathbb{H}$ si consideramos la inyección $a \mapsto a\mathbf{1}$, y de forma similar, podemos ver $\mathbb{R}^3$ como subespacio de $\mathbb{H}$ (concretamente, $\mathbb{H}_p$) con la inyección $(b, c, d) \mapsto b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$.

Definición 3.2 Sea un cuaternión $X = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$. Definimos su conjugado como $\overline{X} = a\mathbf{1} - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$.

Si multiplicamos un cuaternión por su conjugado, obtenemos $X\overline{X} = (a^2 + b^2 + c^2 + d^2)\mathbf{1} = N(X)\mathbf{1}$. La cantidad $N(X)$ se conoce como norma reducida del cuaternión. Si $X \neq 0$, $N(X) \neq 0$ y podemos obtener el inverso multiplicativo de X :

Definición 3.3 Sea $X = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ un cuaternión con $N(X) \neq 0$. Su inverso se define como $X^{-1} = \overline{X}/(a^2 + b^2 + c^2 + d^2)$.

3.2.2 Representacion como matrices reales

Por otra parte, también se pueden expresar los cuaterniones con matrices cuadradas reales de dimensión 4. Para verlo, consideramos primero la representación de los números complejos como matrices cuadradas reales de dimensión 2.

Dado un número complejo $z = a + bi$, podemos obtener su representación matricial mediante la siguiente aplicación:

$$\begin{aligned}\phi : \mathbb{C} &\rightarrow \mathcal{M}_{2 \times 2}(\mathbb{R}) \\ z &\mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}\end{aligned}$$

Teorema 3.4 La aplicación ϕ es un homomorfismo de álgebras.

Demostración

Sea $z = a + bi, w = c + di$ con $a, b, c, d \in \mathbb{R}$.

- $\phi(z + w) = \begin{pmatrix} a + c & -(b + d) \\ b + d & a + c \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} + \begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \phi(z) + \phi(w)$
- $\phi(\lambda z) = \begin{pmatrix} \lambda a & -\lambda b \\ \lambda b & \lambda a \end{pmatrix} = \lambda \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \lambda \phi(z)$
- $\phi(z)\phi(w) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \begin{pmatrix} ac - bd & -(ad + bc) \\ ad + bc & ac - bd \end{pmatrix} = \phi(zw)$

□

Obsérvese que cada columna de la matriz de la imagen de ϕ puede ser vista como las coordenadas reales de $z = (a, b)$ y de $zi = (-b, a)$

Ahora seguimos el mismo proceso para un cuaternión: $X = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, $a, b, c, d \in \mathbb{R}$:

- $X\mathbf{1} = X$
- $X\mathbf{i} = a\mathbf{i} - b - c\mathbf{k} + d\mathbf{j}$
- $X\mathbf{j} = a\mathbf{j} + b\mathbf{k} - c - d\mathbf{i}$
- $X\mathbf{k} = a\mathbf{k} - b\mathbf{j} + c\mathbf{i} - d$

Y con esto llegamos a la aplicación $\varphi : \mathbb{H} \rightarrow \mathcal{M}_{4 \times 4}(\mathbb{R})$ dada por:

$$X \mapsto \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix}$$

Teorema 3.5 *La aplicación φ es un homomorfismo de álgebras.*

Demostración

Sean $X = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, $Y = e\mathbf{1} + f\mathbf{i} + g\mathbf{j} + h\mathbf{k}$ y $\lambda \in \mathbb{R}$

$$\begin{aligned} \bullet \varphi(X + Y) &= \begin{pmatrix} a+e & -(b+f) & -(c+g) & -(d+h) \\ b+f & a+e & -(d+h) & c+g \\ c+g & d+h & a+e & -(b+f) \\ d+h & -(c+g) & b+f & a+e \end{pmatrix} = \\ &= \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} + \begin{pmatrix} e & -f & -g & -h \\ f & e & -h & g \\ g & h & e & -f \\ h & -g & f & e \end{pmatrix} = \varphi(X) + \varphi(Y) \\ \bullet \varphi(\lambda X) &= \begin{pmatrix} \lambda a & -\lambda b & -\lambda c & -\lambda d \\ \lambda b & \lambda a & -\lambda d & \lambda c \\ \lambda c & \lambda d & \lambda a & -\lambda b \\ \lambda d & -\lambda c & \lambda b & \lambda a \end{pmatrix} = \lambda \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} = \lambda \varphi(X) \\ \bullet \varphi(X)\varphi(Y) &= \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \begin{pmatrix} e & -f & -g & -h \\ f & e & -h & g \\ g & h & e & -f \\ h & -g & f & e \end{pmatrix} = \\ &= \begin{pmatrix} ae - bf - cg - dh & -af - be - ch + dg & -ag - ce - bh + df & -ah - de - bg + cf \\ af + be + ch - dg & ae - bf - cg - dh & -ah - de - bg + cf & ag + ce + bh - df \\ ag + ce + bh - df & ah + de + bg - cf & ae - bf - cg - dh & -af - be - ch + dg \\ ah + de + bg - cf & -ag - ce - bh + df & af + be + ch - dg & ae - bf - cg - dh \end{pmatrix} = \\ &\varphi(XY) \end{aligned}$$

Con lo que efectivamente φ es un homomorfismo de álgebras, y la matriz que nos proporciona para un cuaternión X es su representación matricial. $\square$

3.2.3 Representación como matrices complejas

Por último veremos la representación de los cuaterniones como matrices cuadradas complejas de dimensión 2. Hemos visto previamente que $\mathbb{H}$ es equivalente a $\mathbb{R}^4$ como espacio vectorial, y sabemos que los complejos $\mathbb{C}$ son a su vez equivalentes a $\mathbb{R}^2$. Con esto, podemos tomar dos números complejos $z = a + bi$ y $w = c + di$ y expresar un número cuaterniónico en función de ellos ya que $(a + bi) + (c + di)j = a + bi + cj + dk \in \mathbb{H}$. Sabiendo esto, consideramos la siguiente aplicación, que nos permite expresar un número cuaterniónico como matriz compleja de dimensión 2×2 :

$$\Phi : \mathbb{H} \rightarrow \mathcal{M}_{2 \times 2}(\mathbb{C})$$

$$a + bi + cj + dk \mapsto \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix} = \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}$$

Curiosamente, en este caso no ha funcionado el proceso que hemos usado previamente de multiplicar el cuaternión por las coordenadas de la base, quizás por no ser la multiplicación conmutativa para cuaterniones.

