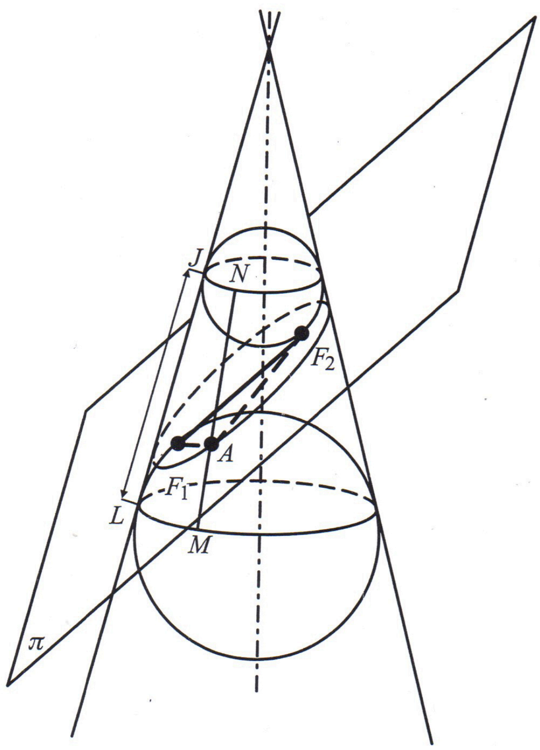
Álgebra lineal y geometría

Miguel González
mgonzalez.contacto@gmail.com
miguelgg.com

Enero de 2019



Revisado en 2022
Apuntes de la asignatura impartida por Yolanda Fuertes
en la Universidad Autónoma de Madrid en Enero de 2019.

Acerca de este documento

Estos apuntes son una versión revisada de los de la asignatura Álgebra lineal y geometría del grado en matemáticas, tomados en Enero de 2019 por Miguel González. La asignatura fue impartida por Yolanda Fuertes. A los apuntes originales se les ha añadido esta página, una imagen de portada, y breves párrafos explicativos en las zonas menos completas. Asimismo se han revisado las erratas y completado los contenidos faltantes.

Este documento es:

- Una recopilación ordenada y directa de las definiciones y resultados más importantes del tema en cuestión, al nivel de los estudios de grado.
- Una colección de demostraciones completas de dichos resultados (salvo en los casos más básicos).
- Una *guía* para revisar de manera rápida las ideas que se han adquirido previamente, o para consultar enunciados puntuales que puedan no haberse comprendido en su totalidad.

Este documento NO es:

- Un libro de texto de la asignatura.
- Una colección de ejercicios para practicar los conceptos adquiridos.
- Un listado de ejemplos para ilustrar las ideas tratadas. A pesar de ello, en ocasiones se incluyen ejemplos puntuales que puedan ser de especial interés o curiosidad, pero se intentan reducir al mínimo en virtud del primer punto de la lista anterior.

Sobre Álgebra lineal y geometría

En esta asignatura se aplica y amplía el álgebra lineal con el fin de estudiar la geometría afín en espacios euclídeos. Los dos principales objetos de estudio son, en primer lugar, los espacios afines, que intuitivamente pueden pensarse como *desplazamientos* de espacios vectoriales y, en segundo lugar, las curvas cónicas y superficies cuádricas en el plano y en el espacio (es decir, los conjuntos de ceros de polinomios de grado dos en dos o tres variables).

Requisitos previos

1. Familiaridad con la notación matemática básica.
2. Conocimientos de la asignatura álgebra lineal, es decir, manejo de los espacios vectoriales de dimensión finita, transformaciones lineales, diagonalización...

Índice

1. Espacios euclídeos y unitarios

1.1. Formas bilineales y sesquilineales. Productos escalares.

Comenzamos definiendo el espacio euclídeo:

Definición 1. Un **espacio euclídeo** es un $\mathbb{R}$-espacio vectorial V junto con una forma bilineal simétrica definida positiva, el **producto escalar**, $\phi : V^2 \to \mathbb{R}$. Se denota $E = (V, \phi)$.

Un **espacio unitario** es el análogo pero con $\mathbb{C}$ el cuerpo de los escalares, en cuyo caso se pide del producto que sea una forma sesquilineal, hermítica y definida positiva.

A continuación definimos los términos que aparecen anteriormente:

Definición 2. Una **forma** es una aplicación que va sobre el cuerpo de un espacio vectorial, es decir, si V es un $\mathbb{K}$-espacio, es una aplicación del tipo $\phi : V^n \to \mathbb{K}$.

Definición 3. Una forma $\phi : V \times V \to \mathbb{K}$ es **bilineal** si es lineal en cada componente, esto es, si $\forall v \in V$, las formas $\phi_v : V \to \mathbb{K}$ dada por $\phi_v(w) = \phi(v, w)$; y $\phi'_v : V \to \mathbb{K}$ dada por $\phi'_v(w) = \phi(w, v)$ son lineales.

A continuación determinaremos cómo expresar con facilidad cualquier forma bilineal en términos de matrices.

Observación 1. Fijada una base $B \subset V$, una matriz $A \in \mathcal{M}_n(\mathbb{K})$ y vectores $x, y \in V$, si denotamos por $[x]_B$ al vector columna con coordenadas de B, tenemos que la aplicación:

$$\phi(x,y) = [x]_B^t A [y]_B = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

es bilineal.

Razón. Se da para el primer factor, siendo análogo para el segundo: $\phi(\lambda x+\beta x', y) = [\lambda x+\beta x']_B^t A[y]_B = (\lambda[x]_B^t + \beta[x']_B^t)A[y]_B = \lambda[x]_B^t A[y]_B + \beta[x']_B^t A[y]_B = \lambda\phi(x,y) + \beta\phi(x',y)$. Se ha usado la distributividad en el producto de matrices y el hecho de que asociar a cada vector sus coordenadas es un isomorfismo.

Ahora veamos que toda forma bilineal se puede expresar como en esta observación.

Proposición 1. *Sea $B = \{v_1, ..., v_n\} \subset V$ una base del $\mathbb{K}$-espacio vectorial V. Sea $\phi : V^2 \to \mathbb{K}$ una forma bilineal cualquiera. Entonces, $\exists M_B(\phi) \in \mathcal{M}_n(\mathbb{K})$, $M_B(\phi) = (a_{ij})_{i,j}$ tal que $\forall u, v \in V$, se tiene que $\phi(x,y) = [x]_B^t M_B(\phi)[y]_B$.*

Dicha matriz se obtiene con las imágenes de la base, poniendo $a_{ij} = \phi(v_i, v_j)$.

Demostración. Sean $u, v \in V$ arbitrarios. Entonces $u = \sum_1^n \lambda_i v_i$ y $v = \sum_1^n \beta_j v_j$. Ahora nos valemos de la bilinealidad para ver que $\phi(u,v) = \phi(\sum_1^n \lambda_i v_i, \sum_1^n \beta_j v_j) = \sum_{i,j=1}^n \lambda_i\beta_j\phi(v_i,v_j) = \sum_{i,j=1}^n \lambda_i\phi(v_i,v_j)\beta_j = [x]_B^t A[y]_B$, donde $A = (a_{ij})_{i,j}$ tiene $a_{ij} = \phi(v_i, v_j)$. Como A no depende de los vectores, solo de la base, basta con hacer $M_B(\phi) = A$. □

Observemos que el coeficiente que acompaña a cada producto de coordenadas $\lambda_i\beta_j$ es precisamente dicho $\phi(v_i, v_j) = a_{ij}$. Por ejemplo, en el producto escalar usual de $\mathbb{R}^2$ dado por $((x_1, x_2), (y_1, y_2)) = x_1y_1 + x_2y_2$, el coeficiente en cada par con coordenadas de mismo índice es 1 y de distinto índice es 0, por lo que su matriz es I_2.

A continuación veremos como **cambiar la base** de la matriz de una forma bilineal.

Proposición 2 (Cambio de base en una matriz de forma bilineal)**.** *Sea V un $\mathbb{K}$-espacio, $\phi : V^2 \to \mathbb{K}$ una forma bilineal, y $B = \{v_1, ..., v_n\}$, $B' = \{v'_1, v'_2, ..., v'_n\}$ sendas bases de V. Supongamos, además, que se conoce $M_B(\phi)$ y la matriz $C_{B'B}$ formada, en columnas, por las coordenadas de los vectores de B' en B, de manera que $[u]_B = C_{B'B}[u]_{B'}$ $\forall u \in V$.*

Entonces, se tiene que $M_{B'}(\phi) = C_{B'B}^t M_B(\phi) C_{B'B}$.

Demostración. En adelante se abrevia $C_{B'B}$ como C. Sean $u, v \in V$. Tenemos que $\phi(u,v) = [u]_B^t M_B(\phi)[v]_B = (C[u]_{B'})^t M_B(\phi) C[v]_{B'} = [u]_{B'}^t C^t M_B(\phi) C[v]_{B'} = [u]_{B'}^t (C^t M_B(\phi) C)[v]_{B'}$, lo que queríamos ver. Hemos usado que $(AB)^t = B^t A^t$. □

Definición 4. Una forma bilineal $\phi : V \times V \to \mathbb{K}$ es **simétrica** si $\forall u, v \in V$, se tiene $\phi(u,v) = \phi(v,u)$.

Proposición 3. *Fijada $B = \{v_i\}_1^n \subset V$ base, se tiene que $\phi : V \times V \to \mathbb{K}$ es simétrica si y solo si $M(\phi)_B = (a_{ij})_{i,j}$ lo es, es decir, si y solo si $M_B(\phi) = M_B(\phi)^t$.*

Demostración. Para $\implies$, veamos que como ϕ es simétrica se tiene en particular que $a_{ij} = \phi(v_i, v_j) = \phi(v_j, v_i) = a_{ji}$. Por otra parte, para $\impliedby$, sean $u, v \in V$, entonces, notando que los escalares no varían al trasponerse, se tiene: $\phi(u,v) = \phi(u,v)^t = ([u]_B^t M_B(\phi)[v]_B)^t = [v]_B^t M_B(\phi)^t [u]_B = [v]_B^t M_B(\phi)[u]_B = \phi(v,u)$. □

Veamos, sin embargo, que en espacios unitarios la bilinealidad no es conveniente. Si queremos que el producto induzca una norma real, la bilinealidad impide que el escalado de vectores sea coherente, puesto que si $\phi(u,u)$ es real, si ϕ fuese bilineal, entonces $\phi(\lambda u, \lambda u) = \lambda^2 \phi(u,u)$ que puede ser complejo. Para evitar esto, haremos que uno de los factores se conjugue, obteniendo $\lambda\overline{\lambda}\phi(u,u) = |\lambda|^2 \phi(u,u)$.

Definición 5. Sea V un $\mathbb{C}$-espacio vectorial. Se dice que $\phi : V \times V \to \mathbb{C}$ es una **forma sesquilineal** si verifica las siguientes propiedades $\forall u_i, v_i \in V$, $\lambda \in \mathbb{C}$:

1. $\phi(u_1 + u_2, v) = \phi(u_1, v) + \phi(u_2, v)$
2. $\phi(u, v_1 + v_2) = \phi(u, v_1) + \phi(u, v_2)$
3. $\phi(\lambda u, v) = \lambda \phi(u, v)$
4. $\phi(u, \lambda v) = \overline{\lambda} \phi(u, v)$

Por ejemplo, las formas

$\phi_1 : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ dada por $\phi_1((x_1, ..., x_n), (y_1, ..., y_n)) = \sum_1^n x_i \overline{y_i}$,

$\phi_2 : C^0[a,b] \times C^0[a,b] \to \mathbb{C}$ dada por $\phi_2(f,g) = \int_a^b f(x)\overline{g(x)}dx$, donde $C^0[a,b] = \{f : [a,b] \to \mathbb{C}$ continuas$\}$ y

$\phi_3 : \mathcal{M}_n(\mathbb{C}) \times \mathcal{M}_n(\mathbb{C}) \to \mathbb{C}$ dada por $\phi_3(A,B) = tr(A^t \overline{B})$

son sesquilineales.

Proposición 4. *Sea $B = \{v_1, ..., v_n\} \subset V$ una base del $\mathbb{K}$-espacio vectorial V. Sea $\phi : V^2 \to \mathbb{K}$ una forma sesquilineal cualquiera. Entonces, $\exists M_B(\phi) \in \mathcal{M}_n(\mathbb{K})$, $M_B(\phi) = (a_{ij})_{i,j}$ tal que $\forall u, v \in V$, se tiene que $\phi(x,y) = [x]_B^t M_B(\phi)\overline{[y]_B}$.*

Dicha matriz se obtiene con las imágenes de la base, poniendo $a_{ij} = \phi(v_i, v_j)$.

Demostración. Completamente análoga a la proposición 1, teniendo en cuenta que las coordenadas del segundo vector se conjugan al salir de la forma. □

Observación. En algunos contextos tiene sentido hablar de vectores conjugados, sin embargo la proposición anterior se refiere exclusivamente a conjugar las coordenadas del segundo vector, no dicho vector.

Proposición 5 (Cambio de base en una matriz de forma bilineal)**.** *Sea V un $\mathbb{K}$-espacio, $\phi : V^2 \to \mathbb{K}$ una forma sesquilineal, y $B = \{v_1, ..., v_n\}$, $B' = \{v'_1, v'_2, ..., v'_n\}$ sendas bases de V. Supongamos, además, que se conoce $M_B(\phi)$ y la matriz $C_{B'B}$ formada, en columnas, por las coordenadas de los vectores de B' en B, de manera que $[u]_B = C_{B'B}[u]_{B'}$ $\forall u \in V$.*

Entonces, se tiene que $M_{B'}(\phi) = C_{B'B}^t M_B(\phi)\overline{C_{B'B}}$.

Demostración. Es análoga a la proposición 2, veamos: sean $u, v \in V$. Tenemos que $\phi(u,v) = [u]_B^t M_B(\phi)\overline{[v]_B} = (C[u]_{B'})^t M_B(\phi)\overline{C[v]_{B'}} = [u]_{B'}^t C^t M_B(\phi)\overline{C}\overline{[v]_{B'}} = [u]_{B'}^t (C^t M_B(\phi)\overline{C})\overline{[v]_{B'}}$, lo que queríamos ver. Hemos usado que $(AB)^t = B^t A^t$, y $\overline{(AB)} = \overline{A} \cdot \overline{B}$. □

Definición 6. Una forma sesquilineal $\phi : V \times V \to \mathbb{C}$ es **hermítica** si y solo si $\forall u, v \in V$, se tiene $\phi(u,v) = \overline{\phi(v,u)}$.

Observemos que en una forma de este estilo, $\phi(u,u) \in \mathbb{R}$ al ser invariante a la conjugación y por tanto la diagonal y la traza de $M_B(\phi)$ son reales.

Proposición 6. *Fijada $B = \{v_i\}_1^n \subset V$ base, se tiene que $\phi : V \times V \to \mathbb{K}$ es hermítica si y solo si $M(\phi)_B = (a_{ij})_{i,j}$ lo es, es decir, si y solo si $M_B(\phi)^t = \overline{M_B(\phi)}$.*

Demostración. Para $\Longrightarrow$, veamos que como ϕ es hermítica se tiene en particular que $a_{ij} = \phi(v_i, v_j) = \overline{\phi(v_j, v_i)} = \overline{a_{ji}}$. Por otra parte, para $\Longleftarrow$, sean $u, v \in V$, entonces, notando que los escalares no varían al trasponerse, se tiene: $\phi(u,v) = \phi(u,v)^t = ([u]_B^t M_B(\phi)\overline{[v]_B})^t = \overline{[v]_B}^t M_B(\phi)^t [u]_B = \overline{[v]_B}^t \overline{M_B(\phi)} [u]_B = \overline{[v]_B^t M_B(\phi) \overline{[u]_B}} = \overline{\phi(v,u)}$ □

Definición 7. Una forma bilineal o sesquilineal $\phi : V \times V \to \mathbb{K}$ es **definida positiva** si $\forall u \in V$, se tiene $\phi(u,u) \geq 0$, con $\phi(u,u) = 0 \iff u = 0$.

Para la siguiente proposición, observemos que $\det(A) \in \mathbb{R}$ si A es hermítica, dado que $\det(A) = \det(A^t) = \det(\overline{A}) = \overline{\det(A)}$.

Proposición 7 (Sylvester)**.** *Sea ϕ sesquilineal o bilineal y $B = \{u_i\}_1^n$ base de cada espacio de partida, con $M_B(\phi)$ la matriz asociada. Sea $M_B(\phi)_r$ cada menor angular de dicha matriz, formado por las primeras r filas y columnas de la misma. Entonces:*
ϕ es definida positiva $\iff \det(M_B(\phi)_r) > 0$, $\forall r \in \{1, 2, ..., n\}$.

