"El saber de mis hijos hará mi grandeza"

UNIVERSIDAD DE SONORA

División de Ciencias Exactas y Naturales

Programa de Licenciado en Matemáticas

El Algebra y la Geometría de los Cuaternios y Algunas de sus Aplicaciones

T E S I S

Que para obtener el título de:

Licenciado en Matemáticas

Presenta:

Jesús Jairo Rodríguez Padilla

Director de Tesis: Dr. Guillermo Dávila Rascón

Hermosillo, Sonora, México, 15 de Octubre, 2010.

SINODALES

Dr. Rubén Flores Espinoza
Universidad de Sonora

Dr. Yuri M. Vorobiev
Universidad de Sonora

Dr. Martín G. García Alvarado
Universidad de Sonora

Dr. Guillermo Dávila Rascón
Universidad de Sonora

Dedicatoria

... a mi papá, Jesús Rodríguez, que sin él no sería lo que soy ni estuviera donde estoy...

Agradecimientos

Estoy muy agradecido con todos mis maestros del Departamento de Matemáticas por sus enseñanzas durante mi estancia en la licenciatura. En especial, con el Dr. Guillermo Dávila Rascón por su gran apoyo, tanto académico como moral, en la realización de este trabajo.

A mis maestros sinodales, Dr. Rubén Flores Espinoza, Dr. Yuri M. Vorobiev y Dr. Martín G. García Alvarado, por sus importantes observaciones que han servido para mejorar sustancialmente este trabajo.

Al Consejo Nacional de Ciencia y Tecnología por apoyarme para la realización de esta tesis como becario del proyecto "Mecanismos de Promedios en Sistemas de Evolución Clásicos y Cuánticos" (Ref. no. 55463).

A mis amigos, ustedes que tantas palabras de aliento me han dado, mismas que me han servido para salir adelante en todo momento ¡Muchas gracias!

Contenido

Introducción

Si bien las ventajas que proporciona el uso de los cuaternios en la navegación aeroespacial han sido aplicadas y estudiadas por la comunidad de la aeronáutica y la astronáutica desde hace bastante tiempo ([4, 6, 9, 11, 16]) , los fundamentos teóricos de los cuaternios y de sus aplicaciones no se estudiaron por la comunidad de ciencias de la computación, específicamente en lo que respecta a la animación y graficación por computadora, sino hasta el importante artículo de Ken Shoemake publicado en 1985, ([20]). La relevancia del artículo de Shoemake es que abordó el concepto de orientación de marcos de referencia para el movimiento de objetos en tres dimensiones y el de la orientación de las cámaras de video en la animación gráfica por computadoras y expuso las deficiencias de los métodos tradicionales utilizados para definir dichas orientaciones, los cuales se basaban en los ángulos de Euler y la cual, era la metodolgía estándar utilizada en ese tiempo. Cabe señalar que el problema de orientación de cámaras en la industria de la animación es de particular relevancia ya que es necesario especificar, en cada momento, orientaciones muy precisas. Shoemake introdujo la representación de esas orientaciones mediante cuaternios como una solución a tales deficiencias. Esto trajo como consecuencia la apertura de una área muy rica de investigación, sobre todo por sus potenciales aplicaciones a la animación por computadoras, la robótica y la simulación del movimiento de cuerpos rígidos.

La herramienta primaria introducida en el artículo de Shoemake fue la fórmula de interpolación para dos puntos en una esfera de dimensión arbitraria conectados por un arco. Shoemake acuñó el término "SLERP", por sus siglas en inglés ("spherical linear interpolation"), para tal interpolación, una terminología aún utilizada.

El objetivo de este trabajo es presentar un estudio sobre los cuaternios, sus propiedades, tanto algebraicas como geométricas, y algunas de sus aplicaciones. En particular, su estrecha relación con las rotaciones en el espacio $\mathbb{R}^3$. Esto último es la motivación principal para este trabajo, pues la meta del mismo es presentar una alternativa para representar rotaciones en $\mathbb{R}^3$ y exhibir la relación existente con la representación de rotaciones mediante matrices.

En los capítulos 1 y 2 se estudian los preliminares que nos darán los fundamentos sobre la teoría subyacente en el estudio de los cuaternios. Específicamente, en el capítulo 1, se estudian conceptos básicos sobre la teoría de curvas en $\mathbb{R}^3$ y se discuten las ventajas y desventajas del aparato de Frenet-Serret. Luego se presentan algunas nociones básicas tales como espacios cubriente y la fibración de Hopf. En esta sección se estudia el espacio proyectivo, el cual nos será de gran utilidad en el capítulo 4. Además, se estudian variedades diferenciables lo cual nos dará las bases para el desarrollo del siguiente capítulo.

Después, en el capítulo 2, se estudian grupos de Lie y sus álgebras de Lie aso-

ciadas, si bien nuestro interés se centrará en estudiar algunas de las propiedades de los grupos de Lie de matrices y sus respectivas álgebras de Lie. Posteriormente, en el capítulo 3 se introducen los cuaternios y se discuten sus propiedades algebraicas así comom distintas representaciones para éstos. Además, se estudia la geometría de los cuaternios unitarios y la conexión de los mismos con el grupo de Lie $SU(2)$.

En el capítulo 4, se presentan algunos resultados importantes con relación a las rotaciones en el espacio tridimensional, específicamente el Teorema de Euler, del cual se proporciona una demostración utilizando herramientas básicas del álgebra lineal. Asimismo se introduce la fórmula de Rodrigues, la cual nos proporciona una manera efectiva para calcular rotaciones de vectores alrededor de un eje por un ángulo dado y se deriva esta fórmula usando conceptos básicos de la geometría de $\mathbb{R}^3$. Posteriormente, se estudia de manera detallada la relación que existe entre rotaciones en $\mathbb{R}^3$ y los cuaternios unitarios, así como también el problema de representar una rotación, por medio de un producto cuaternios, a partir de la matriz asociada con esa rotación. Se discute también en problema inverso

En el capítulo 5, se estudian marcos de referencia, en concreto los marcos de Frenet-Serret y el marco de transporte paralelo. Posteriormente se introduce una formulación equivalente de estos por medio de cuaternios unitarios. Finalmente, en el capítulo 6 se discuten algunas de las aplicaciones de los cuaternios. Se presenta la implementación de una rotación de un objeto en $\mathbb{R}^3$ mediante el uso de cuaternios y se anexa el código, implementado en el programa de cálculo simbólico Maple en el apéndice. Por otra parte, un problema que surge en la aeronáutica cuando se utilizan ángulos de Euler para representar las rotaciones de un giroscopio,se discute cómo con el uso de cuaternios es posible evitar este problema. Por último, se discute el llamado "truco del cinturón" y cómo se puede explicar fácilmente lo que ocurre en este peculiar pasatiempo, utilizando los cuaternios, a la vez que nos proporciona una realización clara de los cuaternios en la vida cotidiana.

Capítulo 1

Preliminares

En este capítulo se presenta el material básico que nos servirá en los desarrollos de los capítulos posteriores. Concretamente, en este capítulo se estudian conceptos básicos sobre espacios euclidianos y curvas en $\mathbb{R}^3$. Después se estudian nociones básicas sobre espacios cubriente. Por último, se consideran variedades diferenciables.

1.1 Espacios Euclidianos

En esta sección se introducen los espacios euclidianos, los cuales serán de gran utilidad en capítulos posteriores, para explicar cuál es la relación existente entre los cuaternios y el espacio vectorial $\mathbb{R}^4$.

Definición 1.1 *Un espacio euclidiano es un espacio vectorial E junto con una forma bilineal simétrica $\varphi : E \times E \to \mathbb{R}$ que satisface lo siguiente:*

$$\begin{aligned}
\varphi(u_1 + u_2, v) &= \varphi(u_1, v) + \varphi(u_2, v),\\
\varphi(u, v_1 + v_2) &= \varphi(u, v_1) + \varphi(u, v_2),\\
\varphi(\lambda u, v) &= \lambda\varphi(u, v),\\
\varphi(u, \lambda v) &= \lambda\varphi(u, v),\\
\varphi(u, v) &= \varphi(v, u),\\
u \neq 0 &\ \textit{implica}\ \varphi(u, u) > 0.
\end{aligned}$$

El número real $\varphi(u, v)$ se llama *producto interior* de u con v. Definimos también la forma cuadrática asociada con φ como la función $\Phi : E \to \mathbb{R}_+$ por

$$\Phi(u) = \varphi(u, u),$$

para todo $u \in E$. También denotamos $\varphi(u, v)$ por $u \cdot v$ o $\langle u, v\rangle$.

La norma inducida por φ se llama la *norma euclidiana* inducida por φ.

Un producto interior en un espacio vectorial nos permite definir la noción de ortogonalidad. A continuación definimos tal concepto.

Definición 1.2 *Dado un espacio euclidiano E, cualesquiera dos vectores $u, v \in E$ son ortogonales si $u \cdot v = 0$. Dada una familia $(u_i)_{i\in I}$ de vectores en E, decimos que la familia $(u_i)_{i\in I}$ es ortogonal si $u_i \cdot u_j = 0$ para toda $i, j \in I$, donde $i \neq j$. Decimos*

que la familia $(u_i)_{i\in I}$ *es ortonormal si* $u_i \cdot u_j = 0$ *para toda* $i, j \in I$*, y* $u_i \cdot u_i = 1$*, para toda* $i \in I$*. Para cualquier subconjunto* F *de* E*, el conjunto*

$$F^{\perp} = \{v \in E \mid u \cdot v = 0, \text{ para todo } u \in F\},$$

se llama complemento ortogonal de F*.*

Ahora pasamos a definir funciones entre espacios euclidianos que preservan la norma. Estas transformaciones, conocidas como *movimientos rígidos*, juegan un papel muy importante en geometría.

Definición 1.3 *Dados dos espacios euclidianos* E *y* F *de la misma dimensión* n*, una función* $f : E \to F$ *es una transformación ortogonal(o isometría lineal), si es lineal y si*

$$\|f(u)\| = \|u\|,$$

para todo $u \in E$*.*

Definición 1.4 *Dado un espacio euclidiano* E *de dimensión* n*, el conjunto de isometrías* $f : E \to E$ *forma un subgrupo de* $GL(E)$ *denotado por* $O(E)$*, o* $O(n)$ *cuando* $E = \mathbb{R}^n$*, llamado el grupo ortogonal. Para cada isometría* f*, se tiene que* $\det f = \pm 1$*, donde* $\det f$ *denota el determinante de* f*. Las isometrías cuyo determinante es* 1 *son llamadas rotaciones o isometrías propias, y forman un subgrupo del grupo especial lineal* $SL(E)$*, denotado por* $SO(E)$*,* $SO(n)$ *cuando* $E = \mathbb{R}^n$*, llamado el grupo especial ortogonal. Las isometrías cuyo determinante es* -1 *son llamadas isometrías impropias.*

1.2 Curvas en $\mathbb{R}^3$

Una *curva* parametrizada en $\mathbb{R}^3$ es una función continua $\alpha : I \to \mathbb{R}^3$ donde I es un intervalo en la recta real $\mathbb{R}$. Escribimos, para $t \in I$

$$\alpha(t) = \left(\alpha^1(t), \alpha^2(t), \alpha^3(t)\right),$$

donde $\alpha^i(t) : I \to \mathbb{R}$. Decimos que α es *diferenciable* o *suave* si cada función coordenada $\alpha^i(t)$ es diferenciable como una función ordinaria de variable real. El vector *velocidad* de α en t_0, el cual es un vector tangente a la curva en el punto $\alpha(t_0)$, se define por

$$\alpha'(t_0) = \left(\frac{\mathrm{d}\alpha^1}{\mathrm{d}t}\Big|_{t_0}, \frac{\mathrm{d}\alpha^2}{\mathrm{d}t}\Big|_{t_0}, \frac{\mathrm{d}\alpha^3}{\mathrm{d}t}\Big|_{t_0}\right)$$

y la *rapidez* del vector velocidad se define por

$$\|\alpha'(t)\| = \sqrt{(\mathrm{d}\alpha^1/\mathrm{d}t)^2 + (\mathrm{d}\alpha^2/\mathrm{d}t)^2 + (\mathrm{d}\alpha^3/\mathrm{d}t)^2}.$$

El vector *aceleración* de $\alpha(t)$ de define por

$$\alpha''(t) = \left(\frac{\mathrm{d}^2}{\mathrm{d}t^2}\alpha^1(t), \frac{\mathrm{d}^2}{\mathrm{d}t^2}\alpha^2(t), \frac{\mathrm{d}^2}{\mathrm{d}t^2}\alpha^3(t)\right).$$

Una curva $\alpha(t)$ es *regular* si $\alpha'(t) \neq 0$ para todo $t \in I$. Usualmente las curvas pueden tener puntos para los que α' se anula o no existe. En tales casos nos restringiremos a los subintervalos en los que la derivada sea distinta de cero y de tal forma tendremos una curva regular por pedazos.

Sea $\alpha : I \to \mathbb{R}^3$, donde $I \subset \mathbb{R}$, una curva parametrizada. Se dice que una curva parametrizada $\beta : J \to \mathbb{R}^3, J \subset \mathbb{R}$, es una *reparametrización* de la curva α si existe una función suave $h : J \to I$ tal que $\beta(s) = \alpha(h(s))$. Lo anterior lo escribimos también como $\beta \circ h$. Notemos que en este caso las imágenes de α y β son el mismo conjunto en $\mathbb{R}^3$ pero cada punto es alcanzado para valores distintos de los parámetros s y t.

Reparametrización por longitud de arco En la mayoría de los casos es conveniente expresar el parámetro de la curva como la distancia que se ha recorrido desde donde inicia la curva. Esto es, cuando escribimos $\alpha(s)$, el parámetro s es exactamente la distancia que hemos viajado a lo largo de la curva. A esto se le llama *parametrización por longitud de arco*. Si una curva α es regular entonces puede ser reparametrizada para que sea de rapidez unitaria. Para probar este hecho procedemos como sigue. Primero definimos la función longitud de arco por

$$s(t) = \int_a^t \|\alpha'(u)\| \mathrm{d}u.$$

Como α es regular, tenemos

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \|\alpha'(t)\| > 0.$$

Por el teorema del valor medio, s es creciente en I y así es uno a uno. Por lo tanto, s tiene inversa, la cual denotaremos por $t(s)$, y sus respectivas derivadas están inversamente relacionadas,

$$\frac{\mathrm{d}t}{\mathrm{d}s}(s) = \frac{1}{\frac{\mathrm{d}s}{\mathrm{d}t}(t(s))} > 0.$$

Sea $\beta(s) = \alpha(t(s))$. Luego, $\beta'(s) = \alpha'(t(s))\frac{\mathrm{d}t}{\mathrm{d}s}(s)$. De esta manera tenemos

$$\begin{aligned}
\|\beta(s)\| &= \|\alpha'(t(s))\| \|\frac{\mathrm{d}t}{\mathrm{d}s}(s)\| \\
&= \frac{\mathrm{d}s}{\mathrm{d}t}(t(s))\frac{\mathrm{d}t}{\mathrm{d}s}(s) \\
&= \frac{\mathrm{d}s}{\mathrm{d}t}(t(s))\frac{1}{\frac{\mathrm{d}s}{\mathrm{d}t}(t(s))} \\
&= 1.
\end{aligned}$$

Supongamos que la curva β está definida en el intervalo $[0, 1]$. Consideremos la longitud de arco de la reparametrización β para un valor s_0,

$$L(s_0) = \int_0^{s_0} \|\beta'(s)\| \mathrm{d}s = \int_0^{s_0} 1 \mathrm{d}s = s_0.$$

Por lo tanto, una curva está parametrizada por longitud de arco exactamente cuando es de rapidez unitaria.

El Aparato de Frenet-Serret para curvas de rapidez unitaria Supongamos que la curva β es de rapidez unitaria ($\|\beta'\| = 1$) y que β está parametrizada por longitud de arco. Los cambios de dirección de la curva son descritos por lo que se conocen como *fórmulas de Frenet*. Denotemos por $T = \beta'$ al vector tangente unitario a la curva β. Como $\|T\| = 1$, la magnitud de T' indica la taza de cambio en la dirección de T. Notemos que $\|T\| = \sqrt{T \cdot T} = 1$, por lo que $T \cdot T = 1$. De aquí obtenemos que $0 = T \cdot T'$ y, por lo tanto, T' es perpendicular a T. Decimos que T' es *normal* a β.

Definimos la función *curvatura* de β por

$$\kappa(s) = \|T'(s)\|.$$

En el caso cuando $\kappa > 0$ definimos el vector *normal* sobre la curva β como el único vector $N(s)$ que satisface

$$N = \frac{1}{\kappa} T'.$$

Se sigue directamente que $\|N\| = 1$. Además T y N son ortogonales.

En estas condiciones definimos el vector *binormal* a la curva β por medio de la relación $B = T \times N$. Notemos que $\|B\| = \|T\|\|N\| \operatorname{sen} \frac{\pi}{2} = 1$, por lo que B es un vector unitario. Además, B es perpendicular a T y a N.

El conjunto $\{T, N, B\}$ se llama el *aparato de Frenet-Serret* de la curva β. La variación de T,N y B conforme nos movemos sobre β nos dice cuál es el comportamiento de la curva, es decir, cómo es que β cambia en las distintas direcciones en el espacio tridimensional.

Tal variación será determinada por las derivadas T', N' y B'. Sabemos que $T' = \kappa N$, por lo que sólo nos resta calcular N' y B'. Ahora, como T, N y B son ortonormales, cualquier vector en $\mathbb{R}^3$ se puede expresar como combinación lineal de ellos. En particular $B' = aT + bN + cB$. Para calcular a, b y c procedemos como sigue:

$$\begin{aligned} T \cdot B' &= aT \cdot T + bT \cdot N + cT \cdot B \\ &= a \cdot 1 + b \cdot 0 + c \cdot 0 \\ &= a. \end{aligned}$$

Similarmente, $N \cdot B' = b$ y $B \cdot B' = c$. Por lo tanto,

$$B' = (T \cdot B')T + (N \cdot B')N + (B \cdot B')B.$$

Dado que $T \cdot B = 0$, entonces $0 = T' \cdot B + T \cdot B'$ y utilizando $N \cdot B = 0$ obtenemos

$$\begin{aligned} T \cdot B' &= -T' \cdot B \\ &= -\kappa N \cdot B \\ &= 0. \end{aligned}$$

Con un razonamiento análogo, obtenemos $B \cdot B' = 0$. De esta manera sólo nos queda un posible término distinto de cero en la expresión para B'. Definimos $\tau = -N \cdot B'$ como la *torsión* de la curva β. Así

$$B' = -\tau N.$$

Para calcular N' procedemos de manera análoga. Como T, N y B forman una base ortonormal, podemos expresar a N' como combinación lineal de ellos, es decir, $N' = aT + bN + cB$. Luego,

$$\begin{aligned} T \cdot N' &= aT \cdot T + bT \cdot N + cT \cdot B \\ &= a \cdot 1 + b \cdot 0 + c \cdot 0 \\ &= a. \end{aligned}$$

Similarmente, $N \cdot N' = b$ y $B \cdot N' = c$. Por lo tanto,

$$N' = (T \cdot N')T + (N \cdot N')N + (B \cdot N')B.$$

Por medio de un cálculo directo se obtiene $T \cdot N' = -\kappa$, $N \cdot N' = 0$ y $B \cdot N' = \tau$, con lo cual concluimos que

$$N' = -\kappa T + \tau B.$$

Todo lo anterior se resume en el siguiente resultado (ver [19]):

Teorema 1.1 ***(Las fórmulas de Frenet-Serret)****. Para una curva parametrizada por longitud de arco con curvatura $\kappa > 0$, se tiene que*

$$\begin{aligned} T' &= \kappa N, \\ N' &= -\kappa T + \tau B, \\ B' &= -\tau N. \end{aligned} \tag{1.1}$$

El vector de Darboux Supongamos que una partícula se mueve a lo largo de una trayectoria dada por una curva $\alpha(s)$ de rapidez unitaria, entonces el movimiento de la partícula consta de traslaciones y rotaciones sobre α. La rotación está determinada por un vector de velocidad angular ω el cual satisface $T' = \omega \times T$, $N' = \omega \times N$ y $B' = \omega \times B$. El vector ω se llama el vector de *Darboux* y está dado por

$$\omega = \tau T + \kappa B.$$

Para obtener esta fórmula primeramente recordemos que $\{T, N, B\}$ es una base ortonormal, por lo que el vector de Darboux, ω, se puede escribir como combinación lineal de ellos, esto es

$$\omega = aT + bN + cB.$$

De esta manera, por medio de un cálculo directo se obtienen las siguientes fórmulas:

$$\begin{aligned} \omega \times T &= aT \times T + bN \times T + cB \times T \\ &= b(-B) + cN \\ T' &= -bB + cN \end{aligned} \tag{1.2}$$

$$\begin{aligned} \omega \times N &= aT \times N + bN \times N + cB \times N \\ &= aB + c(-T) \\ N' &= aB - cT \end{aligned} \tag{1.3}$$

$$\begin{aligned} \omega \times T &= aT \times B + bN \times B + cB \times B \\ &= a(-N) + bT \\ B' &= -aN + bT \end{aligned} \tag{1.4}$$

Del Teorema 1.1 y la ecuación (1.3) obtenemos

$$a = \tau, \qquad c = \kappa.$$

Usando nuevamente el Teorema 1.1 junto con la ecuación (1.2) se llega a

$$b = 0,$$

de lo cual se tiene que

$$\omega = \tau T + \kappa B,$$

como se quería.

Aparato de Frenet-Serret para curvas de rapidez arbitraria Todas las curvas regulares tienen parametrizaciones de rapidez unitaria, pero en algunos casos no es posible encontrar tal parametrización explícitamente. De esta manera, para entender la geometría de estas curvas es necesario modificar las fórmulas de Frenet.

Sea $\alpha(t)$ una curva parametrizada con rapidez $r = \|\alpha'(t)\| = \mathrm{d}s/\mathrm{d}t$. Podemos reparametrizar la curva α para obtener una curva de rapidez unitaria $\widehat{\alpha}(s(t))$, y de esta manera definir la curvatura y torsión de α en términos de su reparametrización por longitud de arco $\widehat{\alpha}(s(t))$. Bajo estas condiciones tenemos el siguiente resultado (ver [19]):

Teorema 1.2 ***(Fórmulas de Frenet-Serret para curvas de rapidez arbitraria.)*** *Sea $\alpha(t)$ una curva parametrizada regular de rapidez $r = \mathrm{d}\alpha/\mathrm{d}t$ y curvatura $\kappa > 0$.Entonces*

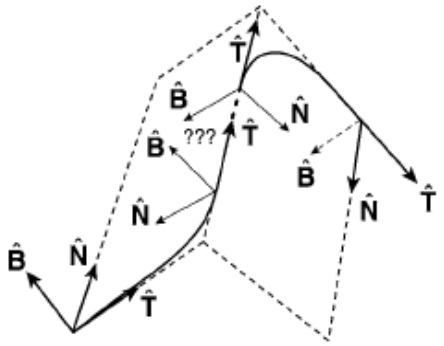
$$\begin{aligned} T' &= \kappa r N, \\ N' &= -\kappa r T + \tau r B, \\ B' &= -\tau r N. \end{aligned}$$

Para una curva de rapidez arbitraria tenemos que $T = \alpha'/|\alpha'|$,$N = T'/|T'|$, $B = T \times N$ y $N = B \times T$. En estas condiciones se tiene el siguiente resultado(ver [19]):

Teorema 1.3 *Para cualquier curva regular α, las siguientes fórmulas se satisfacen*

$$\begin{aligned} B &= \frac{\alpha' \times \alpha''}{\|\alpha' \times \alpha''\|} \\ \kappa &= \frac{\|\alpha' \times \alpha''\|}{\|\alpha'\|^3} \\ \tau &= \frac{(\alpha' \times \alpha'') \times \alpha'''}{\|\alpha' \times \alpha''\|^2} \end{aligned}$$

Al calcular el aparato de Frenet-Serret en un punto de una curva parametrizada $\alpha(t)$ obtenemos un marco ortonormal en dicho punto. Una propiedad característica del marco de Frenet-Serret es la siguiente: si una curva parametrizada, $\alpha(t), t \in [0, L]$, es una curva cerrada ($\alpha(0) = \alpha(L)$) y tiene primera y segunda derivada continuas en cada punto de $[0, L]$, entonces el marco es continuo. Sin embargo, no siempre es utilizado, pues presenta anormalidades cuando la curva en cuestión tiene partes rectas o puntos de inflexión. Esto es debido a que la curvatura de la curva se anula y de esta manera es imposible calcular el vector normal a la curva N. Aún más, cuando la curvatura se anula en un punto t_0, el aparato de Frenet-Serret antes de t_0 y después de t_0 tiene direcciones opuestas. Se presenta una figura a continuación donde se muestra este problema:



Una conclusión inmediata de lo anterior es que al utilizar el marco de Frenet-Serret para calcular marcos ortonormales en puntos sobre una curva α , automáticamente se le imponen restricciones a la curva en cuestión, pues para no tener problemas se necesita que la curva tenga primera y segunda derivadas continua.

Enseguida presentamos una generalización del marco de Frenet([15]), la cual no tiene problema en la mayoría de los casos donde la curvatura se anula. Antes de continuar, diremos que una curva α es plana si está totalmente contenida en un plano. Esta generalización usa un marco un poco distinto para curvas regulares planas. Tenemos dos razones por las cuales tratamos curvas planas y no planas por separado. Primero, el marco modificado para curvas planas se define inclusive en puntos donde la curvatura es cero; y segundo, es más fácil de calcular que el marco de Frenet, específicamente en casos donde la curva no está parametrizada por longitud de arco.

Consideremos una curva parametrizada $\alpha(t)$ plana. Podemos redefinir los vectores normal y binormal del marco de Frenet de la siguiente manera: fijamos un vector normal $\mathbf{m}$ al plano que contiene a la curva α y tomamos $B = \mathbf{m}$ en cada punto $\alpha(t)$ como uno de los vectores ortogonales del marco. También utilizamos el vector tangente T y completamos el marco mediante la relación $N = B \times T$. El marco modificado difiere del marco de Frenet en que está definido en puntos donde la curvatura es cero y que el vector normal N está siempre en el mismo lado de la curva como se puede apreciar en la Figura (1.1).

De esta manera es posible utilizar este marco en curvas que tienen sólo su primera

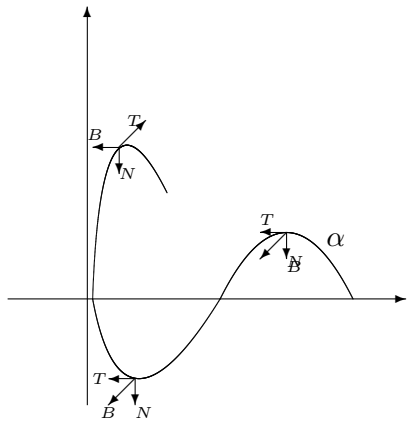


Figura 1.1: El marco de Frenet modificado.

derivada continua y consisten de varios segmentos planos.

1.3 Espacios cubriente

En nuestro trabajo todos los espacios topológicos tendrán la *propiedad de Hausdorff*. Recordemos que un espacio topológico X es de *Hausdorff* si para cualesquiera dos puntos $x, y \in X$, existen abiertos U, V tales que $x \in U$, $y \in V$ y $U \cap V = \emptyset$.

Sea X un espacio topológico. X es un espacio *localmente conexo* si para cada $x \in X$ y cualquier vecindad U de x, existe un subconjunto conexo $V \subset X$ tal que $x \in V \subset U$.

X es un espacio *localmente arco-conexo* si para cada $x \in X$ y cualquier vecindad U de x, existe $V \subset X$ tal que $x \in V \subset U$ y para cualquier par de puntos $y, z \in V$, se pueden unir por una trayectoria totalmente contenida en V.

Definición 1.5 *Sea X un espacio topológico. Un espacio topológico $\widetilde{X}$ se dice ser un* ***espacio cubriente*** *o* ***una cubierta*** *de X si existe una función continua $p : \widetilde{X} \to X$ que cumple*

1. *p es sobreyectiva,*
2. *para cada $x \in X$, existe $U \subset X$ abierto, tal que $p^{-1}(U)$ es unión ajena de conjuntos abiertos de $\widetilde{X}$ y cada uno de estos conjuntos abiertos de $\widetilde{X}$ es homeomorfo bajo p al conjunto U.*

Un ejemplo sencillo que nos ilustra lo anterior es el siguiente. Consideremos la función $f : \mathbb{R} \to \mathbb{S}^1$, donde $\mathbb{S}^1 = \{z \in \mathbb{C} \mid |z| = 1\}$, definida por

$$f(t) = e^{2\pi i t}.$$

La función f es sobreyectiva, pues si $z_0 = (\cos\theta, \operatorname{sen}\theta) \in \mathbb{S}^1$ basta con tomar $t = \left(\frac{\theta}{2\pi}\right)$ y así $f(t) = e^{i\theta} = z_0$. Notemos también que para cada $k \in \mathbb{Z}$, la imagen del conjunto $[k, k+1]$ bajo la función f es $\mathbb{S}^1$ y de esta manera f no es inyectiva. Consideremos ahora el conjunto $U \subset \mathbb{S}^1$ dado por

$$U = \left\{(1,\theta) \mid 0 < \theta < \frac{\pi}{4}\right\} = \left\{z \in \mathbb{C} \mid z = e^{2\pi i\theta}, 0 < \theta < \frac{\pi}{4}\right\}.$$

Es claro que la imagen inversa de U es el conjunto

$$f^{-1}(U) = \bigcup_{k\in\mathbb{Z}} V_k,$$

donde $V_k = (k, k+\frac{1}{8})$. Además, para cada $k \in \mathbb{Z}$, la imagen del conjunto $[k, k+\frac{1}{8}]$ es U. Así, vemos que $\mathbb{R}$ cubre a $\mathbb{S}^1$.

Ejemplo: La Fibración de Hopf. La $n-$esfera unitaria $\mathbb{S}^n$ es el conjunto de puntos $(x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1}$ que satisfacen la ecuación

$$x_0^2 + x_1^2 + \cdots + x_n^2 = 1,$$

es decir, $\mathbb{S}^n$ es el conjunto de puntos en $\mathbb{R}^{n+1}$ cuya distancia al origen es 1.

Una cubierta especial de la esfera $\mathbb{S}^2 \subset \mathbb{R}^3$ por $\mathbb{S}^3 \subset \mathbb{R}^4$ es la llamada *fibración de Hopf* (ver [17]) la cual se define por una función $h : \mathbb{S}^3 \to \mathbb{S}^2$ dada por

$$h(a,b,c,d) = (a^2+b^2-c^2-d^2, 2(ad+bc), 2(bd-ac)). \tag{1.5}$$

Al calcular el cuadrado de cada coordenada en el lado derecho de (1.5) obtenemos

$$\begin{aligned}(a^2+b^2-c^2-d^2)^2 &= ((a^2+b^2)-(c^2+d^2))^2\\ &= a^4+2a^2b^2+b^4-2(a^2c^2+a^2d^2+b^2c^2+b^2d^2)\\ &\quad + c^4+2c^2d^2+d^4,\end{aligned} \tag{1.6}$$

$$(2(ad+bc))^2 = 4(a^2d^2+2abcd+b^2c^2), \tag{1.7}$$

$$(2(bd-ac))^2 = 4(b^2d^2-2abcd+a^2c^2), \tag{1.8}$$

sumando (1.6),(1.7) y (1.8) se llega a

$$a^4+2a^2b^2+b^4+2(a^2c^2+a^2d^2+b^2c^2+b^2d^2)+c^4+2c^2d^2+d^4,$$

que se puede reescribir como

$$((a^2+b^2)+(c^2+d^2))^2 = (a^2+b^2+c^2+d^2)^2 = 1.$$

De esta manera, la imagen de cualquier punto $x \in \mathbb{S}^3 \subset \mathbb{R}^4$ bajo h es un punto en $\mathbb{S}^2 \subset \mathbb{R}^3$. Calculamos ahora la imagen del punto $(-a,-b,-c,-d)$, es decir,

$$\begin{aligned}h(-a,-b,-c,-d) &= ((-a)^2+(-b)^2-(-c)^2-(-d)^2, 2((-a)(-d)+(-b)(-c)),\\ &\quad 2((-b)(-d)-(-a)(-c)))\\ &= (a^2+b^2-c^2-d^2, 2(ad+bc), 2(bd-ac))\\ &= h(a,b,c,d),\end{aligned} \tag{1.9}$$

de manera que un punto y su antípoda son transformados en el mismo punto mediante la fibración de Hopf, por lo que la imagen del hemisferio $U_1 = \{(a,b,c,d) \in \mathbb{S}^3 \mid d > 0\}$ y la del hemisferio $U_2 = \{(a,b,c,d) \in \mathbb{S}^3 \mid d < 0\}$ es, en cada caso, todo $\mathbb{S}^2$. Por lo tanto $\mathbb{S}^3$ es espacio cubriente de $\mathbb{S}^2$.