Podemos expresar dichas matrices complejas en función de una base, definida como sigue:

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \mathbf{I} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \mathbf{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \mathbf{K} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

Y reescribir de nuevo la aplicación como:

$$\Phi : \mathbb{H} \rightarrow \mathcal{M}_{2 \times 2}(\mathbb{C})$$

$$a + bi + cj + dk \mapsto aI_2 + b\mathbf{I} + c\mathbf{J} + d\mathbf{K}$$

Teorema 3.6 *La aplicación Φ es un homomorfismo de álgebras.*

Demostración

Sean $X = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, $Y = e\mathbf{1} + f\mathbf{i} + g\mathbf{j} + h\mathbf{k}$ y $\lambda \in \mathbb{R}$. Entonces:

- $\varphi(X + Y) = ((a + e)I_2 + (b + f)\mathbf{I} + (c + g)\mathbf{J} + (d + h)\mathbf{K}) = (aI_2 + b\mathbf{I} + c\mathbf{J} + d\mathbf{K}) + (eI_2 + f\mathbf{I} + g\mathbf{J} + h\mathbf{K}) = \varphi(X) + \varphi(Y)$
- $\varphi(\lambda X) = \lambda aI_2 + \lambda b\mathbf{I} + \lambda c\mathbf{J} + \lambda d\mathbf{K} = \lambda(aI_2 + b\mathbf{I} + c\mathbf{J} + d\mathbf{K}) = \lambda\varphi(X)$
- Para ver que $\varphi(X)\varphi(Y) = \varphi(XY)$, basta con comprobar que las matrices de la base de $\mathcal{M}_{2 \times 2}(\mathbb{C})$ siguen las mismas relaciones que $i, j, k \in \mathbb{H}$:

$$I^2 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I_2$$

$$\begin{aligned}
J^2 &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I_2 \\
K^2 &= \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I_2 \\
IJK &= \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = \\
&= \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I_2
\end{aligned}$$

Con lo que Φ es un homomorfismo de álgebras cuya matriz es la representación de un cuaternión como matriz compleja. $\square$

Además, el determinante de A será:

$$|A| = x\bar{x} + y\bar{y} = a^2 + b^2 + c^2 + d^2$$

Se define ahora la aplicación $\varphi : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{R}$ como:

$$\varphi(X, Y) = a_1a_2 + b_1b_2 + c_1c_2 + d_1d_2$$

Donde $X = a_1\mathbf{1} + b_1\mathbf{i} + c_1\mathbf{j} + d_1\mathbf{k}$ e $Y = a_2\mathbf{1} + b_2\mathbf{i} + c_2\mathbf{j} + d_2\mathbf{k}$.

Proposición 3.7 *La aplicación φ es un producto escalar en $\mathbb{H}$.*

Demostración

La matriz asociada a esta aplicación es I_4 y por ello la aplicación es bilineal, simétrica y definida positiva. $\square$

Con esto, llamaremos norma del cuaternión A a: $\sqrt{|A|} = \sqrt{x\bar{x} + y\bar{y}} = \sqrt{a^2 + b^2 + c^2 + d^2}$.

Corolario 3.8 *$\mathbb{H}$ es un espacio euclideo con el producto escalar definido por φ .*

3.3 Los cuaterniones y $SO(3)$

Ahora consideramos las aplicaciones $Y \mapsto XY$ para un X dado, y $X \mapsto XY$ para un Y dado. Estas son aplicaciones lineales y se pueden expresar como:

$$\begin{aligned}
XY = L_X Y &= \begin{pmatrix} a_1 & -b_1 & -c_1 & -d_1 \\ b_1 & a_1 & -d_1 & c_1 \\ c_1 & d_1 & a_1 & -b_1 \\ d_1 & -c_1 & b_1 & a_1 \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \\ c_2 \\ d_2 \end{pmatrix} \\
XY = R_Y X &= \begin{pmatrix} a_2 & -b_2 & -c_2 & -d_2 \\ b_2 & a_2 & d_2 & -c_2 \\ c_2 & -d_2 & a_2 & b_2 \\ d_2 & c_2 & -b_2 & a_2 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \\ c_1 \\ d_1 \end{pmatrix}
\end{aligned}$$

Lema 3.9 Las matrices L_X y R_Y de las aplicaciones previamente descritas son ortogonales.

Demostración

Primero, observamos que:

$$\begin{aligned} \bullet L_X^\top &= \begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ -b_1 & a_1 & d_1 & -c_1 \\ -c_1 & -d_1 & a_1 & b_1 \\ -d_1 & c_1 & -b_1 & a_1 \end{pmatrix} = L_{\overline{X}} \\ \bullet R_Y^\top &= \begin{pmatrix} a_2 & b_2 & c_2 & d_2 \\ -b_2 & a_2 & -d_2 & c_2 \\ -c_2 & d_2 & a_2 & -b_2 \\ -d_2 & -c_2 & b_2 & a_2 \end{pmatrix} = R_{\overline{Y}} \end{aligned}$$

Y con ello, consideramos los siguientes productos de matrices:

$$\begin{aligned} \bullet L_X L_{\overline{X}} &= \\ &= \begin{pmatrix} N(X) & a_1 b_1 - b_1 a_1 + c_1 d_1 - d_1 c_1 & a_1 c_1 - b_1 d_1 - c_1 a_1 + d_1 b_1 & a_1 d_1 + b_1 c_1 - d_1 a_1 - c_1 b_1 \\ b_1 a_1 - a_1 b_1 + d_1 c_1 - c_1 d_1 & N(X) & b_1 c_1 + a_1 d_1 - d_1 a_1 - c_1 d_1 & b_1 d_1 - a_1 c_1 - d_1 b_1 + a_1 c_1 \\ c_1 a_1 - d_1 b_1 - a_1 c_1 + b_1 d_1 & c_1 b_1 + d_1 a_1 - a_1 d_1 - b_1 c_1 & N(X) & c_1 d_1 - d_1 c_2 + a_1 b_1 - b_1 a_1 \\ d_1 a_1 + c_1 b_1 - b_1 c_1 - a_1 d_1 & d_1 b_1 - c_1 a_1 - b_1 d_1 + a_1 c_1 & d_1 c_1 - c_1 d_1 + b_1 a_1 - a_1 b_1 & N(X) \end{pmatrix} \\ &= N(X)I \end{aligned}$$

$$\begin{aligned} \bullet R_Y R_{\overline{Y}} &= \\ &= \begin{pmatrix} N(Y) & a_2 b_2 - b_2 a_2 - c_2 d_2 + d_2 c_2 & a_2 c_2 + b_2 d_2 - c_2 a_2 - d_2 b_2 & a_2 d_2 - b_2 c_2 + d_2 a_2 - c_2 b_2 \\ b_2 a_2 - a_2 b_2 - d_2 c_2 + c_2 d_2 & N(Y) & b_2 c_2 - a_2 d_2 + d_2 a_2 - c_2 d_2 & b_2 d_2 + a_2 c_2 - d_2 b_2 - a_2 c_2 \\ c_2 a_2 + d_2 b_2 - a_2 c_2 - b_2 d_2 & c_2 b_2 - d_2 a_2 + a_2 d_2 - b_2 c_2 & N(Y) & c_2 d_2 - d_2 c_2 - a_2 b_2 + b_2 a_2 \\ d_2 a_2 - c_2 b_2 + b_2 c_2 - a_2 d_2 & d_2 b_2 + c_2 a_2 - b_2 d_2 - a_2 c_2 & d_2 c_2 - c_2 d_2 - b_2 a_2 + a_2 b_2 & N(Y) \end{pmatrix} \\ &= N(Y)I \end{aligned}$$

Además, si hacemos el determinante con estos productos llegamos a que:

- $\det(L_X L_{\overline{X}}) = \det(L_X) \det(L_{\overline{X}}) = \det(L_X)^2 = N(X)^4 = \det(N(X)I)$, y con esto llegamos a $\det(L_X) = \pm N(X)^2$. Como el primer término del determinante de L_X es a_1^4 , llegamos a que $\det(L_X) = N(X)^2$.
- $\det(R_Y R_{\overline{Y}}) = \det(R_Y) \det(R_{\overline{Y}}) = \det(R_Y)^2 = N(Y)^4 = \det(N(Y)I)$, y con esto llegamos a $\det(R_Y) = \pm N(Y)^2$. Como el primer término del determinante de R_Y es a_2^4 , llegamos a que $\det(R_Y) = N(Y)^2$.