Demostración. (Requiere proposiciones que aparecen posteriormente). Denotamos $F_r =< u_1, ..., u_r >$ y entonces es fácil observar que $M_B(\phi)_r = M_{\{u_i\}_1^r}(\phi|_{F_r})$. Para $\Longrightarrow$, vamos a construir $B' = \{w_1, ..., w_n\}$ de V ortonormal mediante el Teorema de Gram-Schmidt (Teorema 2). También vamos a denotar por P_r la matriz de cambio de base de $\{u_i\}_1^r$ a $\{w_i\}_1^r$ en F_r. Entonces:

$$\det(M_B(\phi)_r) = \det(M_{\{u_i\}_1^r}(\phi|_{F_r})) = \det(P_r^t M_{\{w_i\}_1^r}(\phi|_{F_r})\overline{P_r}) = \det(P_r^t I_r \overline{P_r}) = \det(P_r)\overline{\det(P_r)} = |\det(P_r)|^2 > 0$$

al ser P de cambio de base.

Para $\Longleftarrow$, vamos a suponer primero todos los menores angulares positivos. Vamos a elaborar $B' = \{w_1, ..., w_n\}$ 'ortonormal' (en el sentido de que los productos distintos dan 0 y los iguales dan 1) a partir de B por un procedimiento similar a Gram-Schmidt. Esto es, a priori, posible, salvo porque $\phi(u_i, u_i)$ podría ser 0 en cuyo caso no podríamos normalizar. Veamos que nunca se da este caso por inducción:

1. En primer lugar, $\phi(u_1, u_1) \neq 0$ dado que si no $M_B(\phi)_1$ sería 0.

2. Suponiendo que ya llevamos $\{w_1, ..., w_r\}$, que se han podido normalizar, creamos $\tilde{w_{r+1}}$ como es usual y tenemos que

$$M_{\{w_1,...,w_r,\tilde{w_{r+1}}\}}(\phi|_{F_{r+1}}) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \phi(\tilde{w_{r+1}}, \tilde{w_{r+1}}) \end{pmatrix}$$

y resulta que $\phi(\tilde{w_{r+1}}, \tilde{w_{r+1}}) = \det(M_{\{w_1,...,w_r,\tilde{w_{r+1}}\}}(\phi|_{F_{r+1}})) = \det(P^t M_{\{v_1,...,v_{r+1}\}}(\phi|_{F_{r+1}})\overline{P}) = |\det(P)|^2 \cdot M_B(\phi)_r > 0$ por hipótesis, por tanto el producto es no nulo y se puede normalizar $\tilde{w_{r+1}}$.

Finalmente, podemos escribir entonces en esta base $\phi(u,v) = [u]_{B'}^t \cdot I_n \cdot \overline{[v]_{B'}}$ por como se ha construido (para que la matriz quede I), y entonces tenemos que $\phi((x_1, ..., x_n)_{B'}, (x_1, ..., x_n)_{B'}) = \sum_1^n |x_i|^2$ y por tanto es definida positiva. □

1.2. Norma en un espacio euclídeo o unitario.

Definición 8. Una **norma** en un espacio vectorial V es una aplicación $\rho : V \to \mathbb{R}$ que verifica:

1. $\rho(x) \geq 0$, $\rho(x) = 0 \iff x = 0 \; \forall x \in V$.
2. $\rho(\lambda x) = |\lambda| \rho(x) \; \forall \lambda \in \mathbb{K}$, $x \in V$.
3. $\rho(x + y) \leq \rho(x) + \rho(y) \; \forall x, y \in V$.

Definición 9. En un espacio euclídeo/unitario, la **norma inducida** por su producto escalar ϕ viene definida como $\|\|_\phi : V \to \mathbb{R}$ dada por $\|v\|_\phi = \sqrt{\phi(v, v)}$.

La propiedad 1 de las normas viene simplemente por la definición positiva del producto escalar. La 2 es fácilmente demostrable del siguiente modo (sin perder generalidad se hace con ϕ sesquilineal, ya que en el caso real todos los escalares son invariantes a la conjugación): $\sqrt{\phi(\lambda x, \lambda x)} = \sqrt{\lambda \overline{\lambda} \phi(x, x)} = \sqrt{|\lambda|^2 \phi(x, x)} = |\lambda| \sqrt{\phi(x, x)}$. La tercera propiedad, que se conoce como **desigualdad triangular**, requiere un resultado previo:

Teorema 1 (Cauchy-Schwarz). *Si $E = (V, \phi)$ es euclídeo/unitario con su norma inducida, se tiene, $\forall u, v \in V$:*

$$|\phi(u, v)| \leq \|u\| \cdot \|v\|$$

Vale la igualdad si y solo si u y v son proporcionales.

Demostración. Sean $u, v \in V$. Denotamos $\phi(u, v) \equiv (u, v)$ en adelante. Tenemos 2 casos:

- Si son proporcionales, entonces $|(u, v)| = |(\lambda v, v)| = |\lambda (v, v)| = |\lambda| \|v\|^2 = \|\lambda v\| \|v\| = \|u\| \|v\|$ y vale la igualdad.
- Si no lo son, generan un subespacio $W =< u, v >$ de base $B = \{u, v\}$. Observemos que:

$$M_B(\phi) = \begin{pmatrix} (u, u) & (u, v) \\ (v, u) & (v, v) \end{pmatrix} = \begin{pmatrix} \|u\|^2 & (u, v) \\ \overline{(u, v)} & \|v\|^2 \end{pmatrix}$$

  Ahora bien, como se trata de un producto escalar, el criterio de Sylvester asegura que $det(M_B(\phi)) = \|u\|^2 \|v\|^2 - |(u, v)|^2 > 0$, conque se tiene la desigualdad estricta $\|u\|^2 \|v\|^2 > |(u, v)|^2$ de donde sigue lo que se quería.

□

Ahora podemos demostrar que se tiene la triangular:

Proposición 8. *Sea $E = (V, \phi)$ un espacio euclídeo/unitario con la norma inducida. $\forall u, v \in V$, se tiene $\|u + v\| \leq \|u\| + \|v\|$.*

Demostración. $\|u + v\|^2 = (u+v, u+v) = (u, u) + (v, v) + (u, v) + (v, u) = \|u\|^2 + \|v\|^2 + (u, v) + \overline{(u, v)} = \|u\|^2 + \|v\|^2 + 2Re((u, v)) \leq \|u\|^2 + \|v\|^2 + 2|(u, v)| \leq \|u\|^2 + \|v\|^2 + 2 \|u\| \|v\| = (\|u\| + \|v\|)^2$ □

Proposición 9 (Ley del Paralelogramo). *Si $E = (V, \phi)$ es euclídeo/unitario con su norma inducida entonces $\forall u, v \in V$, se tiene $\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$.*

Demostración. $\|u + v\|^2 + \|u - v\|^2 = (u + v, u + v) + (u - v, u - v) = (u, u) + (v, v) + (u, v) + (v, u) + (u, u) + (v, v) - (u, v) - (v, u) = 2(u, u) + 2(v, v)$, como queríamos. □

1.3. Ortogonalidad

Definición 10. Sea V un espacio vectorial con producto escalar ϕ. $v \in V$ es **unitario** si y solo si $\phi(v,v) = 1$, es decir, $\|v\| = 1$.

$u, v \in V$ son ortogonales si y solo si $\phi(u,v) = 0$.

Observación 2. Si $u \in V$ no es unitario ni nulo, el vector $\frac{u}{\|u\|}$ lo es, puesto que $\left\|\frac{u}{\|u\|}\right\| = \frac{\|u\|}{\|u\|} = 1$. Esto permite encontrar un vector proporcional a uno dado que sea unitario (y por tanto genera el mismo espacio).

Proposición 10. *En un espacio euclídeo/unitario $E = (V, \phi)$, si los vectores de la familia $\{v_1, ..., v_k\} \subset V$ son ortogonales dos a dos y no nulos, entonces es linealmente independiente.*

Demostración. Supongamos que los escalares $\{\lambda_i\}_1^n$ verifican $\sum_1^n \lambda_i v_i = 0$. Entonces, por la bilinealidad, $\forall j \in \{1, ..., n\}$: $0 = \phi(\sum_1^n \lambda_i v_i = 0, v_j) = \sum_1^n \lambda_i \phi(v_i, v_j) = \lambda_j \phi(v_j, v_j)$, dado que los demás productos de la suma son nulos. Ahora, como $v_j \neq 0$, $\phi(v_j, v_j) \neq 0$ de modo que $\lambda_j = 0$ $\forall j = \{1, ..., n\}$ como se quería demostrar. □

Definición 11 (Bases ortogonales y ortonormales)**.** Una base $B = \{u_i\}_1^n \subseteq V$ es **ortogonal** si y solo si $\phi(u_i, u_j) = 0$ $\forall i \neq j$, es decir, si los vectores son ortogonales dos a dos. Será **ortonormal** si además todos son unitarios, es decir, $\phi(u_i, u_i) = 1$ $\forall i \in \{1, ..., n\}$.

Proposición 11. *Sea $E = (V, \phi)$ un espacio unitario/euclídeo n-dimensional y sea $B = \{v_i\}_1^n \subset V$ base. Entonces:*

- *B es ortogonal $\iff$ $M_B(\phi)$ es diagonal.*
- *B es ortonormal $\iff$ $M_B(\phi) = I_n$.*

Demostración. Ponemos $M_B(\phi) = (a_{ij})_{i,j}$. Es ortogonal $\iff \phi(v_i, v_j) = a_{ij} = 0$ $\forall i \neq j \iff M_B(\phi)$ es diagonal.

Es ortonormal $\iff \phi(v_i, v_j) = a_{ij} = 0$ $\forall i \neq j$, $\phi(v_i, v_i) = 1 = a_{ii}$ $\forall i \iff M_B(\phi) = I_n$. □

Corolario. En una base ortonormal B, $\phi((x_1, ..., x_n)_B, (y_1, ..., y_n)_B) = \sum_1^n x_i \overline{y_i}$, al ser su matriz I.

Corolario. Para toda base existe un producto tal que B es ortonormal en él, en concreto el que se obtiene al definir $M_B(\phi) = I$.

Proposición 12 (Coordenadas en una base ortonormal.)**.** *Sea $B = \{u_1, ..., u_n\} \subset V$ una base ortonormal de $E = (V; (,))$. Sea $x \in V$. Entonces, $[x]_B = ((x, u_1), (x, u_2), ..., (x, u_n))$.*

Demostración. Ponemos $x = \sum_1^n \lambda_i u_i$. Veamos que $(x, u_j) = (x = \sum_1^n \lambda_i u_i, u_j) = \sum_1^n \lambda_i (u_i, u_j) = \lambda_j$ al ser ortonormal, como se quería ver. □

1.3.1. Método de ortogonalización de Gram-Schmidt

Teorema 2 (Gram-Schmidt)**.** *Sea $\{v_1, ..., v_n\} = B$ base de $E = (V; (,))$ cualquiera. Denotamos por $F_k =< v_1, ..., v_k$, así tenemos que $F_1 \subsetneq F_2 \subsetneq ... \subsetneq F_n = V$. Existe una base ortonormal $\{w_1, ..., w_n\}$ tal que $< w_1, ..., w_k >= F_k$ $\forall k \in \{1, ..., n\}$. El método para su obtención se detalla en la demostración.*

Demostración. Haremos, en primer lugar $w_1 = \frac{v_1}{\|v_1\|}$. De este modo es unitario y claramente $< w_1 >= F_1$. Ahora, obtendremos un vector ortogonal a este: $\tilde{w_2} = v_2 - (v_2, w_1)w_1$. Veamos 2 observaciones clave:

1. $\underline{(w_1, \tilde{w_2})} = 0$, dado que $(w_1, \tilde{w_2}) = (w_1, v_2 - (v_2, w_1)w_1) = (w_1, v_2) - \overline{(v_2, w_1)}(w_1, w_1) = (w_1, v_2) - \overline{(v_2, w_1)} = (w_1, v_2) - (w_1, v_2) = 0$.

2. $< w_1, \tilde{w_2} >= F_2$, dado que $\tilde{w_2}$ es combinación lineal de v_2 y w_1 con escalares no nulos. Así, si $v = \lambda v_2 + \beta w_1 = \lambda(\tilde{w_2} + (v_2, w_1)w_1) + \beta w_1$, y análogo a la inversa.

Veamos que ninguna de las observaciones anteriores cambia si lo hacemos unitario: $w_2 = \frac{\tilde{w_2}}{\|\tilde{w_2}\|}$. El resto de vectores se construyen así. Supongamos que ya tenemos $\{w_1, ..., w_k\}$ tal que generan F_k y son ortonormales. Hacemos:

$$\tilde{w_{k+1}} = v_{k+1} - \sum_{i=1}^{k}(v_{k+1}, w_i)w_i$$

Y las dos observaciones anteriores valen de inmediato:

1. Con $j \in \{1, ..., k\}$, $(\tilde{w_{k+1}}, w_j) = 0$, dado que $(\tilde{w_{k+1}}, w_j) = (v_{k+1} - \sum_{i=1}^{k}(v_{k+1}, w_i)w_i, w_j) = (v_{k+1}, w_j) - \sum_{i=1}^{k}(v_{k+1}, w_i)(w_i, w_j) = (v_{k+1}, w_j) - (v_{k+1}, w_j) = 0$.
2. $< w_1, w_2, ..., w_k, \tilde{w_{k+1}} >= F_{k+1}$, dado que hasta el k-ésimo ya generan F_k, y w_{k+1} es combinación lineal de escalares no nulos de los demás.

Basta con hacer $w_{k+1} = \frac{\tilde{w_{k+1}}}{\|\tilde{w_{k+1}}\|}$. Inductivamente hemos formado la base que se quería. $\square$

Definición 12. Dos subespacios W_1, W_2 son ortogonales en $E = (V, \phi)$ si $\forall w \in W_1$, $w' \in W_2$, se tiene que $\phi(w, w') = 0$.

Definición 13. Dado $S \subset V$ no vacío, se define su espacio ortogonal $S^\perp = \{v \in V : \phi(v, s) = 0 \forall s \in S\}$.

Proposición 13 (Propiedades)**.**

1. *$S^\perp$ es subespacio.*
2. *$S \subset R \implies R^\perp \subset S^\perp$*
3. *$S^\perp =< S >^\perp$*
4. *$< S > \cap S^\perp = \{0\}$*
5. *$< S > \subset (S^\perp)^\perp$*

Demostraciones:

1. Dados $s_1, s_2 \in S^\perp$ y $s \in S$, se tiene que $\phi(\lambda s_1 + \beta s_2, s) = \lambda\phi(s_1, s) + \beta\phi(s_2, s) = 0\lambda + 0\beta = 0$, luego $\lambda s_1 + \beta s_2 \in S^\perp$. Además trivialmente $0 \in S^\perp$.
2. Sea $r \in R^\perp$, y $s \in S$ arbitrario. Hay que mostrar que $\phi(r, s) = 0$, lo cual sigue de que $s \in R$ también.
3. Es evidente que $< S >^\perp \subset S^\perp$ por el punto anterior. Ahora, si $s \in S^\perp$ y $s' \in < S >$, hay que mostrar que $\phi(s, s') = 0$. No obstante, si ponemos $s' = \sum_1^n \lambda_i s_i$ donde cada s_i es un elemento de S, entonces $\phi(s, s') = \sum_1^n \lambda\phi(s, s_i) = 0$ al estar s en $S^\perp$.
4. Sea $s \in < S >, S^\perp$. Entonces por definición $\phi(s, s) = 0$, de modo que no queda otra que $s = 0$.
5. Si s $\in < S >$ y $p \in S^\perp$ arbitrario, entonces $\phi(s, p) = 0$ por definición con lo que $s \in (S^\perp)^\perp$.

$\square$

Proposición 14. *Sea W subespacio en $E = (V, \phi)$ con $V < +\infty$. En este caso, $V = W \oplus W^\perp$. A este último se le denomina **complemento ortogonal**.*

Demostración. Por la propiedad 4 anterior tenemos que la suma es directa. Para ver que da el total, sea $B = \{u_1, ..., u_r\} \subset W$ base ortonormal, que completamos a $B_0 = \{u_1, ..., u_r, v_{r+1}, ..., v_n\}$ base de V. Si ahora aplicamos Gram-Schmidt, como los primeros r vectores no cambian, obtenemos $B' = \{u_1, ..., u_n\}$ ortonormal. Lo que se va a demostrar es que $\{u_{r+1}, ..., u_r\}$ es base de $W^\perp$, finalizando la demostración. Sabemos por la proposición 12 que si $v \in W^\perp$, entonces $v = \sum_1^n (v, u_i)u_i$ al ser B' base de V. Como es perpendicular a los de W, los primeros r términos desvanecen: $v = \sum_{r+1}^n (v, u_i)u_i$ y por tanto $W^\perp \subset < u_{r+1}, ..., u_n >$. La otra inclusión es evidente al ser ortonormal la base. Además son vectores linealmente independientes al ser subconjunto de B', con lo que hemos acabado. □

Observación 3. En dimensión infinita sigue aplicando que la suma sea directa, pero no que de el total dado que hemos necesitado bases para esa prueba. Por ejemplo, en el espacio de sucesiones con términos no nulos finitos, con el producto coordenada a coordenada, podemos considerar el subespacio W de sucesiones cuya suma de términos impares iguala a la de los pares. Veamos que si $(a_1, ..., a_N, 0, 0, ...) \in W^\perp$, basta tomar $(0, 0, ..., 0, 1, 1, 0, 0, ...) \in W$, donde el primer 1 está en la N-ésima coordenada, y como su producto es nulo, se tiene $a_N = 0$. Análogamente para el resto de coordenadas, tendremos que $a_i = 0\ \forall i \in \{1, ..., N\}$ luego $W^\perp = \{0\}$ y la suma no da el total.