1.4 Variedades Diferenciables

Podemos decir que el concepto de variedad diferenciable es una generalización de los conceptos de curva y superficie. Una curva es un espacio topológico que localmente se ve como una recta, mientras que una superficie es un espacio topológico que localmente se ve como un plano. Procedemos a la siguiente definición.

Definición 1.6 *Una variedad (topológica), M, de dimensión n, es un espacio topológico Hausdorff que tiene una base numerable y para cada $x \in M$ existe un abierto $U \subset M$ tal que $x \in U$ y U es homeomorfo a un abierto en $\mathbb{R}^n$.*

De esta definición, es claro que si M es una variedad topológica de dimensión n, entonces para cada punto $p \in M$ existe al menos un abierto U tal que $p \in U$ y un homeomorfismo $\varphi : U \to V \subset \mathbb{R}^n$ sobre un abierto V de $\mathbb{R}^n$. A la pareja (U, φ) se le llama *vecindad coordenada* de p, o *carta coordenada*, ya que a cada punto $q \in U$ podemos asignarle las n coordenadas de su imagen bajo φ. Si además, p pertenece a otra carta coordenada (V, ψ), entonces podemos definir el homeomorfismo

$$\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V),$$

el cual establece un cambio de coordenadas de los puntos de $U \cap V$, de manera que podemos obtener fórmulas para cambio de coordenadas de puntos en la intersección de dos cartas coordenadas. A estas funciones se les llama *funciones de transición*. Notemos que éstas son funciones entre subconjuntos de $\mathbb{R}^n$, por lo cual podemos considerar la diferenciabilidad de las mismas.

Definición 1.7 *Diremos que las cartas coordenadas (U, φ) y (V, ψ) en una variedad topológica M de dimensión n son C^∞-compatibles si $U \cap V \neq \emptyset$ implica que las funciones de transición $\varphi \circ \psi^{-1}$ y $\psi \circ \varphi^{-1}$ son funciones de clase C^∞.*

Definición 1.8 *Una estructura diferenciable (o estructura C^∞) en una variedad topológica M de dimensión n, es una familia $\mathcal{U} = \{(U_\alpha, \varphi_\alpha)\}_{\alpha \in A}$ de cartas coordenadas tales que*

1. *La colección $\{U_\alpha\}_{\alpha \in A}$ es una cubierta abierta para M.*
2. *Dados $\alpha, \beta \in A$, las cartas coordenadas $(U_\alpha, \varphi_\alpha)$ y (U_β, φ_β) son C^∞-compatibles.*
3. *Si (V, ψ) es C^∞-compatible con $(U_\alpha, \varphi_\alpha)$ para todo $\alpha \in A$, entonces (V, ψ) pertenece a $\mathcal{U}$.*

Si la colección $\mathcal{U}$ satisface las dos primeras condiciones diremos que $\mathcal{U}$ es un *atlas coordenado* de M.

Definición 1.9 *Una variedad diferenciable (o variedad suave) es una variedad topológica M de dimensión n, con una estructura diferenciable. Diremos en este caso que M es una variedad diferenciable de dimensión n.*

Un ejemplo sencillo de una variedad suave es la esfera unitaria $\mathbb{S}^2 \subset \mathbb{R}^3$ definida por

$$\mathbb{S}^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}.$$

$\mathbb{S}^2$ es una variedad suave $2-$dimensional ya que si consideremos los conjuntos

$$U_i^+ = \{(x_1, x_2, x_3) \in \mathbb{S}^2 : x_i > 0\}, \qquad U_i^- = \{(x_1, x_2, x_3) \in \mathbb{S}^2 : x_i < 0\},$$

para cada $i \in \{1, 2, 3\}$, se tiene que la colección $\mathcal{U} = \{U_i^+, U_i^-\}_{i=1}^3$ es una cubierta para $\mathbb{S}^2$. Luego, las proyecciones $\varphi_i^\pm : U_i^\pm \to \mathbb{R}^2$, definidas de la siguiente manera:

$$\begin{aligned}
\varphi_1^\pm(x_1, x_2, x_3) &= (x_2, x_3),\\
\varphi_2^\pm(x_1, x_2, x_3) &= (x_1, x_3),\\
\varphi_3^\pm(x_1, x_2, x_3) &= (x_1, x_2),
\end{aligned}$$

para cada $i = 1, 2, 3$, son homeomorfismos de U_i sobre el conjunto $\{(y_1, y_2) \in \mathbb{R}^2 \mid y_1^2 + y_2^2 < 1\}$ y, además, son C^∞-compatibles. Por lo tanto, $\mathbb{S}^2$ es una variedad de dimensión 2 y $(U_i^\pm, \varphi_i^\pm)$ es una carta coordenada para $i = 1, 2, 3$.

Si $(M, \{(U_\alpha, \varphi_\alpha)\}_{\alpha\in\Lambda_1})$ y $(N, \{(V_\beta, \psi_\beta)\}_{\beta\in\Lambda_2})$ son dos variedades diferenciables de dimensiones m y n respectivamente, diremos que una función $f : M \to N$ es *diferenciable* si $\psi_\beta \circ f \circ \varphi_\alpha^{-1}$ es diferenciable, como función de $\mathbb{R}^m$ a $\mathbb{R}^n$, para todo (U_α, V_β).

Si $f : M \to N$ es una función diferenciable y $p \in U_\alpha \subset M$, el *rango* de f en el punto p se define como el rango de la matriz Jacobiana de la función $\psi_\beta \circ f \circ \varphi_\alpha^{-1}$ y lo denotaremos por $\text{rango}_p(f)$. Es claro que $\text{rango}_p(f)$ no depende de la carta coordenada que se tome.

Consideremos una función diferenciable $f : M \to N$. Diremos que:

- f es *inmersión* si $\dim(M) \leq \dim(N)$ y $\text{rango}_p(f) = \dim(M)$ para todo $p \in M$.
- f es *submersión* si $\dim(M) \geq \dim(N)$ y $\text{rango}_p(f) = \dim(N)$ para todo $p \in M$.
- f es un *encajamiento* si
  1. f es inmersión,
  2. f es homeomorfismo sobre su imagen $f(M)$

En estas condiciones, decimos que M es una *subvariedad* de N si existe un encajamiento $f : M \to N$.

Espacio tangente. Dos conceptos fundamentales en el estudio de variedades diferenciales son los de vector tangente y de espacio tangente. En el caso de curvas o superficies en $\mathbb{R}^3$, estos conceptos son bastante intuitivos y nuestra intención es generalizarlos para el caso de variedades.

Consideremos a $\mathbb{R}^n$ como variedad de dimensión n y sea $x \in \mathbb{R}^n$ un punto arbitrario. Sea $\gamma : \Delta \to \mathbb{R}^n$, la curva paramertrizada definida por

$$\gamma(t) = (\gamma_1(t), \gamma_2(t), \ldots, \gamma_n(t)),$$

donde $0 \in \Delta \subset \mathbb{R}$ y $\gamma(0) = x$.

El *vector tangente* $\mathbf{u}$ a la curva γ en el punto $x = \gamma(0)$ es el vector

$$\mathbf{u} = \frac{\mathrm{d}}{\mathrm{d}t}\gamma(0) = \left(\frac{\mathrm{d}\gamma_1}{\mathrm{d}t}(0), \frac{\mathrm{d}\gamma_2}{\mathrm{d}t}(0), \ldots, \frac{\mathrm{d}\gamma_n}{\mathrm{d}t}(0)\right) \in \mathbb{R}^n.$$

Si γ_1 y γ_2 son dos curvas parametrizadas, en $\mathbb{R}^n$, con $\gamma_1(0) = \gamma_2(0)$, diremos que γ_1 *está relacionada con* γ_2 sí y sólo si

$$\frac{\mathrm{d}\gamma_1}{\mathrm{d}t}(0) = \frac{\mathrm{d}\gamma_2}{\mathrm{d}t}(0).$$

Esto lo denotaremos por $\gamma_1 \sim \gamma_2$.

Es claro que la relación $\sim$ es una relación de equivalencia en el conjunto de todas las curvas en $\mathbb{R}^n$ que coinciden en el punto $x \in \mathbb{R}^n$.

Consideremos ahora una variedad suave M de dimensión n y sea $p \in M$. Sean $\gamma, \tilde{\gamma} : \Delta \to M$, donde $0 \in \Delta \subset \mathbb{R}$, y (V, φ) una carta, con $p \in V$, $\gamma(0) = p = \tilde{\gamma}(0)$ y $x = \varphi(p)$. Notemos que

$$\begin{aligned} \varphi \circ \gamma &: \Delta \to \mathbb{R}^n, \\ \varphi \circ \tilde{\gamma} &: \Delta \to \mathbb{R}^n. \end{aligned}$$

Diremos que γ está relacionada con $\tilde{\gamma}$, $\gamma \sim \tilde{\gamma}$, sí y sólo si

$$\frac{\mathrm{d}(\varphi \circ \gamma)}{\mathrm{d}t}(0) = \frac{\mathrm{d}(\varphi \circ \tilde{\gamma})}{\mathrm{d}t}(0). \tag{1.10}$$

Esta definición es independiente de las cartas. En efecto, sea $(\tilde{V}, \tilde{\varphi})$ otra carta con $p \in \tilde{V}$, $(p \in V \cap \tilde{V} \neq \emptyset)$. Luego, $\varphi \circ \gamma$ y $\tilde{\varphi} \circ \gamma$ son dos curvas en M. Así,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\tilde{\varphi} \circ \gamma)(0) = \frac{\partial \tilde{\varphi} \circ \varphi^{-1}}{\partial x}(\varphi(m))\frac{\mathrm{d}}{\mathrm{d}t}(\varphi \circ \gamma)(0).$$

Por otro lado, también se tiene que

$$\frac{\mathrm{d}}{\mathrm{d}t}(\tilde{\varphi} \circ \tilde{\gamma})(0) = \frac{\partial \tilde{\varphi} \circ \varphi^{-1}}{\partial x}(\varphi(m))\frac{\mathrm{d}}{\mathrm{d}t}(\varphi \circ \tilde{\gamma})(0).$$

En estas condiciones, la relación $\sim$ es una relación de equivalencia en el conjunto de curvas en M que coinciden en el punto $p \in M$. Se tiene así la definición siguiente.

Definición 1.10 *Un vector tangente en el punto $p \in M$ es una clase de equivalencia, $[\gamma]$, de curvas que coinciden en p y que satisfacen (1.10). Definimos el espacio tangente en el punto $p \in M$ por*

$$\mathrm{T}_p M = \{[\gamma] \mid \gamma : \Delta \to M, \gamma(0) = p\}.$$

El espacio tangente $\mathrm{T}_p M$ es un espacio vectorial y $\dim(\mathrm{T}_p M) = n$. En efecto, si $[\gamma_1], [\gamma_2] \in \mathrm{T}_m M$, entonces

$$[\gamma_1] + [\gamma_2] = [\gamma],$$

donde $[\gamma]$ viene a ser la curva que satisface: existe una carta (V, φ) y γ con la propiedad

$$\frac{\mathrm{d}\varphi \circ \gamma}{\mathrm{d}t}(0) = \frac{\mathrm{d}\varphi \circ \gamma_1}{\mathrm{d}t}(0) + \frac{\mathrm{d}\varphi \circ \gamma_2}{\mathrm{d}t}(0).$$

Además, si $\lambda \in \mathbb{R}$

$$\begin{aligned} \lambda[\gamma] &= \lambda \frac{\mathrm{d}\varphi \circ \gamma}{\mathrm{d}t}(0) \\ &= \frac{\mathrm{d}(\lambda(\varphi \circ \gamma))}{\mathrm{d}t}(0) \\ &= [\lambda\gamma]. \end{aligned}$$

Sea $\mathbf{u} = [\gamma]$ en (V, φ) y $p \in V$. Luego

$$\mathbf{u}_i = \frac{\mathrm{d}}{\mathrm{d}t}(\varphi \circ \gamma)_i(0), \quad i = 1, 2, \ldots, n.$$

Si tenemos otra carta $(\tilde{V}, \tilde{\varphi})$ y $p \in \tilde{V}$, entonces

$$\tilde{\mathbf{u}}_i = \sum_{j=i}^{n} \frac{\partial}{\partial x_j}(\tilde{\varphi} \circ \varphi^{-1})_i(\varphi(p))\mathbf{u}_j.$$

Consideremos una función diferenciable $f : M \to N$, donde M y N son variedades y $p \in M$. La *diferencial* de f en el punto p es la función lineal

$$\mathrm{d}_p f : \mathrm{T}_p M \to \mathrm{T}_{f(p)} N$$

definida por

$$\mathrm{d}_p f([\gamma]_p) = [f \circ \gamma]_{f(p)}.$$

Se define el *haz tangente* de la variedad M por

$$TM \stackrel{\text{def}}{=} \bigcup_{p \in M} T_p M,$$

donde la unión es ajena. TM es también una variedad diferenciable de dimensión el doble que la de M. Notemos que un punto típico en TM es de la forma (p, v_p), donde $p \in M$ y $v_p \in T_p M$. Así, se tiene definida la *proyección natural*

$$\begin{aligned} \pi : TM &\to M \\ (p, v_p) &\mapsto p, \end{aligned} \tag{1.11}$$

la cual define una estructura de *haz vectorial*.

Campos vectoriales. Si M es una variedad diferenciable de dimensión n, diremos que una función

$$\begin{aligned} X : M &\to TM \\ p &\mapsto v_p \in T_pM, \end{aligned}$$

es un *campo vectorial suave* si $X \circ \pi = \mathrm{id}_M$, donde π es la proyección (1.11). Denotaremos por $\mathfrak{X}(M)$ al conjunto de todos los campos vectoriales suaves en la variedad M.

Si $\{\frac{\partial}{\partial x_1}\big|_p, \frac{\partial}{\partial x_2}\big|_p, \ldots, \frac{\partial}{\partial x_n}\big|_p\}$ denota una base para T_pM, entonces en una vecindad del punto $p \in M$, el campo vectorial $X \in \mathfrak{X}(M)$ se escribe como

$$X = \sum_i^n a_i \frac{\partial}{\partial x_i}\big|_p,$$

donde $a_1, a_2, \ldots, a_n$ son funciones suaves definidas en una vecindad de p. Notemos que el campo X *opera* sobre funciones en $C^\infty(M)$ de la siguiente manera:

$$X(f)_q = \sum_{i=1}^n a_i(q) \frac{\partial}{\partial x_i}\big|_p (f)(q) \equiv \mathcal{L}_X(f)(q).$$

Una operación muy importante entre campos vectoriales es el corchete de dos campos. Si $X, Y \in \mathfrak{X}(M)$, se define su *corchete* por

$$[X, Y] \stackrel{\text{def}}{=} XY - YX,$$

de tal manera que si $f \in C^\infty(M)$, entonces $[X, Y](f) = X(Y(f)) - Y(X(f))$ y se tiene que $[X, Y]$ es también un campo vectorial suave en M.

Sean M y N variedades suaves y $f : M \to N$ una función diferenciable. Decimos que un punto $q \in N$ es un *valor regular* de f si para cada $p \in f^{-1}(q)$, d_pf es sobreyectiva.

El siguiente resultado es importante pues nos permite construir subvariedades a partir de los valores regulares de una función. Además, nos será de gran utilidad en el siguiente capítulo, cuando se estudien grupos de Lie y álgebras de Lie. Se omite la demostración del mismo, la cual se puede encontrar en ([2]).

Proposición 1.4 *Sea $f : M \to N$ una submersión y sea $q \in N$. Entonces $f^{-1}(q)$ es una variedad diferenciable y $\dim(f^{-1}(q)) = \dim(M) - \dim(N)$. Además, $f^{-1}(q)$ es cerrada en M y es una subvariedad de M.*

En particular, este resultado nos es de gran utilidada para obtener variedades en $\mathbb{R}^n$ es por medio de la imagen inversa de valores regulares de funciones suaves $F : \mathbb{R}^n \to \mathbb{R}$. Por ejemplo, si consideramos $F : \mathbb{R}^3 \to \mathbb{R}$ definida por $F(x, y, z) = x^2 + y^2 + z^2 - 1$. Notemos que $F^{-1}(0) = \mathbb{S}^2$ y F es una función suave y $\nabla F = (2x, 2y, 2z)$, por lo que $\nabla F(p) \neq 0$ para todo $p \in \mathbb{S}^2$, por lo que $0 \in \mathbb{R}$ es un valor regular de F y se tiene que $\mathbb{S}^2$ es una subvariedad de $\mathbb{R}^3$.

Espacios proyectivos. El siguiente concepto nos será de gran utilidad en capítulos posteriores, precisamente cuando se estudien los grupos $SU(2)$ y $SO(3)$.

Definición 1.11 *Dado un espacio vectorial V sobre un campo $\mathbb{F}$, el* espacio proyectivo $\mathrm{P}(V)$ inducido por V *es el conjunto $(V - \{0\})/\sim$ de clases de equivalencia de vectores distintos de cero en V bajo la relación de equivalencia $\sim$ definida como sigue: para todo $u, v \in V - \{0\}$,*

$$u \sim v \text{ sí y sólo si } v = \lambda u,$$

para algún $\lambda \in \mathbb{F} - \{0\}$.

Para nuestros propósitos, el espacio vectorial será $\mathbb{R}^n$ sobre el campo $\mathbb{R}$. En este caso, el espacio proyectivo se denota por $\mathbb{R}\mathrm{P}^{\mathrm{n}}$.

Consideremos la esfera $\mathbb{S}^n$, es decir, el conjunto de puntos $(x_1, x_2, \ldots, x_n, x_{n+1}) \in \mathbb{R}^{n+1}$ tales que

$$x_1^2 + \cdots + x_n^2 + x_{n+1}^2 = 1.$$

En este caso, cada línea D que pasa por el centro de la esfera intersecta a la misma en dos puntos antípoda, digamos a_+ y a_-. El espacio proyectivo $\mathbb{R}\mathrm{P}^{\mathrm{n}}$ es el espacio cociente que se obtiene al identificar los puntos a_+ y a_- de la esfera $\mathbb{S}^n$.

Capítulo 2

Grupos de Lie de matrices

Aunque en este capítulo nuestro principal interés es introducir los grupos de Lie de matrices y sus ágebras de Lie asociadas, para nuestros propósitos es conveniente introducir el concepto de grupo de Lie y de álgebra de Lie, en general. EL motivo principal por el cual es necesario definir estos conceptos es que en los capítulos siguientes se habla constantemente de las relaciones que existen entre los cuaternios y el grupo de Lie $SU(2)$, y las rotaciones en $\mathbb{R}^3$ y el grupo de Lie $SO(3)$.

El desarrollo de este capítulo es el siguiente. Primero, se define, de manera general, lo que es un grupo de Lie y lo que es una álgebra de Lie asociada a un grupo de Lie. Posteriormente, se estudian los grupos de Lie clásicos; y se termina el capítulo con el estudio de la exponencial de una matriz.

2.1 Grupos de Lie

En un grupo de Lie G están presentes dos estructuras: una algebraica, que está determinada por la operación en el grupo G, y otra diferenciable, la cual define la variedad diferenciable G, que suponemos de dimensión finita. Estas estructuras deben ser compatibles una con la otra. Para lograr esto, es necesario que las funciones que definen la multiplicación en el grupo y la asignación de inverso para cada elemento del grupo

$$\begin{aligned} \mu &: G \times G \to G & \qquad \iota &: G \to G \\ \mu(g,h) &= gh & \qquad \iota(g) &= g^{-1}, \end{aligned} \tag{2.1}$$

respectivamente, sean diferenciables como funciones entre variedades.

Recordemos que los axiomas que definen la estructura algebraica de G, vienen dados por medio de las funciones μ y ι, en los términos siguientes:

(1) Existe $e \in G$, tal que

$$\mu(g,e) = \mu(e,g) = a, \qquad \forall\, g \in G.$$

(2) Para cada $g \in G$, existe $\iota(g) \in G$ tal que

$$\mu(g,\iota(g)) = \mu(\iota(g),g) = e.$$

(3) Para cualesquiera $g_1, g_2, g_3 \in G$

$$\mu(g_1, \mu(g_2, g_3)) = \mu(\mu(g_1, g_2), g_3).$$

Al elemento $e \in G$ de (1) s ele llama la *identidad* (o *neutro*) de G. Al elemento $\iota(g) \in G$ en (2) se le llama el *inverso* de $g \in G$ y es usual denotarlo por g^{-1}, ya que es único. Asimismo, el neutro e del grupo es único. Notemos además que (3) es simplemente pedir que la operación del grupo sea asociativa.

Definición 2.1 *Un grupo de Lie en un par (G, μ) donde G es una variedad diferenciable y $\mu : G \times G \to G$ es la operación que define el producto del grupo, de tal manera que la función*

$$\begin{aligned} G \times G &\to G \\ (g, h) &\mapsto \mu(g, h^{-1}) = gh^{-1}, \end{aligned} \tag{2.2}$$

es diferenciable. La dimensión del grupo es la dimensión de la variedad diferenciable G.

Es claro que esta definición es equivalente a pedir que las funciones deinidas en (2.1) sean diferenciables.

Ejemplos muy conocidos de grupos de Lie son: $\mathbb{R}$ con la suma usual de reales; $\mathbb{S}^1$ con la multiplicación usual de complejos; cualquier grupo discreto es un grupo de Lie, $\mathbb{Z}$ con la suma de enteros, por ejemplo.

En lo que sigue, omitiremos la notación (G, μ) para referirnos a un grupo de Lie y simplemente diremos que G es un grupo de Lie, sobreentendiendo que se tiene una operación binaria en G que define el producto del grupo. Además, tendremos oportunidad de conocer muchos más grupos de Lie y nuestro interés se centrará en ciertos grupos de Lie de matrices.

Sea G un grupo de Lie y sea $g \in G$ un elemento fijo. Se definen la *traslación por la izquierda* y la *traslación por la derecha* para el elemento $g \in G$ como las funciones

$$\begin{aligned} L_g &: G \to G & \qquad R_g &: G \to G \\ L_g(h) &= gh & \qquad R_g(h) &= hg, \end{aligned} \tag{2.3}$$

respectivamente. Notemos que tanto L_g como R_g son difeomorfismos para cada $g \in G$ y

$$L_g^{-1} = L_{g^{-1}} \qquad R_g^{-1} = R_{g^{-1}}.$$

Dado que G es una variedad diferenciable, podemos hablar de campos vectoriales en G y denotaremos por $\mathfrak{X}(G)$ al conjunto de los campos vectoriales suaves en el grupo de Lie G.

Un campo vectorial $X \in \mathfrak{X}(G)$ se dice ser *invariante por la izquierda* si

$$(\mathrm{d}_h L_g)(X(h)) = X(gh) \qquad \forall\, g, h \in G, \tag{2.4}$$

Notemos que $X(h) \in T_h G$ por lo que $\mathrm{d}_h L_g : T_h G \to T_{gh} G$.

Denotaremos por $\mathfrak{X}_L(G)$ al conjunto de campos vectoriales invariantes por la izquierda en G. Es claro que la suma de dos campos invariantes por la izquierda es de nuevo un campo invariante por la izquierda. Ademas, si $\lambda \in \mathbb{R}$ y $X \in \mathfrak{X}_L(G)$, entonces $\lambda X \in \mathfrak{X}_L(G)$. Por lo tanto, $\mathfrak{X}_L(G)$ es un espacio vectorial real.

Otra manera de escribir la ecuación (2.4) es por medio del *pull-back* del campo $X \in \mathfrak{X}(G)$ bajo el difeomorfismo L_g. Así, $X \in \mathfrak{X}(G)$ es invariante por la izquierda si y sólo si

$$L_g^* X = X,$$

donde $(L_g^* X)(h) \stackrel{\text{def}}{=} (\mathrm{d}_{gh} L_g^{-1})(X(L_g(h))) = (\mathrm{d}_{gh} L_{g^{-1}})(X(gh)) = \big(\mathrm{d}_h L_g\big)^{-1}(X(gh))$.

Una propiedad importante de los campos vectoriales invariantes por la izquierda es que su corchete es tambien un campo vectorial invariante por la izquierda. En efecto, si $X, Y \in \mathfrak{X}_L(G)$, entonces

$$L_g^*[X, Y] = [L_g^* X, L_g^* Y] = [X, Y].$$

Esto hace de $(\mathfrak{X}_L(G), [\,,])$ una *álgebra de Lie*. De hecho, esta álgebra de Lie es el álgebra de Lie asociada al grupo de Lie G. Sin embargo, podemos identificar al espacio de los campos vectoriales invariantes por la izquierda con el espacio tangente al grupo G en $e \in G$, T_eG, de la siguiente manera: Si $v \in T_eG$, definimos el campo vectorial $X_v \in \mathfrak{X}(G)$ por

$$X_v(g) \stackrel{\text{def}}{=} (\mathrm{d}_e L_g)(v) \in T_g G. \tag{2.5}$$

Notemos que para cualquier $y \in G$ en una vecindad de $g \in G$,

$$(\mathrm{d}_g L_y)(X_v(g)) = (\mathrm{d}_g L_y) \circ (\mathrm{d}_e L_g)(v) = (\mathrm{d}_e L_{yg})(v) = X_v(yg),$$

lo cual establece que el campo vectorial X_v definido en (2.5) es invariante por la izquierda. Más aún, si $X \in \mathfrak{X}_L(G)$, se tiene que $X(e) \in T_eG$ por lo que la correspondencia

$$\begin{aligned} T_eG &\to \mathfrak{X}_L(G) \\ v &\mapsto X_v, \end{aligned}$$

donde X_v esta definido por (2.5) es una biyección. Luego, $\mathfrak{X}_L(G)$ es un espacio finito dimensional y se tiene así la siguiente

Definición 2.2 *Sea G un grupo de Lie de dimensión finita. El álgebra de Lie de G, que denotaremos por $\mathfrak{g}$, se define como el espacio tangente a G en la identidad $e \in G$.*

Se tiene así que el ágebra de Lie $\mathfrak{g}$ del grupo de Lie G es un espacio vectorial real en el cual está definida un operación

$$[\,,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

definida para cualesquiera $u, v \in \mathfrak{g} = T_eG$ por

$$[u, v] \stackrel{\text{def}}{=} [X_u, X_v](e)$$

Se tiene que esta operación satisface las siguientes propiedades:

1. Bilinealidad

$$[\lambda u + \mu v, \xi w + \zeta z] = \lambda\xi[u, w] + \lambda\zeta[u, z] + \mu\xi[v, w] + \mu\zeta[v, z],$$

para cualesquiera $u, v, w, z \in \mathfrak{g}$, $\lambda, \mu, \xi, \zeta \in \mathbb{R}$.

2. Antisimetría.

$$[u, v] = -[v, u].$$

3. Identidad de Jacobi

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0.$$

2.2 Los grupos de Lie clásicos

Sea V un espacio vectorial de dimensión finita sobre un campo $\mathbb{F}$, donde $\mathbb{F}$ es uno de $\mathbb{R}$ o $\mathbb{C}$. De aquí en adelante cuando hagamos referencia a $\mathbb{F}$, estaremos hablando de $\mathbb{R}$ o $\mathbb{C}$ indistintamente. Recordemos del álgebra lineal que $\mathrm{End}(V)$ denota conjunto de transformaciones lineales de V en sí mismo,

$$\mathrm{End}(V) = \{T : V \to V \mid T \text{ es lineal }\}.$$

Es claro que $\mathrm{End}(V)$ es un espacio vectorial sobre $\mathbb{F}$. Además, si consideramos la composición

$$(T \circ V)(\mathbf{v}) = T(S(\mathbf{v})),$$

para $S, T \in \mathrm{End}(V)$ y $\mathbf{v} \in V$, se tiene que $\mathrm{End}(V)$ es una álgebra sobre $\mathbb{F}$.

Si fijamos una base para el espacio vectorial V, podemos asociar a cada elemento de $\mathrm{End}(V)$ una matriz de $n \times n$ con entradas en $\mathbb{F}$, el cual denotaremos por

$$M_n(\mathbb{F}) = \{(a_{ij}) \mid a_{ij} \in \mathbb{F}, i, j = 1, \ldots, n\}.$$

Notemos que si cambiamos la base para el espacio vectorial V, la matriz asociada a un elemento de $\mathrm{End}(V)$, digamos T, es distinta a la matriz que se tiene con otra base. De esta manera, el conjunto $\mathrm{End}(V)$ es isomorfo a $M_n(\mathbb{F})$ tantas veces como bases tenga el espacio vectorial V.

Al conjunto de los endomorfismos biyectivos del espacio vectorial V, lo denotaremos por $\mathrm{Aut}(V)$ y sus elementos son las transformaciones lineales invertibles, es decir, aquellas cuyo determinante es distinto de cero

$$\mathrm{Aut}(V) = \{A \in \mathrm{End}(V) \mid \det(A) \neq 0\}.$$

El conjunto $\mathrm{Aut}(V)$ es un subconjunto abierto de $\mathrm{End}(V)$ ya que la función determinante, $\det : \mathrm{End}(V) \to \mathbb{R}$, que no depende de la base que se elija ([5]), al ser continua y sabiendo que la imagen inversa de un conjunto abierto es un conjunto abierto bajo funciones continuas, tenemos que $\det^{-1}(\mathbb{R} - \{0\})$ es un conjunto abierto en $\mathrm{End}(V)$. Así, $\mathrm{Aut}(V)$ tiene estructura de variedad diferenciable. De esta manera $\mathrm{Aut}(V)$ tiene una estructura de grupo de Lie, y obtenemos los grupos

$$GL_n(\mathbb{R}) = \mathrm{Aut}_{\mathbb{R}}(\mathbb{R}^n) \qquad \text{y} \qquad GL_n(\mathbb{C}) = \mathrm{Aut}_{\mathbb{C}}(\mathbb{C}^n).$$

Si fijamos una base para $\mathbb{R}^n$, entonces $GL_n(\mathbb{R})$ es isomorfo al grupo de matrices invertibles de $n \times n$, el cual denotaremos por $GL(n, \mathbb{R})$. De esta manera podemos pensar en $GL_n(\mathbb{R})$, junto con sus subgrupos, como grupos de matrices. Sean $A, B \in GL_n(\mathbb{R})$. Sabemos que $\det(AB) = \det(A)\det(B) \neq 0$ y así $AB \in GL_n(\mathbb{R})$. Ahora bien, como $A \in GL_n(\mathbb{R})$, entonces $\det(A) \neq 0$, de esta manera existe la inversa $A^{-1} \in GL_n(\mathbb{R})$. Por lo tanto $GL_n(\mathbb{R})$ es un grupo. Luego, las entradas de la matriz AB constan de polinomios de las entradas de A y B, las cuales son exactamente las expresiones en coordenadas locales de la función producto, de esta manera, el producto AB es C^∞.

La inversa de $A = (a_{ij})$ se puede escribir como $A^{-1} = (1/\det(A))(\tilde{a}_{ij})$, donde $\tilde{a}_{ij}$ son los cofactores de A, es decir, polinomios en las entradas de A y donde $\det(A)$ es un polinomio en las entradas de A el cual no se anula en $GL_n(\mathbb{R})$. Esto nos dice que las entradas de A^{-1} son funciones racionales cuyos denominadores son distintos de cero, de esta manera es C^∞. Por lo tanto $GL_n(\mathbb{R})$ es un grupo de Lie. Un caso especial es $GL_1(\mathbb{R}) = \mathbb{R}^* = \mathbb{R} - \{0\}$, el grupo multiplicativo de números reales distintos de cero.

Siguiendo el mismo razonamiento, $GL_n(\mathbb{C})$ es también un grupo de Lie.

Los **grupos especiales lineales** sobre $\mathbb{R}$ y $\mathbb{C}$ respectivamente se definen por

$$SL(n, \mathbb{R}) = \{A \in GL(n, \mathbb{R}) \mid \det(A) = 1\},$$

y

$$SL(n, \mathbb{C}) = \{A \in GL(n, \mathbb{C}) \mid \det(A) = 1\}.$$

Verifiquemos que son grupos de Lie. Consideremos la función $F : GL_n(\mathbb{R}) \to \mathbb{R}^*$ definida por $F(X) = \det(X)$. De acuerdo a la regla del producto, $\det(XY) = \det(X)\det(Y)$. Así, F es un homomorfismo sobre $\mathbb{R}^*$; también es C^∞ pues está dado por polinomios en las entradas.