Entonces, hemos llegado a que L_X y R_Y son ortogonales, puesto que multiplicandolas por su traspuesta obtenemos la identidad, y que su determinante es $N(X)^2$ y $N(Y)^2$ respectivamente. $\square$

Proposición 3.10 Si tomamos X e Y unitarios, las matrices L_X y R_Y serán matrices de rotación.

Demostración

El lema anterior nos da la ortogonalidad de L_X y R_Y . A lo largo de la demostración de dicho lema se ve que $\det(L_X) = \pm N(X)^2$ y $\det(R_Y) = \pm N(Y)^2$. Como X e Y son unitarios, $N(X) = N(Y) = 1$ y con esto $\det(L_X) = \det(R_Y) = 1$. Como son ortogonales y su determinante vale 1, $L_X, R_Y \in \mathbb{H}$. $\square$

Consideremos ahora la aplicación $\rho_{Y,Z} : \mathbb{H} \rightarrow \mathbb{H}$ que lleva $X \mapsto YXZ$, donde $Y, Z \in \mathbb{H}$ tales que $N(Y)N(Z) = 1$. Como $N(UV) = N(U)N(V) \forall U, V \in \mathbb{H}$, vemos que esta aplicación es una isometría. Además, $\rho_{Y,Z} = \rho_{Y,1} \circ \rho_{1,Z}$, y como sabemos que $\rho_{Y,1}$ y $\rho_{1,Z}$ son rotaciones, $\rho_{Y,Z}$ lo es por ser composición de estas.

Proposición 3.11 *La aplicación $\rho_{Y,Y^{-1}}$ (que denominaremos ρ_Y) es la identidad para $1\mathbb{R}$ y envía $\mathbb{H}_p$ a si mismo, como podemos observar a continuación:*

Demostración

1. $\rho_Y = id_{1\mathbb{R}}$

Sean $X = a\mathbf{1}$ e $Y = e\mathbf{1} + f\mathbf{i} + g\mathbf{j} + h\mathbf{k}$:

$$\begin{aligned} YXY^{-1} &= (ae\mathbf{1} + af\mathbf{i} + ag\mathbf{j} + ah\mathbf{k})(e\mathbf{1} - f\mathbf{i} - g\mathbf{j} - h\mathbf{k})/N(Y) = \\ &= [(a(e^2 + f^2 + g^2 + h^2))\mathbf{1} + (-aef + aef - agh + agh)\mathbf{i} + (-aeg + afh + aeg - afh)\mathbf{j} + (-aeh - afg + afg - aeh)\mathbf{k}]/N(Y) = (a(e^2 + f^2 + g^2 + h^2))\mathbf{1}/N(Y) = a\mathbf{1} \end{aligned}$$

2. ρ_Y envía $\mathbb{H}_p$ a si mismo

Sean $X = b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{H}_p$ e $Y = e\mathbf{1} + f\mathbf{i} + g\mathbf{j} + h\mathbf{k}$:

$$\begin{aligned} YXY^{-1} &= \\ &= [(-fb - gc - hd)\mathbf{1}(eb + gd - hc)\mathbf{i} + (ec - fd + hb)\mathbf{j} + (ed + fc - gb)\mathbf{k}](e\mathbf{1} - f\mathbf{i} - g\mathbf{j} - h\mathbf{k})/N(Y) = \\ &= [(b(f^2 + e^2 - h^2 - g^2) + 2(gcf + hdf + gde - hce))\mathbf{i} + (c(g^2 - h^2 + e^2 - f^2) + 2(fbg + hdg + ebh - fde))\mathbf{j} + (d(h^2 - g^2 - f^2 + e^2) + 2(fbh + gch + ecf - gbe))\mathbf{k}]/N(Y) \end{aligned}$$

Que pertenece a $\mathbb{H}_p$ ya que la parte real se anula.

$\square$

Teorema 3.12 *La aplicación $\rho_Y \in SO(3)$.*

Demostración

Primero, vemos que $\rho_Y(X + Z) = \rho_Y(X) + \rho_Y(Z)$ y $\rho_Y(\overline{X}) = \overline{\rho_Y(X)}$.

1. $\rho_Y(X + Z) = \rho_Y(X) + \rho_Y(Z)$

Sean $X = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, $Y = e\mathbf{1} + f\mathbf{i} + g\mathbf{j} + h\mathbf{k}$ y $Z = m\mathbf{1} + n\mathbf{i} + o\mathbf{j} + p\mathbf{k}$

$$\begin{aligned} \rho_Y(X) + \rho_Y(Z) &= [(a(e^2 + f^2 + g^2 + h^2))\mathbf{1} + (b(e^2 + f^2 - g^2 - h^2) + 2(-hce + gcf + hdf + gde))\mathbf{i} + (c(e^2 - f^2 + g^2 - h^2) + 2(hbe - fde + gbh + hdg))\mathbf{j} + (d(e^2 - f^2 - g^2 + h^2) + 2(fbh + gch + ecf - gbe))\mathbf{k}] + [(m(e^2 + f^2 + g^2 + h^2))\mathbf{1} + (n(e^2 + f^2 - g^2 - h^2) + 2(-hoe + gof + hpf + gpe))\mathbf{i} + (o(e^2 - f^2 + g^2 - h^2) + 2(hne - fpe + gnf + hpg))\mathbf{j} + (p(e^2 - f^2 - g^2 + h^2) + 2(fnh + goh + eof - gne))\mathbf{k}] = \end{aligned}$$

$$= [((a+m)(e^2+f^2+g^2+h^2))\mathbf{1} + ((b+n)(e^2+f^2-g^2-h^2) + 2(-h(o+c)e + g(o+c)f + h(d+p)f + g(d+p)e))\mathbf{i} + ((o+c)(e^2-f^2+g^2-h^2) + 2(h(b+n)e - f(d+p)e + g(b+n)f + h(d+p)g))\mathbf{j} + ((d+p)(e^2-f^2-g^2+h^2) + 2(f(b+n)h + g(o+c)h + e(o+c)f - g(b+n)e))\mathbf{k}] = \rho_Y(X + Z)$$