1.3.2. Proyecciones y proyecciones ortogonales.

Definición 14. Dado V un $\mathbb{K}$-espacio y $W, W' \subset V$ subespacios con $W \oplus W' = V$, entonces la **proyección sobre W en dirección de W'** es $p_W^{W'} : V \to W$ tal que $p_W^{W'}(x) = w$ si $x = w + w'$ de forma única con $w \in W$, $w' \in W$.

Definición 15. La **proyección ortogonal** sobre W en un espacio euclídeo/unitario $E = (V, \phi)$ es $p_W^\perp$ tal que $p_W^\perp = p_W^{W^\perp}$, dado que estaban en suma directa.

Observación 4. Si $W = 0$, $p_W^{W'} = 0$ y si $W = V$, $p_W^{W'} = Id_V$.

Teorema 3 (Definición equivalente de proyección)**.** *Sea $f : V \to V$ lineal tal que $f \circ f = f$, $Im(f) = W$ y $Nuc(f) = W'$. Entonces $f = p_W^{W'}$. El recíproco es evidentemente cierto.*

Demostración. Se tiene que $f^2 - f = 0$, de modo que el polinomio mínimo de f ha de dividir a $X(X-1)$, luego el característico tiene esos factores y por tanto $V = Nuc(f) \oplus Nuc(f - Id)$ (es posible que alguno sea el cero, pero da igual).

Vamos a ver ahora que en este caso $Nuc(f - Id) = Im(f)$. Sea $v \in Nuc(f - Id)$, entonces $f(v) = v$ luego $v \in Im(f)$. Si $v \in Im(f)$, entonces hay un w con $f(w) = v$, luego $f(f(w)) = f(v) = f(w) = v$ así que $f(v) - v = 0$ y $v \in Nuc(f - Id)$.

Así, tenemos hasta ahora que $V = Nuc(f) \oplus Im(f) = W \oplus W'$. Veamos que dado x, se tiene $x - f(x) \in Nuc(f)$ dado que $f(x - f(x)) = f(x) - f^2(x) = 0$, por tanto, $p_W^{W'}(x) = p_W^{W'}(x - f(x) + f(x)) = f(x)$, con lo que $p_W^{W'} = f$. □

Proposición 15. *Si $W \subset V$ es subespacio en $(V, (,))$, si $x = y + z$ con $y \in W$, entonces $\|z\| \geq \left\|p_{W^\perp}^\perp(x)\right\|$.*

Demostración. $\|z\|^2 = \|x - y\|^2 = \left\|p_{W^\perp}^\perp(x) + p_W^\perp(x) - y\right\|^2 = \left\|p_{W^\perp}^\perp(x)\right\|^2 + \left\|p_W^\perp(x) - y\right\|^2 \geq \left\|p_{W^\perp}^\perp(x)\right\|^2$, donde hemos usado el teorema de pitágoras dado que $(p_W^\perp(x) - y) \in W$ por hipótesis y por tanto es ortogonal a la proyección sobre $W^\perp$. □

1.3.3. Cálculo de proyecciones ortogonales.

Comenzamos con un caso preliminar en el que conocemos una **base ortonormal** del espacio sobre el que se proyecta:

Proposición 16. *Sea $W =< u_1, ..., u_r >$ un subespacio de $E = (V; (,))$ donde esos vectores son base ortonormal. Sea B una base de V cualquiera. Denotamos las coordenadas de cada u_i por $(u_{1i}, ..., u_{ni})_B$. Entonces, $M_B(p_W^\perp) = A \cdot \overline{A^t} \cdot \overline{M_B((,))}$, donde $A = ([u_1]_B, ..., [u_r]_B)$ En el caso de B ortonormal, será simplemente $A \cdot \overline{A^t}$.*

Demostración. Sea $x = (x_1, ..., x_n)_B$. Escribimos $p(x)_W^\perp = \sum_{j=1}^r \lambda_j u_j$. Sabemos que $(x - \sum_{j=1}^r \lambda_j u_j, u_i) = 0\ \forall i \in \{1, ..., r\}$, al estar esa resta en $W^\perp$. Si ahora nos valemos de la ortonormalidad de la base de W y desarrollamos, obtenemos: $0 = (x, u_i) - \lambda_i$ de modo que

$$p(x)_W^\perp = \sum_{j=1}^r (x, u_j) u_j = \sum_{j=1}^r u_j \overline{(u_j, x)} = \sum_1^r \begin{pmatrix} u_{1j} \\ \vdots \\ u_{nj} \end{pmatrix} \begin{pmatrix} \overline{u_{1j}} & \dots & \overline{u_{nj}} \end{pmatrix} \overline{M_B((,))} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Por tanto, sacando el factor de las coordenadas de x, tenemos que:

$$M_B(p_W^\perp(x)) = \sum_{j=1}^r u_j \overline{(u_j, x)} = \sum_1^r \begin{pmatrix} u_{1j} \\ \vdots \\ u_{nj} \end{pmatrix} \begin{pmatrix} \overline{u_{1j}} & \dots & \overline{u_{nj}} \end{pmatrix} \overline{M_B((,))} = A\overline{A^t M_B((,))}$$

□

A continuación demostramos con generalidad la proposición anterior.

Teorema 4. *Sea $W =< u_1, ...u_r >$ un subespacio de $E = (V; (,))$ y $B \subset V$ una base. Las coordenadas de cada u_i son $(u_{1i}, ..., u_{ni})_B$. Definimos además las matrices:*

$$A = ([u_1]_B, ..., [u_n]_B)$$

$$G = M_W((,)|_W) = \begin{pmatrix} (u_1, u_1) & \dots & (u_1, u_r) \\ \vdots & \ddots & \vdots \\ (u_r, u_1) & \dots & (u_r, u_r) \end{pmatrix}.$$

Entonces, la matriz de proyección sobre W es:

$$M_B(p_W^\perp) = A \cdot \overline{G^{-1}} \cdot \overline{A^t} \cdot \overline{M_B((,))}$$

Demostración. Sea $x = (x_1, ..., x_n)_B$. Ponemos $p_W^\perp(x) = \sum_1^r \lambda_j u_j$, y entonces, para cada $i \in \{1, ..., r\}$: $0 = (x - \sum_1^r \lambda_j u_j, u_i) = (x, u_i) - (\sum_1^r \lambda_j u_j, u_i) \iff \overline{(u_i, x)} = \sum_{j=1}^r \lambda_j (u_j, u_i)$. Matricialmente:

$$\begin{pmatrix} \overline{(u_1, x)} \\ \vdots \\ \overline{(u_r, x)} \end{pmatrix} = \begin{pmatrix} (u_1, u_1) & \dots & (u_r, u_1) \\ \vdots & \ddots & \vdots \\ (u_1, u_r) & \dots & (u_r, u_r) \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix} = G^t \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix} = \overline{G} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix}.$$

Por otro lado:

$$\begin{pmatrix} \overline{(u_1, x)} \\ \vdots \\ \overline{(u_r, x)} \end{pmatrix} = \begin{pmatrix} \overline{[u_1]_B^t M_B((,))}[x]_B \\ \vdots \\ \overline{[u_r]_B^t M_B((,))}[x]_B \end{pmatrix} = \begin{pmatrix} \overline{[u_1]_B^t} \\ \vdots \\ \overline{[u_r]_B^t} \end{pmatrix} \cdot \overline{M_B((,))} \cdot [x]_B = \overline{A^t} \cdot \overline{M_B((,))} \cdot [x]_B.$$

Es decir, combinando ambas expresiones:

$$\overline{A^t} \cdot \overline{M_B((,))} \cdot [x]_B = \overline{G} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix} \iff \overline{G^{-1}} \cdot \overline{A^t} \cdot \overline{M_B((,))} \cdot [x]_B = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix}.$$

Ahora, observando que $[p_W^{\perp}(x)]_B = \begin{pmatrix} [u_1]_B & \dots & [u_r]_B \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix} = A \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix}$, podemos escribir:

$$[p_W^{\perp}(x)]_B = A \cdot \overline{G^{-1}} \cdot \overline{A^t} \cdot \overline{M_B((,))} \cdot [x]_B.$$

De modo que finalmente $M_B(p_W^{\perp}) = A \cdot \overline{G^{-1}} \cdot \overline{A^t} \cdot \overline{M_B((,))}$. □

1.4. Aplicaciones adjuntas y autoadjuntas

Definición 16. Sea $(V;(,))$ euclídeo/unitario, $f : V \to V$ lineal. Se dice que $g : V \to V$ es la **aplicación adjunta de** f, $adj(f)$ si $\forall u, v \in V$ se tiene $(f(u), v) = (u, g(v))$.

Definición 17. Una aplicación f es **autoadjunta** si $adj(f) = f$.

Lema 1. *Si* $(x, z) = (y, z)$ $\forall z \in V$*, entonces* $x = y$*.*

Demostración. Por linealidad, $(x - y, z) = 0$ $\forall z \in V$. En particular, $(x - y, x - y) = 0$ de modo que $x - y = 0 \implies x = y$. □

Proposición 17. *Sea* $f : V \to V$ *lineal en* $E = (V;(,))$ *euclídeo/unitario. Si* $dimV < +\infty$*, entonces* $\exists! g \in L(V, V)$ *tal que* $adj(f) = g$*, y además*

$$M_B(g) = \overline{M_B((,))^{-1}} \cdot \overline{M_B(f)^t} \cdot \overline{M_B((,))}$$

En particular, si B *es ortonormal,* $M_B(g) = \overline{M_B(f)^t}$*.*

Demostración.

$$(fu, v) = [fu]_B^t \cdot M_B((,)) \cdot \overline{[v]_B} = (M_B(f)[u_B])^t \cdot M_B((,)) \cdot \overline{[v]_B} = [u_B]^t M_B(f)^t \cdot M_B((,)) \cdot \overline{[v]_B}$$

Por otro lado, si definimos g como aquella de matriz $M_B(g) = \overline{M_B((,))^{-1}} \cdot \overline{M_B(f)^t} \cdot \overline{M_B((,))}$, entonces

$$(u, gv) = [u]_B^t \cdot M_B((,)) \cdot \overline{\overline{M_B((,))^{-1}} \cdot \overline{M_B(f)^t} \cdot \overline{M_B((,))} \cdot [v]_B} = [u]_B^t \cdot M_B(f)^t \cdot M_B((,)) \cdot \overline{[v]_B} = (fv, u)$$

con lo cual dicha g es adjunta de f.

Para ver que es única, nos valemos del lema previo. Si g, h son ambas adjuntas de f, entonces $(fu, v) = (u, gv) = (u, hv)$ $\forall u, v \in V$, de modo que $gv = hv$ $\forall v \in V$ con lo que $g = h$. □

Observación. f es autoadjunta si y solo si en B ortonormal, $M_B(f) = \overline{M_B(f)^t}$, es decir, si su matriz es hermítica.

Proposición 18 (Propiedades)**.**

1. $adj(Id_V) = Id_V$
2. $adj(adj(f)) = f$
3. $adj(f + g) = adj(f) + adj(g)$

4. $adj(f \circ g) = adj(g) \circ adj(f)$

5. $adj(f^{-1}) = adj(f)^{-1}$

Demostración. Sean $x, y \in V$:

1. $(Id(x), y) = (x, y) = (x, Id(y))$
2. $(f(x), y) = (x, adj(f)(y)) = \overline{(adj(f)(y), x)} = \overline{(y, adj(adj(f))(x))} = (adj(adj(f))(x), y)$, por tanto $f = adj(adj(f))$.
3. $((f+g)(x), y) = (f(x), y) + (g(x), y) = (x, adj(f)(y)) + (x, adj(g)(y)) = (x, (adj(f) + adj(g))(y))$
4. $(f \circ g(x), y) = (g(x), adj(f)(y)) = (x, adj(g)(adj(f)(y)))$.
5. $adj(f \circ f^{-1}) = adj(f^{-1} \circ f) = adj(Id_V) \implies adj(f^{-1}) \circ adj(f) = adj(f) \circ adj(f^{-1}) = Id$.

□

Proposición 19. *Sea $f : V \to V$ lineal, $(V; (,))$ euclídeo/unitario y $dimV < +\infty$. Entonces, $Nuc(f) = Im(adj(f))^{\perp}$ y $Im(f) = Nuc(adj(f))^{\perp}$.*

Demostración. $u \in Im(adj(f))^{\perp} \iff \forall v \in V,\ (u, adj(f)(v)) = 0 \iff (f(u), v) = 0\ \forall v \in V \iff f(u) = 0 \iff u \in Nuc(f)$.

Para la segunda igualdad usamos la primera: $Nuc(adj(f)) = Im(adj(adj(f)))^{\perp} = Im(f)^{\perp} \implies Im(f) = Nuc(adj(f))^{\perp}$ al estar en dimensión finita. □

Corolario. Si f es autoadjunta, $Nuc(f) = Im(f)^{\perp}$.

1.4.1. Diagonalización en aplicaciones autoadjuntas.

Proposición 20. *Si W es invariante por f, $W^{\perp}$ lo es por $adj(f)$, y por tanto si f es autoadjunta, dado un invariante W lo será también $W^{\perp}$.*

Demostración. Sea $w \in W^{\perp}$. Entonces, $\forall v \in W$, al ser W invariante, se tiene que: $(f(v), w) = 0 \implies (v, adj(f)(w)) = 0 \implies adj(f)(w) \in W^{\perp}$ al verificarse para todo v. □

Veamos que las aplicaciones autoadjuntas siempre diagonalizan, y no solo eso, sino que lo hacen en base ortonormal:

Proposición 21. *Sea $f : V \to V$ autoadjunta. Entonces, todos sus autovalores son reales (es decir, el polinomio característico tiene todas sus raíces reales).*

Demostración. Si V es un $\mathbb{R}$-espacio, vamos a extenderlo brevemente a los complejos para asegurarnos de que existe un autovalor suyo (aunque tenga que ser complejo), y después probaremos que de hecho es real. Para ello, permitimos escalares complejos en V dando lugar a $V_{\mathbb{C}}$, extendemos la función f a $f_{\mathbb{C}}$ definida por $f_{\mathbb{C}}((x_1, ..., x_n)_B) = M_B(f) \cdot (x_1, ..., x_n)^t$ como es usual (solo que ahora permitimos escalares complejos) y finalmente extendemos el producto escalar ϕ a $\phi_{\mathbb{C}}$ dado por $\phi_{\mathbb{C}}(x, y) = [x]_B^t M_B(\phi) \overline{[y]_B}$ como es usual. Seguirá siendo producto escalar al mantenerse la misma matriz.

Una vez hecho esto, obtenemos un autovalor $\lambda \in \mathbb{C}$, que sabemos existe al tener $f_{\mathbb{C}}$ su polinomio característico en los complejos. Ahora bien, en este caso, si v es autovector de este valor: $\phi_{\mathbb{C}}(\lambda v, v) = \phi_{\mathbb{C}}(f_{\mathbb{C}}(v), v) = \phi_{\mathbb{C}}(v, f_{\mathbb{C}}(v)) = \phi_{\mathbb{C}}(v, \lambda v)$, de modo que $\lambda(v, v) = \overline{\lambda}(v, v)$, y como $v \neq 0$ por hipótesis, $\lambda = \overline{\lambda}$ luego $\lambda \in \mathbb{R}$. Como el polinomio de f y el de $f_{\mathbb{C}}$ son el mismo al tener la misma matriz, y λ era arbitrario, sigue que todos los autovalores de f son reales. □

Teorema 5. *Sea $f : V \to V$ autoadjunta en el espacio $(V, (,))$ de dimensión finita. Entonces es diagonalizable en una base ortonormal.*

Demostración. Se procede por inducción en $\dim V$. Si $\dim V = 1$, entonces por la proposición anterior hay un autovalor de vector v, y por tanto $\{\frac{v}{\|v\|}\}$ es la base buscada.

Asumimos que se cumple si $\dim V = n$, y veremos con facilidad que lo hace para $\dim V = n + 1$. Tomamos un autovalor y su vector v, y entonces $< v >^{\perp}$ es invariante por f, de modo que $f|_{<v>^{\perp}}$ está bien definida, sigue siendo adjunta y sobre un espacio de dimensión n, luego diagonaliza en base ortonormal $B = \{v_1, ..., v_n\}$. Tomamos la base $B' = \{v_1, ..., v_n, \frac{v}{\|v\|}\}$. Es evidente que todos son unitarios, y además es ortogonal porque $\frac{v}{\|v\|} \in < v >$ y el resto, que ya eran ortogonales, están en su complemento ortogonal. Como tanto $< v >$ como su ortogonal son invariantes, la matriz resultante es diagonal. □

1.5. Aplicaciones ortogonales

Definición 18. Sea $E = (V, (,))$ un espacio euclídeo/unitario. La aplicación $f : V \to V$ es **ortogonal** (o **unitaria** en espacios unitarios) si $\forall x, y \in V$, se tiene $(f(x), f(y)) = (x, y)$. Es decir, preserva el producto escalar.