Sean $A \in GL_n(\mathbb{R})$, $a = \det A$ y consideremos las traslaciones izquierdas L_X, L_x en $GL_n(\mathbb{R})$ y $GL_1(\mathbb{R}) = \mathbb{R}^*$, respectivamente. Como $a \cdot \det(A^{-1}X) = \det X$, entonces

$$F(X) = L_a \circ F \circ L_{A^{-1}}(X).$$

Luego, utilizando $DL_a = a \neq 0$, y usando el hecho de que $L_{A^{-1}}$ es un difeomorfismo, es decir, que $L_{A^{-1}}$ es no singular, tenemos

$$rangoDF(X) = rango[aDF(A^{-1}X)DL_{A^{-1}}(X)] = rangoDF(A^{-1}X)$$

para todo $A \in GL_n(\mathbb{R})$. En particular, $rangoDF(X) = rangoDF(X^{-1}X) = rangoDF(I)$, y así vemos que el rango es constante. Bajo estas condiciones se sigue que $F^{-1}(1) = SL_n(\mathbb{R})$ es una subvariedad regular cerrada de $GL_n(\mathbb{R})$, por lo tanto es un grupo de Lie.

El **grupo ortogonal**, el cual denotaremos por $O(n)$, es un subgrupo de $GL(n, \mathbb{R})$ y consiste en aquellas matrices que preservan el producto interior usual en $\mathbb{R}^n$

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \cdot y_i,$$

con $x, y \in \mathbb{R}^n$. Es decir, si $A \in O(n)$

$$\langle Ax, Ay \rangle = \langle x, y \rangle. \tag{2.6}$$

De la ecacióna (2.6) obtenemos

$$\langle x, A \cdot A^{\intercal} y \rangle = \langle x, y \rangle,$$

por lo tanto $A \cdot A^{\intercal} = I$. De esta manera, el grupo ortogonal, $O(n)$, se define por

$$O(n) = \{A \in GL(n, \mathbb{R}) \mid A \cdot A^{\intercal} = I\}.$$

Consideremos la función $F : GL_n(\mathbb{R}) \to GL_n(\mathbb{R})$ definida por $F(X) = X^{\intercal} X$. Si $A \in GL_n(\mathbb{R})$, mostraremos que $rangoDF(X) = rangoDF(XA^{-1})$; y como cualquier $Y \in GL_n(\mathbb{R})$ se puede escribir en la forma $Y = XA^{-1}$, tenemos que $rangoDF$ es constante en $GL_n(\mathbb{R})$. Para obtener esta igualdad, notemos que $F(XA^{-1}) = L_{(A^{-1})^{\intercal}} \circ R_{A^{-1}} \circ F(X)$. Luego,

$$DF(XA^{-1}) = DL_{(A^{-1})^{\intercal}} \circ DR_{A^{-1}} \circ DF(X),$$

donde $DR_{A^{-1}}$ y $DL_{(A^{-1})^{\intercal}}$ son evaluadas en $F(X)$ y $R_{A^{-1}}(F(X))$, respectivamente. De esta manera, la igualdad $rangoDF(XA^{-1}) = rangoDF(X)$ se sigue del hecho de que $DL_{(A^{-1})^{\intercal}}$ y $DR_{A^{-1}}$ son no singulares en todo punto. Como $F^{-1}(I) = O(n)$, entonces $O(n)$ es una subvariedad regular cerrada en $GL_n(\mathbb{R})$ y así $O(n)$ es un grupo de Lie.

Los elementos de $O(n)$ son llamados matrices ortogonales. Si consideramos la función determinante $\det : O(n) \to \{+1, -1\}$, ésta separa a $O(n)$ en dos componentes conexas. Una de estas partes tiene estructura de grupo y es conocido como el **grupo especial ortogonal**

$$SO(n) = \{A \in O(n) \mid \det(A) = 1\},$$

cuyos elementos representan rotaciones en el espacio $\mathbb{R}^n$.

Análogamente, para el caso complejo, el **grupo unitario**, que denotaremos por $U(n)$, son las matrices que preservan el producto interior Hermitiano

$$\langle z, w \rangle = \sum_{i=1}^{n} z_i \overline{w}_i,$$

con $w, z \in \mathbb{C}$. Es decir, si $A \in U(n)$

$$\langle Az, Aw \rangle = \langle z, w \rangle. \tag{2.7}$$

De la ecuación (2.7) tenemos

$$\langle z, A \cdot A^{*} w \rangle = \langle z, w \rangle,$$

por lo tanto $A \cdot A^{*} = I$, donde $A^{*} = \overline{A}^{\intercal}$ denota la matriz transpuesta conjugada de A. De esta manera, el grupo unitario, $U(n)$, se define por

$$U(n) = \{A \in GL(n, \mathbb{C}) \mid A \cdot A^{*} = I\}$$

y sus elementos son llamados matrices unitarias. Análogamente a como se vió con el grupo $O(n)$, el grupo $U(n)$ es también un grupo de Lie.

El **grupo especial unitario** se define de manera análoga:

$$SU(n) = \{A \in U(n) \mid \det(A) = 1\}.$$

Los grupos $SO(n)$ y $SU(n)$ son compactos, pues son cerrados y acotados en el espacio vectorial de dimensión finita $\mathrm{End}(V)$ y éste se encuentra dentro de algún espacio $\mathbb{R}^n$ o $\mathbb{C}^n$ según sea el caso que estemos considerando.

2.3 La exponencial de una matriz

Sea $A \in M_n(\mathbb{F}) = \{A = (a_{ij}) \mid a_{ij} \in \mathbb{F}, i, j = 1, \ldots, n\}$. La exponencial de A, $\exp A$ se define por

$$\exp(A) = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots = \sum_{p=0}^{\infty} \frac{A^p}{p!} \tag{2.8}$$

En adelante usaremos $\exp(A)$ o bien e^A para denotar la exponencial de una matriz.

Notemos que, aparentemente, $\exp(A)$ debiera ser también una matriz en $M_n(\mathbb{F})$ ya que en su definición sólo estamos usando las operaciones matriciales básicas en este espacio, por lo que cada término en (2.8) está bien definido. Sin embargo, el problema en este caso es garantizar la convergencia de la serie en (2.8), lo cual nos permitiría estar seguros de que $\exp(A)$ está bien definida para cualquier matriz A. De este problema nos ocupamos en lo que sigue.

La serie definida en (2.8) converge si cada una de las entradas

$$(I)_{ij} + (A)_{ij} + \left(\frac{A^2}{2!}\right)_{ij} + \left(\frac{A^3}{3!}\right)_{ij} + \cdots$$

converge. Esto motiva el siguiente resultado.

Proposición 2.1 *Sea $A \in M_n(\mathbb{F})$ y sea*

$$m = \max\{|a_{ij}|, 1 \leq i, j \leq n\}.$$

Si $A^p = (a_{ij}^{(p)})$, entonces

$$|a_{ij}^{(p)}| \leq (nm)^p$$

para cualesquiera i, j, con $1 \leq i, j, \leq n$. En consecuencia la serie definida en (2.8) converge y la matriz e^A está bien definida.

Demostración. La demostración es por inducción en p. Sea $A = (a_{ij})$. Entonces $A^2 = (a_{ij}^{(2)})$, donde $(a_{ij}^{(2)}) = \sum_{k=1}^{n} a_{ik}a_{kj}$. Luego

$$|(a_{ij}^{(2)})| = \left|\sum_{k=1}^{n} a_{ik}a_{kj}\right| \leq \sum_{k=1}^{n} |a_{ik}||a_{kj}| \leq \sum_{k=1}^{n} m^2 = nm^2 \leq n^2m^2.$$

Ahora, $A^3 = (a_{ij}^{(3)})$ donde $(a_{ij}^{(3)}) = \sum_{k=1}^{n} a_{ik}^{(2)} a_{kj}$ y de esta manera

$$|(a_{ij}^{(3)})| = \left|\sum_{k=1}^{n} a_{ik}^{(2)} a_{kj}\right| \leq \sum_{k=1}^{n} |a_{ik}^{(2)}||a_{kj}| \leq \sum_{k=1}^{n} nm^2 m = n^2 m^3 \leq n^3 m^3.$$

Supongamos que

$$|a_{ij}^{(p)}| \leq (nm)^p$$

para todo i, j, donde $1 \leq i, j, \leq n$. Luego

$$|a_{ij}^{(p+1)}| = \left|\sum_{k=1}^{n} a_{ik}^{(p)} a_{kj}\right| \leq \sum_{k=1}^{n} |a_{ik}^{(p)}||a_{kj}| \leq nm \sum_{k=1}^{n} |a_{ik}^{(p)}| \leq nm(nm)^p = (nm)^{p+1},$$

para cualesquiera i, j con $1 \leq i, j, \leq n$. Ahora, como

$$|a_{ij}^{(p)}| \leq (nm)^p,$$

tenemos que

$$\sum_{p=0}^{\infty} \frac{|a_{ij}^{(p)}|}{p!} \leq \sum_{p=0}^{\infty} \frac{(nm)^p}{p!} = e^{nm}$$

y por lo tanto cada una de las n^2 series $\sum_{p=0}^{\infty} \frac{|a_{ij}^{(p)}|}{p!}$ converge absolutamente. Con esto mostramos que

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

está bien definida. ■

De la definición de $\exp(A)$, es claro que si $A = 0$, la matriz nula, se tiene que $e^0 = I$.

Recordemos que una propiedad importante de la función exponencial en $\mathbb{R}$ es $e^{x+y} = e^x e^y$ y nos gustaría ver bajo qué condiciones esto se cumple para el caso matricial. Sean $A, B \in M_n(\mathbb{F})$ y tratemos de calcular los primeros términos en la expansión de e^{A+B}. Se tiene así que

$$e^{A+B} = I + (A+B) + \frac{(A+B)^2}{2!} + \frac{(A+B)^3}{3!} + \cdots$$

Por otra parte, si desarrollamos los primeros términos del producto $e^A e^B$ se obtiene

$$\begin{aligned} e^A e^B &= (I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots)(I + B + \frac{B^2}{2!} + \frac{B^3}{3!} + \cdots) \\ &= I + (A+B) + \frac{1}{2!}(A^2 + AB + BA + B^2) \\ &+ \frac{1}{3!}(A^3 + 3A^2B + 3AB^2 + B^3) + \cdots \end{aligned}$$

Luego, si las matrices A y B conmutan, $AB = BA$, se sigue que

$$e^A e^B = I + (A+B) + \frac{(A+B)^2}{2!} + \frac{(A+B)^3}{3!} \cdots$$

A continuación demostraremos formalmente este resultado.

Proposición 2.2 *Sean $A, B \in M_n(\mathbb{F})$. Si A y B conmutan, entonces*

$$e^{A+B} = e^A e^B$$

Demostración. Como las matrices A y B conmutan, podemos utilizar la fórmula binomial para calcular $(A+B)^p$:

$$(A+B)^p = \sum_{k=0}^{p} \begin{pmatrix} p \\ k \end{pmatrix} A^{p-k} B^k,$$

de tal manera que

$$\frac{1}{p!}(A+B)^p = \sum_{k=0}^{p} \frac{A^{p-k}B^k}{k!(p-k)!}.$$

Notemos que

$$\begin{aligned}
\sum_{p=0}^{2} \frac{1}{p!}(A+B)^p &= \sum_{p=0}^{2}\sum_{k=0}^{p} \frac{A^{p-k}B^k}{k!(p-k)!} \\
&= I + \sum_{k=0}^{1} \frac{A^{1-k}B^k}{k!(1-k)!} + \sum_{k=0}^{2} \frac{A^{2-k}B^k}{k!(2-k)!} \\
&= I + (A+B) + \left(\frac{A^2}{2!} + AB + \frac{B^2}{2!}\right) = (I+A)(I+B) + \left(\frac{A^2}{2!} + \frac{B^2}{2!}\right).
\end{aligned}$$

Si ahora tomamos la suma desde 0 hasta 4 tenemos

$$\begin{aligned}
\sum_{p=0}^{4} \frac{1}{p!}(A+B)^p &= \sum_{p=0}^{4}\sum_{k=0}^{p} \frac{A^{p-k}B^k}{k!(p-k)!} \\
&= I + \sum_{k=0}^{1} \frac{A^{1-k}B^k}{k!(1-k)!} + \sum_{k=0}^{2} \frac{A^{2-k}B^k}{k!(2-k)!} + \sum_{k=0}^{3} \frac{A^{2-k}B^k}{k!(2-k)!} + \sum_{k=0}^{4} \frac{A^{2-k}B^k}{k!(2-k)!} \\
&= I + (A+B) + \left(\frac{A^2}{2!} + AB + \frac{B^2}{2!}\right) + \left(\frac{A^3}{3!} + \frac{A^2}{2!}B + A\frac{B^2}{2!} + \frac{B^3}{3!}\right) \\
&+ \left(\frac{A^4}{4!} + \frac{A^3}{3!}B + \frac{A^2}{2!}\frac{B^2}{2!} + A\frac{B^3}{3!} + \frac{B^4}{4!}\right) \\
&= (I + A + \frac{A^2}{2!})(I + B + \frac{B^2}{2!}) + \left(\frac{A^3}{3!} + \frac{B^3}{3!} + \frac{A^4}{4!} + \frac{A^3}{3!}B + A\frac{B^3}{3!} + \frac{B^4}{4!}\right).
\end{aligned}$$

En general, para cualquier entero $N \geq 0$, podemos expresar

$$\begin{aligned}
\sum_{p=0}^{2N} \frac{1}{p!}(A+B)^p &= \sum_{p=0}^{2N}\sum_{k=0}^{p} \frac{A^{p-k}B^k}{k!(p-k)!} \\
&= \left(\sum_{p=0}^{N} \frac{A^p}{p!}\right)\left(\sum_{p=0}^{N} \frac{B^p}{p!}\right) + \sum_{\max(k,l)>N}^{k+l\leq 2N} \frac{A^k}{k!}\frac{B^l}{l!},
\end{aligned}$$

donde el segundo término de la suma consta de $N(N+1)$ elementos. Sean

$$r = \max\{|a_{ij}| \mid 1 \leq i, j \leq n\}, s = \max\{|b_{ij}| \mid 1 \leq i, j \leq n\},$$

y $m = \max\{r, s\}$. Notemos que para cada entrada c_{ij} en $\left(\frac{A^k}{k!}\right)\left(\frac{B^l}{l!}\right)$ tenemos

$$|c_{ij}| \leq n\frac{(nm)^k}{k!}\frac{(nm)^l}{l!} \leq \frac{(n^2m)^{2N}}{N!}.$$

Como consecuenca, el valor absoluto de cada entrada en

$$\sum_{\max(k,l)>N}^{k+l\leq 2N} \frac{A^k}{k!}\frac{B^l}{l!}$$

está acotada por

$$N(N+1)\frac{(n^2m)^{2N}}{N!},$$

cuyo límite es cero cuando $N \to \infty$. De aquí se sigue que $e^{A+B} = e^A e^B$ como se quería. ■

De la Proposición (2.2) se obtiene el siguiente resultado

Proposición 2.3 *Para cualquier matriz $A \in M_n(\mathbb{F})$, e^A es no singular.*

Demostración. Las matrices A y $-A$ conmutan. Luego $I = e^0 = e^{A+(-A)} = e^A e^{-A}$, por lo que e^A es una matriz invertible cuya inversa es $(e^A)^{-1} = e^{-A}$, por la unicidad de la inversa. Luego

$$1 = \det(I) = \det(e^A e^{-A}) = \det e^A \det e^{-A},$$

y, por lo tanto, $\det e^A \neq 0$ como se quería. ■

De esta proposición observamos que la imagen de la exponencial de cualquier matriz (2.8) es siempre una matriz invertible. Así, vemos que $\exp : M_n(\mathbb{F}) \to GL(n, \mathbb{F})$, y es una función bien definida. Otro resultado importante es el suguiente

Proposición 2.4 *Si A es una matriz antisimétrica, entonces $e^A e^{A^T} = I$, donde A^T denota la transpuesta de A.*

Demostración. Como A es antisimétrica, entonces $A^T = -A$, es decir, $A + A^T = A + (-A) = 0$. Luego

$$e^0 = I = e^{A+(-A)} = e^{A+A^T} = e^A e^{A^T} = e^A e^{-A}$$

y así, $e^{-A} = e^{A^T}$, por lo tanto $e^A e^{A^T} = I$. ■

Otro resultado sobre la exponencial que nos será útil más adelante es el siguiente,

Proposición 2.5 *Sean $A, U \in M_n(\mathbb{F})$, con U invertible. Entonces*

$$e^{UAU^{-1}} = Ue^A U^{-1}.$$

Demostración. Un cálculo directo nos muestra que

$$UA^pU^{-1} = (UAU^{-1})^p,$$

y de esta manera

$$e^{UAU^{-1}} = \sum_{p=0}^{\infty} \frac{(UAU^{-1})^p}{p!} = \sum_{p=0}^{\infty} \frac{UA^pU^{-1}}{p!} = U\left(\sum_{p=0}^{\infty} \frac{A^p}{p!}\right)U^{-1} = Ue^AU^{-1}.$$

■

2.4 Álgebras de Lie asociadas a los grupos de Lie clásicos

Sea V un espacio vectorial real de dimensión finita. Una *curva* parametrizada γ en V es una función continua $\gamma : (a, b) \to V$ donde (a, b) es un intervalo abierto en $\mathbb{R}$. Decimos que γ es diferenciable en $c \in (a, b)$ si el límite

$$\lim_{h \to 0} \frac{\gamma(c + h) - \gamma(c)}{h}$$

existe. De ser así, es un vector en V el cual denotaremos por $\gamma'(c)$ y lo llamaremos *vector tangente* a γ en $\gamma(c)$.

Ahora, si G es un grupo de matrices en $M_n(\mathbb{F})$, entonces una curva en G es una curva en $M_n(\mathbb{F})$ tal que $\gamma(t) \in G$, para todo $t \in (a, b)$.

Supongamos que tenemos curvas $\gamma, \sigma : (a, b) \to G$. Podemos definir una nueva curva, la curva producto, por

$$(\gamma\sigma)(t) = \gamma(t)\sigma(t).$$

Proposición 2.6 *Sean $\gamma, \sigma : (a, b) \to G$ curvas, ambas diferenciables en $c \in (a, b)$. Entonces la curva producto $\gamma\sigma$ es diferenciable en c y*

$$(\gamma\sigma)'(c) = \gamma(c)\sigma'(c) + \gamma'(c)\sigma(c).$$

Demostración. Sean $\gamma(t) = (\gamma_{ij}(t)), \sigma(t) = (\sigma_{ij}(t))$. Entonces

$$(\gamma\sigma)(t) = \left(\sum_k \gamma_{ik}(t)\sigma_{kj}(t)\right),$$

y así

$$\begin{aligned}(\gamma\sigma)'(t) &= \left(\sum_k \{\gamma'_{ik}(t)\sigma_{kj}(t) + \gamma_{ik}(t)\sigma'_{kj}(t)\}\right) \\ &= \gamma'(t)\sigma(t) + \gamma(t)\sigma'(t).\end{aligned}$$

■

Proposición 2.7 *Sea G un grupo de matrices en $M_n(\mathbb{F})$. Sea T el conjunto de todos los vectores tangente $\gamma'(0)$ a curvas $\gamma : (a,b) \to G$, $\gamma(0) = I$ $(0 \in (a,b))$. Entonces T es un supespacio de $M_n(\mathbb{F})$.*

Demostración. Si $\gamma'(0), \sigma'(0) \in T$, entonces $(\gamma\sigma)(0) = \gamma(0)\sigma(0) = II = I$ y

$$(\gamma\sigma)'(0) = \gamma'(0)\sigma(0) + \gamma(0)\sigma'(0) = \gamma'(0) + \sigma'(0),$$

de esta manera T es cerrado bajo la suma. Luego, si $\gamma'(0) \in T$ y $r \in \mathbb{R}$, sea

$$\sigma(t) = \gamma(rt),$$

entonces $\sigma(0) = \gamma(0) = I$, σ es diferenciable y $\sigma'(0) = r\gamma'(0)$. Como $M_n(\mathbb{F})$ es un espacio vectorial de dimensión finita, también lo es T. ■

Definición 2.3 *Si G es un grupo de matrices, su dimensión es la dimensión del espacio vectorial T (de vectores tangentes a G en la identidad I).*

Definición 2.4 *Sean $B \in M_n(\mathbb{C})$. B se dice ser anti-Hermitiana si*

$$B + \overline{B}^T = 0$$

Proposición 2.8 *Si β es una curva en $O(n)$ que pasa por la identidad ($\beta(0) = I$), entonces $\beta'(0)$ es anti-simétrica. Si β es una curva en $U(n)$ que pasa por la identidad ($\beta(0) = I$), entonces $\beta'(0)$ es anti-Hermitiana.*

Demostración. En cada caso tenemos que la curva producto es constante

$$\beta(t)\overline{\beta}^T(t) = I$$

y por tanto su derivada es cero, y el resultado se sigue usando (2.6). ■

El espacio vectorial de matrices de $n \times n$ con entradas reales cuya traza es cero se denota por $\mathfrak{sl}(n)$, y el espacio vectorial de matrices de $n \times n$ con entradas reales antisimétricas se denota por $\mathfrak{so}(n)$.

Los grupos $GL(n,\mathbb{R}), SL(n,\mathbb{R}), O(n)$ y $SO(n)$ no son sólo grupos. También son grupos topológicos, lo que significa que son espacios topológicos y que las operaciones de multiplicación e inversa son operaciones continuas. Aún más, son variedades reales suaves. Tales objetos son llamados grupos de Lie. Los espacios vectoriales reales $\mathfrak{sl}(n)$ y $\mathfrak{so}(n)$ son llamados *álgebras de Lie*. Sin embargo, aún no hemos definido la estructura de álgebra para $\mathfrak{sl}(n)$ y $\mathfrak{so}(n)$. Dicha estructura está dada por lo que es llamado el *corchete de Lie*, el cual se define por

$$[A,B] = AB - BA,$$

donde $A, B \in \mathfrak{sl}(n)$ o $A, B \in \mathfrak{so}(n)$ y satisface las siguientes propiedades:

- **bilinealidad**: $[\alpha A + \beta B, C] = \alpha[A, C] + \beta[B, C]$ y $[C, \alpha A + \beta B] = \alpha[C, A] + \beta[C, B]$, para cualesquiera A, B, C en el espacio vectorial, y cualesquiera $\alpha, \beta \in \mathbb{F}$.
- **antisimetría**: $[A, B] = -[B, A]$, para cualesquiera A, B en el espacio vectorial.
- **identidad de Jacobi**: $[[A, B], C] + [[C, A], B] + [[B, C], A] = 0$ para todo A, B, C en el espacio vectorial.

Las álgebras de Lie están asociadas con grupos de Lie. Lo que pasa es que el álgebra de Lie de un grupo de Lie es su espacio tangente en la identidad, es decir, el espacio de todos los vectores tangentes a la identidad. De manera intuitiva, el álgebra de Lie logra una cierta linealización del grupo de Lie. La función exponencial es una función que transforma elementos del álgebra de Lie en elementos del grupo de Lie, por ejemplo

$$\exp : \mathfrak{so}(n) \to SO(n)$$

y

$$\exp : \mathfrak{sl}(n) \to SL(n, \mathbb{R}).$$

Esta fución a menudo permite una parametrización de los elementos del grupo de Lie mediante objetos más simples, elementos del álgebra de Lie.

Las propiedades de la función exponencial juegan un papel muy importante en el estudio de grupos de Lie, y dada la estrecha relación de estos con las álgebras de Lie, la función exponencial también resulta muy útil a la hora de estudiar álgebras de Lie. Uno de los resultados más importantes de la función exponencial es el siguiente

Proposición 2.9 *La función exponencial*

$$\exp : \mathfrak{so}(n) \to SO(n)$$

está bien definida y es sobreyectiva.

Demostración. Por la Proposición 2.4 sabemos que si $A \in M_n(\mathbb{R})$ es una matriz antisimétrica, entonces e^A es una matriz ortogonal. También tenemos que

$$\det(e^A) = e^{\mathrm{tr}A},$$

y como $A \in M_n(\mathbb{R})$ es antisimétrica, sus elementos de la diagonal son todos cero, es decir, $\mathrm{tr}A = 0$ y entonces $\det(e^A) = 1$ y de esta manera $e^A \in SO(n)$. Para verificar que es sobreyectiva, usaremos algunos hechos conocidos del álgebra lineal que se pueden consultar en ([9]). Específicamente, para cada matriz antisimétrica A existe una matriz ortogonal P tal que $A = PDP^T$ donde D es una matriz diagonal por bloques de la forma

$$D = \begin{pmatrix} D_1 & & \dots & \\ & D_2 & \dots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \dots & D_p \end{pmatrix}$$

tal que cada bloque D_i es cero o una matriz de 2×2 de la forma

$$D_i = \begin{pmatrix} 0 & -\theta_i \\ \theta_i & 0 \end{pmatrix}$$

donde $\theta_i \in \mathbb{R}$, y $\theta_i > 0$. Otro resultado que utilizaremos establece que para toda matriz ortogonal R existe una matriz ortogonal P tal que $R = PEP^T$, donde E es una matriz diagonal por bloques de la forma

$$E = \begin{pmatrix} E_1 & & \dots & \\ & E_2 & \dots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \dots & E_p \end{pmatrix}$$

tal que el bloque E_i es $1, -1$ o una matriz de 2×2 de la forma

$$E_i = \begin{pmatrix} \cos\theta_i & -\operatorname{sen}\theta_i \\ \operatorname{sen}\theta_i & \cos\theta_i \end{pmatrix}.$$

Si R es una matriz ortogonal, hay un número par de -1 en la matriz E y estos pueden ser agrupados en bloques de 2×2 asociados con $\theta = \pi$. Sea D la matriz por bloques asociada con E donde una entrada igual a 1 en E está asociado con un cero en la matriz D. Entonces por la proposición (2.5)

$$e^A = e^{PDP^{-1}} = Pe^D P^{-1},$$

y como D es una matriz diagonal por bloques, podemos obtener e^D calculando la exponencial de cada bloque. Si $D_i = 0$ tenemos $E_i = e^0 = 1$, y si

$$D_i = \begin{pmatrix} 0 & -\theta_i \\ \theta_i & 0 \end{pmatrix}$$

obtenemos

$$e^D = \begin{pmatrix} \cos\theta_i & -\operatorname{sen}\theta_i \\ \operatorname{sen}\theta_i & \cos\theta_i \end{pmatrix},$$

exactamente el bloque E_i. Así, $E = e^D$, y como consecuencia,

$$e^A = e^{PDP^{-1}} = Pe^D P^{-1} = PEP^{-1} = PEP^T = R,$$

y con esto mostramos que la función exponencial es sobreyectiva. ■

Cuando $n = 3$ y la matriz A antisimétrica, es posible encontrar una fórmula explícita para e^A. Para cualquier matriz antisimétrica real

$$A = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix},$$

sean $\theta = \sqrt{a^2 + b^2 + c^2}$ y

$$B = \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix}.$$

Entonces tenemos el siguiente resultado, conocido como la fórmula de Rodrigues la cual retomaremos en el siguiente capítulo,

Proposición 2.10 *La función exponencial* $e : \mathfrak{so}(3) \to SO(3)$ *está dada por*

$$e^A = \cos\theta I + \frac{\operatorname{sen}\theta}{\theta} A + \frac{1 - \cos\theta}{\theta^2} B,$$

o

$$e^A = I + \frac{\operatorname{sen}\theta}{\theta} A + \frac{1 - \cos\theta}{\theta^2} A^2,$$

Demostración. Daremos un bosquejo de la demostración. Primero, se prueba que $A^2 = -\theta^2 I + B$ y que $AB = BA = 0$. Luego, se deduce que $A^3 = -\theta^2 A$, y que para cualquier $k \geq 0$,

$$\begin{aligned} A^{4k+1} &= \theta^{4k} A, \\ A^{4k+2} &= \theta^{4k} A^2, \\ A^{4k+3} &= -\theta^{4k+2} A, \\ A^{4k+4} &= -\theta^{4k+2} A^2. \end{aligned}$$

Después se expande la serie para e^A y se reagrupan los términos para que aparezcan las series de seno y coseno como se requiere. Una demostración más detallada se puede encontrar en ([9]). ■

Capítulo 3

Cuaternios

En este capítulo se hace un estudio detallado sobre los cuaternios. Se inicia con una reseña histórica sobre los cuaternios y después se define formalmente a los mismos. Luego, se discuten las propiedades más importantes de los cuaternios. Posteriormente, se estudian formas alternativas de representar cuaternios. Por último, se estudia la geometría de los cuaternios unitarios y la relación de éstos con el grupo de Lie $SU(2)$.

3.1 Un poco de historia

Los números complejos han sido objeto de estudio desde los tiempos de Cardano, de quien podemos decir que fue el primero en multiplicar dos números complejos, y al igual que él, varios matemáticos trabajaron con números complejos como raíces de polinomios, aunque sin estar plenamente conscientes de su verdadero significado.

Euler, por otro lado, contribuyó de manera significativa en darles un estatus a los números complejos. Fue él quien introdujo la notación $i = \sqrt{-1}$.

Un paso que también fue muy importante en la comprensión de los números complejos se dio con la representación gráfica de éstos. Gauss fue el que le dio un completo sentido a esa representación gráfica al identificar un número complejo $a+bi$ con el punto (a, b) del plano, y a él le debemos el término "número complejo".

Por otro parte, la estructura algebraica de los números complejos estaba claramente definida pues la suma y el producto de complejos se entendían desde mucho antes de la representación gaussiana. Sabemos que si $z = a + bi$ y $w = c + di$ son dos números complejos entonces su suma es

$$z + w = (a + c) + (b + d)i$$

y su producto viene dado por

$$zw = (ac - bd) + (ad + bc)i.$$

Estas operaciones son asociativas y conmutativas; además, $z(v + w) = zv + zw$, la llamada propiedad distributiva de la suma respecto al producto.

Si vemos cada número complejo como una pareja ordenada de números reales, esto es como un elemento de $\mathbb{R}^2$, resulta natural cuestionarnos cómo se deben realizar las operaciones de suma y producto en $\mathbb{R}^2$ para que se tenga una correspondencia

entre éstas operaciones y las respetctivas en el espacio vectorial $\mathbb{C}$. Fué Sir William Rowan Hamilton quien, en el año de 1833, logró darle una estrucura algebraica a las parejas de números reales. Hamilton efinió la suma y producto en $\mathbb{R}^2$ como sigue:

$$(a,b)+(c,d)=(a+b,c+d), \qquad (a,b)\cdot(c,d)=(ac-bd,ad+bc).$$

Esto vino a darle estructura algebraica a $\mathbb{R}^2$ la cual se corresponde con la del conjunto $\mathbb{C}$ de los números complejos, pues no sólo se tiene una correspondencia entre elementos, $(a,b)\to a+ib$, sino que también se respetan las operaciones de suma y producto entre un espacio y otro.

Esto fue un triunfo del álgebra pues se empezaron a vislumbrar otro tipo de estructuras que no eran números en el sentido usual del término, pero estos nuevos objetos obedecían ciertas reglas de operación, de manera similar a las reglas de operación acostumbradas que se conocían para los números.

Dado el éxito que Hamilton tuvo al darle una estructura algebraica a $\mathbb{R}^2$, los cuaternios surgieron de los intentos de Sir William Rowan Hamilton por generalizar las operaciones (aritmética) de los números complejos de una manera que fuese aplicable en $\mathbb{R}^3$. Como los números complejos tienen dos partes, una real y otra "imaginaria", Hamilton primeramente conjeturó que necesitaba otra componente "imaginaria" adicional. Durante años batalló intentando dar sentido a un insatisfactorio sistema algebraico que contenía una parte real y dos "imaginarias". Sin embargo, es imposible definir una producto en $\mathbb{R}^3$ que generalice la multiplicación de números complejos. En efecto, supongamos que tenemos una base para $\mathbb{R}^3$ digamos $\{1,i,j\}$. Así, cualquier $\mathbf{v}\in\mathbb{R}^3$ se puede expresar como $\mathbf{v}=a1+bi+cj$, donde $a,b,c\in\mathbb{R}$. Supongamos que se tiene un producto en $\mathbb{R}^3$ que extiende al producto de $\mathbb{R}^2$. Luego, $ij=\lambda 1+\mu i+\kappa j$, para escalares $\lambda,\kappa,\mu,\in\mathbb{R}$, y de esta manera

$$\begin{aligned} i(ij) &= \lambda\cdot i\cdot 1+\mu i\cdot i+\kappa i\cdot j,\\ &= \lambda\cdot i-\mu\cdot 1+\kappa(\lambda 1+\mu i+\kappa j),\\ &= (\lambda\kappa-\mu)1+(\lambda+\kappa\mu)i+\kappa^2 j. \end{aligned}$$

Como se quiere que el producto en $\mathbb{R}^3$ sea asociativo, entonces $i(ij)=(ii)j=-1j=-j$. De esta manera,

$$-j=(\lambda\kappa)1+(\lambda+\kappa\mu)i+\kappa^2 j,$$

por lo que $\kappa^2=-1$, lo cual es una contradicción, pues $\kappa\in\mathbb{R}$. Así, las operaciones de $\mathbb{R}^2$ no pueden extenderse a $\mathbb{R}^3$.