$$2. \rho_Y(\overline{X}) = \overline{\rho_Y(X)}$$

$$\rho_Y(\overline{X}) = (a(e^2+f^2+g^2+h^2))\mathbf{1} + (-b(e^2+f^2-g^2-h^2) + 2(hce - gcf - hdf - gde))\mathbf{i} + (-c(e^2-f^2+g^2-h^2) + 2(-hbe + fde - gbf - hdg))\mathbf{j} + (-d(e^2-f^2-g^2+h^2) + 2(-fbh - gch - ecf + gbe))\mathbf{k} =$$

$$= (a(e^2+f^2+g^2+h^2))\mathbf{1} - (b(e^2+f^2-g^2-h^2) + 2(-hce + gcf + hdf + gde))\mathbf{i} - (c(e^2-f^2+g^2-h^2) + 2(hbe - fde + gbf + hdg))\mathbf{j} - (d(e^2-f^2-g^2+h^2) + 2(fbh + gch + ecf - gbe))\mathbf{k} = \overline{\rho_Y(X)}$$

Como consecuencia de estas dos propiedades, se tiene que $\rho_Y(X + \overline{X}) = \rho_Y(X) + \rho_Y(\overline{X}) = \rho_Y(X) + \overline{\rho_Y(X)}$.

También tenemos que $X + \overline{X} = 2a\mathbf{1}$, y concretamente, si $X \in \mathbb{H}_p$, $X + \overline{X} = 0$. Con esto, y con lo que hemos visto previamente llegamos a $\rho_Y(X + \overline{X}) = \rho_Y(X) + \overline{\rho_Y(X)} = 0$. Como conclusión, si tenemos un cuaternion puro X , sabemos que su transformado $\rho_Y(X)$ también lo es, y no solo eso, si no que el opuesto de este es $\overline{\rho_Y(X)} = \rho_Y(\overline{X})$ y con esto queda probado que $\rho_Y \in SO(3)$. $\square$

Ahora que he visto que $\rho_Y \in SO(3)$, me falta ver que todas las rotaciones de $SO(3)$ tienen forma de ρ_Y .

Teorema 3.13 *Toda rotación de $SO(3)$ es de la forma ρ_y con $y \in \mathbb{H}$.*

Demostración

Es fácil ver que $\rho_{YX} = \rho_Y \circ \rho_X$, es decir, la composición de dos rotaciones es una rotación, y sabemos que toda rotación distinta de la identidad es el resultado de componer dos reflexiones. Entonces, ahora vamos a ver que para toda reflexion σ de $\mathbb{H}_p$ respecto de un plano H , existe $Z \in \mathbb{H}_p$, $z \neq 0$ de forma que $\sigma(X) = -ZXZ^{-1}$, $\forall X \in \mathbb{H}_p$:

Si tomamos $Z \in \mathbb{H}_p$ ortogonal a H , sabemos que

$$\sigma(X) = X - 2\frac{X \cdot Z}{Z \cdot Z}Z$$

para todo $X \in \mathbb{H}_p$. Además, para dos cuaterniones puros $X, Y \in \mathbb{H}_p$:

$$2(X \cdot Y)\mathbf{1} = -(XY - YX)$$

Y $(Z \cdot Z)\mathbf{1} = -Z^2$ por ser Z cuaternión puro. Ahora, con todo esto:

$$\sigma(X) = X - 2\frac{X \cdot Z}{Z \cdot Z}Z = X + 2(X \cdot Z)Z^{-1} = X - (XZ + ZX)Z^{-1} = -ZXZ^{-1}$$

Con lo que hemos visto que $\sigma(X) = -ZXZ^{-1} \forall X \in \mathbb{H}_p$ para un $Z \in \mathbb{H}_p$, $Z \neq 0$ que era lo que queríamos. $\square$

3.4 Determinando una rotación a partir de un cuaternión

Ahora vamos a ver como obtener el eje y el ángulo de una rotación a partir del cuaternión que la representa.

3.4.1 Determinando el eje

Si consideramos la rotación representada por $Z = a\mathbf{1} + t$, con $t \in \mathbb{H}_p$, el eje de rotación quedará fijo a través de esta.

Proposición 3.14 *El eje de la rotación dada por el cuaternión $Z = a\mathbf{1} + t$ es el vector t en $\mathbb{R}^3$.*

Demostración

Para ver que t es el eje de rotación, tengo que ver que es invariante por esta. Entonces, tomando $t = b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ consideramos:

$$\begin{aligned}\rho_Z(t) &= ZtZ^{-1} = ((-b^2 - c^2 - d^2)\mathbf{1} + ab\mathbf{i} + ac\mathbf{j} + ad\mathbf{k})Z^{-1} = \\ &= ((a^2b + b^3 + bc^2 + bd^2)\mathbf{i} + (a^2c + b^2c + c^3 + cd^2)\mathbf{j} + (a^2d + b^2d + c^2d + d^3)\mathbf{k}) / (a^2 + b^2 + c^2 + d^2) = \\ &= b\mathbf{i} + c\mathbf{j} + d\mathbf{k} = t\end{aligned}$$

Con lo que efectivamente, t es invariante con la rotación ρ_Z , y por ello es el eje de esta. $\square$

Ahora que conocemos el eje de rotación, vemos como obtener el ángulo.

3.4.2 Determinando el ángulo

Proposición 3.15 *Dados dos cuaterniones puros no nulos X, Y tales que $N(X) = N(Y)$, existirá $W \in \mathbb{H}$ no nulo tal que $\rho_W(X) = WXW^{-1} = Y$*

Demostración

Para cualesquiera vectores $X, Y \in \mathbb{R}^3$ tales que $N(X) = N(Y)$ existe una rotación ρ tal que $\rho(X) = Y$. Si $X = Y$ sería la identidad y en caso contrario tomaríamos como eje de rotación $W = X \times Y$ para rotar X hasta Y a lo largo del plano que contiene a ambos. $\square$

Proposición 3.16 *El ángulo de las rotaciones ρ_Z y $\rho_{WZW^{-1}}$ es el mismo.*

Demostración

Sea $Z = a\mathbf{1} + t$, $t \in \mathbb{H}_p$. Veamos primero que el eje de $\rho_{WZW^{-1}}$ es $WtW^{-1} = \rho_W(t)$:

$$WZW^{-1} = W(a\mathbf{1} + t)W^{-1} = Wa\mathbf{1}W^{-1} + WtW^{-1} = a\mathbf{1} + WtW^{-1}$$

Como ya se ha visto previamente, ρ_W envía $\mathbb{H}_p$ a si mismo, luego $WtW^{-1} \in \mathbb{H}_p$ y concluimos que es el eje de la rotación $\rho_{WZW^{-1}}$.