Ejemplos son los giros centrados en el origen o las simetrías ortogonales.

Proposición 22. *Son equivalentes:*

1. f *es ortogonal en* $E = (V; (,))$
2. $\forall x \in V$, $\|f(x)\| = \|x\|$ *(preserva normas)*
3. *Si* $\{v_1, ..., v_n\}$ *es base ortonormal en* V, $\{f(v_1), ...f(v_n)\}$ *también lo es.*
4. $adj(f) \circ f = Id$, *es decir,* $adj(f) = f^{-1}$.

Demostración. (1) $\implies$ (2). Se tiene $\|f(x)\| = \sqrt{(f(x), f(x))} = \sqrt{(x, x)} = \|x\|$.
(2) $\implies$ (1). Usaremos la siguiente identidad (se demuestra directamente):

$$(x, y) = \frac{1}{2}(\|x + y\|^2 + i\|x + iy\|^2 - (1 + i)\|x\|^2 - (1 + i)\|y\|^2)$$

.

Entonces, tenemos que $(f(x), f(y)) = \frac{1}{2}(\|f(x) + f(y)\|^2 + i\|f(x) + if(y)\|^2 - (1 + i)\|f(x)\|^2 - (1 + i)\|f(y)\|^2) = \frac{1}{2}(\|f(x + y)\|^2 + i\|f(x + iy)\|^2 - (1+i)\|f(x)\|^2 - (1+i)\|f(y)\|^2) = \frac{1}{2}(\|x + y\|^2 + i\|x + iy\|^2 - (1+i)\|x\|^2 - (1+i)\|y\|^2) = (x, y)$. En espacios euclídeos se verifica la identidad prescindiendo de la parte imaginaria, y el procedimiento es análogo.

(1) $\implies$ (3). Se tiene que $(f(v_i), f(v_j)) = (v_i, v_j) = \begin{cases} 1 \ si \ i = j \\ 0 \ si \ i \neq j \end{cases}$ luego son ortonormales (y por tanto base).

(3) $\implies$ (1). Sean $x, y \in V$, entonces en ambas bases y por linealidad de f se tiene: $x = \sum_1^n x_i v_i$, $y = \sum_1^n y_i v_i$, $f(x) = \sum_1^n x_i f(v_i)$, $f(y) = \sum_1^n y_i f(v_i)$. Como ambas son ortonormales, $(x, y) = (x_1, ..., x_n) \cdot I_n \cdot (y_1, ..., y_n)^t = (f(x), f(y))$ al ser las mismas sus coordenadas.

(1) $\implies$ (4). Se tiene $\forall x, y \in V$ que $(x, y) = (f(x), f(y)) = (x, adj(f) \circ f(y))$, por tanto $y = adj(f)(f(y))$ $\forall y \in V$ y por tanto son la misma.

(4) $\implies$ (1) Tenemos que $\forall x, y \in V$, se da $(x, y) = (x, adj(f)(f(y))) = (f(x), f(y))$. □

Observación 5. Si f es ortogonal/unitaria, es biyectiva y además $f^-1 = adj(f)$.

Razón. Si lleva una base en otra base, como hemos visto anteriormente, debe ser biyectiva. Por tanto, como $adj(f) \circ f = Id$, sigue que su inversa es $adj(f)$.

Proposición 23. *Si* f *es ortogonal/unitaria, fijada base ortonormal* B, *se tiene que* $|det(M_B(f))| = 1$. *En particular, si es ortogonal, deberá ser* 1 *o* -1.

Demostración. Tenemos en base ortonormal que $\overline{M_B(f)^t} M_B(f) = I$, por tanto, tomando determinantes, $\overline{det(M_B(f))} det(M_B(f)) = 1 \implies |det(M_B(f))|^2 = 1$. □

Cabe destacar que la condición anterior no es suficiente puesto que hay contraejemplos. Ahora vamos a conocer la estructura de las aplicaciones ortogonales/unitarias:

Proposición 24. *Si $f : V \to V$ es ortogonal/unitaria y λ es un autovalor suyo, entonces $|\lambda| = 1$.*

Demostración. Sea λ un autovalor de la aplicación. Si no tiene, podemos permitir escalares complejos, extendiendo el espacio a $V_\mathbb{C}$, la aplicación a $f_\mathbb{C}$ y el producto a $(,)_\mathbb{C}$ de la manera evidente (misma matriz, permitiendo escalares complejos, conjugando en el producto). Entonces, si v es autovector asociado, se tiene que $(v, v)_\mathbb{C} = (f_\mathbb{C}(v), f_\mathbb{C}(v))_\mathbb{C} = (\lambda v, \lambda v)_\mathbb{C} = \lambda\overline{\lambda}(v, v) = |\lambda|^2(v, v)$ y como $v \neq 0$, se tiene $|\lambda|^2 = 1$. □

Observemos que, a diferencia de las autoadjuntas, no se garantiza que las aplicaciones ortogonales tengan autovalores reales, y puede ser necesario el paso a los complejos. No obstante, también dejan invariantes los ortogonales y esto permitirá repetir el argumento de las autoadjuntas:

Proposición 25. *Sea $f : V \to V$ ortogonal/unitaria y W invariante por f. Entonces $W^\perp$ también lo es.*

Demostración. Observemos que $f|_W : W \to W$ está bien definida, es biyectiva por ser ortogonal y por tanto $f^{-1} : W \to W$ también lo está. Sea $u \in W^\perp$ y $w \in W$ arbitario. Entonces, $(f(u), w) = (u, f^{-1}(w)) = 0$ dado que $f^{-1}(w) \in W$ como hemos visto, por tanto $f(u) \in W^\perp$. □

Teorema 6. *Sea $f : V \to V$ una aplicación unitaria en el espacio unitario $(V, (,))$. Entonces, es diagonalizable en base ortonormal.*

Demostración. Por inducción en $\dim V$. Si $\dim V = 1$, entonces evidentemente es diagonalizable, y si u es uno de los autovectores, entonces $\{\frac{u}{\|u\|}\}$ es la base buscada. Supongamos que lo es cuando $\dim V = n$. Ahora, si $\dim V = n + 1$, sea λ una raíz del polinomio característico, que existe al estar en $\mathbb{C}[X]$, y sea u un autovector de λ. Ahora, $< u >^\perp$ es un invariante n-dimensional y f sigue siendo unitaria allí, luego $f|_{<u>^\perp}$ diagonaliza en base ortonormal B. Ahora $B \cup \{\frac{u}{\|u\|}\}$ es ortonormal y por construcción f diagonaliza en esa base. □

Ahora nos interesa ver qué sucede con las aplicaciones en $\mathbb{R}$. Daremos una 'forma canónica' que todas alcanzan, compuesta por invariantes de orden 1 que se quedan fijos (o se invierten) e invariantes de orden 2 que giran. Después clasificaremos todas las aplicaciones ortogonales posibles de $\mathbb{R}^2$ y $\mathbb{R}^3$.

Lema 2. *Si $f : V \to V$ es una aplicación ortogonal/unitaria y λ, μ son 2 autovalores distintos con autovectores v, w respectivamente, entonces $(v, w) = 0$.*

Demostración. $(v, w) = (f(v), f(w)) = (\lambda v, \mu w) = \lambda\overline{\mu}(v, w)$. Si $(v, w) \neq 0$, entonces debe ser $\lambda\overline{\mu} = 1 \implies \overline{\mu} = \lambda^{-1} = \overline{\lambda}$ al tener módulo 1, lo cual contradice que sean distintos. □

Teorema 7 (Forma canónica de las aplicaciones ortogonales)**.** *Sea $f : V \to V$ una aplicación ortogonal con $(V, (,))$ un espacio euclídeo. Entonces, existe una base ortonormal B tal que:*

$$M_B(f) = \begin{pmatrix} I_k & 0 & 0 & 0 & \dots & 0 \\ 0 & -I_l & 0 & 0 & \dots & 0 \\ 0 & 0 & A_1 & 0 & \dots & 0 \\ 0 & 0 & 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & A_i \end{pmatrix}$$

Donde I_k es la identidad de orden k y cada bloque $A_i = \begin{pmatrix} \cos\alpha_i & -\sin\alpha_i \\ \sin\alpha_i & \cos\alpha_i \end{pmatrix}$ para algún ángulo $\alpha_i \in [0, 2\pi)$.

Demostración. Realizamos cada uno de estos pasos hasta que podamos, es decir, hasta agotar el procedimiento correspondiente o hasta obtener $\dim V$ vectores para B, comenzando por el espacio $W = V$:

1. Si $f|_W$ tiene el autovalor 1, tomamos el autovector v y agregamos a nuestra base al vector $\frac{v}{\|v\|}$. Nuestra base hasta el momento es el conjunto T. Ahora, continuamos tomando como nuevo espacio $W' =< T >^{\perp}$. (Este paso se realiza k veces)

2. Cuando $f|_W$ ya no admita al 1 como autovalor, hacemos lo mismo que el paso anterior para el valor -1. (Este paso se realiza l veces)

3. Hemos llegado a $f|_W$ que no admite autovalores reales. Para obtener los vectores que faltan, tomamos $\lambda \in \mathbb{C}$, $\lambda = a + ib$ raíz del polinomio característico. Como este tiene coeficientes reales, tomamos $\overline{\lambda} = a - ib$ el conjugado de dicho autovalor, que también es autovalor. Sea $w = u - iv$ el autovector de $f_{\mathbb{C}}|_W$ unitario para el valor λ, con $u, v \in V$, y sea $\overline{w} = u + iv$, que por construcción también es unitario ya que sus coordenadas son las conjugadas de las de w.

   No es difícil comprobar ahora que $\overline{w}$ es un autovector de $\overline{\lambda}$, dado que si M es matriz en base cualquiera de $f|_W$, entonces $Mw = \lambda w \implies \overline{M\overline{w}} = M\overline{w} = \overline{\lambda}\overline{w}$. Además, se tiene que $w + \overline{w} = 2u$ y que $\overline{w} - w = 2iv$.

   Ahora veremos que u, v forman un invariante, dado que que $f(\overline{w}) = f(u + iv) = f(u) + if(v)$ pero también $f(u + iv) = (a - ib)(u + iv) = au + bv - ibu + iav$, de modo que:

$$\begin{cases} f(u) = au + bv \\ f(v) = -bu + av \end{cases} \tag{1}$$

   A continuación veamos que son ortogonales, es decir que el invariante es de orden 2 con base ortogonal $\{u, v\}$. Para ello, vemos que $(u, v) = \frac{1}{4i}(w+\overline{w}, \overline{w}-w) = -(w, w)+(\overline{w}, \overline{w})+(w, \overline{w})-(\overline{w}, w) = -1 + 1 + 0 + 0 = 0$, por el lema previo.

   Para finalizar, lo único que debemos hacer es añadir a nuestra base a los vectores $\{\frac{u}{\|u\|}, \frac{v}{\|v\|}\}$. Como $\|w + \overline{w}\|^2 = \|\overline{w} - w\|^2 = 2$ al ser estos ortogonales entre sí (teorema de pitágoras), entonces, dichos vectores se pueden escribir como $\{\sqrt{2}u, \sqrt{2}v\}$, y vemos que la relación (1) se sigue manteniendo en esa base, y por tanto:

$$M_{\{\sqrt{2}u, \sqrt{2}v\}}(f|_{<\sqrt{2}u, \sqrt{2}v>}) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

   Como λ tiene módulo 1, sique que $a^2 + b^2 = 1$ y por tanto se puede expresar $a = \cos\alpha$ y $b = \sin\alpha$, como se quería.

   (Este paso se repite i veces.)

4. Finalmente, es evidente que la base obtenida es ortonormal, dado que los vectores son unitarios y se han ido tomando cada uno en el ortogonal de lo que generaban los anteriores. Los vectores añadidos en el paso 1 y 2 son autovectores de 1 y -1, y las parejas de vectores añadidas en el paso 3 dan lugar a los bloques de 2×2. □

1.5.1. Clasificación de aplicaciones ortogonales

Observación 6. Si $f : \mathbb{R}^2 \to \mathbb{R}^2$ es ortogonal, entonces hay una base ortonormal $B = \{u, v\}$ donde tiene una de las siguientes formas:
$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ o bien $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
El primer caso, de determinante $\det f = 1$, es un **giro de ángulo** α. En el caso de $\alpha = 0$ es la identidad, y si $\alpha = \pi$ es la simetría central o $-Id$.

El segundo caso, de determinante $\det f = -1$ es una **simetría ortogonal con respecto a** $< u >$, donde evidentemente la dirección de la simetría es $< v >$.

Observación 7. Estos son los casos que se pueden dar para $f : \mathbb{R}^3 \to \mathbb{R}^3$ en base ortonormal $B = \{u, v, w\}$:

a) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$, con $\det f = 1$. Es un **giro de ángulo** α **respecto del eje** $< u >$.
Cabe notar que $Tr(f) = 1 + 2\cos\alpha$, que es invariante a la base, y que, en base ortonormal $B = \{u, v, w\}$, se tiene $(fv, w) = \sin\alpha$. Estos dos métodos permiten hallar el ángulo sin calcular explícitamente la matriz en dicha base.
Si $\alpha = 0$ es la identidad y si $\alpha = \pi$ también se puede ver como una simetría respecto del eje $< u >$.

b) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, con $\det f = -1$. Es una **simetría ortogonal respecto del plano** $< v, w >$.

c) $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, con $\det f = -1$. Es una **simetría central**.

d) $\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$, con $\det f = -1$. Es una simetría respecto de $< u >^{\perp} = < v, w >$ compuesta con un giro respecto de $< u >$ de α grados.

Vemos que los giros, dependiendo del orden en que se escoja la base del plano a girar, se realizan en un sentido o en otro. Normalmente, se busca el sentido que induce el vector del eje como es usual, para lo que vamos a definirlo formalmente:

Definición 19. Dos bases de $\mathbb{R}^n$, B y B', tienen la **misma orientación** si $\det C > 0$, donde C es el cambio de base de una a otra. En caso contrario, **no tienen la misma orientación**.

Observamos que esto define una relación de equivalencia en el espacio de bases de $\mathbb{R}^n$. Orientar un espacio será cambiar el orden de la base para que coincida en orientación con una dada.

Definición 20. Diremos que una base está **positivamente orientada** en $\mathbb{R}^n$ si su orientación es la de la canónica, es decir, el determinante de las coordenadas de los vectores es positivo.

Observemos que, por ejemplo, en $\mathbb{R}^3$, dados dos vectores u_1, u_2 linealmente independientes, un tercero que forma base positivamente orientada con ellos es $u_1 \times u_2$.

Definición 21. Un isomorfismo $f : V \to V$ **conserva la orientación** si $\det f > 0$, y la **invierte** si $\det f < 0$.

1.5.2. Obtención de giros y simetrías

Observación 8. Las simetrías se pueden obtener fácilmente dado que $S_W^\perp = P_W^\perp - P_{W^\perp}^\perp = P_W^\perp - (Id - P_W^\perp) = 2P_W^\perp - Id$, y sabemos ya como obtener simetrías.

A continuación vemos como obtener una matriz de giro de $\mathbb{R}^3$ dados sus elementos geométricos. (En $\mathbb{R}^2$ es evidente, en base ortonormal se coloca la forma que ya conocemos con cos y sin.)

Observación 9. Queremos obtener el giro o simetría+giro G de eje $< u_1 >$ y ángulo α.

Sea $B = \{u_1, u_2, u_3\}$ la base ortonormal tomada con orientación positiva, de modo que:

$$M_B(G) = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$$

La orientación se toma positiva para que el giro sea en el sentido inducido por u_1. El primer 1 irá con signo negativo si revierte orientación por una simetría con $< u_1 >^\perp$. Si observamos:

$$M_B(G) = \pm \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \cos\alpha \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \sin\alpha \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} =$$

$$= \pm M_B(P_{<u_1>}^\perp) + \cos\alpha M_B(P_{<u_1>^\perp}^\perp) + \sin\alpha M_B(M_{u_1})$$

Donde $M_{u_1}(x) = u_1 \times x$, dado que como es ortonormal la base, y positivamente orientada, $u_1 \times u_2 = u_3$ y $u_1 \times u_3 = -u_2$.

Finalmente, entonces, podemos decir que:

$$G = \pm P_{<u_1>}^\perp + \cos\alpha P_{<u_1>^\perp}^\perp + \sin\alpha M_{u_1}$$

Y como ya sabemos obtener matrices de proyección (Y ambas están relacionadas fácilmente por $Id = P_{<u_1>}^\perp + P_{<u_1>^\perp}^\perp$), y la matriz de la aplicación de producto vectorial también es sencilla, podemos obtener la matriz de G en cualquier base.