Fue hasta el año de 1843, a la edad de 38 años, que Hamilton, en un chispazo de inspiracón, inventó en un instante un sistema de 3 partes "imaginarias" que se convirtió en el álgebra de los cuaternios. Según las propias palabra de Hamilton, el día 16 de octubre de 1843 iba caminando con su esposa a lo largo del Gran Canal en Dublin, Irlanda, en camino a una reunión de la Real Academia Irlandesa, que él presidía, cuando de repente el pensamiento le llegó. Concerniente a ese momento, él después le escribe una carta a su hijo de la cual citamos un fragmento: (citado en [11])

> *. . . Ni tampoco pude resistir el impulso –irracional como pudo haber sido– de grabar con una navaja en una roca del puente de Brougham, cuando lo cruzamos, la fórmula fundamental con los símbolos, i, j, k, como sigue,*
>
> $$i^2 = j^2 = k^2 = ijk = -1$$
>
> *la cual contiene la solución al problema . . .*

Una característica de los cuaternios (si no es que la más importante) que explotaremos en este trabajo es su estrecha relación con las rotaciones en $\mathbb{R}^3$, un hecho aparentemente casi inmediato para Hamilton pero que fue publicado primeramente por Arthur Cayley en 1845. La regla de multiplicación de cuaternios de Hamilton, la cual contiene 3 subreglas identicas a la multiplicación compleja ordinaria, expresa una profunda conexión entre cuaternios de norma 1 y rotaciones en el espacio euclidiano tridimensional. Curiosamente, la regla pudiera haber sido, en un principio, descubierta directamente mediante la búsqueda de tal conexión, ya que las relaciones con cuaternios aparecen dentro de las fórmulas de rotación que datan de antes del descubrimiento de Hamilton en cuanto a cuaternios se refiere. El primero en encontrar estas relaciones fue, al parecer, Olinde Rodrigues en 1840 al escribir versiones equivalentes de las ecuaciones de Hamilton. La aparición de 3 copias de la multiplicación ordinaria compleja y la inclusión de espacios euclidianos tridimensionales no es accidental. La multiplicación compleja representa rotaciones en dos dimensiones, y cada subregla de multiplicación compleja en la regla de multiplicación de cuaternios se corresponde con una rotación que combina dos de tres posibles ejes ortogonales en $\mathbb{R}^3$.

En la siguiente década, Hamilton publicó su libro clásico *Lectures on Quaternions*, exactamente 10 años después de haber propuesto el álgebra de los cuaternios en 1853. En 1866 se publicó su aún más extenso libro *Elements of Quaternions*, después de su muerte en 1865.

Aún cuando Hamilton luchó incansablemente por hacer de los cuaternios una notación estándar para las operaciones de vectores tridimensionales tales como el producto punto y producto cruz, esta notación siempre fué percibida como rara y así la notación tensorial alternativa propuesta por el matemático estadounidense Josiah Willard Gibbs ($1839 - 1903$), donde un vector $\mathbf{x} \in \mathbb{R}^3$ se denota por $\mathbf{x} = (x, y, z)$, se convirtió en el estándar. Cuando la mecánica cuántica del electrón fue desarrollada en el siglo XX, se pudo observar que los cuaternios estaban relacionados con los objetos matemáticos recientemente desarrollados llamados "spinors", y que los electrones estaban relacionados con ambos. Una vez más, una notación basada en cuaternios era posible, pero fue rechazada debido a una notación alternativa; la notación equivalente a cuaternios la cual es estándar en aplicaciones físicas se basa en matrices de 2×2, conocidas como matrices de Pauli.

La teoría física de rotaciones de partículas elementales y algunos elementos de la teoría de grupos incorporan cuaternios, pero el intenso interés de la era de Hamilton se fue apagando debido a la falta de importancia en aplicaciones prácticas. Sin embargo, con el desarrollo tecnológico de las computadoras, los cuaternios han resurgido para tener un papel importante en el área de graficación y animación computacional.

3.2 El álgebra de los cuaternios

Antes de definir de manera precisa lo que son los cuaternios, los introduciremos de manera intuitiva y discutiremos algunas de sus propeidades. Así, un cuaternio es un elemento $(\alpha, x, y, z) \in \mathbb{R}^4$, pero considerando a este espacio no solamente como un espacio vectorial sino con una estructura algebraica que lo hace ser una álgebra de división, como explicamos a continuación.

Dados dos elementos $p, q \in \mathbb{R}^4$, la suma $p + q$ será la suma usual en $\mathbb{R}^4$, es decir, si $p = (\alpha, x, y, z)$ y $q = (\beta, u, v, w)$, entonces

$$p + q = (\alpha + \beta, x + u, y + v, z + w).$$

Asimismo, definimos el producto de dos elementos $p, q \in \mathbb{R}^4$ por

$$p \cdot q = (\alpha\beta - xu - yv - zw, \alpha u + \beta x + yw - zv, \alpha v + \beta y + zu - xw, \alpha w + \beta z + xv - yu). \quad (3.1)$$

Es fácil verificar que $(\mathbb{R}^4, +, \cdot)$ es un anillo asociativo con 1, donde $1 = (1, 0, 0, 0)$. Más aún, la base usual para el espacio vectorial real $\mathbb{R}^4$,

$$\{(1, 0, 0, 0), \quad (0, 1, 0, 0), \quad (0, 0, 1, 0), \quad (0, 0, 0, 1)\} \subset \mathbb{R}^4,$$

es de particular importancia ya que bajo el producto (3.1) se tiene:

$$\begin{aligned}
(0, 1, 0, 0) \cdot (0, 1, 0, 0) &= (-1, 0, 0, 0) = -1, \\
(0, 0, 1, 0) \cdot (0, 0, 1, 0) &= (-1, 0, 0, 0) = -1, \\
(0, 0, 0, 1) \cdot (0, 0, 0, 1) &= (-1, 0, 0, 0) = -1.
\end{aligned}$$

Por otra parte, si se introducen los símbolos

$$\begin{aligned}
1 &= (1, 0, 0, 0), \\
i &= (0, 1, 0, 0), \\
j &= (0, 0, 1, 0), \\
k &= (0, 0, 0, 1),
\end{aligned}$$

se tiene que el conjunto $\{1, i, j, k\}$ es una base para $\mathbb{R}^4$ y un cálculo directo nos muestra que $ij = k$, mientras que $ji = -k$. Asimismo, $jk = i, \quad kj = -i$. Finalmente, $ki = j$, mientras que $ik = -j$.

De esta manera, representaremos un cuaternio por medio de una expresión de la forma

$$q = \alpha + xi + yj + zk, \quad (3.2)$$

donde α, x,y y z son números reales. Los "símbolos" o "unidades imaginarias" i, j, k satisfacen las propiedades:

$$i^2 = j^2 = k^2 = -1, \quad (3.3)$$

$$ij = k, \quad jk = i, \quad ki = j, \quad (3.4)$$

$$ji - k, \quad kj = -i, \quad ik = -j \tag{3.5}$$

Al conjunto $\mathbb{H}$ de todas las expresiones formales dadas por (3.2):

$$\mathbb{H} = \{\alpha + xi + yj + zk \mid \alpha, x, y, z \in \mathbb{R}\}.$$

lo llamaremos el conjunto de los *cuaternios reales* o *números de Hamilton*.

La *suma* de dos expresiones formales del tipo (3.2), digamos $q = \alpha + xi + yj + zk \in \mathbb{H}$ y $r = \beta + ui + vj + wk \in \mathbb{H}$ se define por

$$\begin{aligned} q + r &= (\alpha + xi + yj + zk) + (\beta + ui + vj + wk) \\ &\stackrel{\text{def}}{=} (\alpha + \beta) + (x + u)i + (y + v)j + (z + w)k. \end{aligned} \tag{3.6}$$

Asimismo, el *producto* de $q = \alpha + xi + yj + zk \in \mathbb{H}$ y $r = \beta + ui + vj + wk \in \mathbb{H}$ se define por

$$\begin{aligned} q \cdot r &= (\alpha + xi + yj + zk) \cdot (\beta + ui + vj + wk) \\ &\stackrel{\text{def}}{=} (\alpha\beta - xu - yv - zw) + (\alpha u + \beta x + yw - zv)i \\ &\quad + (\alpha v + \beta y + zu - xw)j + (\alpha w + \beta z + xv - yu)k. \end{aligned} \tag{3.7}$$

Si comparamos este producto con (3.1) podemos notar que son iguales. Además, notemos que en $\mathbb{H}$ se tiene una imagen de $\mathbb{R}$ por medio de la inclusión $\mathbb{R} \hookrightarrow \mathbb{H}$, $\mathbb{R} \ni a \mapsto a + 0i + 0j + 0k \in \mathbb{H}$. Así, por medio de esta identificación se tiene que $\mathbb{R} \subset \mathbb{H}$.

Con las operaciones (3.6) y (3.7), se tiene que $(\mathbb{H}, +, \cdot)$ es un anillo asociativo con 1. Es claro también que de la definición del producto de cuaternios se sigue que si $a \in \mathbb{R}$ y $q \in \mathbb{H}$, se obtiene $aq = qa$. En general, los cuaternios que conmutan con todos los otros cuaternios son preciamente los números reales, por lo que el *centro* del anillo $\mathbb{H}$ es $\mathbb{R}$: $Z(\mathbb{H}) = \mathbb{R}$.

Similarmente, la correspondencia $\mathbb{C} \ni a + bi \mapsto a + bi + 0j + 0k \in \mathbb{H}$, identifica al conjunto de los números complejos $\mathbb{C}$ con un subonjunto de $\mathbb{H}$, por lo que $\mathbb{C} \subset \mathbb{H}$.

Por otra parte, $\mathbb{H}$ es un espacio vectorial real donde la multiplicación por escalar está definida por

$$\lambda \cdot q = \lambda(\alpha + xi + yj + zk) = \lambda\alpha + (\lambda x)i + (\lambda y)j + (\lambda z)k, \tag{3.8}$$

donde $\lambda \in \mathbb{R}$.

Así, al cuaternio $\alpha + xi + yj + zk$ le hacemos corresponder el vector $(\alpha, x, y, z) \in \mathbb{R}^4$, por lo que $\mathbb{H}$ y $\mathbb{R}^4$ son isomorfos como espacios vectoriales reales. Definiremos también una norma en $\mathbb{H}$ que se relaciona con la norma euclidiana en $\mathbb{R}^4$ y veremos que esto nos dará una manera de relacionar los cuaternios unitarios (de longitud 1) con las rotaciones en $\mathbb{R}^3$ ya que un cuaternio unitario se representa como un punto en la esfera $\mathbb{S}^3$:

$$\mathbb{S}^3 = \{(\alpha, x, y, z) \in \mathbb{R}^4 \mid \alpha^2 + x^2 + y^2 + z^2 = 1\} \subset \mathbb{R}^4.$$

Un cuaternio de la forma $xi + yj + zk$ es llamado *cuaternio puro*. Denotaremos por $\mathbb{H}_{\mathrm{p}}$ al conjunto de todos los cuaternios puros

$$\mathbb{H}_{\mathrm{p}} = \{xi + yj + zk \mid x, y, z \in \mathbb{R}\}.$$

Una definición no intuitiva de los cuaternios. Sea V un espacio vectorial sobre un campo $\mathbb{F}$, con $\dim_{\mathbb{F}}(V) = 2$, y $\{e_1, e_2\}$ una base para V. Consideremos la forma bilineal simétrica, $B_{\mathbb{H}} : V \times V \to \mathbb{F}$, definida por

$$\begin{aligned} B_{\mathbb{H}}(e_1, e_1) &= B_{\mathbb{H}}(e_2, e_2) = -1 \\ B_{\mathbb{H}}(e_1, e_2) &= B_{\mathbb{H}}(e_2, e_1) = 0. \end{aligned}$$

Consideremos ahora el álgebra tensorial $\mathcal{T}(V) = \oplus_{r=0}^{\infty} \mathcal{T}^r(V)$, donde $\dim_{\mathbb{F}}(\mathcal{T}^r(V)) = 2^r$. Sea I el ideal de $\mathcal{T}(V)$ generado por los elementos de la forma

$$v \otimes w + w \otimes v - 2B_{\mathbb{H}}(v, w),$$

para cualesquiera $v, w \in V$. El espacio cociente $\mathcal{T}(V)/I$ es isomorfo, de manera canónica, al conjunto $\mathbb{H}$. Tal correspondencia es como sigue

$$1 + I \mapsto 1, \quad e_1 + I \mapsto i, \quad e_2 + I \mapsto j, \quad e_1 \otimes e_2 + I \mapsto k.$$

Verifiquemos que esta correspondencia también satisface las propiedades (3.3),(3.4),(3.5). En efecto, primero

$$\begin{aligned} & e_1 \otimes e_1 + e_1 \otimes e_1 - 2B_{\mathbb{H}}(e_1, e_1) \in I, \\ & \rightsquigarrow 2e_1 \otimes e_1 + 2 \in I, \\ & \rightsquigarrow e_1 \otimes e_1 + 1 \in I, \\ & \rightsquigarrow e_1 \otimes e_1 - (-1) \in I, \\ & \rightsquigarrow e_1 \otimes e_1 \in (-1) + I \\ & \rightsquigarrow (e_1 \otimes e_1) + I = (-1) + I, \end{aligned}$$

de esta manera

$$(e_1 + I)(e_1 + I) = (e_1 \otimes e_1) + I = -1 + I.$$

Análogamente,

$$\begin{aligned} & e_2 \otimes e_2 + e_2 \otimes e_2 - 2B_{\mathbb{H}}(e_2, e_2) \in I, \\ & \rightsquigarrow 2e_2 \otimes e_2 + 2 \in I, \\ & \rightsquigarrow e_2 \otimes e_2 + 1 \in I, \\ & \rightsquigarrow e_2 \otimes e_2 - (-1) \in I, \\ & \rightsquigarrow e_2 \otimes e_2 \in (-1) + I \\ & \rightsquigarrow (e_2 \otimes e_2) + I = (-1) + I. \end{aligned}$$

Luego, notemos que $e_1 \otimes e_2 + e_2 \otimes e_1 \in I$, y así

$$e_1 \otimes e_2 + I = -(e_2 \otimes e_1) + I. \tag{3.9}$$

Además,

$$\begin{aligned}(e_1 \otimes e_2 + I)(e_1 \otimes e_2 + I) &= (e_1 \otimes e_2) \otimes (e_1 \otimes e_2) + I \\ &= e_1 \otimes (e_2 \otimes e_1) \otimes e_2 + I \\ &= (e_1 + I)((e_2 \otimes e_1) \otimes e_2 + I) \\ &= (e_1 + I)(e_2 \otimes e_1 + I)(e_2 + I) \\ &= (e_1 + I)((-1)(e_1 \otimes e_2) + I)(e_2 + I) \\ &= -(e_1 + I)(e_1 \otimes e_2 + I)(e_2 + I) \\ &= -(e_1 \otimes (e_1 \otimes e_2) \otimes e_2) + I \\ &= -(e_1 \otimes e_1) \otimes (e_2 \otimes e_2) + I \\ &= -(-1 + I) \otimes (-1 + I) + I \\ &= -((-1) \otimes (-1)) + I \\ &= -1 + I,\end{aligned}$$

es decir $(e_1 \otimes e_2 + I)^2 = -1 + I$. Continuando con los productos, tenemos que

$$(e_1 + I) \otimes (e_2 + I) = (e_1 \otimes e_2 + I),$$

de igual manera,

$$\begin{aligned}(e_2 + I) \otimes (e_1 \otimes e_2 + I) &= (e_2 \otimes e_1 \otimes e_2 + I) \\ &\quad - (e_2 \otimes e_2 \otimes e_1 + I) = -(-1 \otimes e_1 + I) = e_1 + I.\end{aligned}$$

Los productos restantes se realizan de manera análoga y efectivamente cumplen las propiedades (3.3),(3.4) y (3.5). Así, tenemos que $\mathcal{T}(V)/I \cong \mathbb{H}$.

Nuestra intención de presentar a los cuaternios como el espacio cociente $\mathcal{T}(V)/I$ tiene el único propósito de enfatizar su existencia y resaltar el hecho de que los podemos realizar por medio de una construcción algebraica y no solamente como expresiones formales sujetos a las relaciones (3.3)–(3.5).

Los cuaternios como una álgebra de división y como espacio euclidiano. Las operaciones definidas en $\mathbb{H}$ hacen de este espacio una de las tres álgebras de división sobre $\mathbb{R}$ (Teorema de Frobenius)([3]), la cual es asociativa pero que no es conmutativa (por ejemplo $ij = k \neq ji = -k$).

El *conjugado* del cuaternio $q = \alpha + xi + yj + zk$ se define por

$$\overline{q} \stackrel{\text{def}}{=} \alpha - xi - yj - zk,$$

y se tiene que

$$q \cdot \overline{q} = \alpha^2 + x^2 + y^2 + z^2 \geq 0,$$

por lo que es posible definir la *norma* de q.

Definición 3.1 *Sea $q \in \mathbb{H}$. Definimos la norma de q, $\|q\|$, por*

$$\|q\| \stackrel{\text{def}}{=} \sqrt{q\overline{q}} = \sqrt{\alpha^2 + x^2 + y^2 + z^2}.$$

De esta manera, si $q \neq 0$, entonces existe el *inverso multiplicativo* de q, que resulta ser

$$q^{-1} = \frac{\overline{q}}{\|q\|^2}. \tag{3.10}$$

Definición 3.2 *Sea $q \in \mathbb{H}$. Definimos la traza de q, $\mathrm{tr}(q)$, por*

$$\mathrm{tr}(q) = q + \overline{q} = 2\alpha. \tag{3.11}$$

Lo anterior nos permite afirmar lo siguiente: un cuaternio q es un cuaternio puro ($q \in \mathbb{H}_p$) sí y sólo si $\overline{q} = -q$ o equivalentemente sí y sólo si $\mathrm{tr}(q) = 0$.

Resumimos varias propiedades de la conjugación, la norma y la traza de los cuaternios en el resultado siguiente.

Proposición 3.1 *Para cualesquiera $q, r \in \mathbb{H}$ se satisfacen las siguientes identidades*

1. $\overline{qr} = \overline{r}\ \overline{q}$,
2. $\mathrm{tr}(qr) = \mathrm{tr}(rq)$,
3. $\|qr\| = \|q\|\ \|r\|$,
4. $\mathrm{tr}(rqr^{-1}) = \mathrm{tr}(q)$, *siempre que* $r \neq 0$.

Demostración. Las demostraciones de 1 y 2 son mediante cálculos directos. Sean $q = \alpha + xi + yj + zk$ y $r = \beta + ui + vj + wk$. Luego

$$\begin{aligned}\overline{qr} &= (\alpha\beta - xu - yv - zw) - (\alpha u + \beta x + yw - zv)i \\ &\quad - (\alpha v + \beta y + zu - xw)j - (\alpha w + \beta z + xv - yu)k,\end{aligned} \tag{3.12}$$

ahora,

$$\begin{aligned}\overline{r}\ \overline{q} &= (\beta - ui - vj - wk)(\alpha - xi - yj - zk) \\ &= (\alpha\beta - xu - yv - zw) + (-\beta x - \alpha u + zv - wy)i \\ &\quad + (-\beta y - \alpha v + wx - uz)j + (-\beta z - \alpha w + uy - vx)k \\ &= (\alpha\beta - xu - yv - zw) - (\beta x + \alpha u + wy - zv)i \\ &\quad - (\beta y + \alpha v + uz - wx)j - (\beta z + \alpha w + vx - uy)k,\end{aligned} \tag{3.13}$$

finalmente, de las ecuaciones (3.12) y (3.13) tenemos que $\overline{qr} = \overline{rq}$. Luego

$$\mathrm{tr}(qr) = 2(\alpha\beta - xu - yv - zw), \tag{3.14}$$

además

$$\begin{aligned}rq &= (\beta\alpha - ux - vy - wz) + (\beta x + \alpha u + vz - wy)i \\ &\quad + (\beta y + \alpha v + wx - uz)j + (\beta z + \alpha w + uy - vx)k,\end{aligned}$$

de aquí

$$\mathrm{tr}(rq) = 2(\beta\alpha - ux - vy - wz). \tag{3.15}$$

Finalmente, utilizando las ecuaciones (3.14) y (3.15) concluimos que $\mathrm{tr}(qr) = \mathrm{tr}(rq)$.

Para demostrar 3, de la definición de norma sabemos que

$$\|qr\| = \sqrt{qr\overline{qr}},$$

utilizando 1, obtenemos

$$\|qr\| = \sqrt{qr\overline{r}\overline{q}} = \sqrt{q}\|r\|\sqrt{\overline{q}},$$

ahora, como la norma es un número real podemos conmutar el producto $\|r\|\sqrt{\overline{q}}$ y de esta manera

$$\|qr\| = \sqrt{q}\sqrt{\overline{q}}\|r\| = \sqrt{q\overline{q}}\|r\| = \|q\|\|r\|,$$

por lo tanto $\|qr\| = \|q\|\|r\|$ como queríamos.

Finalmente, para probar 4, recordemos que el producto de cuaternios es asociativo, en particular, $r(qr^{-1}) = rq(r^{-1})$, entonces

$$\mathrm{tr}(r(qr^{-1})) = \mathrm{tr}(rq(r^{-1})),$$

esta expresión se puede expresar, utilizando 2, como

$$\mathrm{tr}(r^{-1}(rq)) = \mathrm{tr}((r^{-1}r)q) = \mathrm{tr}(q),$$

por lo tanto $\mathrm{tr}(rqr^{-1}) = \mathrm{tr}(q)$. ■

Definimos la función $\varphi : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$ por

$$\varphi(p, q) = \frac{1}{2}\mathrm{tr}(p\overline{q}) = \alpha\beta + xu + yv + zw.$$

Esta función es bilineal, simétrica y satisface los axiomas de espacio euclidiano ([9]). Así, los cuaternios forman un espacio euclidiano bajo el producto interior definido por φ. Como tal, $\mathbb{H}$ es isomorfo a $\mathbb{R}^4$. También, el subespacio $\mathbb{H}_{\mathrm{p}}$ de cuaternios puros es ortogonal al espacio de los cuaternios de la forma $\alpha+0i+0j+0k$ que es isomorfo a $\mathbb{R}$. Luego, el subespacio $\mathbb{H}_{\mathrm{p}}$ de cuaternios puros hereda una estructura euclidiana, y de esta manera este subespacio es isomorfo al espacio euclidiano $\mathbb{R}^3$. Así, sus grupos de rotaciones $SO(\mathbb{H}_{\mathrm{p}})$ y $SO(3)$ son isomorfos. En el siguiente capítulo se discute más a fondo este hecho.

Propiedades de los cuaternios unitarios. Si $q \in \mathbb{H}$ y $\|q\| = 1$,

$$q^{-1} = \overline{q}$$

Recordemos que la forma polar de un número complejo unitario es $\alpha = \cos\phi + i\,\mathrm{sen}\,\phi$ donde ϕ es un ángulo en el intervalo $[0, 2\pi)$. Un cuaternio unitario tiene una representación similar,

$$q = \cos\phi + \widehat{d}\,\mathrm{sen}\,\phi$$

donde $\widehat{d} = xi + yj + zk$ y $\|\widehat{d}\| = x^2 + y^2 + z^2 = 1$. Sin embargo, observemos que el producto $\widehat{d}\widehat{d} = -1$. Note la similitud en con los números complejos unitarios

$\cos\phi + i\,\mathrm{sen}\,\phi$. Aún más, la identidad de Euler para para números complejos se generaliza para cuaternios ([7]), esto es,

$$\exp(\widehat{d}\phi) = \cos\phi + \widehat{d}\,\mathrm{sen}\,\phi,$$

donde la exponencial en el lado izquierdo se evalúa en la representación en series de potencias para $\exp(x)$, sustituyendo los productos $\widehat{d}\widehat{d}$ por -1. De esta identidad, es posible definir la *potencia* ([7]) de un cuaternio unitario,

$$q^t = (\cos\phi + \widehat{d}\,\mathrm{sen}\,\phi)^t = \exp(\widehat{d}t\phi) = \cos(t\phi) + \widehat{d}\,\mathrm{sen}(t\phi). \tag{3.16}$$

También es posible definir el *logaritmo* de un cuaternio unitario ([7]),

$$\log(q) = \log(\cos\phi + \widehat{d}\,\mathrm{sen}\,\phi) = \log(\exp(\widehat{d}\phi)) = \widehat{d}\phi. \tag{3.17}$$

Es importante notar que como el producto de cuaternios no es conmutativo, se tiene que las identitades estándar para la exponencial y el logaritmo de un cuaternio unitario no siempre se satisfacen. Esto es, si $p, q \in \mathbb{H}$, con $\|p\| = \|q\| = 1$, los cuaternios $\exp(p)\exp(q)$ y $\exp(p+q)$ no son necesariamente iguales. De igual manera, los cuaternios $\log(pq)$ y $\log(p) + \log(q)$ no son necesariamente iguales.

3.3 Formas alternativas de representar a los cuaternios

Abordaremos ahora otras maneras de representar los cuaternios, las cuales nos permiten ver de forma mucho más natural al conjunto $\mathbb{H}$ y así, estar seguros de que estos "números" realmente son realizables.

Los cuaternios como matrices reales . Consideremos ahora el conjunto de las matrices de la forma

$$\widetilde{\mathbb{H}} = \left\{ \begin{pmatrix} \alpha & -x & -y & -z \\ x & \alpha & -z & y \\ y & z & \alpha & -x \\ z & -y & x & \alpha \end{pmatrix} \mid \alpha, x, y, x \in \mathbb{R} \right\} \subset M_4(\mathbb{R}). \tag{3.18}$$

Es claro que si $p, q \in \widetilde{\mathbb{H}}$, su suma, $p + q \in \widetilde{\mathbb{H}}$. Lo que no resulta evidente es que el producto también está en $\tilde{\mathbb{H}}$. En efecto, si $p, q \in \widetilde{\mathbb{H}}$,

$$\begin{aligned} pq &= \begin{pmatrix} \alpha & -x & -y & -z \\ x & \alpha & -z & y \\ y & z & \alpha & -x \\ z & -y & x & \alpha \end{pmatrix} \begin{pmatrix} \beta & -u & -v & -w \\ u & \beta & -w & v \\ v & w & \beta & -u \\ w & -v & u & \beta \end{pmatrix} \\ &= \begin{pmatrix} \alpha\beta - xu - yv - zw & -(x\beta + \alpha u - zv + yw) & -(y\beta + zu + \alpha v - xw) & -(z\beta - yu + xv + \alpha w) \\ x\beta + \alpha u - zv + yw & \alpha\beta - xu - yv - zw & -(z\beta - yu + xv + \alpha w) & y\beta + zu + \alpha v - xw \\ y\beta + zu + \alpha v - xw & z\beta - yu + xv + \alpha w & \alpha\beta - xu - yv - zw & -(x\beta + \alpha u - zv + yw) \\ z\beta - yu + xv + \alpha w & -(y\beta + zu + \alpha v - xw) & x\beta + \alpha u - zv + yw & \alpha\beta - xu - yv - zw \end{pmatrix}. \end{aligned}$$

De esta manera, $pq \in \widetilde{\mathbb{H}}$. Con cada cuaternio $\alpha + xi + yj + zk \in \mathbb{H}$, asociamos una matriz en $\tilde{\mathbb{H}}$ de la siguiente manera:

$$\alpha + xi + yj + zk \mapsto \begin{pmatrix} \alpha & -x & -y & -z \\ x & \alpha & -z & y \\ y & z & \alpha & -x \\ z & -y & x & \alpha \end{pmatrix}.$$

Esta correspondencia nos define una función $\mathbb{H} \to \tilde{\mathbb{H}}$ tal que

- es biyectiva,
- $(\alpha+xi+yj+zk)+(\beta+ui+vj+wk) \mapsto \begin{pmatrix} \alpha & -x & -y & -z \\ x & \alpha & -z & y \\ y & z & \alpha & -x \\ z & -y & x & \alpha \end{pmatrix} + \begin{pmatrix} \beta & -u & -v & -w \\ u & \beta & -w & v \\ v & w & \beta & -u \\ w & -v & u & \beta \end{pmatrix}$,
- $(\alpha+xi+yj+zk)(\beta+ui+vj+wk) \mapsto \begin{pmatrix} \alpha & -x & -y & -z \\ x & \alpha & -z & y \\ y & z & \alpha & -x \\ z & -y & x & \alpha \end{pmatrix} \begin{pmatrix} \beta & -u & -v & -w \\ u & \beta & -w & v \\ v & w & \beta & -u \\ w & -v & u & \beta \end{pmatrix}$.

Esto nos muestra que $\mathbb{H}$ y $\tilde{\mathbb{H}}$ son anillos isomorfos.

Consideremos las siguientes matrices en $\widetilde{\mathbb{H}}$,

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

Es fácil verificar que las propiedades (3.3),(3.4) y (3.5) se cumplen para estas matrices (véase el apéndice 1).

Es claro que todo elemento $\begin{pmatrix} \alpha & -x & -y & -z \\ x & \alpha & -z & y \\ y & z & \alpha & -x \\ z & -y & x & \alpha \end{pmatrix} \in \tilde{\mathbb{H}}$ se expresa como

$$\begin{pmatrix} \alpha & -x & -y & -z \\ x & \alpha & -z & y \\ y & z & \alpha & -x \\ z & -y & x & \alpha \end{pmatrix} = \alpha\mathbf{1} + x\mathrm{I} + y\mathrm{J} + z\mathrm{K},$$

lo cual hace más clara la correspondencia entre $\mathbb{H}$ y $\tilde{\mathbb{H}}$. En particular

$$\mathbf{1} \mapsto 1, \qquad \mathrm{I} \mapsto i, \qquad \mathrm{J} \mapsto j, \qquad \mathrm{K} \mapsto k.$$

Esto nos da una manera de representar los cuaternios como matrices reales de 4×4.

Los cuaternios como matrices complejas. Otra manera de representar los cuaternios es por medio de matrices complejas de 2×2. Consideremos el conjunto de las matrices complejas de 2×2 de la forma

$$\widehat{\mathbb{H}} = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \mid a, b \in \mathbb{C} \right\} \subset M_2(\mathbb{C}), \tag{3.19}$$

donde $a = \alpha + ix$ y $b = y + iz$. Es claro que si $p, q \in \widehat{\mathbb{H}}$, su suma $p+q \in \widehat{\mathbb{H}}$. También el producto $pq \in \widehat{\mathbb{H}}$. En efecto,

$$\begin{aligned} pq &= \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} \begin{pmatrix} c & d \\ -\overline{d} & \overline{c} \end{pmatrix}, \\ &= \begin{pmatrix} ac - b\overline{d} & ad + b\overline{c} \\ -\overline{b}c - \overline{a}\overline{d} & -\overline{b}d + \overline{ac} \end{pmatrix}, \\ &= \begin{pmatrix} ac - b\overline{d} & ad + b\overline{c} \\ -\overline{(ad + b\overline{c})} & -\overline{(ac - b\overline{d})} \end{pmatrix}, \end{aligned}$$

así, el producto $pq \in \widehat{\mathbb{H}}$. Con cada cuaternio $\alpha + xi + yj + zk \in \mathbb{H}$ le hacemos corresponder una matriz de la forma (3.19) de la siguiente manera

$$\alpha + xi + yj + zk \mapsto \begin{bmatrix} \alpha + xi & y + zi \\ -(y - zi) & \alpha - xi \end{bmatrix},$$

así, tenemos una correspondencia entre los conjuntos $\mathbb{H}$ y $\widehat{\mathbb{H}}$ la cual nos define una función $\mathbb{H} \mapsto \widehat{\mathbb{H}}$ tal que

- es biyectiva,
- $(\alpha + xi + yj + zk) + (\beta + ui + vj + wk) \mapsto \begin{bmatrix} \alpha + xi & y + zi \\ -(y - zi) & \alpha - xi \end{bmatrix} + \begin{bmatrix} \beta + ui & v + wi \\ -(v - wi) & \beta - ui \end{bmatrix}$,
- $(\alpha + xi + yj + zk)(\beta + ui + vj + wk) \mapsto \begin{bmatrix} \alpha + xi & y + zi \\ -(y - zi) & \alpha - xi \end{bmatrix} \begin{bmatrix} \beta + ui & v + wi \\ -(v - wi) & \beta - ui \end{bmatrix}$.