Por otra parte, para $X \in \mathbb{H}_p$ no nulo y ortogonal a t , el ángulo de la rotación ρ_Z es el ángulo que forman X y $\rho_Z(X)$. Como las rotaciones son isometrías, si $X \cdot t = 0$, cosa que sabemos porque son ortogonales, entonces $\rho_W(X) \cdot \rho_W(t) = 0$. Además, también sabemos que el ángulo de la rotación $\rho_{WZW^{-1}}$ es el ángulo entre $\rho_W(X)$ y $\rho_{WZW^{-1}}(\rho_W(X))$ por lo que hemos visto previamente. Con todo esto:

$$\rho_{WZW^{-1}}(\rho_W(X)) = (\rho_W \circ \rho_Z \circ (\rho_W)^{-1} \circ \rho_W)(X) = \rho_W(\rho_Z(X))$$

Con lo que el ángulo de $\rho_{WZW^{-1}}$ es el ángulo entre $\rho_Z(X)$ y $\rho_W(\rho_Z(X))$, que será el mismo que el ángulo entre X y $\rho_Z(X)$ puesto que las rotaciones conservan los ángulos, es decir, el ángulo de la rotación ρ_Z , que era lo que queríamos demostrar.

Una vez visto esto, y sabiendo que para todo $Z = a\mathbf{1} + t$ existe $W \neq 0$ tal que $WtW^{-1} = \sqrt{N(t)}\mathbf{i}$ y $WZW^{-1} = a\mathbf{1} + \sqrt{N(t)}\mathbf{i}$, es suficiente con hayar el ángulo de rotación para un $Z = a\mathbf{1} + b\mathbf{i}$. Para ello, calculo $\rho_Z(\mathbf{j})$ puesto que $\mathbf{i} \perp \mathbf{j}$. Se diferencian dos casos:

- Si $a = 0$

$$\rho_Z(\mathbf{j}) = (b\mathbf{i})\mathbf{j}(b\mathbf{i})^{-1} = \frac{1}{b^2}b\mathbf{k}(-b\mathbf{i}) = \frac{1}{b}b(-\mathbf{j}) = -\mathbf{j}$$

Y entonces el ángulo será $\theta = \pi$

- Si $a \neq 0$

$$\rho_Z(\mathbf{j}) = (a\mathbf{1} + b\mathbf{i})\mathbf{j}(a\mathbf{1} + b\mathbf{i})^{-1} = \frac{1}{a^2 + b^2}(a\mathbf{1} + b\mathbf{i})\mathbf{j}(a\mathbf{1} - b\mathbf{i}) = \frac{a^2 - b^2}{a^2 + b^2}\mathbf{j} + \frac{2ab}{a^2 + b^2}\mathbf{k}$$

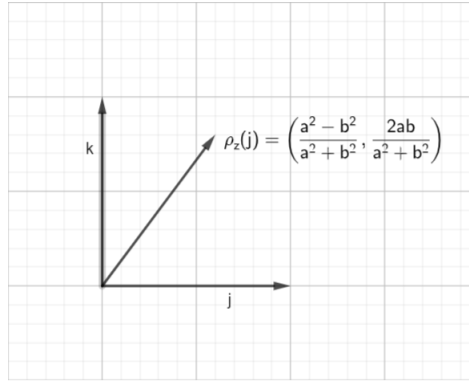


Figure 1: En la imagen se puede observar de forma más visual qué estamos obteniendo con la cuenta anterior.

Y con esto, llego a que

$$\tan(\theta) = \frac{2ab}{a^2 - b^2} = \frac{2ab}{1 - \left(\frac{b}{a}\right)^2}$$

Ahora, con la fórmula del ángulo doble obtengo lo siguiente:

$$\tan(\theta) = \frac{2\tan(\theta/2)}{1 - \tan^2(\theta/2)}$$

Y de aquí obtengo $\tan\left(\frac{\theta}{2}\right) = \frac{b}{|a|} = \frac{\sqrt{N(t)}}{|a|}$.

Además, de la fórmula de la tangente del ángulo doble $\tan\left(\frac{\theta}{2}\right) = \pm \sqrt{\frac{1 - \cos(\theta)}{1 + \cos(\theta)}}$,

despejando $\cos(\theta)$ llego a que $\cos(\theta) = \frac{1 - \tan^2(\theta/2)}{1 + \tan^2(\theta/2)}$. Como hemos visto que

$$\tan\left(\frac{\theta}{2}\right) = \frac{b}{|a|} \text{ llegamos a que } \cos(\theta) = \frac{a^2 - b^2}{a^2 + b^2} = \frac{a^2 - N(t)}{a^2 + N(t)}.$$

Usando lo que acabamos de ver, si $Z \in \mathbb{H}$ fuese unitario, tendríamos $1 = N(Z) = a^2 + N(t)$ y por ello $-N(t) = a^2 - 1$. Entonces, $\cos(\theta) = \frac{a^2 - N(t)}{a^2 + N(t)} = a^2 - N(t) = 2a^2 - 1$.

Para acabar, usando que $\cos(\theta) = 2\cos^2(\theta/2) - 1$ llegamos a que $\cos(\theta) = |a|$.

4 $SO(3)$ como grupo de Lie

Los resultados de esta sección se han extraído de [11], [1], [5], [8] y [9].

4.1 El grupo de Lie $SO(3)$

En esta sección estudiaremos $SO(3)$ a través de la topología diferencial, y nuestro objetivo es hallar el álgebra de Lie asociada al grupo especial ortogonal.

Para ello, primero tenemos que conocer cómo son los grupos de Lie, y ver que estamos frente a uno de ellos.

Definición 4.1 *Un grupo de Lie es un grupo que también tiene estructura de variedad diferenciable, y en el que las funciones multiplicación e inversión son diferenciables.*

Proposición 4.2 *En la definición 4.1, la condición de que los operadores multiplicación e inversión son diferenciables es equivalente en que lo sea el operador definido como $(x, y) \mapsto x^{-1}y$.*

Demostración

Tenemos que ver que $(x, y) \mapsto x^{-1}y$ es diferenciable si y sólo si la multiplicación y la inversión son diferenciables.

Veamos primero que si el operador $(x, y) \mapsto x^{-1}y$ es diferenciable, entonces también lo son la multiplicación y la inversión. Para la inversión, consideramos el operador $(x, e) \mapsto x^{-1}$, que nos da la inversión. Este puede verse como el operador que teníamos pero tomando siempre el neutro como segundo elemento, con lo que será diferenciable por ser una restricción del operador que ya sabemos que lo es. Por otra parte, para la multiplicación, consideramos el siguiente operador $(x, y) \mapsto (x^{-1}, y) \mapsto xy$ que nos daría como resultado la multiplicación. Como la primera parte es la inversión, que acabamos de ver que es diferenciable, no nos supondrá un problema, y la segunda es el operador que teníamos en un principio que es diferenciable. Como la composición de aplicaciones diferenciables es diferenciable, concluimos que la multiplicación también es diferenciable.