Observación 10. Esto da lugar a otra manera de obtener $\sin\alpha$, sin necesidad de obtener bases ortonormales. La idea es que los dos primeros sumandos anteriores conforman una matriz autoadjunta g, y el último sumando es una matriz antiautoadjunta h. Es decir, si denotamos al giro f, tenemos que $f = g + h$ luego $adj(f) = g - h$, y por tanto $h = \frac{f - adj(f)}{2}$ (esto vale para descomponer cualquier otra función en suma de autoadjunta y antiautoadjunta).

Es decir, dada la matriz A en canónica del giro, podemos obtener $B = \frac{1}{2}(A - A^t)$ (con el producto usual para que canónica sea ortonormal), y esta matriz B debe verificar $B = \sin\alpha M_c(M_{u_1})$, que de hecho, si $u_1 = (a, b, c)$ (unitario, como era necesario en la demostración anterior), entonces:

$$B = \sin\alpha \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}$$

Como el vector es unitario, se tiene que, para la matriz B, $b_{21}^2 + b_{31}^2 + b_{32}^2 = \sin^2\alpha(a^2 + b^2 + c^2) = \sin^2\alpha$.

En resumen, **si se toman esas entradas de $\frac{1}{2}(A - A^t)$ y se elevan al cuadrado, se obtiene el seno al cuadrado. Si luego se dividen por $\sin\alpha$, se obtienen las componentes del eje u_1.**

2. Geometría Afín

Hasta ahora, la teoría de álgebra lineal aplicada a la geometría tiene una importante restricción: solo son subespacios aquellos elementos que pasan por el origen. El objetivo ahora es poder extender de manera natural todo a un espacio en el que no haya puntos distinguidos como el origen:

Definición 22. El conjunto no vacío A es un **espacio afín** sobre un $\mathbb{K}$-espacio vectorial V si está dotado de la operación $\varphi : A \times A \to V$, que se suele denotar $\varphi(p,q) = \overrightarrow{pq}$, tal que verifica las propiedades:

1. $\forall p \in A$, se tiene que $\varphi_p : A \to V$, dada por $\varphi_p(q) = \overrightarrow{pq}$, es biyectiva.
2. $\forall p,q,r \in A$, se tiene $\overrightarrow{pq} + \overrightarrow{qr} = \overrightarrow{pr}$.

La **dimensión** del espacio afín es $\dim V$.

A los elementos del espacio afín se les suele denominar puntos. El espacio V es la **dirección** del espacio afín. El primer axioma se puede entender como que cada punto encierra un espacio vectorial, V, centrado en él. El segundo axioma es bastante intuitivo si se visualiza φ como la función que proporciona un vector con origen y extremo en un punto dado.

Un ejemplo de espacio afín es (V,V,φ) donde V es un $\mathbb{K}$-espacio que además actúa como conjunto de puntos, con el operador $\varphi(v,w) = w - v$, independientemente de la base.

Las siguientes propiedades siguen de inmediato:

Proposición 26. *Se verifican en un espacio afín A:*

1. $\forall p,q \in A$, $\overrightarrow{pq} = 0 \iff p = q$
2. $\forall p,q \in A$, $-\overrightarrow{pq} = \overrightarrow{qp}$
3. $\forall p,q,r,s \in A$, $\overrightarrow{pq} = \overrightarrow{rs} \implies \overrightarrow{pr} = \overrightarrow{qs}$ *(Ley del paralelogramo)*

Demostración:

1. $\Longleftarrow$ se tiene por el segundo axioma: $\overrightarrow{pp} + \overrightarrow{pp} = \overrightarrow{pp} \implies \overrightarrow{pp} = 0$. Para $\implies$, veamos que si $\overrightarrow{pq} = 0$, entonces $\overrightarrow{pp} = \overrightarrow{pq}$ como hemos visto antes, y por el primer axioma $p = q$.
2. Por el segundo axioma, $\overrightarrow{pq} + \overrightarrow{qp} = \overrightarrow{pp} = 0$.
3. $\overrightarrow{pq} = \overrightarrow{rs} \implies \overrightarrow{pr} + \overrightarrow{rq} = \overrightarrow{rq} + \overrightarrow{qs} \implies \overrightarrow{pr} = \overrightarrow{qs}$

□

Cuando $\overrightarrow{pq} = u \in V$, se suele denotar $p + u = q$. A continuación generalizamos el concepto de subespacio a espacio afín:

Definición 23 (Variedad lineal)**.** Sea $a \in A$ y $F \subset V$ un subespacio en el espacio afín (A,V,φ). Se define la **variedad lineal** de a en dirección de F por:

$a + F = \{b \in A : \overrightarrow{ab} \in F\}$

Su dimensión es $\dim(a + F) := \dim F$.

Observación 11. Las variedades lineales son subespacios afines sobre el espacio F con la aplicación φ_{a+F}, que ahora lleva los vectores exclusivamente sobre F y hereda las propiedades de φ. Ejemplos son el punto $a + \{0\}$, la recta $a+ < u >$, etcétera.

Veamos que cualquier punto de la variedad sirve para representarla:

Proposición 27. *Si* $b \in a + F$*, entonces* $a + F = b + F$*.*

Demostración. Sea $x \in a+F$. Entonces $\overrightarrow{ax} \in F$. Como también $\overrightarrow{ab} \in F$, entonces $\overrightarrow{ba} \in F$ luego $\overrightarrow{ba}+\overrightarrow{ax}=\overrightarrow{bx} \in F$, así que $x \in b+F$. A la inversa, se tiene que si $x \in b+F$, entonces $\overrightarrow{bx} \in F$ luego $\overrightarrow{ab}+\overrightarrow{bx}=\overrightarrow{ax} \in F$. □

Corolario. Si $p, q \in a+F$, entonces $\overrightarrow{pq} \in F$, dado que $\overrightarrow{pa}+\overrightarrow{aq} \in F$.

A continuación veremos como estudiar la posición relativa de dos variedades lineales:

Proposición 28. *Dos variedades $a+F$ y $b+G$ del espacio afín (A, V, φ) se cortan (es decir, su intersección es no vacía) $\iff \overrightarrow{ab} \in F+G$.*

Demostración. Si se cortan, sea $c \in (a+F) \cap (b+G)$. Entonces, $\overrightarrow{ac} \in F$ y $\overrightarrow{bc} \in G \implies \overrightarrow{cb} \in G$. Por lo tanto, $\overrightarrow{ab}=\overrightarrow{ac}+\overrightarrow{cb} \in F+G$. Para el recíproco, si $\overrightarrow{ab} \in F+G$, entonces $\overrightarrow{ab}=u+v$ con $u \in F$, $v \in G$. Como ϕ_a es biyectiva, entonces $\exists c \in A$ tal que $\overrightarrow{ac}=u$, de modo que $c \in a+F$. Veremos que también está en la otra: $\overrightarrow{ab}-\overrightarrow{ac}=v \in G \implies \overrightarrow{cb} \in G \implies \overrightarrow{bc} \in G \implies c \in b+G$, con lo que se cortan en c. □

Definición 24. Dos variedades $a+F$ y $b+G$ son **paralelas** si $F \subseteq G$ o $G \subseteq F$.

Proposición 29. *Sean $a+F$ y $b+G$ dos variedades paralelas con $F \subseteq G$ sin perder en generalidad. Entonces, o bien $(a+F) \cap (b+G)=\emptyset$, o bien $(a+F) \subseteq (b+G)$.*

Demostración. Supongamos que su intersección es no vacía. Entonces, si $c \in (a+F) \cap (b+G)$, se cortan, con lo que $\overrightarrow{ab} \in F+G=G$, con lo que $\overrightarrow{ba} \in G$, es decir $a \in b+G$ luego $a+G=b+G$. Naturalmente, $a+F \subseteq a+G$, dado que si $x \in a+F$, $\overrightarrow{ax} \in F \implies \overrightarrow{ax} \in G \implies x \in a+G$, luego se tiene lo que se quería. □

Definición 25. Dos variedades lineales se **cruzan** si ni se cortan ni son paralelas.

2.1. Intersección y suma de variedades lineales. Variedad generada.

Proposición 30. *Sean $a+F$, $b+G$ variedades lineales en (A, V, φ) que se cortan, con $c \in (a+F) \cap (b+G)$. Entonces, $(a+F) \cap (b+G)=c+(F \cap G)$ y por tanto es otra variedad lineal.*

Demostración. Observemos primero que $\overrightarrow{ac} \in F$ y $\overrightarrow{bc} \in G$. Sea $x \in (a+F) \cap (b+G)$. Entonces $\overrightarrow{xc}=\overrightarrow{xa}+\overrightarrow{ac} \in F$ al estar ambos en F ($\overrightarrow{xa}=-\overrightarrow{ax}$ que está en F), y análogamente $\overrightarrow{xc}=\overrightarrow{xb}+\overrightarrow{bc} \in G$, luego $\overrightarrow{xc} \in F \cap G$ así que $x \in c+F \cap G$. Por otra parte, si $x \in c+F \cap G$, entonces $\overrightarrow{xc} \in F, G$ y entonces $\overrightarrow{xa}=\overrightarrow{xc}-\overrightarrow{ac} \in F$ y $\overrightarrow{xb}=\overrightarrow{xc}-\overrightarrow{bc} \in G$ luego $x \in (a+F) \cap (b+G)$. □

Por otra parte y al igual que sucedía en subespacios, la unión de 2 variedades no es variedad lineal, dado que debería contener muchos puntos más que se generan. Para aplicar este concepto a variedades, definimos:

Definición 26. Dados los puntos $\{a_0, ...a_k\} \subset A$, se define la **variedad lineal generada** por tales puntos como $L(a_0, ..., a_k)=a_0+<\overrightarrow{a_0a_1}, \overrightarrow{a_0a_2}, ..., \overrightarrow{a_0a_k}>$.

Proposición 31. *La variedad L generada por los puntos $\{a_0, ..., a_k\}$ es la menor que los contiene.*

Demostración. Claramente contiene a los puntos, dado que como cada $\overrightarrow{a_0a_i}$ está en el subespacio, se tiene por definición que $a_i \in L$. Por otra parte, sea $a+F$ una variedad tal que $\{a_0, ..., a_k\} \subset a+F$. Entonces, $a+F=a_0+F$ y por tanto $\overrightarrow{a_0a_i} \in F \ \forall i=1, ..., k$, lo que quiere decir que al ser subespacio $<\overrightarrow{a_0a_1}, ...\overrightarrow{a_0a_k}> \subset F$, o lo que es lo mismo, $L \subset a+F$. □

Definición 27. Se define la **suma** de las variedades lineales $L_1=a+F$, $L_2=b+G$ como la variedad lineal $L_1+L_2 := a+(F+G+<\overrightarrow{ab}>)$.

Esta definición es la que garantiza que contiene a ambas variedades lineales y que es la mínima con tal propiedad. Veámoslo:

Proposición 32. *La variedad $L_1 + L_2$ es la mínima que contiene a L_1 y a L_2.*

Demostración. Manteniendo la notación de la definición. Si $x \in L_1$, $\overrightarrow{ax} \in F$ con lo que $\overrightarrow{ax} \in F + G + < \overrightarrow{ab} >$ y $x \in L_1 + L_2$. Por otro lado, veamos que $b \in L_1 + L_2$, ya que $\overrightarrow{ab} \in < \overrightarrow{ab} >$, luego $L_1 + L_2 = b + (F + G + < \overrightarrow{ab} >)$, y análogamente al anterior ya se puede ver que $L_2 \subset L1 + L2$.

Para ver que es la menor, sea $c + H$ tal que $L_1 \subset c + H$, $L_2 \subset c + H$. Si se tiene eso, $c + H = a + H$ y entonces, como $a + F \subset c + H$, debe ser que $F \subset H$. Análogamente $G \subset H$. Como $a, b \in c + H$, también $\overrightarrow{ab} \in H$ luego $< \overrightarrow{ab} > \subset H$. Como la suma de subespacios es el menor que los contiene, debe ser que $F + G + \overrightarrow{ab} \subset H$, y se tiene lo que se quería ya que $L_1 + L_2$ y $c + H$ pasan por el mismo punto. □

Proposición 33 (Grassmann)**.** *Sean $L_1 = a + F$, $L_2 = b + G$ variedades lineales. Si se cortan, se tiene:*

$$\dim L_1 + L_2 = \dim L_1 + \dim L_2 - \dim L_1 \cap L_2$$

Si no se cortan no tiene sentido hablar de variedad intersección y la fórmula varía:

$$\dim L_1 + L_2 = \dim L_1 + \dim L_2 - \dim F \cap G + 1$$

.

Demostración. Si se cortan, $\overrightarrow{ab} \in F + G$ con lo que $(F + G + \overrightarrow{ab}) = F + G$, luego $\dim L_1 + L_2 = \dim F + \dim G - \dim F \cap G$ por Grassmann de espacios vectoriales, de donde sigue lo que se quería. Si no, entonces $\dim(F + G + < \overrightarrow{ab} >) = \dim(F + G) + \dim(< \overrightarrow{ab} >) - \dim((F + G) \cap < \overrightarrow{ab} >) = \dim F + \dim G - \dim F \cap G + 1 - 0$, como se quería ver. Que $\dim((F + G) \cap < \overrightarrow{ab} >)$ es cero se sabe porque al no cortarse, $\overrightarrow{ab} \notin F + G$. □

2.2. Sistemas de Referencia en Espacios Afines

Buscamos generalizar el concepto de base a espacios afines.

Definición 28. Un **sistema de referencia cartesiano** en un espacio afín n-dimensional (A, E, φ) es el conjunto $R = \{p; e_1, ..., e_n\} = \{p; B\}$ donde $p \in A$ es el 'origen' y $B = \{e_1, ..., e_n\}$ es una base de E. Para cualquier $x \in A$, sus coordenadas son $(x)_R = (\overrightarrow{px})_B = (x_1, ..., x_n)_B$. Son únicas por el primer axioma de espacio afín y la unicidad de coordenadas dada una base.

Observación 12. Si $b = a + \vec{w}$, es decir, $\overrightarrow{ab} = \vec{w}$, entonces $(b)_R = (a)_R + (\vec{w})_B$ de manera muy conveniente, dado que $(\overrightarrow{pb})_B = (\overrightarrow{pa})_B + (\overrightarrow{ab})_B$.

Para definir el otro tipo de sistema (baricéntrico) se requieren conceptos previos:

Definición 29. Los puntos $\{a_1, ..., a_k\} \subset A$ son **afínmente independientes** si $\{\overrightarrow{a_1a_2}, \overrightarrow{a_1, a_3}, ..., \overrightarrow{a_1a_k}\}$ es un conjunto linealmente independiente de vectores.

Proposición 34. *Son equivalentes:*

1. *$\{a_1, ..., a_k\}$ son afínmente independientes.*
2. *$\forall 1 \leq i \leq k$, $\{\overrightarrow{a_ia_h} : h \neq i\}$ son linealmente independientes.*
3. *$\forall p \in A$, si se tiene $\sum_1^k \lambda_j \overrightarrow{pa_j} = 0$ y además $\sum_1^k \lambda_j = 0$, entonces $\lambda_1 = ... = \lambda_k = 0$.*

Demostración. (1) $\implies$ (2) se tiene porque si $0 = \sum_{h\neq i} \lambda_h \overrightarrow{a_i a_h} = \sum_{h\neq i} \lambda_h (a_i a_1 + a_1 a_h) = -(\sum_{h\neq i} \lambda_h) a_1 a_i + \sum_{h\neq i} \lambda_h a_1 a_h$, y como es una combinación nula de los vectores que según (1) son independientes, se tiene $\lambda_h = 0\ \forall i \neq h$, como se quería ver.

(2) $\implies$ (1) se tiene particularizando (2) con $i = 1$.

(1) $\implies$ (3) se tiene porque $0 = \sum_1^k \lambda_j \overrightarrow{pa_j} = \sum_1^k \lambda_j (\overrightarrow{pa_1} + \overrightarrow{a_1 a_j}) = (\sum_1^k \lambda_j)\overrightarrow{pa_1} + \sum_2^k \lambda_j \overrightarrow{a_1 a_j} = 0 + \sum_2^k \lambda_j \overrightarrow{a_1 a_j}$, y por (1) se tiene que $\lambda_j = 0\ \forall j = 2, ..., k$, no obstante, como $\lambda_1 = -\sum_2^k \lambda_j$ por hipótesis, entonces también $\lambda_1 = 0$.