Esto nos muestra que $\mathbb{H}$ y $\widehat{\mathbb{H}}$ son anillos isomorfos.

Sean $\mathbf{1}, \mathbf{i}, \mathbf{j}$ y $\mathbf{k}$ las siguientes matrices:

$$\mathbf{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{i} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad \mathbf{j} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \text{ y } \mathbf{k} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

Resulta obvio que cualquier elemento $\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} \in \widehat{\mathbb{H}}$ se expresa como

$$\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} = \alpha\mathbf{1} + x\mathbf{i} + y\mathbf{j} + z\mathbf{k},$$

lo cual hace más clara la correspondencia entre $\mathbb{H}$ y $\widehat{\mathbb{H}}$. En particular,

$$\mathbf{1} \mapsto 1 \quad \mathbf{i} \mapsto i \quad \mathbf{j} \mapsto j \quad \mathbf{k} \mapsto k.$$

Esto nos da una manera de representar cuaternios como matrices complejas de 2×2.

Consideremos a $\mathbb{C}$ como un subcampo de $\mathbb{H}$ mediante la inclusión $\mathbb{C} \to \mathbb{H}$ dada por

$$c \mapsto \begin{bmatrix} c & 0 \\ 0 & \overline{c} \end{bmatrix},$$

y de esta manera vemos a $\mathbb{C}$, y por lo tanto también a $\mathbb{R}$, como subcampos de $\mathbb{H}$.

$\mathbb{H}$ es un espacio vectorial sobre los complejos, con $\mathbb{C}$ actuando como multiplicación por la izquierda. Esto quiere decir que cuando multiplicamos elementos de $\mathbb{H}$ por escalares de $\mathbb{C}$, la operación se lleva a cabo siempre por la izquierda, ya que no siempre la multiplicación es conmutativa. Como espacio vectorial complejo tiene una base estándar que consta de dos elementos

$$\mathbf{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ y } \mathbf{j} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

con las siguientes reglas de multiplicación

$$zj = j\overline{z} \text{ para todo } z \in \mathbb{C} \text{ y } j^2 = -1.$$

Con esta base obtenemos un isomorfismo de espacios vectoriales complejos

$$\mathbb{C}^2 \to \mathbb{H}, \qquad (a, b) \mapsto a + bj = \begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix}.$$

Como un espacio vectorial real, $\mathbb{H}$ tiene una base estándar que consta de cuatro elementos

$$\mathbf{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{i} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad \mathbf{j} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \text{ y } \mathbf{k} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix},$$

que cumplen las reglas (3.3),(3.4) y (3.5)(ver apéndice 1).

Con el isomorfismo estándar de espacios vectoriales reales $\mathbb{R}^4 \to \mathbb{H}$ que asocia $(a, b, c, d) \mapsto a+bi+cj+dk$, la norma en $\mathbb{H}$ corresponde a la norma euclidiana en $\mathbb{R}^4$. Con el isomorfismo estándar $\mathbb{C}^2 \cong \mathbb{H}$, la norma corresponde a la norma Hermitiana en $\mathbb{C}^2$.

3.4 Geometría y Topología de los cuaternios

Para entender claramente nuestras opciones para conocer visualizaciones gráficas de cuaternios, necesitamos examinar las maneras en las cuales los puntos en las esferas pueden ser vistos en dimensiones más pequeñas. La geometría de los cuaternios unitarios es la geometría de la esfera $\mathbb{S}^3$. De esta manera, podemos entender la geometría de los cuaternios unitarios estudiando la geometría de $\mathbb{S}^1, \mathbb{S}^2$ y finalmente $\mathbb{S}^3$. El truco básico para ver una esfera se basa en el hecho de que si tenemos cualquier vector unitario, la esfera que describe sus grados de libertad puede ser vista, cualitativamente, en una dimensión menor.

Empezamos con la esfera $\mathbb{S}^1$ y examinamos una de sus parametrizaciones dada por las dos raíces

$$\begin{aligned} \alpha &= t, \\ x &= \pm\sqrt{1 - t^2}. \end{aligned}$$

La raíz positiva representa el hemisferio norte de la esfera, mientras que la raíz negativa hace lo propio con el hemisferio sur. Si $t = -1$, tenemos el punto $(-1, 0)$

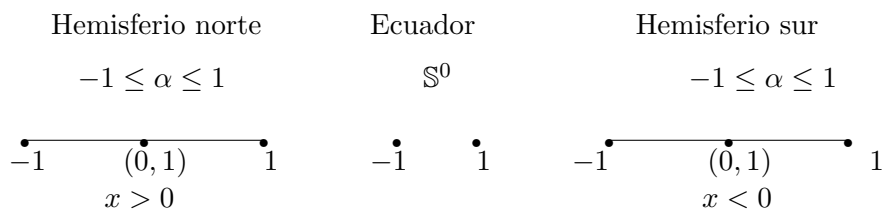


Figura 3.1: Visualización de la esfera $\mathbb{S}^1$.

y si $t = 1$ obtenemos el punto $(1, 0)$. Observe que el conjunto $\{(-1, 0), (1, 0)\} = \mathbb{S}^0$ no es otra cosa que el ecuador de la esfera $\mathbb{S}^1$. Así, podemos ver a $\mathbb{S}^1$ como dos segmentos de recta unidos por el ecuador ($\mathbb{S}^0$) como se muestra en la Figura 3.1.

Procedemos ahora a estudiar la esfera $\mathbb{S}^2$,

$$\mathbb{S}^2 = \{(\alpha, x, y) \in \mathbb{R}^3 \mid \alpha^2 + x^2 + y^2 = 1\},$$

si conocemos α y x, podemos parametrizar la esfera nuevamente mediante dos raíces

$$\begin{aligned} \alpha &= t_1, \\ x &= t_2, \\ y &= \pm\sqrt{1 - t_1^2 - t_2^2}, \end{aligned}$$

donde la raíz positiva representa el hemisferio norte de la esfera $\mathbb{S}^2$ y la raíz negativa el hemisferio sur. Notemos ahora que el ecuador de la esfera $\mathbb{S}^2$ es el conjunto $\{(\alpha, x, y) \in \mathbb{S}^2 \mid y = 0\} = \mathbb{S}^1$, de manera que podemos visualizar a $\mathbb{S}^2$ como dos discos rellenos en $\mathbb{R}^2$ unidos mediante el ecuador, como se muestra en la Figura 3.2, que en este caso es la esfera $\mathbb{S}^1$.

Siguiendo el mismo razonamiento, dados α, x, y podemos parametrizar la esfera $\mathbb{S}^3$ como sigue

$$\begin{aligned} \alpha &= t_1, \\ x &= t_2, \\ y &= t_3, \\ z &= \pm\sqrt{1 - t_1^2 - t_2^2 - t_3^2}, \end{aligned}$$

de tal manera que la raíz positiva es una representación del hemisferio norte de la esfera $\mathbb{S}^3$, mientras que la raíz negativa es una representación del hemisferio sur de $\mathbb{S}^3$.

El ecuador de la esfera $\mathbb{S}^3$ es el conjunto $\{(\alpha, x, y, z) \in \mathbb{S}^3 \mid z = 0\} = \mathbb{S}^2$. Así, es posible visualizar la esfera $\mathbb{S}^3 \subset \mathbb{R}^4$ como dos bolas sólidas, cada una representada mediante la ecuación $\alpha^2 + x^2 + y^2 < 1$, unidas por el ecuador como se muestra en la Figura 3.3.

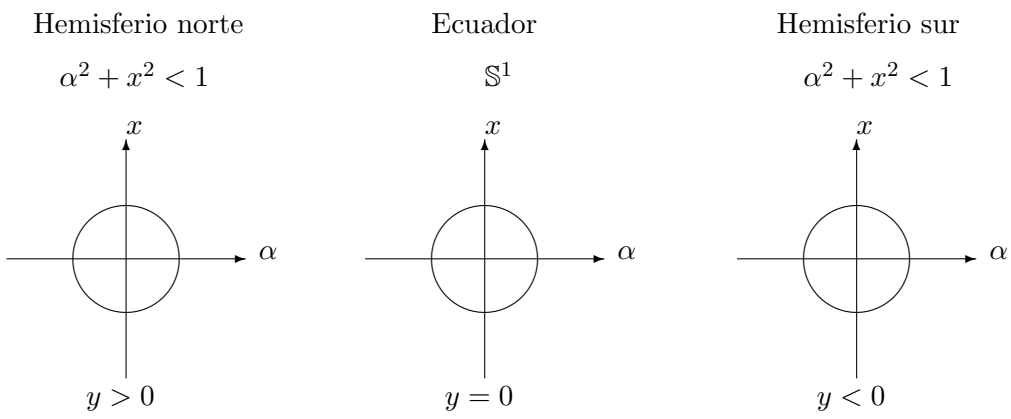


Figura 3.2: Visualización de la esfera $\mathbb{S}^2$.

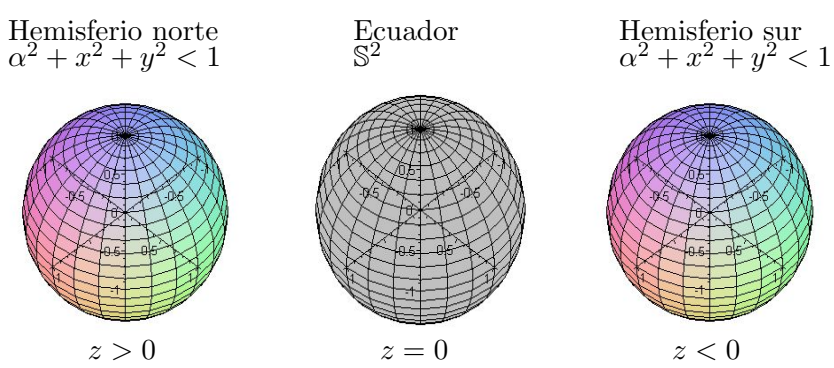


Figura 3.3: Visualización de la esfera $\mathbb{S}^3$

3.5 El grupo simpléctico

De manera análoga a como se hizo con $\mathbb{R}$ y $\mathbb{C}$ podemos identificar el grupo

$$GL(n, \mathbb{H}) = \mathrm{Aut}_{\mathbb{H}}(\mathbb{H}^n)$$

con el grupo de matrices de $n \times n$ invertibles con entradas en $\mathbb{H}$ pero teniendo siempre en cuenta que la multiplicación por escalares de $\mathbb{H}$ es siempre por la izquierda ya que no siempre al operar con elementos de $\mathbb{H}$ se da la conmutatividad.

En $\mathbb{H}^n$ existe un producto interior, el producto escalar simpléctico estándar: si $h = (h_1, \ldots, h_n) \in \mathbb{H}^n$ y $k = (k_1, \ldots, k_n) \in \mathbb{H}^n$, entonces

$$\langle h, k \rangle = \sum_{v=1}^{n} h_v \overline{k}_v.$$

La norma correspondiente está dada por $\langle h, h \rangle = \sum_v h_v \overline{h}_v = \sum_v N(h_v) \geq 0$. El **grupo simpléctico** $Sp(n)$, es el grupo cuyos elementos son automorfismos que preservan la norma en $\mathbb{H}^n$:

$$Sp(n) = \{\phi \in GL(n, \mathbb{H}) \mid N(\phi(h)) = N(h) \text{ para todo } h \in \mathbb{H}^n\}.$$

El grupo

$$Sp(1) = \{h \in \mathbb{H} \mid N(h) = 1\}$$

es llamado el grupo de cuaternios unitarios. En notación matricial $Sp(1)$ consiste en las matrices

$$\begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix}, \qquad a, b, \in \mathbb{C}, \qquad |a|^2 + |b|^2 = 1,$$

y así es el mismo que $SU(2)$. El isomorfismo estándar $\mathbb{H} \cong \mathbb{R}^4$ identifica a $Sp(1)$ con la esfera $\mathbb{S}^3$.

Capítulo 4

Rotaciones y Cuaternios

El presente capítulo es el más importante en nuestro trabajo. Se divide dos partes: en la primera se discuten algunos resultados importantes sobre rotaciones en tres dimensiones, más concretamente el teorema de Euler, el cual es demostrado utilizando herramientas del álgebra lineal, y la fórmula de Rodrigues. En la sección correspondiente a la fórmula de Rodrigues, se discute su derivación además de la relación que tiene con matrices de rotación ya conocidas en la teoría de rotaciones.

En la segunda parte del capítulo se aborda la relación entre rotaciones en el espacio tridimensional y los cuaternios. En esta sección se discute también la relación existente en entre la fórmula de Rodrigues y las rotaciones mediante cuaternios. Tal relación nos dará pie para poder llevar a cabo una de las primeras aplicaciones que se discuten en capítulos más delante.

4.1 Euler, Rodrigues y rotaciones en $\mathbb{R}^3$

Aún cuando Hamilton fue el primero en construir los cuaternios como una álgebra, éstos tienen una historia anterior, empezando con el descubrimiento de Euler conocido como la identidad de los cuatro cuadrados, la cual dice que el producto de dos números, cada uno de los cuales es una suma de cuatro cuadrados, es, en sí, una suma de cuatro cuadrados.

Más aun, un tratamiento estrictamente geométrico de las rotaciones en el espacio euclidiano tridimensional, lleva necesariamente, a una caraterización de las rotaciones en $\mathbb{R}^3$ que está muy cerca del trabajo de Hamilton para representar estas rotaciones por medio de cuaternios. Este trabajo lo realizó el matemático francés Olinde Rodrigues (1795-1851) en 1840, antes del descubrimiento de los cuaternios por Hamilton en 1843.

Consideremos rotaciones de la esfera unitaria centrada en el origen. El primer problema a resolver es probar que si aplicamos una rotación seguida de otra, el resultado final es una rotación de la esfera alrededor de algún eje por un ángulo. (Esto es crucial para establecer la estructura de grupo de el conjunto de todas las posibles rotaciones en la esfera). Euler ([1]) resolvió un problema aún más general que este, pues el consideró la composición de dos transformaciones afines (traslaciones-rotaciones) y mostró que la orientación de los ejes resultantes de tal composición dependían de seis parámetros angulares, tres de los cuales podían ser eliminados algebraicamente, dejando simplemente tres parámetros solamente, y así

determinaba una rotación. Cabe mencionar que el enfoque de Euler es algebraico, no geométrico, y que no es constructivo. Esto es, que no provee expresiones para determinar el ángulo y eje de la rotación resultante. Sin embargo, Euler, es a menudo acreditado por la solución existencial, geométrica, y problemas constructivos concernientes a la composición de dos rotaciones.

El problema que se acaba de describir está relacionado con otro aún más general: probar que el movimiento general de una esfera con centro fijo es una rotación, es decir, que cualquier movimiento de la esfera puede ser expresado como la composición o el producto de dos rotaciones. Éste resultado es el conocido como *Teorema de Euler*

Fue en su tesis doctoral, publicada en 1840 ([1]), que Rodrigues resolvió todos los tres aspectos del problema de Euler concerniente al producto de dos rotaciones. En tal artículo, Rodrigues describe una construcción geométrica la cual, dados los ángulos y ejes de dos rotaciones, determina la orientación del eje resultante de rotación y el valor de su ángulo de rotación. Esta construcción geométrica es llamada por varios autores la "construcción de Euler-Rodrigues".

A pesar de que la construcción de Euler-Rodrigues contiene la llave geométrica para el grupo de rotaciones, ha sido ignorada en los estudios sobre el grupo de rotaciones y momento angular. En 1848, ocho años después del artículo de Rodrigues, Stokes hace incapié en el hecho de que no existía un método geométrico para la composición de dos rotaciones. Sylvester, en 1850, produce una buena figura de la construcción de Rodrigues -la primera en ser publicada- con una clara explicación sin hacer referencia alguna a Rodrigues ni a Euler. Hamilton, en 1853, redescubrió geométricamente los resultados de la construcción de Rodrigues.

Una vez resueltos los problemas de existencia y geométricos debido al producto de dos rotaciones, Rodrigues provee, en su artículo, fórmulas para determinar el ángulo y el eje de la rotación resultante. Para llevar a cabo esto, él parametriza una rotación con cuatro parámetros. Si ϕ es el ángulo de rotación y (n_x, n_y, n_z) son las componentes del vector unitario que denota el eje de rotación, sus parámetros son:

$$\cos\frac{1}{2}\phi, \quad \operatorname{sen}\frac{1}{2}\phi n_x, \quad \operatorname{sen}\frac{1}{2}\phi n_y, \quad \operatorname{sen}\frac{1}{2}\phi n_z,$$

Si hacemos corresponder estos parámetros con los números α, x, y, z de la ecuación (3.2), entonces la fórmula para la multiplicación que propuso Rodrigues es precisamente la regla de multiplicación de Hamilton para cuaternios. Esto nos dice que Rodrigues fue, en cierta manera, precursor de Hamilton.

El teorema de Euler. Uno de los resultados más importantes sobre rotaciones, el cual enunciamos a continuación, es el teorema de Euler. Este teorema nos asegura que toda rotación por un cierto ángulo en cualquier espacio deja fija una línea recta que es el eje de rotación.

Teorema 4.1 ***(Euler)*** *Si R es una matriz que representa una rotación en $\mathbb{R}^3$, entonces R tiene un vector propio $\mathbf{n} \in \mathbb{R}^3$ tal que*

$$R\,\mathbf{n} = \mathbf{n},$$

esto es, $\mathbf{n}$ *es un vector propio con valor propio* 1.

Demostración. Como R es una rotación, $R \in SO(3)$ y $\det(R) = 1$. Además, $RR^T = I = RR^{-1}$ y $R^T = R^{-1} \in SO(3)$.

Por otro lado

$$\begin{aligned}\det[R-I] &= \det[(R-I)^T] = \det[R^T - I] = \det[R^{-1} - I] = \det[-R^{-1}(R-I)] \\ &= (-1)\det[R^{-1}(R-I)] = -\det[R^{-1}]\det[R-I] = -\det[R-I],\end{aligned}$$

es decir, $\det[R-I] = -\det[R-I]$, por lo que $\det(R-I) = 0$. Luego, $R-I$ tiene núcleo distinto de cero y por lo tanto, existe $\mathbf{n} \in \mathbb{R}^3$ tal que $(R-I)\mathbf{n} = \mathbf{0}$. De esto se sigue que $R\,\mathbf{n} = \mathbf{n}$. ■

Este teorema nos permite ver una secuencia de rotaciones sobre distintos ejes como una sola rotación alrededor de un eje pues cada rotación, al tener asociada una matriz, hace que la secuencia de ellas tenga, a su vez, asociada una sola matriz y usando el teorema de Euler sabemos que tiene un eje de rotación.

4.2 Rotaciones

De los cursos de geometría analítica recordamos que al rotar un vector $(x, y) \in \mathbb{R}^2$ alrededor del origen por un ángulo θ obtenemos otro vector $(x', y') \in \mathbb{R}^2$ cuyas coordenadas son

$$x' = \cos(\theta)x - \operatorname{sen}(\theta)y, \quad y' = \operatorname{sen}(\theta)x + \cos(\theta)y,$$

lo cual expresamos por

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = R(\theta)\begin{bmatrix} x \\ y \end{bmatrix},$$

donde

$$R(\theta) = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}.$$

Observemos que

$$R(\theta) \cdot R^T(\theta) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

por lo que $R \in SO(2)$.

La forma de mitad de ángulo. Podemos obtener la matriz de rotación para un vector en el plano como sigue. Consideremos:

$$A = a^2 - b^2, \qquad B = 2ab.$$

Si $a = \cos(\frac{\theta}{2})$ y $b = \operatorname{sen}(\frac{\theta}{2})$ tenemos que

$$\begin{aligned} A &= \cos^2\left(\frac{\theta}{2}\right) - \operatorname{sen}^2\left(\frac{\theta}{2}\right) \\ &= 1 - 2\operatorname{sen}^2\left(\frac{\theta}{2}\right) \\ &= \cos\theta, \end{aligned}$$

$$\begin{aligned} B &= 2\cos\left(\frac{\theta}{2}\right)\operatorname{sen}\left(\frac{\theta}{2}\right) \\ &= \operatorname{sen}\theta, \end{aligned}$$

de modo que la matriz de rotación no cambia

$$R(\theta) = \begin{bmatrix} a^2 - b^2 & -2ab \\ 2ab & a^2 - b^2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix},$$

En el espacio $\mathbb{R}^3$ no es tan inmediato especificar la matriz que rota un vector por un ángulo α alrededor de un eje arbitrario. Para ello utilizaremos las siguientes matrices conocidas como matrices fundamentales, las cuales son rotaciones en el plano que determina cada par de ejes ortogonales,

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\operatorname{sen}\alpha \\ 0 & \operatorname{sen}\alpha & \cos\alpha \end{bmatrix},$$

$$R_y(\phi) = \begin{bmatrix} \cos\phi & 0 & \operatorname{sen}\phi \\ 0 & 1 & 0 \\ -\operatorname{sen}\phi & 0 & \cos\phi \end{bmatrix},$$

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

En este caso se tiene:

- $R_x(\alpha)$ rota el plano yz alrededor del origen por un ángulo α.
- $R_y(\phi)$ rota el plano xz alrededor del origen por un ángulo ϕ.
- $R_z(\theta)$ rota el plano xy alrededor del origen por un ángulo θ.

Consideremos ahora un vector unitario $\mathbf{n} = (n_1, n_2, n_3) \in \mathbb{R}^3, \|\mathbf{n}\| = 1$. En coordenadas esféricas (ρ, θ, ϕ) se tiene que $\mathbf{n} = (\cos\theta\operatorname{sen}\phi, \operatorname{sen}\theta\operatorname{sen}\phi, \cos\phi)$. Una rotación por un ángulo α, teniendo como eje de rotación el vector $\mathbf{n}$, queda representada mediante la matriz

$$R(\alpha, \mathbf{n}) = R_z(\theta) \cdot R_y(\phi) \cdot R_z(\alpha) \cdot R_y^T(\phi) \cdot R_z^T(\theta) \tag{4.1}$$

Realizando los cálculos en el lado derecho de (4.1) se obtiene

$$R(\alpha, \mathbf{n}) = \begin{pmatrix} c + n_1^2(1-c) & n_1 n_2(1-c) - s n_3 & n_1 n_3(1-c) + s n_2 \\ n_2 n_1(1-c) + s n_3 & c + n_2^2(1-c) & n_2 n_3(1-c) - s n_1 \\ n_3 n_1(1-c) - s n_2 & n_3 n_2(1-c) + s n_1 & c + n_3^2(1-c) \end{pmatrix} \tag{4.2}$$

donde $c = \cos\alpha$ y $s = \operatorname{sen}\alpha$.

Podemos pensar la fórmula (4.1) de la siguiente manera: si $\mathbf{z} = (0,0,1)^T$, entonces es fácil ver que $\mathbf{n} = R_z(\theta) \cdot R_y(\phi)\,\mathbf{z}$. Así, para construir la matriz de rotación que representa un giro por un ángulo α sobre el eje $\mathbf{n}$, transformamos este vector al vector $\mathbf{z}$ invirtiendo la fórmula previa, luego giramos alrededor de $\mathbf{z}$ por el ángulo α aplicando la matriz $R_z(\alpha)$ y nos regresamos con $\mathbf{z}$ al vector $\mathbf{n}$, donde comenzamos. Este proceso está dado precisamente por la fórmula (4.1) y se puede apreciar en la Figura 4.1:

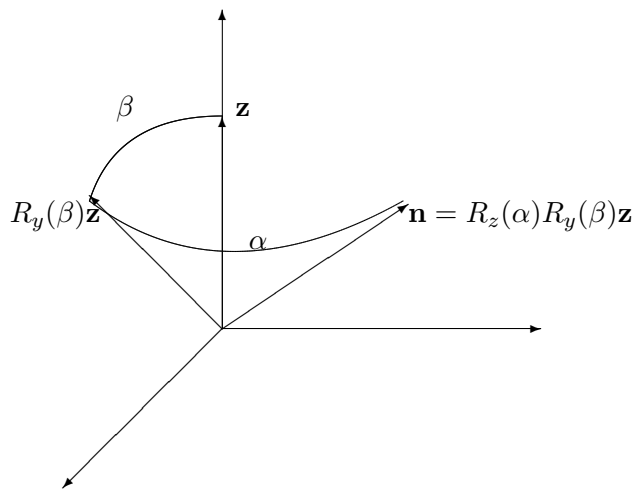


Figura 4.1: Derivación de la matriz de rotación en $\mathbb{R}^3$ alrededor de un eje unitario $\mathbf{n}$ por un ángulo θ.

La fórmula de Rodrigues. Otra forma de calcular rotaciones en el espacio $\mathbb{R}^3$, equivalente a la fórmula (4.1), es mediante la llamada *fórmula de Rodrigues*, en honor al matemático del mismo nombre, Olinde Rodrigues.

Sea $\mathbf{v}$ un vector cualquiera en $\mathbb{R}^3$. Si $\mathbf{v}_{\text{rot}}$ es el vector que se obtiene al rotar el vector $\mathbf{v}$ alrededor del vector unitario $\mathbf{n} \in \mathbb{R}^3$ por un ángulo α, entonces la fórmula de Rodrigues nos da las coordenadas de este vector:

$$\mathbf{v}_{\text{rot}} = \cos\alpha\,\mathbf{v} + \operatorname{sen}\alpha\,(\mathbf{n} \times \mathbf{v}) + (1 - \cos\alpha)\,\langle \mathbf{n}, \mathbf{v} \rangle \mathbf{n}, \tag{4.3}$$

donde $\times$ indica el producto vectorial en $\mathbb{R}^3$ y $\langle\,,\,\rangle$ es el producto punto en este espacio.

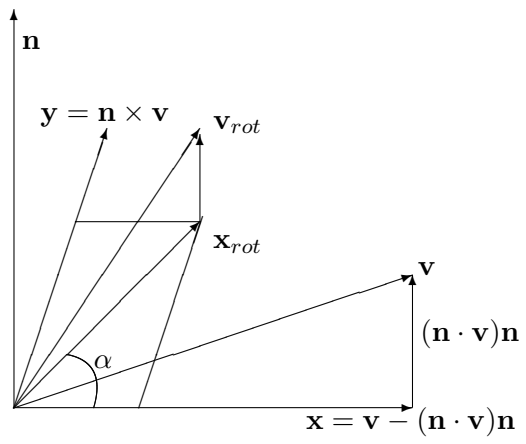


Figura 4.2: Derivación de la fórmula de Rodrigues

Para derivar esta fórmula se procede como sigue. El vector $\mathbf{x} = \mathbf{v} - \mathbf{n}\langle \mathbf{n}, \mathbf{v}\rangle$, es la proyección del vector $\mathbf{v}$ en el plano ortogonal a $\mathbf{n}$, como se muestra en la Figura 4.2. Ahora, sea $\mathbf{y} = \mathbf{n} \times \mathbf{v}$. Por medio de un poco de trigonometría, vemos que al rotar el vector $\mathbf{x}$ alrededor de $\mathbf{n}$, por un ángulo α, se obtiene la proyección del vector $\mathbf{v}_{\text{rot}}$:

$$\mathbf{x}_{\text{rot}} = \mathbf{x} \cos \alpha + \mathbf{y} \operatorname{sen} \alpha.$$

De esto se tiene que

$$\mathbf{x}_{\text{rot}} = \big(\mathbf{v} - \mathbf{n}\langle \mathbf{n}, \mathbf{v}\rangle\big) \cos \alpha + (\mathbf{n} \times \mathbf{v}) \operatorname{sen} \alpha.$$

Si ϕ denota el ángulo formado por $\mathbf{n}$ y $\mathbf{v}$, y tomamos un vector unitario $\mathbf{u}$, ortogonal a ambos, entonces podemos calcular el producto escalar y el producto vectorial de $\mathbf{n}$ y $\mathbf{v}$, por medio de las fórmulas

$$\langle \mathbf{n}, \mathbf{v}\rangle = \|\mathbf{n}\| \, \|\mathbf{v}\| \cos \phi, \tag{4.4}$$

$$\mathbf{n} \times \mathbf{v} = \|\mathbf{n}\| \, \|\mathbf{v}\| \operatorname{sen} \phi \, \mathbf{u}. \tag{4.5}$$

De esto se sigue que los vectores $\mathbf{x}$ y $\mathbf{y}$ tienen la misma longitud. Como $\mathbf{n}$ es unitario, de (4.5) se tiene

$$\|\mathbf{y}\| = \|\mathbf{n} \times \mathbf{v}\| = \|\mathbf{v}\| \operatorname{sen} \phi.$$

Si usamos (4.4) podemos calcular la longitud de $\mathbf{n}\langle \mathbf{n}, \mathbf{v}\rangle$ y este vector junto con $\mathbf{v}$ y $\mathbf{x}$ forman los lados de un triángulo rectángulo. De esto se sigue que $|\langle \mathbf{n}, \mathbf{v}\rangle| \, \|\mathbf{n}\|$ es también la longitud de $\mathbf{x}$.

Finalmente, para obtener el vector $\mathbf{v}_{\text{rot}}$ tenemos que sumarle al vector $\mathbf{x}_{\text{rot}}$ la componente de $\mathbf{v}$ que es paralela al vector $\mathbf{n}$: $\langle \mathbf{n}, \mathbf{v}\rangle \mathbf{n}$. El resultado que se obtiene es

$$\mathbf{v}_{\text{rot}} = (\mathbf{v} - \mathbf{n}\langle \mathbf{n}, \mathbf{v}\rangle) \cos \alpha + (\mathbf{n} \times \mathbf{v}) \operatorname{sen} \alpha + (\mathbf{n} \cdot \mathbf{v})\mathbf{n},$$

la cual claramente se expresa como

$$\mathbf{v}_{\text{rot}} = \mathbf{v}\cos\alpha + (\mathbf{n}\times\mathbf{v})\operatorname{sen}\alpha + \mathbf{n}\langle\mathbf{n},\mathbf{v}\rangle(1-\cos\alpha), \tag{4.6}$$

que es la fórmula de Rodrigues (4.3).

4.3 Relación entre rotaciones y cuaternios

Para entender la relación entre rotaciones en $\mathbb{R}^3$ y los cuaternios, es conveniente utilizar la notación $q = [\lambda, \mathbf{a}]$ para el cuaternio $q = \lambda + xi + yj + zk$, donde $\mathbf{a} = xi + yj + zk$. La separación del cuaternio q en dos partes nos permite distinguir su parte real λ y su "*parte imaginaria*" $\mathbf{a}$; además, identificamos el cuaternio puro $\mathbf{a} = xi + yj + zk$ con el vector $\mathbf{a} = (x, y, z) \in \mathbb{R}^3$. De manera recíproca, dado un vector $\mathbf{b} = (u, v, w) \in \mathbb{R}^3$, le hacemos corresponder el cuaternio $q_{\mathbf{b}} = [0, \mathbf{b}]$, con $\mathbf{b} = ui + vj + wk \in \mathbb{H}_{\text{p}}$.

De esta manera, dados los cuaternios $q = [\lambda, \mathbf{a}]$ y $r = [\mu, \mathbf{b}]$ definimos dos operaciones entre las partes imaginarias de q y r: El *producto punto* y el *producto cruz* de $\mathbf{a}$ y $\mathbf{b}$ los cuales se corresponden con el producto punto y el producto cruz de los vectores $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$.

Notemos que con esta definición, $\langle\mathbf{a}, \mathbf{b}\rangle \in \mathbb{R}$ y $\mathbf{a}\times\mathbf{b}$ es otro cuaternio con parte real igual a cero y su "parte imaginaria" está dada por las componentes del vector $\mathbf{a}\times\mathbf{b}$. Usando esta notación se tiene que el producto (3.7) de los cuaternios $q = [\lambda, \mathbf{a}]$ y $r = [\mu, \mathbf{b}]$ se expresa como

$$q \cdot r = [\lambda, \mathbf{a}] \cdot [\mu, \mathbf{b}] = [\lambda\mu - \langle\mathbf{a}, \mathbf{b}\rangle,\ \lambda\,\mathbf{b} + \mu\,\mathbf{a} + \mathbf{a}\times\mathbf{b}], \tag{4.7}$$

y esta fórmula nos será muy útil para describir la relación entre los cuaternios unitarios y las rotaciones en $\mathbb{R}^3$.

En efecto, supongamos que un vector $\mathbf{v} \in \mathbb{R}^3$ se rota un ángulo α alrededor de un eje, determinado por un vector unitario $\mathbf{n} = (n_1, n_2, n_3) \in \mathbb{R}^3$. Suponemos también que el giro es positivo, es decir, contrario a las manecillas del reloj. Sea $\mathbf{u}$ el vector que se obtiene de esta rotación.