Ahora vemos que si la multiplicación y la inversión son diferenciables, también lo es $(x, y) \mapsto x^{-1}y$. Tomamos $(x, y) \mapsto (x^{-1}, y) \mapsto x^{-1}y$ que resulta en $(x, y) \mapsto x^{-1}y$. La primera parte es una inversión, y la segunda una multiplicación, la cuales sabemos que son diferenciables. Como la composición de aplicaciones diferenciables es diferenciable, concluimos que el operador $(x, y) \mapsto x^{-1}y$ también es diferenciable.

Hemos visto que ambas implicaciones son ciertas, luego concluimos que $(x, y) \mapsto x^{-1}y$ es diferenciable si y sólo si la multiplicación y la inversión son diferenciables. $\square$

Una vez visto esto, veamos que $SO(3)$ es un grupo de Lie:

Teorema 4.3 *El grupo $SO(3)$ es un grupo de Lie.*

Demostración

La proposición 2.8 ya demostrada nos dice que $SO(3)$ es grupo con la multiplicación de matrices. Por otra parte, para ver que es una variedad diferenciable usaremos que

$SO(3) \cong \mathbb{P}^3$, que ya hemos visto en la sección 2.4.2. Sabiendo esto, nos bastará con ver que el espacio proyectivo $\mathbb{P}^3$ es variedad diferenciable. Veámoslo:

Sabemos que $\mathbb{P}^3$ es el resultado del cociente de la esfera $S^3/\mathbb{Z}_2$. Para ver que tenemos una variedad diferenciable usamos la siguiente propiedad:

Sea G un grupo finito o numerable de difeomorfismos de una variedad M tales que para cada $p \in M$, existe un entorno U de p tal que $U \cap g(U) = \emptyset$ para todo $g \in G$ distinto de la identidad. Entonces M/G es también una variedad diferenciable de la misma dimensión que M . Esto se conoce como un grupo que actúa de modo discontinuo.

En este caso, tenemos $M = S^3$ y $G = \mathbb{Z}_2$. Con lo que si vemos que la esfera S^3 es variedad diferenciable habremos demostrado esta parte. Pero ver que S^3 es una variedad diferenciable es mucho más sencillo que lo que teníamos que demostrar en un primer momento. Basta considerar el siguiente atlas $\mathcal{A} = \{\alpha, \beta\}$:

$$\begin{aligned}\alpha : S^3 - \{(1, 0, 0, 0)\} &\rightarrow \mathbb{R}^3 \\ (x, y, z, t) &\mapsto \left(\frac{y}{1-x}, \frac{z}{1-x}, \frac{t}{1-x} \right) \\ \beta : S^3 - \{(-1, 0, 0, 0)\} &\rightarrow \mathbb{R}^3 \\ (x, y, z, t) &\mapsto \left(\frac{y}{1+x}, \frac{z}{1+x}, \frac{t}{1+x} \right)\end{aligned}$$

Veamos primero que $\mathcal{A}$ es un atlas. Las aplicaciones α y β son dos proyecciones estereográficas de la esfera S^3 . Son inyectivas puesto que el único punto que no tendría imagen para cada una está excluido de sus dominios, y su imagen es un abierto puesto que es todo $\mathbb{R}^3$. Con lo que estas dos aplicaciones son cartas. Como la unión de sus dominios cubre toda S^3 tenemos un atlas. Falta ver que este atlas es C^∞ :

Llamamos U y V al dominio de α y β respectivamente. Tenemos que $\alpha(U \cap V) = \beta(U \cap V) = \mathbb{R}^3 - \{(0, 0, 0)\}$, que es abierto por ser el complementario de un cerrado. Veamos que el cambio de cartas es un difeomorfismo:

$$\begin{aligned}\alpha(U \cap V) &\xrightarrow{\alpha^{-1}} U \cap V \xrightarrow{\beta} \beta(U \cap V) \\ (x, y, z) &\mapsto \left(\frac{x^2+y^2+z^2-1}{1+x^2+y^2+z^2}, \frac{2x}{1+x^2+y^2+z^2}, \frac{2y}{1+x^2+y^2+z^2}, \frac{2z}{1+x^2+y^2+z^2} \right) \mapsto \\ &\quad \left(\frac{x}{x^2+y^2+z^2}, \frac{y}{x^2+y^2+z^2}, \frac{z}{x^2+y^2+z^2} \right)\end{aligned}$$

Las funciones de cada componente son continuas, puesto que el denominador no se va a anular ya que $(0, 0, 0) \notin \alpha(U \cap V)$. Lo mismo sucede con sus derivadas. Además, es biyectiva por construcción. Luego efectivamente $\beta \circ \alpha^{-1}$ es un difeomorfismo. Un proceso similar también nos permite llegar a que $\alpha \circ \beta^{-1}$ también es un difeomorfismo. Con esto hemos visto que el cambio de cartas es siempre un difeomorfismo y por ello concluimos que $\mathcal{A}$ es un atlas C^∞ . Como todo atlas C^∞ está contenido en uno maximal, concluimos que S^3 con el atlas maximal dado por $\mathcal{A}$ tiene estructura de variedad diferenciable.

Por último, consideramos las siguientes aplicaciones:

$$\cdot : SO(3) \times SO(3) \rightarrow SO(3)$$

$$(A, B) \mapsto AB$$

Esta aplicación será diferenciable puesto que las entradas de la matriz AB serán expresiones polinómicas.

$$\cdot : SO(3) \rightarrow SO(3)$$

$$A \mapsto A^{-1}$$

Como $A^{-1} = A^\top$ ya que $A \in SO(3)$, la inversa siempre estará definida y las entradas de la matriz A^{-1} serán polinómicas de nuevo, y con ello la aplicación será diferenciable.

Con todo esto, hemos visto que $SO(3)$ es variedad diferenciable, grupo, y está dotado de las funciones multiplicación e inversión, las cuales son diferenciables. Es por ello que concluimos que es un grupo de Lie. $\square$

Además, el teorema 3.25 de [11], dice que todo grupo de Lie conexo admite un recubrimiento simplemente conexo, que es a su vez grupo de Lie y tal que la aplicación del recubrimiento es un homomorfismo de grupos.

Si lo aplicamos a nuestra situación, hemos visto que $SO(3)$ es un grupo de Lie y sabemos que es conexo, y además tenemos que $\pi_1(S^3) \cong \{e\}$, con lo que S^3 es recubrimiento simplemente conexo de $SO(3)$ y S^3 es grupo de Lie.

Ahora que hemos visto que $SO(3)$ es un grupo de Lie, podemos pasar a ver que es un álgebra de Lie y como se relaciona con un grupo de Lie.

4.2 El álgebra de Lie de $SO(3)$

Definición 4.4 *Un álgebra de Lie es un espacio vectorial $\mathfrak{g}$ sobre un cuerpo F que tiene definida una operación binaria $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ que cumple las siguientes propiedades:*

1. $[ax + by, z] = a[x, z] + b[y, z]$
2. $[x, ay + bz] = a[x, y] + b[x, z]$
3. $[x, y] = -[y, x]$
4. $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$

para todo $x, y, z \in \mathfrak{g}$, $a, b \in F$.