(3) $\implies$ (1) se tiene porque si $p \in A$ arbitrario y $\sum_2^k \lambda_j \overrightarrow{a_1 a_j} = 0$ y definimos $\lambda_1 = -\sum_2^k \lambda_j$, entonces $0 = \sum_2^k \lambda_j \overrightarrow{a_1 a_j} = \sum_2^k \lambda_j (\overrightarrow{a_1 p} + \overrightarrow{pa_j}) = -\lambda_1 \overrightarrow{a_1 p} + \sum_2^k \lambda_j \overrightarrow{pa_j} = \lambda_1 \overrightarrow{pa_1} + \sum_2^k \lambda_j \overrightarrow{pa_j}$, donde además $\sum_1^k \lambda_j = 0$, luego por (3) sigue lo que se quería: $\lambda_1 = ... = \lambda_k = 0$ $\square$

A continuación vamos a ver que se puede saber cuántos puntos afínmente independientes hay como máximo. Evidentemente, en un espacio de dimensión n, un conjunto de $n+2$ puntos nunca serán independientes ya que dan lugar a $n+1$ vectores. No obstante:

Proposición 35. *Sea (A, E, φ) espacio afín n-dimensional. Sean $a_0, ..., a_k$ $(k \leq n)$ afínmente independientes. $\exists\{a_{k+1}, ..., a_n\}$ tales que $\{a_0, ..., a_n\}$ son $n+1$ puntos afínmente independientes.*

Demostración. Completamos el sistema $\{\overrightarrow{a_0 a_1}, ..., \overrightarrow{a_0 a_k}\}$ a base de V: $B = \{\overrightarrow{a_0 a_1}, ..., \overrightarrow{a_0 a_k}, v_{k+1}, ..., v_n\}$. Ahora basta con poner a_j como aquel tal que $a_0 a_j = v_j\ \forall j \in \{k+1, ..., n\}$. $\square$

El siguiente lema prepara los sistemas baricéntricos:

Lema 3. *Sean $p_0, ..., p_k$ puntos de (A, E, φ) un espacio afín n-dimensional. Sea $x \in A$. Si $\exists p \in A$ tal que:*

$$\begin{cases} \overrightarrow{px} = \sum_{i=0}^k x_i \overrightarrow{pp_i} & (1) \\ \sum_{i=0}^k x_i = 1 & (2) \end{cases}$$

Para algunos $x_0, ..., x_k \in \mathbb{K}$, entonces:

1. *Dado $q \in A$, vale la expresión (1) reemplazando p por q.*

2. *Si $\{p_0, ..., p_k\}$ son afínmente independientes, los escalares x_i son únicos, es decir, no hay otros tales que esa suma de $\overrightarrow{px}$.*

Demostración. Para 1, sea $q \in A$. Entonces $\overrightarrow{qx} = \overrightarrow{qp} + \overrightarrow{px} = \overrightarrow{qp} + \sum_0^k x_i \overrightarrow{pp_i} = \overrightarrow{qp} + \sum_0^k x_i (\overrightarrow{pq} + \overrightarrow{qp_i}) = (-1 + \sum_0^k x_i)\overrightarrow{pq} + \sum_0^k x_i \overrightarrow{qp_i} = 0 + \sum_0^k x_i \overrightarrow{qp_i}$, como se quería.

Para 2, supongamos que son independientes, y que $\{y_0, ..., y_n\}$ también cumplen lo que se dice. Entonces, $\sum_0^k x_i \overrightarrow{pp_i} = \sum_0^k y_i \overrightarrow{pp_i} \implies \sum_0^k (x_i - y_i)\overrightarrow{pp_i} = 0$, y como además $\sum_0^k (x_i - y_i) = 1 - 1 = 0$, entonces por las equivalencias vistas anteriormente, se tiene que $x_i - y_i = 0\ \forall i \in \{0, ..., k\}$, de manera que $x_i = y_i$ para toda i. $\square$

Ahora veremos si se tienen justo $n+1$ puntos independientes, **todo punto puede ser expresado como en el lema anterior**, y como hemos visto, de forma única. Esto constituye el **sistema baricéntrico o afín**:

Definición 30 (Sistema de referencia baricéntrico o afín)**.** Un sistema afín del espacio n-dimensional (A, E, φ) es un conjunto $R = \{a_0, ..., a_n\}$ de $n+1$ puntos afínmente independientes. Las coordenadas de un punto $x \in A$ se forman de esta manera: por la proposición 34, y como $\{x, a_0, ..., a_n\}$ ya no son independientes, $\exists p$ tal que:

$$0 = \lambda \overrightarrow{px} + \sum_0^n \lambda_i \overrightarrow{pa_i}$$

con $\lambda + \sum_0^n \lambda_i = 0$ y no todos los escalares nulos. Además, no puede ser que $\lambda = 0$ porque en ese caso se contradice la proposición 34. Por tanto, despejando y multiplicando por el inverso de λ, queda que existen x_i con:

$$\overrightarrow{px} = \sum_0^n x_i \overrightarrow{pa_i}$$

Estas son las **coordenadas baricéntricas**: $(x)_R = (x_0, ..., x_n)$. Además, $\sum_0^n x_i = \frac{-\sum_0^n \lambda_i}{\lambda} = \frac{\lambda}{\lambda} = 1$, luego por el lema previo, son únicas y además no dependen del punto p elegido.

Es decir, finalmente, las coordenadas de $x \in A$ son $(x_0, ..., x_n)_R$ tales que $\overrightarrow{px} = \sum_0^n x_i \overrightarrow{pa_i}$, sin importar $p \in A$, que además cumplan $\sum_0^k x_i = 1$.

En este caso, **se denota** $x = x_0 a_0 + x_1 a_1 + ... + x_n a_n$.

Observación 13. Dada $R = \{a_0, ..., a_n\}$ una referencia baricéntrica y $x \in A$ tal que $(x)_R = (x_0, ..., x_n)$, entonces, si $R_c = \{a_i, \overrightarrow{a_i a_h} : h \neq i\}$ es la referencia cartesiana con origen en a_i y base los vectores que parten de a_i, las coordenadas de x son: $(x)_{R_c} = (x_0, ..., x_{i-1}, x_{i+1}, ..., x_n)$.

Esto es así porque $\overrightarrow{a_i x} = \sum_{j=0}^n x_j \overrightarrow{a_i a_j} = \sum_{j \neq i} x_j \overrightarrow{a_i a_j}$.

Observación 14. Dado $R = \{p_0; v_1, ..., v_n\}$ referencia cartesiana con $(x)_R = (x_1, ..., x_n)$, entonces en el sistema $R_b = \{p_0, p_1 = p_0 + v_1, ..., p_n = p_0 + v_n\}$, que, dado que $\overrightarrow{p_0 p_i} = v_i$, son puntos independientes, se tiene que $(x)_{R_b} = (1 - \sum_1^n x_i, x_1, x_2, ..., x_n)$, dado que $\overrightarrow{p_0 x} = \sum_1^n x_i \overrightarrow{p_0 p_i}$, y por lo tanto solo hace falta cuadrar la primera coordenada para que sumen 1.

2.2.1. Cambio de coordenadas

Proposición 36 (Cambio en sistemas afines)**.** *Dados dos sistemas afines $R = \{p_0, ..., p_n\}$ y $R' = \{p'_0, ..., p'_n\}$, podemos cambiar de R' a R si conocemos las coordenadas del sistema R' en R. En este caso, si $(p'_j)_R = (\lambda_{0j}, ..., \lambda_{nj})$ y consideramos la matriz de paso $P = (\lambda_{ij})_{i,j}$, se tiene, $\forall x \in A$:*

$$[x]_R = P \cdot [x]_{R'}$$

Donde, como hemos dicho, P es:

$$P = \begin{pmatrix} [p'_0]_R & [p'_1]_R & \dots & [p'_n]_R \end{pmatrix}$$

Demostración. Supongamos que $(x)_R = (x_0, ..., x_n)$ y $(x)_{R'} = (x'_0, ..., x'_n)$. Entonces, fijado $p \in A$, $\overrightarrow{px} = \sum_0^n x'_i \overrightarrow{pp'_i} = \sum_{i=0}^n (x'_i \sum_{j=0}^n \lambda_{ji} \overrightarrow{pp_j}) = \sum_{i,j=0}^n x'_i \lambda ji \overrightarrow{pp_j}$, es decir, que $x_j = x'_i \lambda ji$, o, en notación matricial:

$$\begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda_{00} & \dots & \lambda_{0n} \\ \vdots & \ddots & \vdots \\ \lambda_{n0} & \dots & \lambda_{nn} \end{pmatrix} \cdot \begin{pmatrix} x'_0 \\ x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

□

Proposición 37 (Cambio en sistemas cartesianos)**.** *Dados dos sistemas cartesianos, $R = \{p, B = \{e_1, ..., e_n\}\}$ y $R' = \{p', B' = \{e'_1, ..., e'_n\}\}$, podemos cambiar de R' a R si conocemos el cambio de base de B' a B y p' en R. Suponiendo que $P = (\lambda_{ij})$ es el cambio de base, es decir, que $e'_j = \sum_1^n \lambda_{ij} e_i$, y que $(p')_R = (p_1, ..., p_n)$, entonces, dado $x \in A$:*

$$[x]_R = P \cdot [x]_{R'} + \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}$$

Donde, como hemos dicho, P es:

$$P = \begin{pmatrix} [e'_1]_B & [e'_2]_B & \dots & [e'_n]_B \end{pmatrix}$$

Demostración. Pongamos que $(x)_R = (x_1, ..., x_n)$ y $(x)'_R = (x'_1, ..., x'_n)$. Entonces, $\overrightarrow{p'x} = \sum_{j=1}^n x'_j e'_j = \sum_{j=1}^n x'_j \sum_{i=1}^n \lambda_{ij} e_i = \sum_{i=1}^n \sum_{j=1}^n (x'_j \lambda_{ij}) e_i$. Por otra parte, $\overrightarrow{p'x} = \overrightarrow{p'p} + \overrightarrow{px} = \sum_1^n (x_i - p_i) e_i$. Igualando los coeficientes anteriores, queda que $\sum_{j=1}^n x'_j \lambda_{ij} = (x_i - p_i)\ \forall i \in \{1, ..., n\}$, o, matricialmente, que:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda_{11} & \dots & \lambda_{1n} \\ \vdots & \ddots & \vdots \\ \lambda_{n1} & \dots & \lambda_{nn} \end{pmatrix} \cdot \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} + \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}$$

□

2.3. Ecuaciones de variedades lineales

Proposición 38 (Ecuaciones paramétricas)**.** *Dada $a + F$ una variedad lineal k-dimensional en (A, E, φ) n-dimensional, con base de F: $\{v_1, ..., v_k\}$, y dado un sistema cartesiano $R = \{p; B = \{e_1, ..., e_n\}\}$, donde $[v_j]_B = (v_{1j}, v_{2j}, ..., v_{nj})$, y $(a)_R = (a_1, ..., a_n)$, entonces, dado $x \in A$ con $(x)_R = (x_1, ..., x_n)$, se tiene que $x \in a + F \iff \overrightarrow{ax} \in F$, y como $\overrightarrow{ax} = \overrightarrow{ap} + \overrightarrow{px}$, entonces $[\overrightarrow{ax}]_B = (x_1 - a_1, ..., x_n - a_n)$. Por tanto,*

$$x \in a + F \iff \begin{pmatrix} x_1 - a_1 \\ x_2 - a_2 \\ \vdots \\ x_n - a_n \end{pmatrix} = \sum_1^n \lambda_i \begin{pmatrix} v_{1i} \\ v_{2i} \\ \vdots \\ v_{ni} \end{pmatrix}$$

*para algunos escalares λ_i. Esto se conoce como **ecuaciones paramétricas** de la variedad.*

Eliminando los parámetros, o bien con el método siguiente, se obtienen las ecuaciones cartesianas:

Proposición 39 (Ecuaciones cartesianas)**.** *Con la notación anterior, $x \in a + F$ si y solo si $\overrightarrow{ax}$ es combinación de la base de F, es decir:*

$$x \in a + F \iff rg \begin{pmatrix} v_{11} & \dots & v_{1k} \\ \vdots & \ddots & \vdots \\ v_{n1} & \dots & v_{nk} \end{pmatrix} = rg \begin{pmatrix} v_{11} & \dots & v_{1k} & x_1 - a_1 \\ \vdots & \ddots & \vdots & \vdots \\ v_{n1} & \dots & v_{nk} & x_n - a_n \end{pmatrix}$$

Suponiendo que el rango de la primera matriz (la dimensión del subespacio) es r, bastará con tomar la segunda matriz y escalonarla con el método de gauss, imponiendo que las $n - r$ filas finales sean nulas para que mantenga rango r. Esto nos proporciona r ecuaciones:

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{r1}x_1 + \dots + a_{rn}x_n = b_r \end{cases}$$

*Que son precisamente las **ecuaciones cartesianas.** Alternativamente se pueden obtener a través de los determinantes menores: el mayor menor no nulo debe ser de tamaño r, luego se forman $n-r$ menores independientes de tamaño $r+1$, y se igualan a cero.*

Análogamente, dadas ecuaciones cartesianas como las anteriores, describen una variedad lineal de subespacio dirección:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \vdots \\ a_{r1}x_1 + \cdots + a_{rn}x_n = 0 \end{cases}$$

Y que pasa por un punto solución al sistema.

El siguiente teorema proporciona la dimensión de la variedad generada por una serie de puntos. Para el mismo, necesitaremos un sencillo resultado previo:

Lema 4. *Dado $R = \{p_0, ..., p_n\}$ un sistema baricéntrico y $q_0, ..., q_m$ puntos tales que $(q_j)_R = (q_{0j}, ..., q_{nj})$. Se tiene que en base $\{\overrightarrow{p_0p_1}, ..., \overrightarrow{p_0p_n}\}$, las coordenadas de cualquier $\overrightarrow{q_0q_j}$ son $(q_{1j} - q_{10}, ..., q_{nj} - q_{n0}) = (q_j)_R - (q_0)_R$. Es decir, en la 'base inducida por el sistema baricéntrico', las coordenadas se pueden obtener restando.*

Demostración: $\overrightarrow{q_0q_j} = \overrightarrow{p_0q_j} - \overrightarrow{p_0q_0} = \sum_0^n (q_{ij} - q_{i0})\overrightarrow{p_0p_i}$, y como $\overrightarrow{p_0p_0}$ es nulo se tiene lo que se quería. □

Teorema 8. *Sea $R = \{p_0, ..., p_n\}$ sistema baricéntrico y$q_0, ..., q_m$ puntos tales que $(q_j)_R = (q_{0j}, ..., q_{nj})$. Si $V = \mathcal{L}(q_0, ..., q_m)$, entonces:*

$$\dim V = rg \begin{pmatrix} q_{00} & \cdots & q_{0m} \\ \vdots & \ddots & \vdots \\ q_{n0} & \cdots & q_{nm} \end{pmatrix} - 1 = rg \begin{pmatrix} (q_0)_R & \cdots & (q_m)_R \end{pmatrix} - 1$$

Demostración.

$$rg \begin{pmatrix} q_{00} & \cdots & q_{0m} \\ \vdots & \ddots & \vdots \\ q_{n0} & \cdots & q_{nm} \end{pmatrix} = rg \begin{pmatrix} q_{00} & q_{01} - q_{00} & \cdots & q_{0m} - q_{00} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n0} & q_{n1} - q_{n0} & \cdots & q_{nm} - q_{n0} \end{pmatrix}$$

Ahora, sumamos a la primera fila todas las demás, y como son baricéntricas:

$$= rg \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ q_{n0} & q_{n1} - q_{n0} & q_{nm} - q_{n0} & \end{pmatrix} = 1 + rg \begin{pmatrix} q_{11} - q_{10} & \cdots & q_{1m} - q_{10} \\ \vdots & \ddots & \vdots \\ q_{n1} - q_{n0} & \cdots & q_{nm} - q_{n0} \end{pmatrix} = 1 + \dim V$$

Dado que esta última matriz, por el Lema 4, contiene las coordenadas de los vectores de la base de V escritas en una base. □

Observación 15. Las ecuaciones implícitas en coordenadas baricéntricas también se pueden hallar con el razonamiento de antes: $x \in A$, con $(x)_R = (x_1, ..., x_n)$, está en $V = \mathcal{L}(q_0, ..., q_m)$ si y solo si $rg([\overrightarrow{q}_0q_1]_B| \dots |[\overrightarrow{q}_0q_m]_B) = rg([\overrightarrow{q}_0q_1]_B| \dots |[\overrightarrow{q}_0q_m]_B|[x_i - q_{i0}]_1^n)$, donde B es la base inducida $B = \{\overrightarrow{q_0q_1}, ..., \overrightarrow{q_0q_n}\}$, lo que por el Lema 4 nos garantiza que las coordenadas de $\overrightarrow{q_0x}$ sean esas. Imponiendo esa condición se obtienen $n-m$ ecuaciones, y la que falta (dado que $\dim V = m-1$ y por tanto debe haber $n-m+1$ ecuaciones, suponiendo siempre que los m puntos sean independientes), se obtiene imponiendo $\sum_0^n x_i = 1$ para obtener coordenadas baricéntricas.

Si ya se tienen las ecuaciones en cartesianas se obtienen fácilmente las del sistema de referencia baricéntrico asociado agregando la condición $\sum_0^n x_i = 1$.

2.4. Afinidades

El concepto de afinidad generaliza el de aplicación lineal:

Definición 31. Dados dos espacios afines, (A_1, E_1, φ_1) y (A_2, E_2, φ_2), se dice que la aplicación $f : A_1 \to A_2$ es **aplicación afín o afinidad** si $\exists \tilde{f} : E_1 \to E_2$ lineal tal que $\forall a, b \in A_1$, se tiene $\varphi_2(f(a), f(b)) = \tilde{f}(\varphi_1(a, b))$, es decir, $\overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab})$. Se dice que $\tilde{f}$ es la **aplicación lineal asociada** a la afinidad.