Con cada vector $\mathbf{v}$, $\mathbf{n}$ y $\mathbf{u}$ asociamos el cuaternio correspondiente:

$$q_{\mathbf{v}} = [0, \mathbf{v}], \qquad q_{\mathbf{n}} = [0, \mathbf{n}], \qquad q_{\mathbf{u}} = [0, \mathbf{u}] \tag{4.8}$$

Sea q el cuaternio dado por

$$q = [\gamma, \sigma\mathbf{n}], \qquad \text{con } \gamma = \cos(\alpha/2), \quad \sigma = \operatorname{sen}(\alpha/2). \tag{4.9}$$

Notemos que $q = \gamma + (\sigma n_1)i + (\sigma n_2)j + (\sigma n_3)k$ y $\|q\| = 1$. Calculando ahora el producto $q \cdot q_{\mathbf{v}} \cdot q^{-1}$ por medio de (4.7) y utilizando varias propiedades muy conocidas del producto cruz y del producto punto, se obtiene:

$$q \cdot q_{\mathbf{v}} \cdot q^{-1} = \mathbf{v} + 2\sigma\gamma\mathbf{n}\times\mathbf{v} + 2\sigma^2\mathbf{n}\times(\mathbf{n}\times\mathbf{v}). \tag{4.10}$$

Si utilizamos las identidades $\operatorname{sen}\alpha = 2\operatorname{sen}\left(\frac{\alpha}{2}\right)\cos\left(\frac{\alpha}{2}\right) = 2\sigma\gamma$ y $1-\cos\alpha = 2\operatorname{sen}^2\left(\frac{\alpha}{2}\right) = 2\sigma^2$, la ecuación (4.10) se escribe como

$$q \cdot q_{\mathbf{v}} \cdot q^{-1} = \mathbf{v} + (\operatorname{sen}\alpha)\mathbf{n} \times \mathbf{v} + (1-\cos\alpha)\mathbf{n} \times (\mathbf{n} \times \mathbf{v}). \tag{4.11}$$

El lado derecho de la ecuación (4.11) es equivalente a aplicar la siguiente matriz al vector $\mathbf{v}$,

$$A = I + \operatorname{sen}\theta\Lambda_{\mathbf{n}} + (1-\cos\theta)(\Lambda_{\mathbf{n}})^2, \tag{4.12}$$

donde

$$\Lambda_{\mathbf{n}} = \begin{pmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{pmatrix}.$$

Utilizando ahora la identidad $\mathbf{n} \times (\mathbf{n} \times \mathbf{v}) = \langle \mathbf{n}, \mathbf{v}\rangle\mathbf{n} - \langle \mathbf{n}, \mathbf{n}\rangle\mathbf{v} = \langle \mathbf{n}, \mathbf{v}\rangle\mathbf{n} - \mathbf{v}$ y tomando en cuenta que $\langle \mathbf{n}, \mathbf{n}\rangle = 1$, la ecuación (4.11) se escribe como

$$q \cdot q_{\mathbf{v}} \cdot q^{-1} = \cos\alpha\,\mathbf{v} + \operatorname{sen}\alpha\,(\mathbf{n} \times \mathbf{v}) + (1-\cos\alpha)\,\langle \mathbf{n}, \mathbf{v}\rangle\mathbf{n}. \tag{4.13}$$

Así, $q \cdot q_{\mathbf{v}} \cdot q^{-1}$ es un cuaternio con parte real igual a cero y por medio de la correspondencia definida arriba, de (4.13) se tiene

$$q \cdot q_{\mathbf{v}} \cdot q^{-1} \longmapsto \cos\alpha\,\mathbf{v} + \operatorname{sen}\alpha\,(\mathbf{n} \times \mathbf{v}) + (1-\cos\alpha)\,\langle \mathbf{n}, \mathbf{v}\rangle\mathbf{n} \in \mathbb{R}^3. \tag{4.14}$$

Al comparar el lado derecho de (4.14) con la fórmula de Rodrigues (4.6), vemos que son exactamente lo mismo. Por lo tanto,

$$q \cdot q_{\mathbf{v}} \cdot q^{-1} \equiv q_{\mathbf{u}} \longmapsto \mathbf{u} = \cos\alpha\,\mathbf{v} + \operatorname{sen}\alpha\,(\mathbf{n} \times \mathbf{v}) + (1-\cos\alpha)\,\langle \mathbf{n}, \mathbf{v}\rangle\mathbf{n}. \tag{4.15}$$

Con esto se prueba que el producto de cuaternios $q \cdot q_{\mathbf{v}} \cdot q^{-1}$, con $q_{\mathbf{v}}$ y q dados por (4.8) y (4.9), respectivamente, representa una rotación, por un ángulo α, del vector $\mathbf{v}$ alrededor del eje $\mathbf{n}$.

Consideremos el cuaternio $q = [q_0, \mathbf{q}]$, con $\|q\| = 1$. Si $\mathbf{v} \in \mathbb{R}^3$, utilizando la ecuación (4.7) llegamos a una expresión equivalente a la ecuación (4.15), que será de gran utilidad más adelante, la cual está dada por

$$qq_{\mathbf{v}}q^{-1} = [0, (q_0^2 - \|q\|^2)\mathbf{v} + 2(\mathbf{q} \cdot \mathbf{v})\mathbf{q} + 2q_0(\mathbf{q} \times \mathbf{v})]. \tag{4.16}$$

La matriz de rotación R correspondiente al cuaternio

$$\begin{aligned} q = [\gamma, \sigma\mathbf{n}] &= \gamma + \sigma\, n_1\, i + \sigma\, n_2\, j + \sigma\, n_3\, k \\ &= \cos(\alpha/2) + \operatorname{sen}(\alpha/2)n_1 i + \operatorname{sen}(\alpha/2)n_2 j + \operatorname{sen}(\alpha/2)n_3 k \\ &\equiv \gamma + xi + yj + zk, \end{aligned} \tag{4.17}$$

se calcula del lado derecho de (4.15) para obtener $\mathbf{u} = R\,\mathbf{v}$, donde

$$R = \begin{pmatrix} 1-2y^2-2z^2 & 2xy-2\gamma z & 2xz+2\gamma y \\ 2xy+2\gamma z & 1-2x^2-2z^2 & 2yz-2\gamma x \\ 2xz-2\gamma y & 2yz+2\gamma x & 1-2x^2-2y^2 \end{pmatrix}, \tag{4.18}$$

la cual es precisamente la matriz de rotación $R(\alpha, \mathbf{n})$ obtenida en (4.2), como se puede ver fácilmente al sustituir γ, x, y, z en (4.18) por sus valores dados por (4.17) y tomando en cuenta que $\gamma^2 + x^2 + y^2 + z^2 = 1$.

Otra representación de la matriz R es la siguiente. Al sustituir $q = (q_0, q_1, q_2, q_3)$ por

$$\begin{aligned} q_0 &= \cos(\theta/2), \\ q_1 &= n_1 \operatorname{sen}(\theta/2), \\ q_2 &= n_2 \operatorname{sen}(\theta/2), \\ q_1 &= n_3 \operatorname{sen}(\theta/2), \end{aligned}$$

y tomando en cuenta que $n_1^2 + n_2^2 + n_3^2 = 1$, tenemos que

$$R(q) = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2q_1q_2 - 2q_0q_3 & 2q_1q_3 + 2q_0q_2 \\ 2q_1q_2 + 2q_0q_3 & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2q_2q_3 - 2q_0q_1 \\ 2q_1q_3 - 2q_0q_2 & 2q_2q_3 + 2q_0q_1 & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix}. \tag{4.19}$$

Por otra parte, la composición de rotaciones se determina fácilmente mediante el producto de cuaternios. En efecto, si p y q son cuaternios unitarios que representan rotaciones y si $q_{\mathbf{v}}$ es el cuaternio identificado con el vector $\mathbf{v}$, entonces la rotación que define q se logra mediante la identificación del cuaternio $q_{\mathbf{u}} = q \cdot q_{\mathbf{v}} \cdot q^{-1}$ con el vector $\mathbf{u}$, como se mostró con anterioridad. Este, a su vez, es modificado por la rotación representada por p:

$$p \cdot q_{\mathbf{u}} \cdot p^{-1} = p \cdot (q \cdot q_{\mathbf{v}} \cdot q^{-1}) \cdot p^{-1} = (pq) \cdot q_{\mathbf{v}} \cdot (q^{-1}p^{-1}) = (p \cdot q) \cdot q_{\mathbf{v}} \cdot (p \cdot q)^{-1}. \tag{4.20}$$

Esta ecuación muestra que la composición de rotaciones en $\mathbb{R}^3$ se puede representar por medio del producto de cuaternios unitarios $p \cdot q$.

A continuación enunciamos un resultado que nos será de gran utilidad en el capítulo siguiente:

Lema 4.2 *Sea $q = [q_0, \mathbf{q}] \in \mathbb{H}$ con $\|q\| = 1$. Si $p = [0, \mathbf{p}] = q * [0, e_i] * q^{-1}$, donde e_i es cualquier elemento de la base estándar de $\mathbb{R}^3$, entonces*

$$\mathrm{d}p = \mathrm{d}q * [0, e_i] * q^{-1} + q * [0, e_i] * \mathrm{d}q^{-1}$$

Demostración. Procedemos a realizar los cálculos,

$$\begin{aligned} p = [0, \mathbf{p}] &= q * [0, e_i] * q^{-1} \\ &= [q_0, \mathbf{q}] * [0, e_i] * [q_0, -\mathbf{q}] \\ &= [(-\mathbf{q} \cdot e_i)q_0 + \mathbf{q} \cdot (q_0 e_i + \mathbf{q} \times e_i), (\mathbf{q} \cdot e_i)\mathbf{q} + q_0^2 e_i + 2q_0 \mathbf{q} \times e_i], \end{aligned}$$

luego $\mathrm{d}p = [\mathrm{d}A, \mathrm{d}B]$ donde

$$\mathrm{d}A = \mathrm{d}q_0(-\mathbf{q} \cdot e_i) + q_0(-e_i \cdot \mathrm{d}\mathbf{q}) + \mathrm{d}\mathbf{q} \cdot (q_0 e_i + \mathbf{q} \times e_i) + \mathbf{q} \cdot (\mathrm{d}q_0 e_i + \mathrm{d}\mathbf{q} \times e_i)$$

y

$$\mathrm{d}B = (\mathrm{d}\mathbf{q} \cdot e_i)\mathbf{q} + (\mathbf{q} \cdot e_i)\mathrm{d}\mathbf{q} + 2q_0 \mathrm{d}q_0 e_i + 2\mathrm{d}q_0 \mathbf{q} \times e_i + 2q_0 \mathrm{d}\mathbf{q} \times e_i.$$

Ahora, calculamos $\mathrm{d}q * [0, e_i] * q^{-1}$,

$$\begin{aligned}\mathrm{d}q * [0, e_i] * q^{-1} = [&q_0(-\mathrm{d}\mathbf{q} \cdot e_i) + \mathbf{q} \cdot (\mathrm{d}q_0 e_i + \mathrm{d}\mathbf{q} \times e_i),\\ &(-\mathrm{d}\mathbf{q} \cdot e_i)(-\mathbf{q}) + q_0(\mathrm{d}q_0 e_i + \mathrm{d}\mathbf{q} \times e_i) +\\ &(\mathrm{d}q_0 e_i + \mathrm{d}\mathbf{q} \times e_i) \times (-\mathbf{q})].\end{aligned} \tag{4.21}$$

Por último, calculamos $q * [0, e_i] * \mathrm{d}q^{-1}$,

$$\begin{aligned}q * [0, e_i] * \mathrm{d}q^{-1} = [&(-e_i \cdot \mathbf{q})\mathrm{d}q_0 + \mathrm{d}\mathbf{q} \cdot (q_0 e_i + \mathbf{q} \times e_i),\\ &(-e_i \cdot \mathbf{q})(\mathrm{d}\mathbf{q}) + \mathrm{d}q_0(q_0 e_i + \mathbf{q} \times e_i) +\\ &(q_0 e_i + \mathbf{q} \times e_i) \times (-\mathrm{d}\mathbf{q})].\end{aligned} \tag{4.22}$$

Sumando las partes escalares de (4.21) y (4.22) obtenemos dA y al sumar las partes escalares de (4.21) y (4.22) obtenemos dB como queríamos. ■

4.4 El operador cuaterniónico de rotación

Hemos observado que si $q \in \mathbb{H}$ con $\|q\| = 1$ y $v \in \mathbb{H}_{\mathrm{p}}$, el producto

$$qvq^{-1},$$

representa una rotación en $SO(3)$. Esta asignación es muy importante y es necesario tratarla de manera especial.

Definimos el *operador cuaterniónico de rotación*, ρ_q, asociado con el cuaternio q y el cual se aplica a un vector $\mathbf{v} \in \mathbb{R}^3$ correspondiente al cuaternio puro v, por la ecuación

$$w = \rho_q(v) = qvq^{-1}. \tag{4.23}$$

En esta sección, observaremos dos propiedades algebraicas del operador ρ_q. La primera de ellas es que el operador ρ_q es *lineal*, como lo indica el siguiente resultado.

Proposición 4.3 *El operador ρ_q es lineal. Esto es, para cualesquiera $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ y para cualquier escalar $k \in \mathbb{R}$, tenemos que*

$$\rho_q(ka + b) = k\rho_q(a) + \rho_q(b),$$

donde a y b son los cuaternios puros asociados a los vectores $\mathbf{a}$ y $\mathbf{b}$, respectivamente.

Demostración. Para probar este resultado procedemos como sigue. Recordemos que el producto de cuaternios es asociativo, así

$$\begin{aligned}\rho_q(ka + b) &= q(ka + b)q^{-1}\\ &= (kqa + qb)q^{-1}\\ &= kqaq^{-1} + qbq^{-1}\\ &= k\rho_q(a) + \rho_q(b).\end{aligned}$$

■

La segunda propiedad algebraica es que la norma o longitud de un vector es invariante bajo el operador ρ_q, esto es, que $\|\rho_q(v)\| = \|\mathbf{v}\|$, con $\mathbf{v} \in \mathbb{R}^3$. Esta propiedad se concluye directamente utilizando el Teorema 3.1.

La siguiente proposición establece que el operador ρ_q representa una rotación en $\mathbb{R}^3$.

Proposición 4.4 *Para cualquier cuaternio unitario*

$$q = [q_0, \mathbf{q}] = \cos\theta + \mathbf{u}\operatorname{sen}\theta$$

y cualquier vector $\mathbf{v} \in \mathbb{R}^3$*, la acción del operador*

$$\rho_q(v) = qvq^{-1},$$

donde v *es el cuaternio puro asociado al vector* $\mathbf{v}$*, sobre el vector* $\mathbf{v}$*, se interpreta geométricamente como una rotación del vector* $\mathbf{v}$ *por un ángulo* 2θ *alrededor del vector* $\mathbf{q}$*.*

Demostración. Escribimos el vector $\mathbf{v}$ en la forma

$$\mathbf{v} = \mathbf{a} + \mathbf{n},$$

donde $\mathbf{a}$ es la componente de $\mathbf{v}$ que se encuentra sobre la parte vectorial del cuaternio q, y $\mathbf{n}$ es la componente de $\mathbf{v}$ la cual es normal a la parte vectorial de q.

Como el vector $\mathbf{a}$ está sobre el vector $\mathbf{q}$, $\mathbf{a}$ es un múltiplo de $\mathbf{q}$, esto es

$$\mathbf{a} = k\mathbf{q}$$

para algún escalar $k \in \mathbb{R}$. Luego, utilizando la ecuación (4.16),

$$\rho_q(\mathbf{a}) = \rho_q(k\mathbf{q}) = k\mathbf{q} = \mathbf{a}.$$

Sólo nos falta mostrar que el operador ρ_q rota la componente $\mathbf{n}$ un ángulo 2θ alrededor de $\mathbf{q}$. Esto se verifica utilizando la ecuación (4.16), reemplazando $\mathbf{v}$ por $\mathbf{n}$. De esta manera

$$\begin{aligned}\rho_q(\mathbf{n}) &= (q_0^2 - \|\mathbf{q}\|^2)\mathbf{n} + 2(\mathbf{q}\cdot\mathbf{n})\mathbf{q} + 2q_0(\mathbf{q}\times\mathbf{n})\\ &= (q_0^2 - \|\mathbf{q}\|^2)\mathbf{n} + 2q_0(\mathbf{q}\times\mathbf{n})\\ &= (q_0^2 - \|\mathbf{q}\|^2)\mathbf{n} + 2q_0\|\mathbf{q}\|(\mathbf{u}\times\mathbf{n}).\end{aligned}$$

Aquí utilizamos el hecho de que $\mathbf{u} = \mathbf{q}/\|\mathbf{q}\|$. Si escribimos $\mathbf{u}\times\mathbf{n} = \mathbf{n}_\perp$, tenemos que

$$\rho_q(\mathbf{n}) = (q_0^2 - \|\mathbf{q}\|^2)\mathbf{n} + 2q_0\|\mathbf{q}\|\mathbf{n}_\perp. \tag{4.24}$$

Ahora veamos que $\|\mathbf{n}\| = \|\mathbf{n}_\perp\|$. Como el ángulo entre $\mathbf{n}$ y $\mathbf{n}_\perp$ es $\pi/2$, entonces

$$\|\mathbf{n}_\perp\| = \|\mathbf{n}\times\mathbf{u}\| = \|\mathbf{n}\|\|\mathbf{u}\|\operatorname{sen}\pi/2 = \|\mathbf{n}\|. \tag{4.25}$$

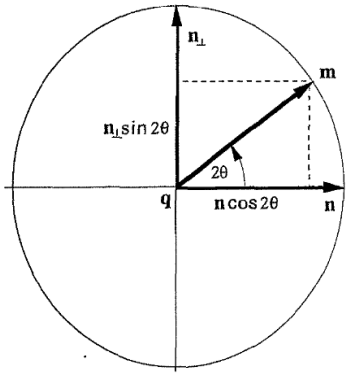


Figura 4.3: Componentes del vector $\rho_q(\mathbf{n})$.

Finalmente, utilizando la forma trigonométrica del cuaternio q, podemos escribir la ecuación (4.24) en la forma

$$\begin{aligned}\rho_q(\mathbf{n}) &= (\cos^2\theta - \operatorname{sen}^2\theta)\mathbf{n} + (2\cos\theta\operatorname{sen}\theta)\mathbf{n}\\ &= \cos 2\theta\mathbf{n} + \operatorname{sen} 2\theta\mathbf{n}_\perp.\end{aligned}$$

Las componentes de $\rho_q(\mathbf{n})$ se ilustran en la Figura 4.3. Hasta este punto, hemos probado que

$$\begin{aligned}\mathbf{w} = q\mathbf{v}q^{-1} &= \rho_q(\mathbf{v})\\ &= \rho_q(\mathbf{a}+\mathbf{n})\\ &= \rho_q(\mathbf{a}) + \rho_q(\mathbf{n})\\ &= \mathbf{a}+\mathbf{m},\end{aligned}$$

donde $\mathbf{m} = \cos 2\theta\mathbf{n} + \operatorname{sen} 2\theta\mathbf{n}_\perp$. Utilizando la ecuación (4.25), obtenemos $\|\mathbf{m}\| = \|\mathbf{n}\| = \|\mathbf{n}_\perp\|$. Esto nos dice que $\mathbf{m}$ es una rotación de $\mathbf{n}$ por un ángulo 2θ. Como $\mathbf{w} = \mathbf{a}+\mathbf{m}$, es claro que $\mathbf{w} = q\mathbf{v}q^{-1}$ se puede interpretar como la rotación del vector $\mathbf{v}$ por un ángulo 2θ alrededor de $\mathbf{q}$. ■

Notemos, en la Figura 4.4, que el vector $\mathbf{v}$ y su imagen $\mathbf{w}$, bajo el operador ρ_q, se pueden ver como generadores del cono circular cuyo eje es el vector $\mathbf{q}$ y cuya base circular en esta instancia contiene los vectores $\mathbf{n}$ y $\mathbf{m}$. Así, el vector $\mathbf{v}$ y su imagen $\mathbf{w}$ están relacionados por la rotación descrita en la proposición anterior.

Recordemos que el conjunto de los cuaternios unitarios, $\mathbb{S}^3$, es isomorfo a $SU(2)$. Definimos una función $\rho : SU(2) \to SO(3)$ por $\rho(q) = \rho_q$. Así, mediante el operador ρ_q podemos definir una transformación entre $SU(2)$ y $SO(3)$. Además esta función es sobreyectiva ya que cualquier rotación puede ser representada mediante esta función, es decir, cualquier punto en $SO(3)$ le corresponde un cuaternio unitario en $\mathbb{H}$. Lo que no resulta natural es que esta función es también un homomorfismo de los grupos $SU(2)$ y $SO(3)$. Esto motiva la siguiente proposición.

Proposición 4.5 *La función $\rho(q) = \rho_q$ es sobreyectiva y define un homomorfismo de grupos entre $SU(2)$ y $SO(3)$.*

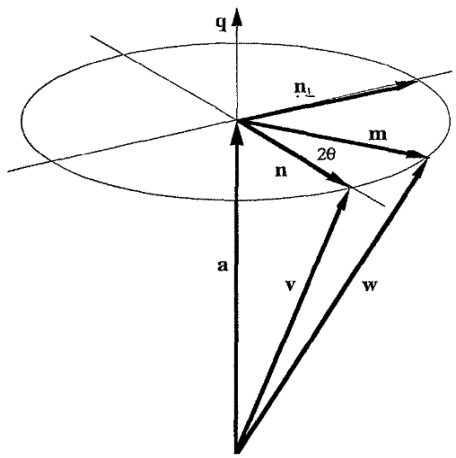


Figura 4.4: Geometría del operador ρ_q.

Demostración. En la sección anterior hemos observado que a cualquier elemento de $SO(3)$ le podemos asociar un elemento en $\mathbb{H}$ mediante la función ρ_q, por lo tanto ρ_q es sobreyectiva, de tal manera que sólo nos resta ver que es un homomorfismo. En efecto, sean $q, s \in \mathbb{H}$ con $\|q\| = \|s\| = 1$ y $t \in \mathbb{H}_{\mathrm{p}}$. Luego,

$$\begin{aligned}
\rho_{sq}(t) &= (sq)t(sq)^{-1} \\
&= (sq)t(q^{-1}s^{-1}) \\
&= s(qtq^{-1})s^{-1} \\
&= \rho_s(t) \circ \rho_q(t).
\end{aligned}$$

■

El siguiente resultado muestra la importancia de la función ρ_q.

Proposición 4.6 *$SU(2)$ es una doble cubierta de $SO(3)$.*

Demostración. Para probar este resultado tenemos que encontrar una función sobreyectiva que transforme el espacio $SU(2)$ en $SO(3)$. Dicha función es precisamente ρ. Para verificar que $SU(2)$ es una doble cubierta, consideremos la función $\rho(-q) = \rho_{-q} = (-q)x(-q)^{-1}$. Recordemos que el cuaternio $-q$ puede ser expresado como $-1 \cdot q$. Luego, dado que $-1 \in \mathbb{R}$ tenemos que $-1 \cdot q = q \cdot (-1)$. Ahora, si $x \in \mathbb{H}_{\mathrm{p}}$ tenemos

$$\begin{aligned}
\rho_{-q}(x) &= (-q)x(-q)^{-1} \\
&= q(-1)x(-1)q^{-1} \\
&= q(-1)(-1)xq^{-1} \\
&= qxq^{-1}
\end{aligned}$$

por lo tanto $\rho_q = \rho_{-q}$, es decir

$$\rho(q) = \rho(-q).$$

Esto es, una rotación por el cuaternio q representa la misma rotación que el cuaternio $-q$. Luego, como $SU(2) \cong \mathbb{S}^3$, con lo anterior mostramos que $SO(3)$ es homeomorfo a la esfera $\mathbb{S}^3$ módulo la función antípoda (la cual identifica a un punto en la esfera $\mathbb{S}^3$ con su antípoda). De esta manera $SU(2)$ es una doble cubierta de $SO(3)$. ■

Recordemos que el espacio proyectivo $\mathbb{R}\mathrm{P}^3$ es el espacio cociente que se obtiene al identificar los puntos antípoda en la esfera $\mathbb{S}^3$. De esta manera, por el resultado anterior tenemos que $SO(3)$ es homeomorfo a $\mathbb{R}\mathrm{P}^3$. Tal homeomorfismo se puede utilizar para transferir la estructura de grupo existente en $SO(3)$ a $\mathbb{R}\mathrm{P}^3$, y con esto se tiene que $\mathbb{R}\mathrm{P}^3$ es un grupo topológico. Más aún, se puede mostrar que $SO(3)$ y $\mathbb{R}\mathrm{P}^3$ son variedades difeomorfas. Así, $SO(3)$ y $\mathbb{R}\mathrm{P}^3$ son, al mismo tiempo, grupos, espacios topológicos, variedades y por último, grupos de Lie.

De la Sección 1.3, observamos que la función ρ tiene un comportamiento similar a la fibración de Hopf. Esto no es una coincidencia, pues la función ρ y la fibración de Hopf son equivalentes.

Como se ha observado, la función ρ_q representa una rotación. Sea $\mathbf{x} \in \mathbb{R}^3$ y $x \in \mathbb{H}$ el cuaternio asociado a tal vector de $\mathbb{R}^3$. Al ángulo que forman el vector $\mathbf{x}$ y el vector asociado al cuaternio $\rho_q(x)$ le llamaremos el *ángulo de* $\rho_q(x)$.

Recuperación del eje de rotación y el ángulo de giro. Ya sabemos cómo obtener la matriz asociada a una rotación por un eje unitario y un ángulo dados mediante cuaternios, pero surge ahora el problema de cómo obtener el cuaternio asociado a una matriz de rotación dada. Para resolver este problema se procede como sigue.

Siempre que $\operatorname{sen}\theta$ sea distinto de cero y lo suficientemente grande como para evadir errores numéricos, el eje de rotación $\mathbf{n}$ puede calcularse utilizando la ecuación (4.2) y restando su traspuesta para encontrar

$$R - R^{\intercal} = \begin{bmatrix} 0 & -2n_3 \operatorname{sen}\theta & 2n_2 \operatorname{sen}\theta \\ 2n_3 \operatorname{sen}\theta & 0 & -2n_1 \operatorname{sen}\theta \\ -2n_2 \operatorname{sen}\theta & 2n_1 \operatorname{sen}\theta & 0 \end{bmatrix} = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix},$$

calculamos $d = \sqrt{a^2 + b^2 + c^2}$ y normalizamos para obtener el resultado

$$\mathbf{n} = \left(\frac{a}{d}, \frac{b}{d}, \frac{c}{d}\right).$$

Para obtener el ángulo de giro, volvemos a la ecuación (4.2) y al examinar su traza encontramos que

$$\operatorname{tr}(R) = 1 + 2\cos\theta. \tag{4.26}$$

De esta manera, el ángulo lo podemos obtener de la ecuación

$$\cos\theta = \frac{1}{2}(\mathrm{tr}(R) - 1),$$

así

$$\theta = \arccos\left[\frac{\mathrm{tr}(R) - 1}{2}\right].$$

Esta forma de obtener el ángulo no siempre es tan precisa, por lo que es necesario hacer los cálculos de una manera más rigurosa. Por ejemplo, si $\theta = \pi$, es decir, al rotar 180 grados, la forma de la matriz $R - R^T$ ya no es útil, pues en este caso es una matriz cuyos elementos son todos cero.

El procedimiento básico general para extraer el cuaternio $q = (q_0, q_1, q_2, q_3)$ asociado a una matriz de rotación R depende del hecho, citado en [11], de que siempre habrá al menos un elemento de la diagonal de R que será "grande". Utilizando la ecuación (4.26) y la fórmula de mitad de ángulo, tenemos que el valor de q_0 es

$$q_0 = \cos\frac{\theta}{2} = \frac{1}{\sqrt{2}}\sqrt{\cos\theta + 1} = \frac{1}{2}\sqrt{\mathrm{tr}(R) + 1}. \tag{4.27}$$

Como un cuaternio y su negativo producen la misma matriz de rotación, podemos escoger la raíz positiva de la ecuación (4.27), siempre que se mantenga la parte vectorial del cuaternio con este signo.

Supongamos que $\mathrm{tr}(R) > 0$. Entonces, como $q_0 > \frac{1}{2}$, no tenemos problema al dividir por q_0, de esta manera si usamos la ecuación (4.19) obtenemos las relaciones

$$\begin{aligned} R_{32} - R_{23} &= 4q_0q_1, \\ R_{13} - R_{31} &= 4q_0q_2, \\ R_{21} - R_{12} &= 4q_0q_3, \end{aligned}$$

así,

$$\begin{aligned} q_1 &= \frac{R_{32} - R_{23}}{4q_0}, \\ q_2 &= \frac{R_{13} - R_{31}}{4q_0}, \\ q_3 &= \frac{R_{21} - R_{12}}{4q_0}, \end{aligned}$$

Por lo que el cuaternio $q = (q_0, q_1, q_2, q_3)$ queda totalmente determinado.

Supongamos ahora que $\mathrm{tr}(R) \leq 0$. En este caso, no sabemos con certeza si q_0 es un buen divisor, pues este podría ser un número cercano a cero. Bajo estas condiciones, examinamos las entradas de la diagonal de la matriz de la ecuación (4.19). Si R_{11} es el elemento mayor de las entradas R_{ii}, entonces q_1 es el mayor de las componentes de la parte vectorial de q. Repitiendo este razonamiento para cada caso, es decir, para R_{22} y R_{33}, podemos obtener los valores para q_i de los elementos

de la diagonal de R, de tal manera que

$$q_1 = \frac{1}{2}\sqrt{1 + R_{11} - R_{22} - R_{33}},$$
$$q_2 = \frac{1}{2}\sqrt{1 + R_{22} - R_{33} - R_{11}},$$
$$q_3 = \frac{1}{2}\sqrt{1 + R_{33} - R_{11} - R_{22}}.$$

De esta información, finalmente podemos determinar los valores restantes mediante las siguientes condiciones:

- Si q_1 es el elemento de mayor magnitud, entonces
$$q_0 = \frac{R_{32} - R_{23}}{4q_1}, \quad q_2 = \frac{R_{21} + R_{12}}{4q_1}, \quad q_3 = \frac{R_{13} + R_{31}}{4q_1}.$$
- Si q_2 es el elemento de mayor magnitud, entonces
$$q_0 = \frac{R_{13} - R_{31}}{4q_2}, \quad q_1 = \frac{R_{12} + R_{21}}{4q_2}, \quad q_3 = \frac{R_{23} + R_{32}}{4q_2}.$$
- Si q_3 es el elemento de mayor magnitud, entonces
$$q_0 = \frac{R_{21} - R_{12}}{4q_3}, \quad q_1 = \frac{R_{31} + R_{13}}{4q_3}, \quad q_2 = \frac{R_{23} + R_{32}}{4q_3}.$$

Los siguientes resultados nos ayudarán justificar el cálculo del ángulo de rotación.

Lema 4.7 *Dados dos cuaternios z, w, no nulos, entonces el ángulo de la rotación ρ_z es igual al ángulo de la rotación $\rho_{wzw^{-1}}$.*

Demostración. Para probar este resultado procedemos como sigue. Primero, sea $z = a\mathbf{1} + r$ con $a \in \mathbb{R}, \mathbf{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in SU(2), r \in \mathbb{H}_p, r \neq 0$, probaremos que el eje de la rotación $\rho_{wzw^{-1}}$ es $wrw^{-1} = \rho_w(r)$. En efecto, ya sabemos que un cuaternio de la forma wrw^{-1} es puro, y

$$wzw^{-1} = w(a\mathbf{1} + r)w^{-1} = w(a\mathbf{1})w^{-1} + w(r)w^{-1} = a\mathbf{1} + w(r)w^{-1}.$$

Luego, dado cualquier cuaternio no nulo x, ortogonal a r, el ángulo de la rotación z es el ángulo entre x y $\rho_z(x)$. Como las rotaciones preservan la orientación (pues preservan el producto cruz), el ángulo θ entre dos vectores $\mathbf{x}$ y $\mathbf{y}$ es preservado bajo rotación. Ahora, como el producto interior también es preservado bajo rotaciones, si $x \cdot r = 0$, tenemos $\rho_w(x) \cdot \rho_w(r) = 0$, y el ángulo de la rotación $\rho_{wzw^{-1}} = \rho_w \circ \rho_z \circ (\rho_w)^{-1}$ es el ángulo entre los dos vectores $\rho_w(x)$ y $\rho_{wzw^{-1}}(\rho_w(x))$. Luego

$$\rho_{wzw^{-1}}(\rho_w(x)) = (\rho_w \circ \rho_z \circ (\rho_w)^{-1} \circ \rho_w)(x) \qquad = (\rho_w \circ \rho_z)(x) = \rho_w(\rho_z(x)),$$

el ángulo de la rotación $\rho_{wzw^{-1}}$ es el ángulo entre los dos vectores $\rho_z(x)$ y $\rho_w(\rho_z(x))$. Como las rotaciones preservan ángulos, éste también es el ángulo entre los dos vectores x y $\rho_z(x)$, el cual es el ángulo de la rotación ρ_z, como se había afirmado. ■

Con este hecho estamos en posición de probar el siguiente resultado,

Proposición 4.8 *Para cualquier cuaternio* $z = a\mathbf{1} + r$, *donde* $\mathbf{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in SU(2), a \in \mathbb{R}$ *y* $r \in \mathbb{H}_p, r \neq 0$, *el eje de la rotación* ρ_z *asociado con* z *está determinado por el vector en* $\mathbb{R}^3$ *correspondiente a* r, *y el ángulo de rotación* θ *es igual a* π *cuando* $a = 0$, *o cuando* $a \neq 0$ *el ángulo de rotación está dado por*

$$\tan\frac{\theta}{2} = \frac{\|r\|}{|a|},$$

con $0 < \theta \leq \pi$.