Consideremos ahora el siguiente conjunto:

$$\mathfrak{so}(3) = \{A \in \mathcal{M}_{3 \times 3}(\mathbb{R}) : A^\top = -A\}$$

El corchete de Lie de $\mathfrak{so}(3)$ viene dado por el conmutador de matrices, que dadas $A, B \in \mathfrak{so}(3)$, se tiene $[A, B] = AB - BA$.

Teorema 4.5 *$\mathfrak{so}(3)$ con el conmutador de matrices es un álgebra de Lie.*

Demostración

Como las matrices cuadradas de dimensión n forman un espacio vectorial y el conmutador de matrices satisface las condiciones de la definición 4.4, son un álgebra de Lie que denominamos como $\mathfrak{gl}(n, \mathbb{R})$. Es por ello que también podríamos demostrar este teorema viendo que $\mathfrak{so}(3)$ es subálgebra de Lie de $\mathfrak{gl}(n, \mathbb{R})$.

Veamos primero que $\mathfrak{so}(3)$ es un espacio vectorial real.

Primero tenemos que ver que la suma es conmutativa y asociativa, y que tiene neutro e inverso. Claramente la suma de matrices es conmutativa y asociativa. Tenemos el elemento neutro para la suma de matrices en $\mathfrak{so}(3)$, que es la matriz cuyas entradas son todas 0, ya que esta es antisimétrica. Dada una matriz $A \in \mathfrak{so}(3)$, sabemos que $A^\top = -A$. Como $A \in \mathfrak{so}(3)$, entonces $A^\top \in \mathfrak{so}(3)$, y por tanto $-A \in \mathfrak{so}(3)$.

Por otra parte, en cuanto a la multiplicación por escalar tenemos que ver que sea asociativa, que tenga neutro, y la propiedad distributiva respecto de la suma tanto escalar como vectorial. Veámoslo. Sea $A, B \in \mathfrak{so}(3)$. Tenemos elemento neutro puesto que $1A = A$. Además, dados $a, b \in \mathbb{R}$, claramente $(ab)A = a(bA)$ con lo que es asociativa. En cuanto a las propiedades distributivas, se cumple que $a(A+B) = aA + aB$ y $(a+b)A = aA + bA$ ya que funciona para matrices en general.

Ahora nos queda ver que el conmutador de matrices verifica las condiciones de la definición 4.4. Veámoslo. Sean $a, b \in \mathbb{R}$ y $A, B, C \in \mathfrak{so}(3)$:

1. Veámos que $[aA + bB, C] = a[A, C] + b[B, C]$:

$$[aA + bB, C] = (aA + bB)C - C(aA + bB) = aAC + bBC - aCA - bCB = a[A, C] + b[B, C]$$

2. Veámos que $[C, aA + bB, C] = a[C, A] + b[C, B]$:

$$[C, aA + bB] = C(aA + bB) - (aA + bB)C = aCA + bCB - aAC - bBC = a[C, A] + b[C, B]$$

3. Comprobemos que es anticonmutativa, esto es, $[A, B] = -[B, A]$

$$\text{Tenemos que } [A, B] = AB - BA = -(BA - AB) = -[B, A]$$

4. Por último, comprobemos que verifica la identidad de Jacobi. Para ello calculamos primero cada término por separado:

$$[A, [B, C]] = [A, BC - CB] = ABC - ACB - BCA + CBA$$

$$[B, [C, A]] = [B, CA - AC] = BCA - BAC - CAB + ACB$$

$$[C, [A, B]] = [C, AB - BA] = CAB - CBA - ABC + BAC$$

$$\text{Y con esto se ve que } [A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$$

Con todo esto, concluimos que $\mathfrak{so}(3)$ es un álgebra de Lie. $\square$

Ahora bien, podemos decir más que $\mathfrak{so}(3)$ es un álgebra de Lie. A cada grupo de Lie se le puede asociar un álgebra de Lie, luego queremos ver que $\mathfrak{so}(3)$ es el álgebra de Lie que le corresponde al grupo de Lie que es $SO(3)$. Para ello necesitamos definir ciertos conceptos previamente:

Definición 4.6 Sea G un grupo de Lie, y sea X un campo vectorial en G . Se dice que X es invariante por la izquierda si:

$$(L_g)_*(X_e) = X_g$$

donde $L_g(x) = gx$, $\forall x \in G$, $(f)_*$ denota la aplicación tangente:

$$(L_g)_* : T_e G \rightarrow T_g G$$

$$X_e \mapsto (L_g)_*(X_e)$$

y para todo $g, h \in G$, $X \in \mathfrak{X}(G)$.

Teorema 4.7 El conjunto $\mathcal{L}(\mathfrak{X})$ definido como:

$$\mathcal{L}(\mathfrak{X}) = \{X \in \mathfrak{X}(G) : X \text{ es invariante por la izquierda}\}$$

es un álgebra de Lie, con la operación $[,]$ dada por el corchete de Lie de campos vectoriales.

Demostración

Vemos solo una idea de la demostración, ya que para verla con todo detalle tendríamos que tratar conceptos que se escapan de los límites del trabajo.

La propiedad anticonmutativa y la identidad de Jacobi se verifican porque las cumple el corchete de Lie de campos vectoriales para todo $X \in \mathfrak{X}(G)$.

Me falta ver que dados $X, Y \in \mathfrak{X}(G)$, entonces $[X, Y] \in \mathcal{L}(\mathfrak{X})$. Para ello tengo que ver que $[X, Y]$ es invariante por la izquierda, es decir, que $(L_g)_*([X, Y]_e) = [X, Y]_g$. Veámoslo:

Como ya hemos visto que el producto de elementos de G es diferenciable, podemos usar que $(L_g)_*([X, Y]_p)(f) = [X, Y](f \circ L_g)$.

Sean $X, Y \in \mathfrak{X}(G)$, $f \in \mathfrak{F}(G)$, $g \in G$ y $e \in G$ elemento neutro de G :

$$\begin{aligned} [X, Y]_g(f) &= X_g(Y(g)) - Y_g(X(g)) = (L_g)_*(X_e)(Y(f)) - (L_g)_*(Y_e)(X(f)) = X_e(Y(f) \circ L_g) - Y_e(X(f) \circ L_g) \\ &= X_e(Y(f \circ L_g)) - Y_e(X(f \circ L_g)) = [X, Y]_e(f \circ L_g) = (L_g)_*([X, Y]_e)(f) \end{aligned}$$

que es lo que queríamos ver. Con todo esto, concluimos que $\mathcal{L}(\mathfrak{X})$ es un álgebra de Lie con la operación definida por el corchete de Lie de campos vectoriales. $\square$

Ahora que hemos probado el teorema anterior, consideramos la siguiente aplicación:

$$\eta : \mathcal{L}(\mathfrak{X}) \rightarrow T_e G$$

$$X \mapsto X_e = \vec{v} \text{ con lo que será } X_g = (L_g)_*(\vec{v})$$

Proposición 4.8 La aplicación η es una biyección de espacios vectoriales.

Demostración

Se ve trivialmente que es lineal.