En adelante, si no se indica lo contrario, las aplicaciones afines irán entre (A_1, E_1, φ_1) y (A_2, E_2, φ_2). La siguiente caracterización es importante:

Proposición 40. *Son equivalentes:*

1. *f es aplicación afín.*
2. *$f(a + \vec{u}) = f(a) + \tilde{f}(\vec{u})$ $\forall a \in A_1$ y $\vec{u} \in E_1$. Es decir, $f(b) = f(a) + \tilde{f}(\overrightarrow{ab})$ $\forall a, b \in A_1$.*

Demostración. Supongamos que se tiene 1, y sea $b \in A_1$ tal que $a + \vec{u} = b$. Entonces sabemos de la definición de afinidad que $f(b) = f(a) + \tilde{f}(\overrightarrow{ab})$, es decir, $f(a + \vec{u}) = f(a) + \tilde{f}(\vec{u})$. Por otra parte, si se tiene 2, sean a, $b \in A_1$ arbitrarios, entonces $f(b) = f(a) + \tilde{f}(\overrightarrow{ab})$ (particularizamos para $\vec{u} = \overrightarrow{ab}$, con lo cual $\overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab})$. $\square$

Una aplicación afín queda completamente determinada por su parte lineal y la imagen de un punto:

Proposición 41. *Sean f, g afinidades tales que $f(p) = g(p)$ para un $p \in A_1$, y $\tilde{f} = \tilde{g}$. Entonces, $f = g$.*

Demostración. Dado $a \in A_1$ arbitrario, $f(a) = f(p) + \tilde{f}(\overrightarrow{pa}) = g(p) + \tilde{g}(\overrightarrow{pa}) = g(a)$, luego son iguales. $\square$

Se comportan más o menos como aplicaciones lineales en tanto que basta con dar las imágenes de una 'base' para obtenerlas. Para ver cuándo existen aplicaciones afines usaremos el siguiente lema previo:

Lema 5. *Sean $q \in A_1$, $u, v \in E_1$. Entonces, $\overrightarrow{(q + u)(q + v)} = v - u$.*

Demostración. Evidentemente, $q + v = (q + u) + v - u$, pero esto es equivalente a que $\overrightarrow{(q + u)(q + v)} = v - u$.

Proposición 42. *Sea $\Phi : E_1 \to E_2$ una aplicación lineal, y sean $p \in A_1$ y $q \in A_2$. Entonces, $\exists! f : A_1 \to A_2$ afín tal que $f(p) = q$ y $\tilde{f} = \Phi$.*

Demostración. Dicha aplicación es la definida por $f(a) = f(p) + \Phi(\overrightarrow{pa}) = q + \Phi(\overrightarrow{pa})$, $\forall a \in A_1$. Efectivamente es afín, dado que para cualesquiera $a, b \in A_1$, se tiene $\overrightarrow{f(a)f(b)} = \overrightarrow{(q + \Phi(\overrightarrow{pa}))(q + \Phi(\overrightarrow{pb}))} = \Phi(\overrightarrow{pb}) - \Phi(\overrightarrow{pa}) = \Phi(\overrightarrow{ap} + \overrightarrow{pb}) = \Phi(\overrightarrow{ab})$, luego es afín de aplicación lineal Φ, y efectivamente $f(p) = q + \Phi(\overrightarrow{pp}) = q + \Phi(0) = q$. Además es única por la proposición 41. $\square$

Proposición 43. *Dado el sistema afín $\{a_0, ..., a_n\} \subset A_1$ de puntos independientes, con $n = \dim E_1$, y dados $\{b_0, ..., b_n\} \subset A_2$ puntos cualesquiera del espacio de llegada, $\exists! f : A_1 \to A_2$ afín tal que $f(a_j) = b_j$ $\forall j \in \{0, ..., n\}$. Es decir, queda determinada por las imágenes de un sistema afín.*

Demostración. Como $\{\overrightarrow{a_0a_1}, ..., \overrightarrow{a_0a_n}\}$ es base de E_1, definimos $\tilde{f}$ como la única aplicación lineal tal que $\tilde{f}(\overrightarrow{a_0a_j}) = \overrightarrow{b_0b_j}$ $\forall j \in \{1, ..., n\}$. Ahora, definimos f como la única afín de componente lineal $\tilde{f}$ y tal que $f(a_0) = b_0$. Por la proposición anterior existe, es única y está dada por $f(a) = b_0 + \tilde{f}(\overrightarrow{a_0a})$. Ahora veamos que satisface lo pedido: $f(a_j) = b_0 + \tilde{f}(a_0a_j) = b_0 + \overrightarrow{b_0b_j} = b_j$ $\forall j \in \{1, ..., n\}$, y además $f(a_0) = b_0 + 0 = b_0$. $\square$

2.4.1. Afinidades importantes

Definición 32. Sea (A, E, φ) espacio afín y sea $v \in E$. Se define la **traslación de vector v** como la aplicación $T_v : A \to A$ dada por $T_v(a) = a + \vec{v}\ \forall a \in A$, es decir, $\overrightarrow{aT_v(a)} = v$.

Proposición 44. *La traslación de vector $\vec{v}$ es afín, de aplicación asocidada Id_E.*

Demostración. Sean $a, b \in A$. Se tiene $\overrightarrow{T_v(a)T_v(b)} = \overrightarrow{T_v(a)a} + \overrightarrow{ab} + \overrightarrow{bT_v(b)} = -v + \overrightarrow{ab} + v = \overrightarrow{ab}$, luego es una afinidad de aplicación lineal asociada Id_E. □

Proposición 45. *Sea $f : A \to A$ una afinidad tal que $\tilde{f} = Id_E$. Entonces es una traslación para algún vector v.*

Demostración. Como $\overrightarrow{f(a)f(b)} = \overrightarrow{ab}$, para todos a y b, se tiene por la identidad del paralelogramo que $\overrightarrow{f(a)a} = \overrightarrow{f(b)b}$. Como a y b son arbitrarios, sigue que $\exists v_0$ tal que $\overrightarrow{f(a)a} = v_0\ \forall a \in A$. Es decir, es la traslación de vector $-v_0$, dado que $f(p) = p + \overrightarrow{pf(p)} = p - v_0\ \forall p \in A$. □

Definición 33. La afinidad $f : A \to A$ es una **proyección** si $f^2 = f$.

Proposición 46. *Dada una proyección afín, su aplicación asociada cumple $\tilde{f}^2 = \tilde{f}$, es decir, es una proyección sobre $Im(\tilde{f})$.*

Demostración. Fijado $a \in A$, todo vector de E admite una expresión como $\overrightarrow{ab}$. Como $\vec{f}^2(\overrightarrow{ab}) = \vec{f}(\overrightarrow{f(a)f(b)}) = \overrightarrow{f^2(a)f^2(b)} = \overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab})$. □

Observación 16. Los puntos fijos de una proyección f son exactamente su imagen, es decir, $Im(f) = \{a \in A : f(a) = a\}$.

Demostración. Claramente si a es punto fijo está en la imagen dado que $f(a) = a$. Por otra parte, si $a \in Im(f)$, $\exists b \in A$ con $f(b) = a$, luego $a = f(b) = f^2(b) = f(a)$, con lo que a es punto fijo. □

Observación 17 (Clasificación de proyecciones). Si $m(X)$ es el polinomio mínimo de $\tilde{f}$ con f una proyección, entonces sabemos que $m(X)|X^2 - X$. Distinguimos tres casos:

1. Si $m(X) = X$, entonces, $\tilde{f} = 0$, luego $\overrightarrow{(}f(a)f(b)) = 0\ \forall a, b \in A$, es decir, $f(a) = f(b)\ \forall a, b \in A$, luego se trata de la **proyección sobre el punto** $f(a)$. (Es una función constante).

2. Si $m(X) = X - 1$, entonces $\tilde{f} = Id$, con lo que se trata de una traslación, pero como ha de fijar puntos, o de lo contrario no tendría imagen, no queda otra que sea la identidad. $f = Id$.

3. Si $m(X) = X(X - 1)$ es una proyección propiamente dicha, con $E = S_0 \oplus S_1$ donde $S_0 = Nuc(\tilde{f})$ y $S_1 = Nuc(\tilde{f} - Id)$.

   Observemos que si $a \in A$ es un punto cualquiera, y $p \in Im(f)$ es punto fijo, y escribimos $\overrightarrow{pa} = u_1 + u_0$ cada uno en un autoespacio, entonces $f(a) = f(p) + \tilde{f}(u_1 + u_0) = p + u_1$, luego $f(a) \in p + S_1$.

   Por otro lado, como $\overrightarrow{f(a)a} = \overrightarrow{f(a)p} + \overrightarrow{pa} = -u_1 + u_1 + u_0 = u_0$, sigue que $f(a) \in a + S_0$.

   Es decir, $f(a) \in (p + S_1) \cap (a + S_0)$. Además como $\dim S_1 + \dim S_2 = \dim E$, y estaban en suma directa los autoespacios, estas variedades intersecan solo en el punto $f(a)$, así que se obtiene la siguiente proposición.

Proposición 47. *Sea $f : A \to A$ proyección afín, con S_0 y S_1 los autoespacios de núcleo y de puntos fijos de $\tilde{f}$. Entonces, si $p \in Im(f)$ es un punto fijo de f, se tiene:*

$$f(a) = (p + S_1) \cap (a + S_0)$$

Demostración. Se vio en la observación anterior. $\square$

Proposición 48. *Sea f afinidad de parte lineal $\tilde{f}$. Entonces:*

1. *f es inyectiva $\iff$ $\tilde{f}$ lo es.*
2. *f es sobreyectiva $\iff$ $\tilde{f}$ lo es.*

Demostración. 1. $\implies$. Supongamos que $\tilde{f}(u) = \tilde{f}(v)$. Entonces, fijado $a \in A_1$, se tiene $\overrightarrow{f(a)f(a+u)} = \overrightarrow{f(a)f(a+v)} \implies f(a+u) = f(a+v) \implies a+u = a+v \implies u = v$. Para $\impliedby$, supongamos que $f(a) = f(b)$. Entonces, $0 = \overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab}) \implies \overrightarrow{ab} = 0 \implies a = b$.

2. $\implies$. Sea $v \in A_2$. Fijado $a_2 \in A_2$, sea b_2 tal que $\overrightarrow{a_2b_2} = v$. Entonces $\exists a_1, b_1 \in A_1$ tales que $f(a_1) = a_2$ y $f(b_1) = b_2$. Por tanto, $v = \overrightarrow{f(a_1)f(b_1)} = \tilde{f}(\overrightarrow{a_1b_1})$, luego ese vector lleva a v. Para $\impliedby$, Sea $p_2 \in Im(f)$ con $p_2 = f(p_1)$, y sea $a_2 \in A_2$ arbitario. Entonces, $\exists u_1 \in E_1$ con $\tilde{f}(u_1) = \overrightarrow{p_2a_2}$. Sea a_1 el que garantiza $p_1a_1 = u_1$. Entonces $f(a_1) = f(p_1) + \tilde{f}(u_1) = p_2 + \overrightarrow{p_2a_2} = a_2$. $\square$

Definición 34. Una afinidad f es una **homotecia** de razón r, donde $r \neq 0, 1$, si $\tilde{f} = rId_E$, es decir, $f(a) = f(p) + r\overrightarrow{pa}$ fijado $p \in A_1$.

Es biyectiva al serlo rId_E.

Proposición 49. *Una homotecia tiene un único punto fijo, c, llamado* ***centro de la homotecia****. Se verifica que $c = p + \frac{1}{1-r}\overrightarrow{pf(p)}$ $\forall p \in A_1$.*

Demostración. Sus puntos fijos son los a tales que $a = f(a) = f(p) + r\overrightarrow{pa} \iff \overrightarrow{f(p)a} = r\overrightarrow{pa} \iff \overrightarrow{f(p)p} + \overrightarrow{pa} = r\overrightarrow{pa} \iff \overrightarrow{pa} = \frac{1}{1-r}\overrightarrow{pf(p)}$, como queríamos ver, sin importar p. Si hubiera otro punto fijo, b, entonces $\overrightarrow{ab} = \overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab}) = r\overrightarrow{ab}$, lo cual no puede ser dado que $r \neq 1$. $\square$

Definición 35. Dados r puntos $a_1, ..., a_r$, su **baricentro** es $b \in A$ tal que $b = \sum_1^r \frac{1}{r} a_i$.

Definición 36. Una afinidad f es una **simetría** si $f^2 = Id_A$.

Proposición 50. *Si f es simetría, entonces $\tilde{f}$ también lo es.*

Demostración. $\forall a, b \in A$, se tiene $\tilde{f}^2(\overrightarrow{ab}) = \tilde{f}(\overrightarrow{f(a)f(b)}) = \overrightarrow{f^2(a)f^2(b)} = \overrightarrow{ab}$. $\square$

Proposición 51. *$\forall a \in A$, el punto $m = \frac{1}{2}a + \frac{1}{2}f(a)$ es fijo bajo la simetría f.*

Demostración. Dado $p \in A$, sabemos que $\overrightarrow{pm} = \frac{1}{2}\overrightarrow{pa} + \frac{1}{2}\overrightarrow{pf(a)}$. En particular, si $p = a$, $\overrightarrow{am} = \frac{1}{2}\overrightarrow{af(a)}$. Aplicando $\tilde{f}$ a ambos lados, queda $\overrightarrow{f(a)f(m)} = \frac{1}{2}\overrightarrow{f(a)f^2(a)} = \frac{1}{2}\overrightarrow{f(a)a} + \frac{1}{2}\overrightarrow{f(a)f(a)}$, luego $f(m) = \frac{1}{2}a + \frac{1}{2}f(a) = m$. $\square$

Observación 18 (Clasificación de simetrías). Si $m(X)$ es el polinomio mínimo de $\tilde{f}$ con f una simetría, entonces sabemos que $m(X)|X^2 - 1$. Distinguimos tres casos:

1. Si $m(X) = X + 1$, entonces, f es una homotecia de razón -1, y con lo que ya hemos visto, su centro (único punto fijo) es $c = p + \frac{1}{2}\overrightarrow{pf(p)}$ $\forall p \in A$.
2. Si $m(X) = X - 1$, entonces $\tilde{f} = Id$, con lo que se trata de una traslación, pero como ha de fijar puntos, como hemos visto antes, no queda otra que sea la identidad. $f = Id$.

3. Si $m(X) = X(X-1)$ es una simetría propiamente dicha, con $E = S_{-1} \oplus S_1$ donde $S_{-1} = Nuc(\tilde{f}+Id)$ y $S_1 = Nuc(\tilde{f} - Id)$.

   Observemos que si $a \in A$ es un punto cualquiera, y $p \in Im(f)$ es punto fijo, y escribimos $\overrightarrow{pa} = u_1 + u_{-1}$ cada uno en un autoespacio, entonces $f(a) = f(p) + \tilde{f}(u_1 + u_{-1}) = p + u_1 - u_{-1}$. En particular, si a es otro punto fijo, entonces $f(a) = p + u_1 - u_{-1} = p + u_1 + u_{-1} = a$, luego $u_{-1} = 0$ y $a \in p + S_1$. Es decir, **los puntos fijos son la variedad** $p + S_1$.

   Por otro lado, sin importar a, se tiene $\overrightarrow{af(a)} = -2u_{-1}$ al ser $a = p + u_1 + u_{-1}$ y $f(a) = p + u_1 - u_{-1}$. Es decir, $f(a) \in a + S_{-1}$. $\forall a \in A$.

   En resumen:

   $$\begin{cases} \frac{1}{2}a + \frac{1}{2}f(a) \in p + S_1 \\ f(a) \in a + S_{-1} \end{cases}$$

   Lo que debería permitirnos hallar $f(a)$ en cualquier caso.