Demostración. Realizando los cálculos uno llega a que la recta que define r es invariante bajo la rotación ρ_z, y de esta manera es el eje de rotación. Note que para cualesquiera dos vectores no nulos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ tales que $\|\mathbf{x}\| = \|\mathbf{y}\|$, existe alguna rotación ρ tal que $\rho(\mathbf{x}) = \mathbf{y}$. Si $\mathbf{x} = \mathbf{y}$ usamos la identidad, y si $\mathbf{x} \neq \mathbf{y}$, usamos la rotación cuyo eje lo determina $\mathbf{x} \times \mathbf{y}$ y rotamos $\mathbf{x}$ a $\mathbf{y}$ en el plano que contiene a $\mathbf{x}$ y $\mathbf{y}$. Así, dados cualesquiera dos cuaternios puros x, y tales que sus normas coinciden, existe un cuaternio no nulo w tal que $y = wxw^{-1}$. Por el Lema 4.7, dados dos cuaternios no nulos arbitrarios z, w, afirmamos que el ángulo de la rotación ρ_z es el mismo que el ángulo de la rotación $\rho_{wzw^{-1}}$. Así, dado cualquier cuaternio $z = a\mathbf{1}+r$, donde $r \in \mathbb{H}_p, r \neq 0$, como existe algún cuaternio no nulo w tal que $wrw^{-1} = \|r\|i$ e $wzw^{-1} = a\mathbf{1} + \|r\|i$, es suficiente encontrar el ángulo de rotación para un cuaternio z de la forma $a\mathbf{1} + bi$ (una rotación del eje i). Es suficiente encontrar el ángulo de rotación entre j y $\rho_z(j)$, y como

$$\rho_z(j) = (a\mathbf{1} + bi)j(a\mathbf{1} + bi)^{-1},$$

obtenemos

$$\rho_z(j) = \frac{1}{a^2 + b^2}(a\mathbf{1} + bi)j(a\mathbf{1} - bi) = \frac{a^2 - b^2}{a^2 + b^2}j + \frac{2ab}{a^2 + b^2}k.$$

Entonces si $a \neq 0$, debemos tener que

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

y como

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

bajo una cierta orientación del plano ortogonal al eje de rotación, obtenemos

$$\tan\frac{\theta}{2} = \frac{b}{|a|} = \frac{\|r\|}{|a|}.$$

Si $a = 0$, obtenemos

$$\rho_z(j) = -j,$$

y $\theta = \pi$. ■

Capítulo 5

Marcos de referencia en $\mathbb{R}^3$ y cuaternios

La teoría de marcos en $\mathbb{R}^3$ es área de estudio básica en geometría diferencial. Sin embargo, la teoría clásica no hace uso de las ventajas de los cuaternios, las cuales pueden ser utilizadas para clarificar la naturaleza de los marcos de referencia. En este capítulo estudiaremos dos de los marcos de referencia clásicos, el marco de Frenet-Serret y el marco de transporte paralelo, y después introduciremos la formulación equivalente en cuaternios.

Marco de referencia para una curva en general La evolución de todos los posibles marcos de referencia para una curva $\alpha(t)$ en $\mathbb{R}^3$ pueden ser escritos en un modelo unificado. La idea básica consiste en considerar la representación de un marco arbitrario en forma de columnas de una matriz ortonormal de rotación de 3×3:

$$\textbf{Marco} = [\widehat{T} \quad \widehat{N}_1 \quad \widehat{N}_2],$$

donde $\widehat{T}(t) = \frac{\alpha'(t)}{\|\alpha'(t)\|}$ es el vector tangente unitario de la curva en el punto t, el cual queda determinado directamente de la geometría de la curva y de esta manera es inalterable. $(\widehat{N}_1(t), \widehat{N}_2(t))$ es un par de vectores ortonormales que generan el plano perpendicular al vector tangente en cada punto de la curva. Como $\{\widehat{T}, \widehat{N}_1, \widehat{N}_2\}$ forman un conjunto ortonormal, cualquier cambio en un vector de este conjunto debe ser ortogonal a él mismo y por lo tanto se puede expresar como combinación lineal de los otros dos vectores. Así, forma general para la ecuación que describe la evolución del marco está dada por

$$[\widehat{T}'(t) \quad \widehat{N}_1'(t) \quad \widehat{N}_2'(t)] = [\widehat{T}(t) \quad \widehat{N}_1(t) \quad \widehat{N}_2(t)]v(t) \begin{bmatrix} 0 & -k_y(t) & k_x(t) \\ k_y(t) & 0 & -k_z(t) \\ -k_x(t) & k_z(t) & 0 \end{bmatrix}, \tag{5.1}$$

donde $v(t) = \|\alpha'(t)\|$ es la velocidad de la curva si no estamos utilizando una parametrización de rapidez unitaria.

Las propiedades naturales del marco de referencia de la curva α son también vistas usando la forma de Darboux de las ecuaciones

$$\begin{aligned} \widehat{T}' &= v(t)F \times \widehat{T}, \\ \widehat{N}_1' &= v(t)F \times \widehat{N}_1, \\ \widehat{N}_2' &= v(t)F \times \widehat{N}_2, \end{aligned} \tag{5.2}$$

donde F generaliza el campo vectorial de Darboux:

$$F = k_x\widehat{N}_1 + k_y\widehat{N}_2 + k_z\widehat{T}.$$

La norma al cuadrado del total de la "fuerza" que actúa sobre el marco es $\|F\|^2 = k_x^2+k_y^2+k_z^2$, y más adelante veremos que esta es mínima para el marco de transporte paralelo. La arbitrariedad de la base $(\widehat{N}_1(t), \widehat{N}_2(t))$ para el plano perpendicular a $\widehat{T}(t)$ se puede explotar para eliminar cualesquiera de las cantidades (k_x, k_y, k_z). Por ejemplo, si

$$\begin{aligned}\widehat{M_1} &= \widehat{N}_1\cos\theta - \widehat{N}_2\operatorname{sen}\theta,\\ \widehat{M_2} &= \widehat{N}_1\operatorname{sen}\theta + \widehat{N}_2\cos\theta,\end{aligned}$$

al derivar $\widehat{M_1}$ y haciendo uso de la ecuación (5.1) tenemos

$$\begin{aligned}\widehat{M_1}' &= \widehat{N}_1'\cos\theta - \widehat{N}_1\operatorname{sen}\theta\frac{\mathrm{d}\theta}{\mathrm{d}t} - \widehat{N}_2'\operatorname{sen}\theta - \widehat{N}_2\cos\theta\frac{\mathrm{d}\theta}{\mathrm{d}t}\\ &= -(\widehat{N}_1\operatorname{sen}\theta + \widehat{N}_2\cos\theta)\frac{\mathrm{d}\theta}{\mathrm{d}t} - k_yv\cos\theta\widehat{T} + k_zv\cos\theta\widehat{N}_2 - k_xv\operatorname{sen}\theta\widehat{T} + k_zv\operatorname{sen}\theta\widehat{N}_1\\ &= -\widehat{M_2}\frac{\mathrm{d}\theta}{\mathrm{d}t} + k_zv(\widehat{N}_1\operatorname{sen}\theta + \widehat{N}_2\cos\theta) - (k_yv\cos\theta + k_xv\operatorname{sen}\theta)\widehat{T}\\ &= \widehat{M_2}(k_zv - \theta') - \widehat{T}(k_yv\cos\theta + k_xv\operatorname{sen}\theta).\end{aligned}$$

Ahora, al derivar $\widehat{M_2}$ y haciendo uso nuevamente de la ecuación (5.1) llegamos a la relación

$$\widehat{M_2}' = \widehat{T}(k_xv\cos\theta - k_yv\operatorname{sen}\theta) - \widehat{M_1}(k_zv - \theta').$$

Así, el ángulo $\theta(t)$ se puede escoger de tal manera que cancele la velocidad angular k_z en el plano generado por $\widehat{N}_1$ y $\widehat{N}_2$. Este mismo argumento se mantiene para cualquier otro par.

El marco de Frenet-Serret. El marco clásico de Frenet-Serret, comúnmente referido como el marco de Frenet, está definido para todo punto de la curva excepto para cuando la curvatura se hace cero, es decir, cuando la curva es una línea recta o tiene un punto de inflexión. Para el marco de Frenet, tomamos

$$\begin{aligned}k_x &= 0,\\ k_y &= \kappa(t),\\ k_z &= \tau(t),\end{aligned}$$

donde $\kappa(t)$ es la curvatura (la cual geométricamente puede ser entendida como el inverso del radio de curvatura en un punto) y $\tau(t)$ es la torsión, la cual mezcla los dos vectores normales en su plano local. Esta forma de identificar produce las ecuaciones de Frenet-Serret

$$[\widehat{T}'(t) \quad \widehat{N}'(t) \quad \widehat{B}'(t)] = [\widehat{T}(t) \quad \widehat{N}(t) \quad \widehat{B}(t)]v(t)\begin{bmatrix} 0 & -\kappa(t) & 0\\ \kappa(t) & 0 & -\tau(t)\\ 0 & \tau(t) & 0\end{bmatrix}. \tag{5.3}$$

Note que la norma al cuadrado del vector de Darboux es $\|F\|^2 = \kappa^2 + \tau^2 \geq k^2$.

Si $\alpha(t)$ es una parametrización de una curva en $\mathbb{R}^3$, entonces $\widehat{T}, \widehat{N}, \widehat{B}, \kappa(t)$ y $\tau(t)$ quedan determinadas utilizando los Teoremas 1.2 y 1.3.

Cuando la segunda derivada se anula en algún intervalo, el marco de Frenet queda indefinido. Para resolver este problema se utiliza el marco de transporte paralelo, el cual se describe a continuación.

Marco de transporte paralelo. La principal característica por la cual destaca el marco de transporte paralelo es el hecho de que usa la mínima rotación posible en cada punto sobre la curva para alinear el vector tangente actual con el siguiente vector tangente. La orientación actual del plano normal al vector tangente depende de la historia de la curva, empezando con un marco arbitrario inicial, y así esencialmente se está integrando una ecuación diferencial para el cambio del marco sobre la curva. El marco depende de las condiciones iniciales, y a diferencia del marco de Frenet no puede ser determinado localmente en la curva. El algoritmo que mejor calcula este marco ([13],[11]) involucra la determinación de la dirección normal $\widehat{N} = \widehat{T}_i \times \widehat{T}_{i+1}/\|\widehat{T}_i \times \widehat{T}_{i+1}\|$ al plano de dos tangentes sucesivas sobre la curva, encontrar el ángulo $\theta = \arccos(\widehat{T}_i \cdot \widehat{T}_{i+1})$ y rotar el marco actual al marco siguiente utilizando la matriz de rotación $R(\theta, \widehat{N})$ o su cuaternio correspondiente,

$$q(\theta, \widehat{N}) = (\cos\frac{\theta}{2}, \widehat{N}\operatorname{sen}\frac{\theta}{2}).$$

Si el siguiente vector tangente es colineal, se deja el marco igual. Si las tangentes son anticolineales, se puede obtener un resultado pero no queda únicamente determinado. Para identificar el marco de transporte paralelo con la ecuación (5.1), tomamos

$$\begin{aligned} k_y &= k_1(t), \\ -k_x &= k_2(t), \\ k_z &= 0, \end{aligned}$$

para evadir relaciones entre las componentes normales. Con esta elección se producen las ecuaciones:

$$[\widehat{T}'(t) \quad \widehat{N}_1'(t) \quad \widehat{N}_2'(t)] = v(t)[\widehat{T}(t) \quad \widehat{N}_1(t) \quad \widehat{N}_2(t)] \begin{bmatrix} 0 & -k_1(t) & -k_2(t) \\ k_1(t) & 0 & 0 \\ k_2(t) & 0 & 0 \end{bmatrix}. \quad (5.4)$$

Esta elección del marco se comporta completamente distinta al marco de Frenet pues permanece continua y bien definida para una curva con curvatura cero en un segmento.

Como $\|\widehat{T}'\|^2 = (k_1)^2 + (k_2)^2$ es un invariante independientemente de la elección de $\widehat{N}_1$ y $\widehat{N}_2$, entonces la curvatura, orientación, y velocidad angular quedan deter-

minadas por las relaciones

$$\kappa(t) = \sqrt{(k_1)^2 + (k_2)^2},$$
$$\theta(t) = \arctan\left(\frac{k_1}{k_2}\right),$$
$$\omega(t) = \frac{\mathrm{d}\theta(t)}{\mathrm{d}t}.$$

De esta manera, k_1 y k_2 corresponden a un sistema de coordenadas cartesiano para las coordenadas polares de la curvatura (κ, θ) con $\theta = \theta_0 + \int \omega(t)\mathrm{d}t$ y $\omega(t)$ es efectivamente la torsión $\tau(t)$ que aparece en las ecuaciones de Frenet. Así, una ambiguedad fundamental en el marco de transporte paralelo comparado con el marco de Frenet surge de la constante de integración arbitraria para θ, la cual desaparece de τ debido al proceso de diferenciación. Notemos que la norma al cuadrado del vector de Darboux $\|F\|^2 = \|\widehat{T}'\|^2 = k_1^2 + k_2^2 = \kappa^2$ es ahora invariante bajo el marco. No tiene la componente de torsión presente en las ecuaciones de Frenet, y de esta manera *alcanza su valor mínimo.*

Un algoritmo para calcular el marco de transporte paralelo con las propiedades deseadas se presenta a continuación

1	Calcular un marco de referencia en α_{i-1}.
2	Calcular dos vectores tangente unitarios $\widehat{T}_i = \frac{T_i}{\|T_i\|}$ y $\widehat{T}_{i-1} = \frac{T_{i-1}}{\|T_{i-1}\|}$
3	Encontrar el ángulo entre ellos $\theta = \arccos(\widehat{T}_i \cdot \widehat{T}_{i-1})$
4	Encontrar la perpendicular al plano de las tangentes dado por $\mathbf{V} = (\widehat{T}_{i-1} \times \widehat{T}_i)$
5	Rotar el marco en α_{i-1} un ángulo θ alrededor de $\mathbf{V}$ para obtener el marco en α_i

Ecuaciones genéricas para marcos cuaterniónicos. Como un marco de referencia consta de 3 vectores ortonormales en $\mathbb{R}^3$, la representación visual de la evolución marcos de referencia no es muy práctica cuando se tiene una cantidad muy grande de estos. A continuación se propone una forma alternativa de representarlos mediante cuaternios. Esto es, que se hace uso de la representación de marcos de referencia en $\mathbb{R}^3$ en cuatro dimensiones utilizando cuaternios unitarios, que de esta manera se corresponden con puntos en la esfera $\mathbb{S}^3$. Así, la evolución de una curva en $\mathbb{R}^3$ se puede transformar en una curva en el espacio de los cuaternios unitarios que se corresponde punto a punto con la curva en $\mathbb{R}^3$.

Empezamos con una fórmula que nos describe la relación entre una matriz de 3×3 R_{ij} y cuaternios (unitarios), q,

$$[R(q)\mathbf{V}]^i = \sum_j R_{ij} V_j = q * [0, \mathbf{V}_i] * q^{-1}.$$

Expresaremos cada componente ortonormal del marco como una columna de la matriz R([11],[12]) utilizando un cuaternio arbitrario q para rotar cada uno de los

tres ejes de referencia cartesianos a una nueva orientación arbitraria:

$$\begin{aligned}[0,\widehat{T}] &= q * [0,e_1] * q^{-1},\\ [0,\widehat{N}_1] &= q * [0,e_2] * q^{-1},\\ [0,\widehat{N}_2] &= q * [0,e_3] * q^{-1}.\end{aligned}$$

Abusando de notación, escribiremos $\widehat{T}$ para denotar el cuaternio $[0,\widehat{T}]$. De igual manera haremos lo propio con $\widehat{N}_1$ y $\widehat{N}_2$.

Todo esto se puede transformar en la siguiente representación explícita de los vectores del marco como columnas de una matriz correspondiente a un cuaternio:

$$[[\widehat{T}] \quad [\widehat{N}_1] \quad [\widehat{N}_2]] = \begin{bmatrix} q_0^2+q_1^2-q_2^2-q_3^2 & 2q_1q_2-2q_0q_3 & 2q_1q_3+2q_0q_2 \\ 2q_1q_2+2q_0q_3 & q_0^2-q_1^2+q_2^2-q_3^2 & 2q_2q_3-2q_0q_1 \\ 2q_1q_3-2q_0q_2 & 2q_2q_3+2q_0q_1 & q_0^2-q_1^2-q_2^2+q_3^2 \end{bmatrix}. \tag{5.5}$$

Antes de continuar, realizaremos algunos cálculos que nos serán de utilidad para determinar los cambios en el marco, es decir, las derivadas de $\widehat{T}, \widehat{N}_1$ y $\widehat{N}_2$.

Consideremos el cuaternio $q = [q_0, \mathbf{q}]$ tal que $\|q\| = 1$. Entonces

$$\begin{aligned}\mathrm{d}q &= [1,\mathbf{0}] * \mathrm{d}q\\ &= (q * q^{-1}) * \mathrm{d}q = q * (q^{-1} * \mathrm{d}q)\\ &= [q_0,\mathbf{q}] * ([q_0,-\mathbf{q}] * [\mathrm{d}q_0,\mathrm{d}\mathbf{q}])\\ &= [q_0,\mathbf{q}] * [q_0\mathrm{d}q_0 + \mathbf{q}\cdot\mathrm{d}\mathbf{q}, q_0\mathrm{d}\mathbf{q} - \mathrm{d}q_0\mathbf{q} - \mathbf{q}\times\mathrm{d}\mathbf{q}]\\ &= [q_0,\mathbf{q}] * [q\cdot\mathrm{d}q, q_0\mathrm{d}\mathbf{q} - \mathrm{d}q_0\mathbf{q} - \mathbf{q}\times\mathrm{d}\mathbf{q}]\\ &= [q_0,\mathbf{q}] * [0, q_0\mathrm{d}\mathbf{q} - \mathrm{d}q_0\mathbf{q} - \mathbf{q}\times\mathrm{d}\mathbf{q}]\\ &= q * \left[\frac{1}{2}[0,\mathbf{k}]\right]\end{aligned} \tag{5.6}$$

donde $[0,\mathbf{k}] = [0, 2(q_0\mathrm{d}\mathbf{q} - \mathrm{d}q_0\mathbf{q} - \mathbf{q}\times\mathrm{d}\mathbf{q})]$. Continuando con más cálculos,

$$\begin{aligned}\mathrm{d}q^{-1} * q &= [\mathrm{d}q_0^{-1}, \mathrm{d}\mathbf{q}^{-1}] * [q_0,\mathbf{q}]\\ &= [\mathrm{d}q_0, -\mathrm{d}\mathbf{q}] * [q_0,\mathbf{q}]\\ &= [q_0\mathrm{d}q_0 + \mathbf{q}\cdot\mathrm{d}\mathbf{q}, \mathrm{d}q_0\mathbf{q} - q_0\mathrm{d}\mathbf{q} - \mathrm{d}\mathbf{q}\times\mathbf{q}]\\ &= [0, \mathrm{d}q_0\mathbf{q} - q_0\mathrm{d}\mathbf{q} + \mathbf{q}\times\mathrm{d}\mathbf{q}],\end{aligned} \tag{5.7}$$

por otro lado

$$\begin{aligned}-(q^{-1} * \mathrm{d}q) &= -[q_0,-\mathbf{q}] * [\mathrm{d}q_0,\mathrm{d}\mathbf{q}]\\ &= -[q_0\mathrm{d}q_0 + \mathbf{q}\cdot\mathrm{d}\mathbf{q}, q_0\mathrm{d}\mathbf{q} - \mathrm{d}q_0\mathbf{q} - \mathbf{q}\times\mathrm{d}\mathbf{q}]\\ &= [0, \mathrm{d}q_0\mathbf{q} - q_0\mathrm{d}\mathbf{q} + \mathrm{d}\mathbf{q}\times\mathbf{q}].\end{aligned} \tag{5.8}$$

De (5.7) y (5.8) concluimos

$$\mathrm{d}q^{-1} * q = -(q^{-1} * \mathrm{d}q). \tag{5.9}$$

Al utilizar las ecuaciones (5.6) y (5.9) obtenemos la siguiente relación

$$\begin{aligned}
\mathrm{d}q^{-1} &= \mathrm{d}q^{-1} * [1, \mathbf{0}] = \mathrm{d}q^{-1} * (q * q^{-1}) \\
&= (\mathrm{d}q^{-1} * q) * q^{-1} \\
&= -(q^{-1} * \mathrm{d}q) * q^{-1} \\
&= -q^{-1} * q * \left[\frac{1}{2}[0, \mathbf{k}]\right] * q^{-1} \\
&= -\left[\frac{1}{2}[0, \mathbf{k}]\right] * q^{-1}. \qquad (5.10)
\end{aligned}$$

Procedemos ahora a describir la evolución del marco de referencia, es decir, calcularemos las derivadas $\mathrm{d}\widehat{T}, \mathrm{d}\widehat{N}_1$ y $\mathrm{d}\widehat{N}_2$. Para llevar a cabo esto, nos apoyaremos en el Lema 4.2 y las ecuaciones (5.6) y (5.10). De esta manera

$$\begin{aligned}
\mathrm{d}\widehat{T} &= \mathrm{d}(q * [0, e_1] * q^{-1}) = \mathrm{d}((q * [0, e_1]) * q^{-1}) \\
&= \mathrm{d}q * [0, e_1] * q^{-1} + q * [0, e_1] * \mathrm{d}q^{-1} \\
&= q * \left[\frac{1}{2}[0, \mathbf{k}]\right] * [0, e_1] * q^{-1} + q * [0, e_1] * \left(-\left[\frac{1}{2}[0, \mathbf{k}]\right] * q^{-1}\right) \\
&= q * \left(\left[\frac{1}{2}[0, \mathbf{k}]\right] * [0, e_1] - \left[\frac{1}{2}[0, \mathbf{k}]\right] * [0, e_1]\right) * q^{-1} \\
&= \frac{1}{2} q * ([-\mathbf{k} \cdot e_1, \mathbf{k} \times e_1] - [-\mathbf{k} \cdot e_1, e_1 \times \mathbf{k}]) * q^{-1} \\
&= \frac{1}{2} q * (2[0, \mathbf{k} \times e_1]) * q^{-1} = q * [0, \mathbf{k} \times e_1] * q^{-1}. \qquad (5.11)
\end{aligned}$$

De manera análoga, $\mathrm{d}\widehat{N}_1$ y $\mathrm{d}\widehat{N}_2$ quedan determinadas por

$$\begin{aligned}
\mathrm{d}\widehat{N}_1 &= q * [0, \mathbf{k} \times e_2] * q^{-1}, \\
\mathrm{d}\widehat{N}_2 &= q * [0, \mathbf{k} \times e_3] * q^{-1}. \qquad (5.12)
\end{aligned}$$

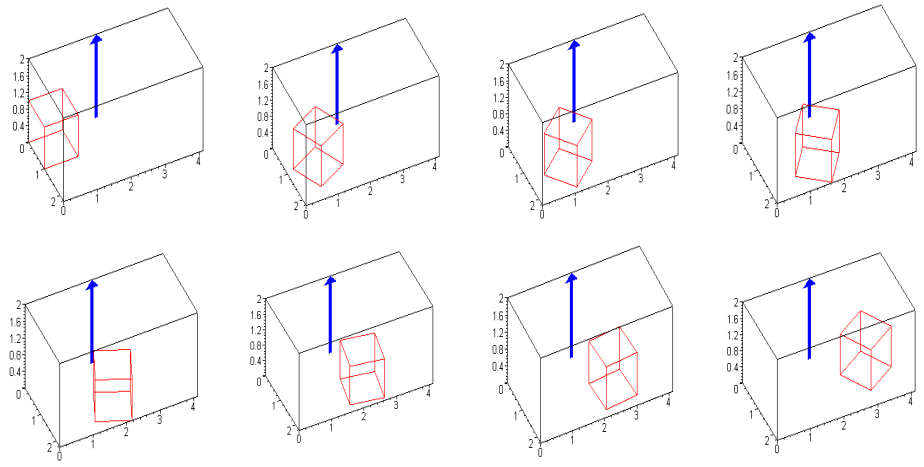
Capítulo 6

Aplicaciones

6.1 Una primera aplicación de los cuaternios: rotación de un cuerpo rígido

Como hemos visto en secciones anteriores, una característica muy importante de los cuaternios es que con éstos es posible representar rotaciones en el espacio $\mathbb{R}^3$, la cual es una alternativa a hacer lo propio con matrices de 3×3. Una de las principales ventajas al utilizar cuaternios en vez de matrices, es que se reducen las operaciones que se llevan a cabo para llegar al mismo resultado. Además, se obtiene una mayor precisión utilizando cuaternios, lo cual es de vital importancia cuando se implementa en computación.

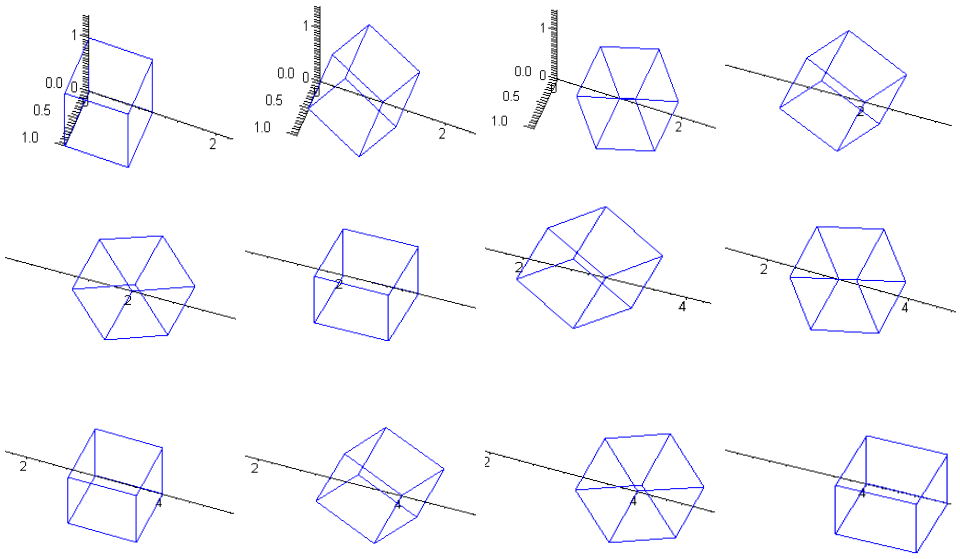
A continuación se presenta un programa implementado en Maple, el cual rota un cuerpo rígido (un cubo en este caso) alrededor de un eje dado utilizando cuaternios, como se muestra en las siguientes figuras



La animación se llevó a cabo mediante el siguiente algoritmo([14]):

1	Definir los vértices.
2	Definir el eje de rotación.
3	Definir el ángulo de rotación.
4	Crear funciones para obtener el cuaternio asociado a un vector, sumar cuaternios, restar cuaternios, multiplicar cuaternios, obtener la i-ésima coordenada de un cuaternio, calcular el inverso de un cuaternio, obtener la norma de un vector, generar el cuaternio asociado a la rotación del ángulo dado.
5	Dividir el ángulo de rotación en partes y almacenarlos en un arreglo.
6	Para cada vértice: Calcular la nueva posición del vértice con la función $R(V-P)R^{-1}+P$ para cada componente del arreglo recién creado. Crear la figura con los nuevos vértices y almacenar cada figura en un arreglo.
7	Mostrar en secuencia las figuras del arreglo.

El código del programa se anexa en el apéndice. Como ya sabemos como generar una rotación de un objeto alrededor de un eje, en el apéndice se anexa el código para generar una secuencia de rotaciones de un objeto, un cubo en este caso, como se muestra en las siguientes figuras:



6.2 Superposición de ejes

Uno de los problemas clásicos que surgen cuando se utilizan matrices de 3×3 para representar rotaciones en $\mathbb{R}^3$ es la posibilidad de que dos o mas ejes de rotación se empalmen. Tal efecto resulta cuando se tienen ejes coplanares de rotación. Cuando intentamos controlar cambios continuos de la orientación de un objeto en $\mathbb{R}^3$, podemos encontrar, esencialmente, este hecho. Si tenemos una sucesión de orientaciones

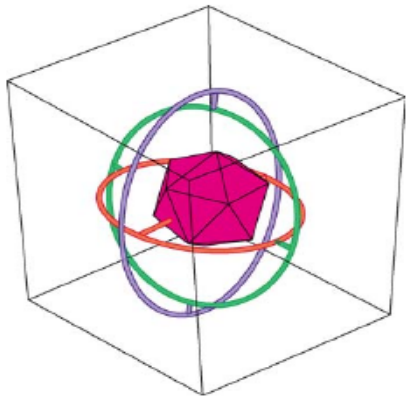


Figura 6.1: Configuración básica en los sistemas de navegación, con tres ejes de rotación ortogonales.

que cambian continuamente, y la sucesión llega a un punto en el cual los ejes de rotación se alinean en un mismo plano, no existe manera alguna de realizar una rotación alrededor del eje perpendicular a tal plano.

En los sistemas de navegación, hay un problema mecánico que ocurre precisamente cuando dos ejes se alinean. El problema es crítico cuando se tienen tres aros concéntricos de rotación (o ejes), el mínimo para simular los sistemas de navegación. En el momento en el que dos de los aros son coplanares, cualquier rotación alrededor del eje perpendicular genera una torca en el sistema de navegación, como se muestra en la figura (6.2). Este es un problema serio, pues los aros son diseados para moverse libremente y prevenir la acción de cualquier fuerza sobre el sistema de navegación. Una vez llegado este punto, o bien el sistema de navegación o el objeto que se está simulando se ve afectado por una fuerza destructiva, pues el sistema de navegación no puede cambiar la dirección sin aplicar una torca que equilibre a la fuerza externa que actúa sobre él.

Si utilizamos cuaternios para representar rotaciones en lugar de matrices, el efecto de superposición de ejes no se presenta. Esto es debido a que la rotación mediante cuaternios sólo utiliza un eje de rotación, mientras que las matrices utilizan tres. De esta manera, cuando usamos cuaternios no hay manera de que ocurra la superposición de ejes, pues sólo tenemos uno.

Así, la simulación de rotaciones mediante cuaternios es más precisa que llevarla a cabo mediante matrices.

6.3 El truco del cinturón

Un entretenimiento popular bastante conocido es el llamado *truco del cinturón*, el cual consiste en fijar un extremo de un cinturón y rotar el otro extremo un ángulo de 360 o bien de 720 grados sobre su eje más largo. La idea aquí es desdoblar

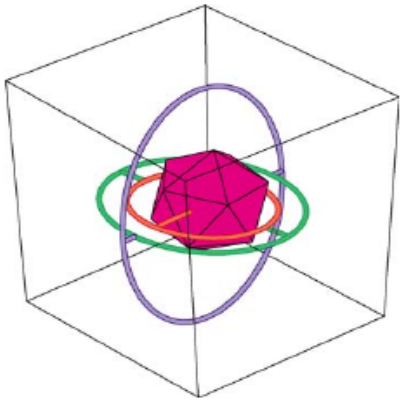


Figura 6.2: Un efecto de superposición de ejes ocurre cuando dos aros de rotación son coplanares.