Para ver que es inyectiva, como para todo $g \in G$, $X_g = (L_g)_*(X_e)$, con lo que si $X_e = 0$, entonces tendremos que $X = 0$. Por otra parte, es sobreyectiva puesto que dado $\vec{v} = X_e \in T_e G$, solo hay un único campo que para e nos da X_e , que es el campo tal que para cada $g \in G$, $X_g = (L_g)_*(X_e)$. Con lo que efectivamente es una biyección. $\square$

Ahora que hemos visto este resultado, podemos pasar a ver cual es el álgebra de Lie de $SO(3)$, ya que al haber visto la relación entre $\mathcal{L}(\mathfrak{X})$ y $T_e G$, podemos construirla viendo cómo es $T_e SO(3)$. En realidad, podríamos tomar el espacio tangente en cualquier punto, ya que podemos pasar de $T_e G$ a $T_g G$ para cualquier $g \in G$ grupo de Lie a través de un isomorfismo. Pero tomamos el neutro porque este es el elemento que siempre va a aparecer en todo G .

Ahora que hemos visto esto, podemos considerar que todo grupo de Lie es paralelizable, es decir, admite una base global de campos. Esto se ve tomando una base $\{v_1, v_2, \dots, v_n\}$ y construyendo los campos invariantes a la izquierda $\{X_1, X_2, \dots, X_n\}$ tales que $(X_i)_e = v_i$ para cualquier $i \in \{1, 2, \dots, n\}$. Es por esto que podemos decir que tanto $SO(3)$ es paralelizable.

Teorema 4.9 *El álgebra de Lie de $SO(3)$ es $\mathfrak{so}(3)$.*

En cuanto a la demostración de este teorema, se queda fuera del trabajo ya que para poder verla tendríamos que tratar conceptos que quedan fuera de los límites del trabajo.

5 Conclusión

Para finalizar el trabajo, vamos a resumir todo lo que hemos visto.

Primero de todo, hemos visto cómo es $SO(3)$ como estructura algebraica. Hemos visto que es subgrupo del grupo ortogonal $O(3)$, y que es isomorfo al grupo de isometrías lineales que preservan la orientación, que a su vez es subgrupo de el grupo de isometrías lineales del espacio.

Después, también hemos estudiado el cómo hallar la rotación del espacio que determina una matriz de $SO(3)$, y también el proceso contrario, es decir, conociendo una rotación por medio de su eje y su ángulo de rotación, hemos visto un proceso para determinar la matriz de $SO(3)$ que representa dicha rotación.

Además hemos visto la relación entre $SO(3)$ y el grupo especial unitario $SU(2)$, para lo cual primero hemos estudiado $SU(2)$ y hemos visto como representa las isometrías lineales que conservan la orientación de $\mathbb{C}^2$. Una vez visto esto, hemos visto como es la relación entre $SO(3)$ y $SU(2)$ desde un punto de vista topológico. Primero vimos que $SU(2)$ es homeomorfo a S^3 , y después que $SO(3)$ es homeomorfo a $S^3/\mathbb{Z}_2$, es decir, el espacio proyectivo $\mathbb{P}^3$, y con ello concluimos que $SO(3)$ es homeomorfo a $SU(2)/\mathbb{Z}_2$. Además, la aplicación que nos daba el homeomorfismo nos daba un doble cubrimiento de $SU(2)$ en $SO(3)$. Esto nos permite deducir propiedades topológicas de $SO(3)$ que serían más complicadas de ver como grupo de matrices. Por ejemplo, como es homeomorfo al espacio proyectivo real $\mathbb{P}^3$, se tiene que $SO(3)$ es compacto y conexo. Además, hemos visto que S^3 es recubridor universal de $SO(3)$ por ser S^3 homeomorfo a $SU(2)$, con lo que resulta que $\pi(SO(3)) = \mathbb{Z}_2$.

Por otra parte, hemos introducido los cuaterniones y hemos visto algunas de sus propiedades básicas, para después ver su relación con el grupo especial ortogonal: un cuaternión determina una rotación y por ello una matriz de $SO(3)$. Además, estudiamos cómo un cuaternión define una rotación, y por ello una matriz de $SO(3)$ y viceversa.

Por último, hemos estudiado la estructura de $SO(3)$ como grupo de Lie, para lo cual también hemos visto que es una variedad diferenciable, y también hemos visto su álgebra de Lie $\mathfrak{so}(3)$ compuesta por las matrices antisimétricas de dimensión 3.

Como conclusión, podemos afirmar que el grupo especial ortogonal es un objeto matemático con propiedades muy interesantes, sin importar a través de que rama matemática lo estudiemos. A lo largo de este trabajo hemos visto algunas de estas propiedades, y nos quedan fuera de este otras tantas igualmente interesantes.

Bibliografía

- [1] ALMIRA, J. M., & GONZÁLEZ, P. D.: (2001). Una demostración constructiva de un homeomorfismo entre P^3 y $SO(3)$. *Gaceta de la Real Sociedad Matemática Española*, 4(2), 366-374.
- [2] ARBEA, MIGUEL: *Variedades casi-complejas*. Universidad de Cantabria. Accesible en <https://repositorio.unican.es/xmlui/bitstream/handle/10902/23304/ArbeaGomezMiguel-TFG-Matematicas.pdf?sequence=1&isAllowed=y>
- [3] BRUSCHI, L. (s.f.). Los grupos $SU(2)$, $SO(3)$ y sus representaciones. *Universidad de Buenos Aires*.
- [4] COLE, I.R. (2015). *Modelling CPV* (Doctoral dissertation, Loughborough University).
- [5] ETAYO, FERNANDO: *Variedades Diferenciables.*, (345), Universidad de Cantabria.
- [6] GALLIER, J.: *Geometric Methods and Applications: For Computer Science and Engineering.*, 38, (2004), 281-300.
- [7] *Cuaterniones y Rotación espacial*. (s.f.). Hmong. https://hmong.es/wiki/Quaternions_and_spatial_rotation
- [8] HUGHES, J.(10 de mayo de 2017). *Re: Lie algebra and tangent space at the identity are isomorphic*. [Comentario en el foro Lie algebra and tangent space at the identity are isomorphic]. Math.stackexchange. <https://math.stackexchange.com/questions/2274555/lie-algebra-and-tangent-space-at-the-identity-are-isomorphic>
- [9] JIMENEZ, E.A.R. (2018). Introducción a los grupos de Lie. *Universidad Nacional Autónoma de México*.
- [10] TORRES DEL CASTILLO G.F. (1999). La representación de rotaciones mediante cuaterniones. *Miscelánea matemática*, 29, 43-50. https://miscelaneamatematica.org/download/tbl_articulos.pdf2.b8d4415d314c3f7d.746f727265735f632e706466.pdf
- [11] WARNER, F. W. (1983). *Foundations of differentiable manifolds and Lie groups* (Vol. 94). Springer Science & Business Media.
- [12] WESTRA, D. B. (2008). $Su(2)$ and $so(3)$. *University of Groningen*.