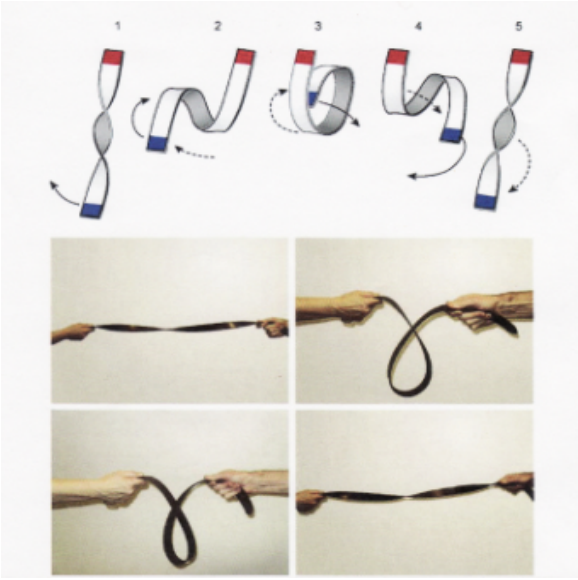


Figura 6.3: El truco del cinturón con un giro de 360 grados.

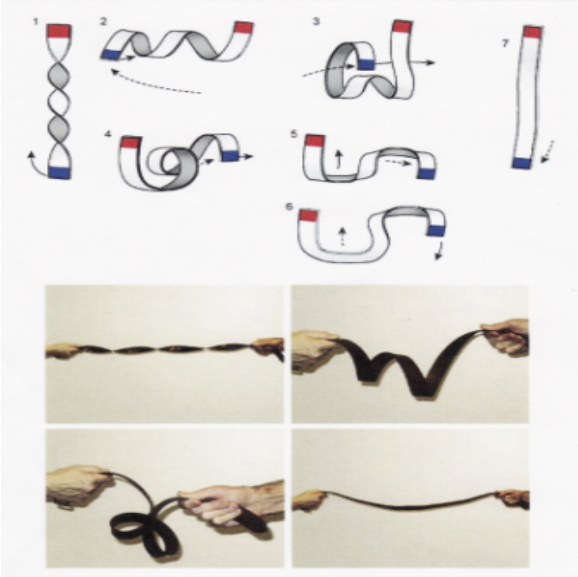


Figura 6.4: El truco del cinturón con un giro de 720 grados.

el cinturón con movimientos que no cambien la orientación los extremos. Lo que resulta interesante es que para el giro de 360 grados es imposible desdoblarlo, como se muestra en la figura (6.3). Sin embargo, para el giro de 720 grados sí es posible hacerlo, como se muestra en la figura (6.4).

El hecho de que podamos resolver el truco para 720 pero no para 360 grados se basa en los siguientes hechos. El grupo correspondiente a los cuaternios unitarios es el grupo $\mathbf{SU}(2)$, con espacio topológico $\mathbb{S}^3$, el cual es un espacio simplemente conexo; las rotaciones de 720 grados corresponden a trayectorias cerradas en el grupo $\mathbf{SU}(2)$ que pueden ser deformadas de manera suave a la identidad. Las rotaciones ordinarias corresponden al grupo $SO(3)$, con espacio topológico el *espacio proyectivo* $\mathbf{RP}^3$, el cual no es simplemente conexo. Esto hace que no todas las trayectorias cerradas en $SO3$ puedan deformarse de manera continua a un punto. Sabemos que $\mathbf{SU}(2)$ es una *doble cubierta* de $SO(3)$; esto es, que hay dos cuaternios para cada marco de orientación distinto en el espacio tridimensional. El truco del cintuón refleja esta doble relación, distinguiendo una rotación de 360 grados de una equivalente de 720 grados. En lo siguiente, veremos qué es lo que pasa realmente.

Utilizando cuaternios, podemos construir una visualización del cinturon y del truco del cinturón la cual resulta interesante y precisa matemáticamente.

La idea básica es la siguiente. Como un pedazo pequeño del cinturón (una línea dibujada en el cinturón en su dirección más corta) y el vector perpendicular al cinturón forman un marco en $\mathbb{R}^3$, y como cada marco es un punto en el espacio de cuaternios, todo el cinturón se puede representar como una *trayectoria conexa de puntos* en el espacio de cuaternios $\mathbb{S}^3$. De esta manera, la deformación inicial del cinturón y los movimientos llevados a cabo por los que sujetan el cinturón intentando deformarlo a su posición inicial, no son nada mas que curvas en $\mathbb{S}^3$, y se pueden

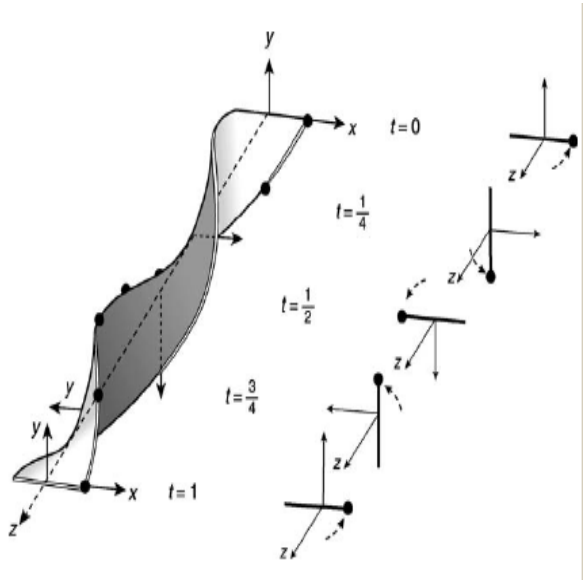


Figura 6.5: Cada sección del cinturón define un marco en $\mathbb{R}^3$ anclado en la línea transversal del cinturón y la normal de la superficie del cinturón en ese punto.

visualizar usando proyecciones apropiadas de las curvas.

Cuando el cinturón no está deformado, todos sus marcos son el marco identidad, y por lo tanto están acomodados como una pila de cuaternios iguales, $w = 1$, que se corresponden con los marcos del cinturón cuando no está deformado. Si t representa el parámetro que describe al cinturón, con $t = 0$ el inicio del cinturón y $t = 1$ el extremo del mismo, entonces

$$q(t) = (1, 0, 0, 0)$$

y no hay cambio alguno en el marco cuando t varía. Cuando rotamos el cinturón alrededor de su eje más largo, digamos el eje z, una rotación por un ángulo θ en $\mathbb{R}^3$ queda determinada por

$$q(t) = \left(\cos\frac{t\theta}{2}, 0, 0, \operatorname{sen}\frac{t\theta}{2}\right).$$

En la figura (6.5), se muestra el cinturón como una secuencia de marcos, junto con marcos aislados que están determinados por

$$\begin{bmatrix} \cos t\theta & -\operatorname{sen} t\theta & 0 \\ \operatorname{sen} t\theta & \cos t\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

De esta manera, si deformamos el cinturón 2π o 360 grados, entonces

$$\begin{aligned} q(0) &= (1, 0, 0, 0), \\ q(t) &= (\cos t\pi, 0, 0, \operatorname{sen} t\pi), \\ q(1) &= (-1, 0, 0, 0), \end{aligned}$$

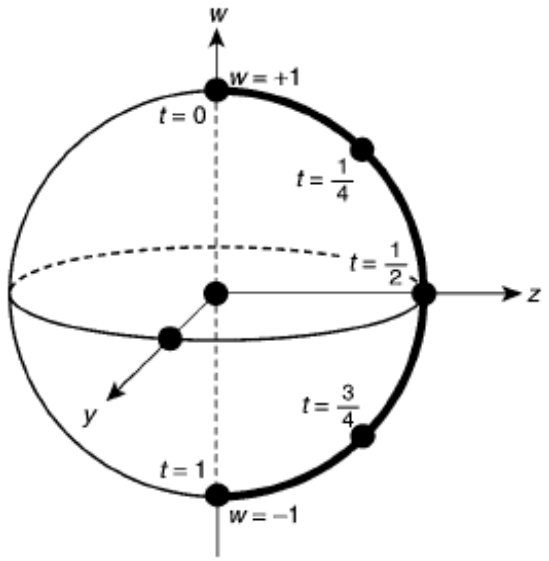


Figura 6.6: El conjunto de marcos de un cinturón con un giro de 360 grados sobre su eje más largo.

y los marcos del cinturón se corresponden con la curva de cuaternios que se muestra en la figura (6.6), la cual inicia en el polo norte y termina en el polo sur.

Si consideramos ahora una rotación de 4π o 720 grados, tenemos

$$\begin{aligned} q(0) &= (1, 0, 0, 0), \\ q(t) &= (\cos 2t\pi, 0, 0, \operatorname{sen} 2t\pi), \\ q(1) &= (1, 0, 0, 0). \end{aligned}$$

En este caso los marcos del cinturón se corresponden con una curva cerrada, como se muestra en la figura (6.7).

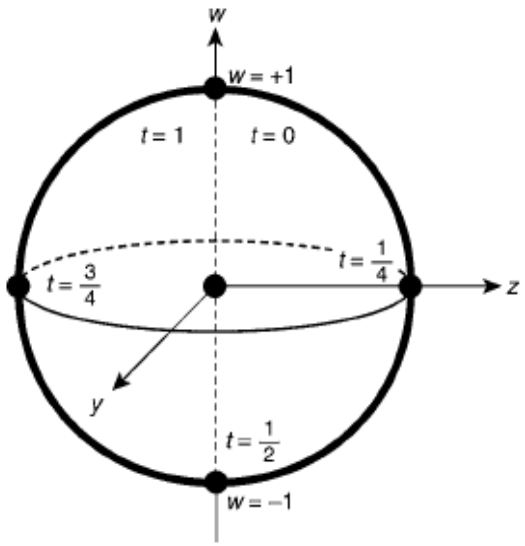


Figura 6.7: El conjunto de marcos de un cinturón con un giro de 720 grados sobre su eje más largo.

Apéndice 1

A continuación se verifica que las matrices de 4×4 con entradas reales definidas en la sección 3.3 satisfacen las propiedades (3.3),(3.4) y (3.5).

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

$$J^2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = -\mathbf{1}$$

$$K^2 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = -\mathbf{1}$$

$$IJ = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = K$$

$$JK = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = I$$

$$KI = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} = J$$

$$JI = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} = -K$$

$$IK = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = -J$$

$$KJ = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} = -I$$

A continuación se verifica que las matrices de 2×2 con entradas complejas definidas en la sección 3.3 cumplen las propiedades (3.3),(3.4) y (3.5).

$$\mathbf{i}^2 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -\mathbf{1}$$

$$\mathbf{j}^2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -\mathbf{1}$$

$$\mathbf{k}^2 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -\mathbf{1}$$

$$\mathbf{ij} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = \mathbf{k}$$

$$\mathbf{jk} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = \mathbf{i}$$

$$\mathbf{ki} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = \begin{bmatrix} 0 & -i^2 \\ i^2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \mathbf{j}$$

$$\mathbf{ik} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 0 & i^2 \\ -i^2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = -\mathbf{j}$$

$$\mathbf{kj} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} = -\mathbf{i}$$

$$\mathbf{ji} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} = -\mathbf{k}$$

Apéndice 2

Programa para generar una animación en de un cuerpo rígido en Maple

```
restart:
with(plots):
with(plottools):

v1:=(0,0,0):
v2:=(1,0,0):
v3:=(1,1,0):
v4:=(0,1,0):
V1:=(0,0,1):
V2:=(1,0,1):
V3:=(1,1,1):
V4:=(0,1,1):

# Puntos que definen el eje de rotacion

p1:=(0,2,0):

p2:=(0,2,2):
#multiplicacion de un cuaternio por un escalar

multcuat:=proc(q1::list,q2::list)::list;
local q,a,b,c,d,e,f,g,h;

a:=q1[1];
b:=q1[2];
c:=q1[3];
d:=q1[4];
e:=q2[1];
f:=q2[2];
g:=q2[3];
h:=q2[4];

q[1]:=a*e-b*f-c*g-d*h;
q[2]:=a*f+b*e+c*h-d*g;
q[3]:=a*g-b*h+c*e+d*f;
```

```
q[4]:=a*h+b*g-c*f+d*e;
return [q[1],q[2],q[3],q[4]];
end proc:

# Funciones que regresan la i-esima coordenada de un cuaternio

coordx:=proc(q::list)
local a;
a:=q[2];
return a;
end proc:

coordy:=proc(q::list)
local a;
a:=q[3];
return a;
end proc:

coordz:=proc(q::list)
local a;
a:=q[4];
return a;
end proc:

# Funciones que regresan la suma, resta de cuaternios, una lista de cuaternios y
otra de cuaternios conjugados dado el angulo de rotacion y el eje de rotacion
(el eje debe ser unitario. Para la animacion)

suma:=proc(q1::list,q2::list)::list;
return [q1[1]+q2[1],q1[2]+q2[2],q1[3]+q2[3],q1[4]+q2[4]];
end proc:
resta:=proc(q1::list,q2::list)::list;
return [q2[1]-q1[1],q2[2]-q1[2],q2[3]-q1[3],q2[4]-q1[4]];
end proc:
rot:=proc(a,c::list,d::list)::list;
local q,b,j,e,eje;
b:=a/2;
e:=resta(c,d);
eje:=[e[1]/tamano(e),e[2]/tamano(e),e[3]/tamano(e),e[4]/tamano(e)];
for j from 0 to 10 do
q[j]:=(cos(j*b/10),eje[2]*sin(j*b/10),eje[3]*sin(j*b/10),eje[4]*sin(j*b/10)):
end do;
return q;
end proc:

rotinv:=proc(a,c::list,d::list)::list;
```

```
local q,b,j,e,eje;
b:=a/2;
e:=resta(c,d);
eje:=[e[1]/tamano(e),e[2]/tamano(e),e[3]/tamano(e),e[4]/tamano(e)];
for j from 0 to 10 do
q[j]:=(cos(j*b/10),-sin(j*b/10)*eje[2],-sin(j*b/10)*eje[3],-sin(j*b/10)*eje[4]):
end do;
return q;
end proc:
R:=rot(3*Pi/4,[0,p1],[0,p2]):
Rinv:=rotinv(3*Pi/4,[0,p1],[0,p2]):

tamano:=proc(a::list);
return sqrt(a[1]^2+a[2]^2+a[3]^2+a[4]^2);
end proc:

for i from 0 to 10 do
    nv1:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v1])),
    [Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
    resta([0,p1],[0,v1])),[Rinv[i]]),[0,p1])),
    coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v1])),
    [Rinv[i]]),[0,p1]))):

nv2:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v2])),
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,v2])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v2])),
[Rinv[i]]),[0,p1]))):

nv3:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v3])),
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,v3])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v3])),
[Rinv[i]]),[0,p1]))):

nv4:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v4])),
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,v4])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v4])),
[Rinv[i]]),[0,p1]))):

nV1:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V1])),
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,V1])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V1])),
[Rinv[i]]),[0,p1]))):
```

```
nV2:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V2])),
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,V2])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V2])),
[Rinv[i]]),[0,p1]))):

nV3:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V3])),
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,V3])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V3])),
[Rinv[i]]),[0,p1]))):

nV4:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V4])),
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,V4])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V4])),
[Rinv[i]]),[0,p1]))):


cubos[i]:=display({line([nv1],[nv2]),line([nv2],[nv3]),
line([nv3],[nv4]),line([nv1],[nv4]),line([nV1],[nV2]),
line([nV2],[nV3]),line([nV3],[nV4]),line([nV1],[nV4]),
line([nv1],[nV1]),line([nv2],[nV2]),line([nv3],[nV3]),
line([nv4],[nV4])}):
end do:
animacion:=display(seq(cubos[i],i=0..10),insequence=true,axes=box,color=red):
fondo:=display(arrow([p1],[p2],.1,.4,.1,color=blue)):
display(animacion,fondo,scaling=constrained);
```

Programa para generar una secuencia de rotaciones en Maple.

```
restart:
with(plots):
with(plottools):


v1:=(0,0,0):
v2:=(1,0,0):
v3:=(1,1,0):
v4:=(0,1,0):
V1:=(0,0,1):
V2:=(1,0,1):
V3:=(1,1,1):
V4:=(0,1,1):

# Puntos que definen el eje de rotacion

p1:=(0,1,0):

p2:=(1,1,0):


multcuat:=proc(q1::list,q2::list)::list;
local q,a,b,c,d,e,f,g,h;

a:=q1[1];
b:=q1[2];
c:=q1[3];
d:=q1[4];
e:=q2[1];
f:=q2[2];
g:=q2[3];
h:=q2[4];

q[1]:=a*e-b*f-c*g-d*h;
q[2]:=a*f+b*e+c*h-d*g;
q[3]:=a*g-b*h+c*e+d*f;
q[4]:=a*h+b*g-c*f+d*e;
return [q[1],q[2],q[3],q[4]];
end proc:

# Funciones que regresan la i-esima coordenada de un cuaternio

coordx:=proc(q::list)
local a;
```

```
a:=q[2];
return a;
end proc:

coordy:=proc(q::list)
local a;
a:=q[3];
return a;
end proc:

coordz:=proc(q::list)
local a;
a:=q[4];
return a;
end proc:

# Funciones que regresan la suma, resta de cuaternios, una lista de cuaternios y
 otra de cuaternio conjugados dado el angulo de rotacion y el eje de rotacion
 el eje debe ser unitario. Para la animacion


suma:=proc(q1::list,q2::list)::list;
return [q1[1]+q2[1],q1[2]+q2[2],q1[3]+q2[3],q1[4]+q2[4]];
end proc:
resta:=proc(q1::list,q2::list)::list;
return [q2[1]-q1[1],q2[2]-q1[2],q2[3]-q1[3],q2[4]-q1[4]];
end proc:
rot:=proc(a,c::list,d::list)::list;
local q,b,j,e,eje;
b:=a/2;
e:=resta(c,d);
eje:=[e[1]/tamano(e),e[2]/tamano(e),e[3]/tamano(e),e[4]/tamano(e)];
for j from 0 to 10 do
q[j]:=(cos(j*b/10),eje[2]*sin(j*b/10),eje[3]*sin(j*b/10),eje[4]*sin(j*b/10)):
end do;
return q;
end proc:
rotinv:=proc(a,c::list,d::list)::list;
local q,b,j,e,eje;
b:=a/2;
e:=resta(c,d);
eje:=[e[1]/tamano(e),e[2]/tamano(e),e[3]/tamano(e),e[4]/tamano(e)];
for j from 0 to 10 do
q[j]:=(cos(j*b/10),-sin(j*b/10)*eje[2],-sin(j*b/10)*eje[3],-sin(j*b/10)*eje[4]):
end do;
return q;
```

```
end proc:
R:=rot(-Pi/2,[0,p1],[0,p2]):
Rinv:=rotinv(-Pi/2,[0,p1],[0,p2]):

tamano:=proc(a::list);
return sqrt(a[1]^2+a[2]^2+a[3]^2+a[4]^2);
end proc:
for i from 0 while i<11 do

nv1:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v1])),
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,v1])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v1])),
[Rinv[i]]),[0,p1]))):

nv2:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v2])),
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,v2])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v2])),
[Rinv[i]]),[0,p1]))):

nv3:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v3])),
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,v3])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v3])),
[Rinv[i]]),[0,p1]))):

nv4:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v4])),
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,v4])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,v4])),
[Rinv[i]]),[0,p1]))):

nV1:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V1])),
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,V1])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V1])),
[Rinv[i]]),[0,p1]))):

nV2:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V2])),
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,V2])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V2])),
[Rinv[i]]),[0,p1]))):

nV3:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V3])),
```

```
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,V3])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V3])),
[Rinv[i]]),[0,p1]))):

nV4:=(coordx(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V4])),
[Rinv[i]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i]],
resta([0,p1],[0,V4])),[Rinv[i]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i]],resta([0,p1],[0,V4])),
[Rinv[i]]),[0,p1]))):

cubos[i]:=display({line([nv1],[nv2]),line([nv2],[nv3]),
line([nv3],[nv4]),line([nv1],[nv4]),line([nV1],[nV2]),
line([nV2],[nV3]),line([nV3],[nV4]),line([nV1],[nV4]),
line([nv1],[nV1]),line([nv2],[nV2]),line([nv3],[nV3]),
line([nv4],[nV4])}):

end do:

v1:=(v1[1],v1[2]+1,v1[3]):
v2:=(v2[1],v2[2]+1,v2[3]):
v3:=(v3[1],v3[2]+1,v3[3]):
v4:=(v4[1],v4[2]+1,v4[3]):
V1:=(V1[1],V1[2]+1,V1[3]):
V2:=(V2[1],V2[2]+1,V2[3]):
V3:=(V3[1],V3[2]+1,V3[3]):
V4:=(V4[1],V4[2]+1,V4[3]):

p1:=(p1[1],p1[2]+1,p1[3]):
p2:=(p2[1],p2[2]+1,p2[3]):
R:=rot(-Pi/2,[0,p1],[0,p2]):
Rinv:=rotinv(-Pi/2,[0,p1],[0,p2]):

for i from 11 while i<22 do

nv1:=(coordx(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,v1])),
[Rinv[i-11]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-11]],
resta([0,p1],[0,v1])),[Rinv[i-11]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,v1])),
[Rinv[i-11]]),[0,p1]))):

nv2:=(coordx(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,v2])),
[Rinv[i-11]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-11]],
resta([0,p1],[0,v2])),[Rinv[i-11]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,v2])),
[Rinv[i-11]]),[0,p1]))):
```

```
nv3:=(coordx(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,v3])),
[Rinv[i-11]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-11]],
resta([0,p1],[0,v3])),[Rinv[i-11]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,v3])),
[Rinv[i-11]]),[0,p1]))):

nv4:=(coordx(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,v4])),
[Rinv[i-11]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-11]],
resta([0,p1],[0,v4])),[Rinv[i-11]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,v4])),
[Rinv[i-11]]),[0,p1]))):

nV1:=(coordx(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,V1])),
[Rinv[i-11]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-11]],
resta([0,p1],[0,V1])),[Rinv[i-11]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,V1])),
[Rinv[i-11]]),[0,p1]))):

nV2:=(coordx(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,V2])),
[Rinv[i-11]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-11]],
resta([0,p1],[0,V2])),[Rinv[i-11]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,V2])),
[Rinv[i-11]]),[0,p1]))):

nV3:=(coordx(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,V3])),
[Rinv[i-11]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-11]],
resta([0,p1],[0,V3])),[Rinv[i-11]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,V3])),
[Rinv[i-11]]),[0,p1]))):

nV4:=(coordx(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,V4])),
[Rinv[i-11]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-11]],
resta([0,p1],[0,V4])),[Rinv[i-11]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-11]],resta([0,p1],[0,V4])),
[Rinv[i-11]]),[0,p1]))):

cubos[i]:=display({line([nv1],[nv2]),line([nv2],[nv3]),
line([nv3],[nv4]),line([nv1],[nv4]),line([nV1],[nV2]),
line([nV2],[nV3]),line([nV3],[nV4]),line([nV1],[nV4]),
line([nv1],[nV1]),line([nv2],[nV2]),line([nv3],[nV3]),
line([nv4],[nV4])}):

end do:

v1:=(v1[1],v1[2]+1,v1[3]):
```

```
v2:=(v2[1],v2[2]+1,v2[3]):
v3:=(v3[1],v3[2]+1,v3[3]):
v4:=(v4[1],v4[2]+1,v4[3]):
V1:=(V1[1],V1[2]+1,V1[3]):
V2:=(V2[1],V2[2]+1,V2[3]):
V3:=(V3[1],V3[2]+1,V3[3]):
V4:=(V4[1],V4[2]+1,V4[3]):

p1:=(p1[1],p1[2]+1,p1[3]):
p2:=(p2[1],p2[2]+1,p2[3]):
R:=rot(-Pi/2,[0,p1],[0,p2]):
Rinv:=rotinv(-Pi/2,[0,p1],[0,p2]):

for i from 22 while i<33 do

nv1:=(coordx(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,v1])),
[Rinv[i-22]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-22]],
resta([0,p1],[0,v1])),[Rinv[i-22]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,v1])),
[Rinv[i-22]]),[0,p1]))):

nv2:=(coordx(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,v2])),
[Rinv[i-22]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-22]],
resta([0,p1],[0,v2])),[Rinv[i-22]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,v2])),
[Rinv[i-22]]),[0,p1]))):

nv3:=(coordx(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,v3])),
[Rinv[i-22]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-22]],
resta([0,p1],[0,v3])),[Rinv[i-22]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,v3])),
[Rinv[i-22]]),[0,p1]))):

nv4:=(coordx(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,v4])),
[Rinv[i-22]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-22]],
resta([0,p1],[0,v4])),[Rinv[i-22]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,v4])),
[Rinv[i-22]]),[0,p1]))):

nV1:=(coordx(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,V1])),
[Rinv[i-22]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-22]],
resta([0,p1],[0,V1])),[Rinv[i-22]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,V1])),
[Rinv[i-22]]),[0,p1]))):

nV2:=(coordx(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,V2])),
```

```
[Rinv[i-22]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-22]],
resta([0,p1],[0,V2])),[Rinv[i-22]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,V2])),
[Rinv[i-22]]),[0,p1])):

nV3:=(coordx(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,V3])),
[Rinv[i-22]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-22]],
resta([0,p1],[0,V3])),[Rinv[i-22]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,V3])),
[Rinv[i-22]]),[0,p1])):

nV4:=(coordx(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,V4])),
[Rinv[i-22]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-22]],
resta([0,p1],[0,V4])),[Rinv[i-22]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-22]],resta([0,p1],[0,V4])),
[Rinv[i-22]]),[0,p1])):

cubos[i]:=display({line([nv1],[nv2]),line([nv2],[nv3]),
line([nv3],[nv4]),line([nv1],[nv4]),line([nV1],[nV2]),
line([nV2],[nV3]),line([nV3],[nV4]),line([nV1],[nV4]),
line([nv1],[nV1]),line([nv2],[nV2]),line([nv3],[nV3]),
line([nv4],[nV4])}):

end do:

v1:=(v1[1],v1[2]+1,v1[3]):
v2:=(v2[1],v2[2]+1,v2[3]):
v3:=(v3[1],v3[2]+1,v3[3]):
v4:=(v4[1],v4[2]+1,v4[3]):
V1:=(V1[1],V1[2]+1,V1[3]):
V2:=(V2[1],V2[2]+1,V2[3]):
V3:=(V3[1],V3[2]+1,V3[3]):
V4:=(V4[1],V4[2]+1,V4[3]):

p1:=(p1[1],p1[2]+1,p1[3]):
p2:=(p2[1],p2[2]+1,p2[3]):
R:=rot(-Pi/2,[0,p1],[0,p2]):
Rinv:=rotinv(-Pi/2,[0,p1],[0,p2]):

for i from 33 while i<44 do

nv1:=(coordx(suma(multcuat(multcuat([R[i-33]],resta([0,p1],[0,v1])),
[Rinv[i-33]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-33]],
resta([0,p1],[0,v1])),[Rinv[i-33]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-33]],resta([0,p1],[0,v1])),
[Rinv[i-33]]),[0,p1])):
```

```
nv2:=(coordx(suma(multcuat(multcuat([R[i-33]],resta([0,p1],[0,v2])),
[Rinv[i-33]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-33]],
resta([0,p1],[0,v2])),[Rinv[i-33]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-33]],resta([0,p1],[0,v2])),
[Rinv[i-33]]),[0,p1]))):

nv3:=(coordx(suma(multcuat(multcuat([R[i-33]],resta([0,p1],[0,v3])),
[Rinv[i-33]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-33]],
resta([0,p1],[0,v3])),[Rinv[i-33]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-33]],resta([0,p1],[0,v3])),
[Rinv[i-33]]),[0,p1]))):

nv4:=(coordx(suma(multcuat(multcuat([R[i-33]],resta([0,p1],[0,v4])),
[Rinv[i-33]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-33]],
resta([0,p1],[0,v4])),[Rinv[i-33]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-33]],resta([0,p1],[0,v4])),
[Rinv[i-33]]),[0,p1]))):

nV1:=(coordx(suma(multcuat(multcuat([R[i-33]],resta([0,p1],[0,V1])),
[Rinv[i-33]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-33]],
resta([0,p1],[0,V1])),[Rinv[i-33]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-33]],resta
([0,p1],[0,V1])),[Rinv[i-33]]),[0,p1]))):

nV2:=(coordx(suma(multcuat(multcuat([R[i-33]],resta([0,p1],[0,V2])),
[Rinv[i-33]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-33]],
resta([0,p1],[0,V2])),[Rinv[i-33]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-33]],resta
([0,p1],[0,V2])),[Rinv[i-33]]),[0,p1]))):

nV3:=(coordx(suma(multcuat(multcuat([R[i-33]],resta([0,p1],[0,V3])),
[Rinv[i-33]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-33]],
resta([0,p1],[0,V3])),[Rinv[i-33]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-33]],resta
([0,p1],[0,V3])),[Rinv[i-33]]),[0,p1]))):

nV4:=(coordx(suma(multcuat(multcuat([R[i-33]],resta([0,p1],[0,V4])),
[Rinv[i-33]]),[0,p1])),coordy(suma(multcuat(multcuat([R[i-33]],
resta([0,p1],[0,V4])),[Rinv[i-33]]),[0,p1])),
coordz(suma(multcuat(multcuat([R[i-33]],resta
([0,p1],[0,V4])),[Rinv[i-33]]),[0,p1]))):


cubos[i]:=display({line([nv1],[nv2]),line([nv2],[nv3]),
line([nv3],[nv4]),line([nv1],[nv4]),line([nV1],[nV2]),
```

```
line([nV2],[nV3]),line([nV3],[nV4]),line([nV1],[nV4]),
line([nv1],[nV1]),line([nv2],[nV2]),line([nv3],[nV3]),
line([nv4],[nV4])}):

end do:
animacion:=display(seq(cubos[i],i=0..43),color=blue,insequence=true,
axes=normal,scaling=constrained):
animacion;
```

Bibliografía

[1] Altman, S., Rotations, Quaternions, and Double Groups. Oxford Science Publications, 1986.

[2] Boothby, W., An Introduction to Differentiable Manifolds And Riemannian Geometry. Academic Press, 1986.

[3] Conway, J., Smith, D., On Quaternions and Octonions: their geometry, arithmetic, and symmetry. A K Peters, 2003.

[4] Diebel, J., Representing Attitude: Euler Angles, Unit Quaternions, and Rotation Vectors. Stanford University. 20 October 2006.

[5] Dávila Rascón G., Flores Espinoza R., Vorobiev Yu., Álgebra Lineal: Teoría y problemas. Editorial Unison, 2006.

[6] Eberly, D. H., *Game Physics*. Morgan Kauffman Publishers, 2004.

[7] Eberly, D. H., Quaternion Algebra and Calculus. Magic Software, Inc, September 27, 2002.

[8] Francis, C. K., Kauffman, L. H., Air on Dirac strings. In *Mathematical Legacy of Wilhelm Magnus*. AMS, Providence, R.I., 1994

[9] Gallier, J., Geometric Methods And Applications For Computer Science and Engineering, Springer, 2000.

[10] Hamilton, W. R., *Elements of Quaternions*. Longmans Green, London, 1866.

[11] Hanson, A. J., *Visualizing Quaternions*. Morgan Kauffman Publishers, 2006.

[12] Hanson, A. J., Quaternion Frenet Frames: Making Optimal Tubes and Ribbons from curves. Computer Science Department. Indiana University.

[13] Hanson, A. J., Ma H., Parallel Transport Approach to Curve Framing. Department of Computer Science. Indiana University. January 11,1995.

[14] Julstrom, B. A., Using Real Quaternions to Represent Rotations in Three Dimensions. Department of Computer Science. St. Cloud State University, 1992.

[15] Klok, F., Two moving coordinate frames for sweeping along a 3D trajectory. Department of Mathematics and Informatics, Delft University of Technology, Delft, The Netherlands. 2 February 1986.

[16] Kuipers, J., Quaternions and Rotation Sequences A primer with Applications to Orbits, Aerospace, and Virtual Reality. Princeton University Press, 1998.

[17] Lyons, D., An Elementary Introduction to the Hopf Fibration. Department of Mathematical Sciences. Lebanon Valley College., April 2003.

[18] Rodrigues, O., "Des Lois Géométriques Qui Régissent les Déplacements d'un Systéme Dans L'espace, et la Variation des Coordonnées Provenant de Ses Déplacements Consideérés Indépendamment des Causes Qui Peuvent les Produire", *Journal de Mathématiques Pure at Appliquées* 5:380-440, 1840.

[19] Oprea, J., Differential Geometry and its applications. Prentice-Hall, 1997.

[20] Shoemake, K., Animating rotation with quaternion curves. Comput. Graph. 19, 3 (July), 1985.