Proposición 52. *Sean $f : A_1 \to A_2$ y $g : A_2 \to A_3$ afines. Entonces también es afín $g \circ f$ y $\widetilde{g \circ f} = \tilde{g} \circ \tilde{f}$.*

Demostración. $\overrightarrow{g \circ f(a) g \circ f(b)} = \tilde{g}(\overrightarrow{f(a)f(b)}) = \tilde{g}\tilde{f}(\overrightarrow{ab})$. □

Proposición 53. *Sea $f : A_1 \to A_2$ biyectiva y afín. Entonces también lo es f^{-1} y $\widetilde{f^{-1}} = (\tilde{f})^{-1}$.*

Demostración. Sean $a_2, b_2 \in A_2$ y a_1, b_1 sus preimágenes. Entonces $\tilde{f}(\overrightarrow{f^{-1}(a_2)f^{-1}(b_2)}) = \overrightarrow{f(a_1)f(b_1)} = \overrightarrow{a_2 b_2}$. Aplicando la inversa de $\tilde{f}$ a ambos lados, queda que $\overrightarrow{f^{-1}(a_2)f^{-1}(b_2)} = \tilde{f}^{-1}(a_2 b_2)$. □

Proposición 54. *Sea $f : A_1 \to A_2$ afín y $a + F \subset A_1$ variedad lineal. Entonces $f(a + F)$ es variedad lineal y $f(a + F) = f(a) + \tilde{f}(F)$.*

Demostración. Está claro que como $a \in a + F$, $f(a) \in f(a + F)$. Ahora, sea $b \in f(a + F)$, con $f(c) = b$. $\overrightarrow{f(a)b} = \tilde{f}(\overrightarrow{ac}) \in \tilde{f}(F)$ dado que $\overrightarrow{ac} \in F$ al ser $a \in a + F$ y $c \in a + F$. □

Proposición 55. *Si $b + F \subset A_2$ es variedad lineal, con $a \in f^{-1}(b)$, entonces se tiene la variedad $f^{-1}(b + F) = a + \tilde{f}^{-1}(F)$.*

Demostración. $x \in f^{-1}(b + F) \iff f(x) \in b + F \iff \overrightarrow{bf(x)} \in F \iff \overrightarrow{f(a)f(x)} = \tilde{f}(\overrightarrow{ax}) \in F \iff \overrightarrow{ax} \in \tilde{f}^{-1}(F) \iff x \in a + \tilde{f}^{-1}(F)$. □

Proposición 56 (Corolario de lo anterior)**.** *Dada $f : A_1 \to A_2$ afín, se tiene que* $\dim Im(f) = rg(\tilde{f}) = \dim \Im(\tilde{f})$, *y que lleva variedades lineales paralelas en otras paralelas.*

Demostración. En primer lugar, está claro que $Im(f) = f(A_1) = f(a) + \tilde{f}(E_1) = f(a) + Im(\tilde{f})$. Ahora, dadas dos variedades paralelas $a + F$ y $b + G$ con $F \subset G$ sin perder en generalidad, tenemos que $f(a + F) = f(a) + \tilde{f}(F)$ y $f(b + G) = f(b) + \tilde{f}(G)$, y como $F \subset G \implies \tilde{f}(F) \subset \tilde{f}(G)$, son paralelas también. □

Proposición 57. *Si el conjunto de puntos fijos de $f : A \to A$ es no vacío, con $f(a) = a$, entonces ese conjunto es $a + Nuc(\tilde{f} - Id)$.*

Demostración. b es fijo $\iff f(b) = b \overset{*}{\iff} \tilde{f}(ab) = \overrightarrow{f(a)f(b)} = \overrightarrow{ab} \iff \overrightarrow{ab} \in Nuc(\tilde{f} - Id) \iff b \in a + Nuc(\tilde{f} - Id)$. En el paso marcado con $*$, la implicación a la izquierda viene dada porque $\overrightarrow{ab} = \tilde{f}(\overrightarrow{ab}) = \overrightarrow{f(a)f(b)} = \overrightarrow{af(b)}$, luego ha de ser $b = f(b)$. □

Proposición 58. *Si $f : A_1 \to A_2$ es afinidad y $x = \sum x_i a_i$ para el sistema afín de puntos $\{a_i\}_1^r$, entonces $f(x) = \sum x_i f(a_i)$. Es decir, los puntos imagen 'generan' la imagen de x con las mismas coordenadas.*

Demostración. $\vec{px} = \sum x_i \vec{pa_i} \implies \tilde{f}(\vec{px}) = \sum x_i \tilde{f}(\vec{pa_i}) \implies \overrightarrow{f(p)f(x)} = \sum x_i \overrightarrow{f(p)f(a_i)} \implies f(x) = \sum x_i f(a_i)$. □

Proposición 59 (Expresión de una afinidad en coordenadas cartesianas.)**.** *Sea $f : A_1 \to A_2$ afinidad, con $\dim A_1 = n$, $\dim A_2 = n$, $R_1 = O_1; B_1$ y $R_2 = \{O_2; B_2\}$ sistemas cartesianos de cada espacio. Se pueden hallar las coordenadas de la salida en función de las de la entrada como:*

$$[f(x)]_{R_2} = [\overrightarrow{O_2 f(O_1)}]_{B_2} + \mathcal{M}_{B_1}^{B_2}(\tilde{f}) \cdot [x]_{R_1}$$

es decir:

$$\begin{pmatrix} f(x)_1 \\ f(x)_2 \\ \vdots \\ f(x)_m \\ 1 \end{pmatrix} = \begin{pmatrix} \mathcal{M}_{B_1}^{B_2}(\tilde{f}) & [\overrightarrow{O_2 f(O_1)}]_{B_2} \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ 1 \end{pmatrix}.$$

A la matriz por la que se multiplica en la expresión de arriba se le denomina ***matriz de la afinidad f en los sistemas R_1 y R_2****, es decir, $\mathcal{M}_{R_1}^{R_2}(f)$.*

Demostración. Si trabajamos exclusivamente con vectores como sabemos, entonces tenemos que $\overrightarrow{O_2 f(x)} = \overrightarrow{O_2 f(O_1)} + \overrightarrow{f(O_1) f(x)} = \overrightarrow{O_2 f(O_1)} + \tilde{f}(\overrightarrow{O_1 x})$. Por el isomorfismo que conecta un espacio vectorial con su base, tenemos, entonces, que: $(\overrightarrow{O_2 f(x)})_{B_2} = \overrightarrow{O_2 f(O_1)}_{B_2} + (\tilde{f}\overrightarrow{O_1 x})_{B_2}$, o, lo que es exactamente lo mismo, $(f(x))_{R_2} = \overrightarrow{O_2 f(O_1)}_{B_2} + \mathcal{M}_{B_1}^{B_2} \cdot (\overrightarrow{O_1 x})_{B_1}$, es decir:

$$(f(x))_{R_2} = \overrightarrow{O_2 f(O_1)}_{B_2} + \mathcal{M}_{B_1}^{B_2} \cdot (x)_{R_1}$$

□

Proposición 60. *Dadas $f_1 : A_1 \to A_2$ y $f_2 : A_2 \to A_3$ afinidades con R_1, R_2 y R_3 sistemas de referencia cartesianos de cada espacio, entonces*

$$\mathcal{M}_{R_1}^{R_3}(f_2 \circ f_1) = \mathcal{M}_{R_2}^{R_3}(f_2) \cdot \mathcal{M}_{R_1}^{R_2}(f_1) = \begin{pmatrix} \mathcal{M}_{B_2}^{B_3}(\tilde{f}_2)\mathcal{M}_{B_1}^{B_2}(\tilde{f}_1) & \mathcal{M}_{B_2}^{B_3} \cdot [\overrightarrow{O_2 f(O_1)}]_{B_2} + [\overrightarrow{O_3 f(O_2)}]_{B_3} \\ 0 & 1 \end{pmatrix}$$

Además, dada $f : A \to A$ biyectiva, entonces $\mathcal{M}_B(f^{-1}) = (\mathcal{M}_B(f))^{-1}$

Demostración. Como $f_1(x)_{R_2} = \mathcal{M}_{B_1}^{B_2}(\tilde{f}_1)(x)_{R_1} + [\overrightarrow{O_2 f(O_1)}]_{B_2}$, entonces $f_2(f_1(x))_{R_3} = \mathcal{M}_{B_2}^{B_3}(\tilde{f}_2)(\mathcal{M}_{B_1}^{B_2}(\tilde{f}_1)(x)_{R_1} + [\overrightarrow{O_2 f(O_1)}]_{B_2}) + [\overrightarrow{O_3 f(O_2)}]_{B_3} = \mathcal{M}_{B_2}^{B_3}(\tilde{f}_2)\mathcal{M}_{B_1}^{B_2}(\tilde{f}_1)(x)_{R_1} + \mathcal{M}_{B_2}^{B_3}(\tilde{f}_2)[\overrightarrow{O_2 f(O_1)}]_{B_2}) + [\overrightarrow{O_3 f(O_2)}]_{B_3}$, como queríamos ver.

Además, como $f \circ f^{-1} = Id$, ha de ser $\mathcal{M}_B(f) \cdot (\mathcal{M}_B(f^{-1})) = Id$ y al ser cuadradas se tiene lo que se quería. □

3. Espacio afín euclídeo

Definición 37. Sea (A, E, φ) espacio afín real con (E, ϕ) espacio euclídeo. Denotamos por (A, E, ϕ) a lo que se denomina **espacio afín eucídeo** (espacio afín donde la dirección es un espacio eucídeo).

Definición 38. Sea X un conjunto. Una **distancia en X** $d : X \times X \to \mathbb{R}$ es una aplicación tal que $\forall p, q, r \in X$:

1. $d(p,q) \geq 0$, $d(p,q) = 0 \iff p = q$.
2. $d(p,q) = d(q,p)$
3. $d(p,r) \leq d(p,q) + d(q,r)$.

Proposición 61. *En el espacio afín eucídeo (A, E, ϕ), se tiene una distancia dada por $d(p,q) = \|\overrightarrow{pq}\|_\phi$.*

Demostración. Vemos que tiene las propiedades de norma:

1. Sale de la definición positiva de la norma. Valdrá cero solo si $\overrightarrow{pq} = 0 \iff p = q$.
2. $d(p,q) = \|\overrightarrow{pq}\| = |-1| \, \|\overrightarrow{qp}\| = \|\overrightarrow{qp}\| = d(q,p)$.
3. $d(p,r) = \|\overrightarrow{pq} + \overrightarrow{qr}\| \leq \|\overrightarrow{pq}\| + \|\overrightarrow{qr}\| = d(p,q) + d(q,r)$.

□

Proposición 62 (Teorema de pitágoras)**.** *Dados puntos $p, q, r \in A$ afín euclídeo con $\phi(\overrightarrow{pq}, \overrightarrow{pr}) = 0$, entonces $d(p,q)^2 + d(p,r)^2 = d(q,r)^2$.*

Demostración. $d(q,r)^2 = \phi(\overrightarrow{qr}, \overrightarrow{qr}) = \phi(\overrightarrow{pr} - \overrightarrow{pq}, \overrightarrow{pr} - \overrightarrow{pq}) = \phi(\overrightarrow{pr}, \overrightarrow{pr}) + \phi(\overrightarrow{pq}, \overrightarrow{pq}) - 2\phi(\overrightarrow{pq}, \overrightarrow{pr}) = d(p,q)^2 + d(p,r)^2 + 0$. □

Definición 39. Las variedades $a + F$ y $b + G$ en un espacio afín euclídeo son **ortogonales** si se tiene alguna de las siguientes:

1. $\phi(u,v) = 0 \ \forall u \in F, v \in G$, en cuyo caso $dimF + dimG \leq n$, al ser $F \subset G^\perp$.
2. $dimF + dimG \geq n$ y además son ortogonales $a + F^\perp$ y $b + G^\perp$.

Definición 40 (Distancia entre dos variedades)**.** Dadas $L, L' \subset A$ dos variedades lineales en un espacio afín euclídeo, se define $d(L, L') = \text{mín}\{d(p,q) : p \in L, q \in L'\}$.

A continuación veremos que tal mínimo se alcanza y cómo calcularlo:

Proposición 63. *Si $L = a + W$ y $L' = b + W'$, se verifica:*

$$d(L, L') = \left\| P^\perp_{(W+W')^\perp}(\overrightarrow{ab}) \right\|$$

Demostración. Lo primero que vamos a ver es que ese valor está bien definido, es decir, $P^\perp_{(W+W')^\perp}(\overrightarrow{ab})$ es constante sin importar la elección de a y b. Sea $U = W + W'$ para abreviar. Si ponemos $a_1, a_2 \in L$ y $b_1, b_2 \in L'$, entonces $a_2 = a_1 + w$ y $b_2 = b_1 + w'$ con $w \in W$ y $w' \in W'$, de modo que $\overrightarrow{a_2b_2} = \overrightarrow{a_1b_1} + (w' - w)$, y como $P^\perp_U(w' - w) = 0$ dado que $w' - w \in W + W'$, y la proyección es lineal, entonces $P^\perp_U(\overrightarrow{a_2b_2}) = P^\perp_U(\overrightarrow{a_1b_1})$.

Ahora, como $\left\| P^\perp_U(\overrightarrow{ab}) \right\| \leq \left\| \overrightarrow{ab} \right\| = d(a,b)$, con igualdad si $\overrightarrow{ab} \in U$, lo único que basta con ver es que hay $a_0 \in L$, $b_0 \in L$ tales que $d(a_0, b_0) = \left\| \overrightarrow{a_0b_0} \right\| = \left\| P^\perp_U(\overrightarrow{a_0b_0}) \right\|$ o, lo que es lo mismo, tales que $\overrightarrow{a_0b_0} \in U$. Definimos la variedad $M = a + (W + U)$, y está claro que $L \subset M$. Como $M + L' = a + (W + (W + W')^\perp + W' + < \overrightarrow{ab} >) = a + E = A$, entonces $\overrightarrow{ab} \in \overrightarrow{M + L'}$ al ser el total, luego $M \cap L' \neq \emptyset$. Sea $b_0 \in M \cap L'$. Como está en M, existen $w \in W$ y $u \in U$ tales que $b_0 = a + w + u$. Sea $a + w = a_0 \in L$. Entonces $\overrightarrow{a_0b_0} = u \in U$ y hemos acabado. □

Definición 41. Una aplicación $f : A_1 \to A_2$ es una **isometría** si $d(f(a), f(b)) = d(a,b) \ \forall a, b \in A_1$. Es decir, preserva las distancias.

Lema 6. *Dada una aplicación ortogonal f entre dos espacios vectoriales euclídeos, f siempre será lineal.*

Razón. Desarrollando con las propiedades de espacio euclídeo y ortogonalidad, es fácil probar que $\phi(f(u+v)-f(u)-f(v), f(u+v)-f(u)-f(v))=0$ y que $\phi(f(\lambda u)-\lambda f(u), f(\lambda u)-\lambda f(u))=0$, lo que prueba que es lineal por la definición positiva de ϕ.

Proposición 64. *La aplicación $f : A_1 \to A_2$ es isometría $\iff$ f es afinidad y $\tilde{f}$ es ortogonal.*

Demostración. $\impliedby$. Sean $a, b \in A_1$. Entonces, $d(f(a), f(b))^2 = \left\|\overrightarrow{f(a)f(b)}\right\|^2 = \left\|\tilde{f}(\overrightarrow{ab})\right\|^2 = (\tilde{f}(\overrightarrow{ab}), \tilde{f}(\overrightarrow{ab})) = (\overrightarrow{ab}, \overrightarrow{ab}) = d(a,b)^2$, como se quería.

Para $\implies$, vamos a fijar $p \in A_1$ y definimos $\psi E_1 \to E_2$, con $\psi(\overrightarrow{pa}) = \overrightarrow{f(p)f(a)}$. Está bien definida por la primera propiedad de espacio afín. Ahora veremos que además es ortogonal. Dados $a, b \in A_1$, tenemos que $d(a,b)^2 = (\overrightarrow{ap} + \overrightarrow{pb}, \overrightarrow{ap} + \overrightarrow{pb}) = \|ap\|^2 + 2(ap, pb) + \|pb\|^2$, es decir, $(ap, pb) = \frac{1}{2}(d(a,b)^2 - d(a,p)^2 - d(p,b)^2)$. Por el mismo motivo, $(f(a)f(p), f(p)(b)) = \frac{1}{2}(d(f(a), f(b))^2 - d(f(a), f(p))^2 - d(f(p), f(b))^2)$, pero como es isometría, las distancias del lado derecho de ambas expresiones son las mismas, luego: $(\overrightarrow{ap}, \overrightarrow{pb}) = (\overrightarrow{f(a)f(p)}, \overrightarrow{f(p)f(b)}) = (\psi(\overrightarrow{ap}, \psi(\overrightarrow{pb}))$, y como a, b son arbitrarios cubren todos los vectores y ψ es ortogonal. Por el lema previo, ψ ha de ser lineal. Así, $\psi(\overrightarrow{ab}) = \psi(\overrightarrow{pb} - \overrightarrow{pa}) = \psi(\overrightarrow{pb}) - \psi(\overrightarrow{pa}) = \overrightarrow{f(p)f(b)} - \overrightarrow{f(p)f(a)} = \overrightarrow{f(a)f(b)}$, luego ψ es la aplicación lineal de f, que es afinidad, y es ortogonal. □

Definición 42. Dos espacios afines son **isomorfos** si existe f afinidad biyectiva entre ambos.

Observación 19. $A_1 \simeq A_2 \iff \dim A_1 = \dim A_2$

Razón. Si sus dimensiones son iguales, sabemos que existe isomorfismo $\tilde{f}$ entre sus direcciones. Toda afinidad de parte lineal $\tilde{f}$ será entonces biyectiva entre ambos, como ya vimos.

Se tiene una definición más fuerte en espacio afín euclídeo:

Definición 43. Dos espacios afines euclídeos son **isomorfos** si existe isometría biyectiva f entre ambos.

Proposición 65. *Los espacios afines euclídeos A_1 y A_2 son isomorfos* $\iff$ $\dim A_1 = \dim A_2$.

Demostración. $\implies$ es evidente, dado que si lo son existe f isometría biyectiva, y por tanto afinidad biyectiva, entre ambos, y $\tilde{f}$ es biyección lineal entre sus espacios. Para $\impliedby$, debemos ver que si $\psi : E_1 \to E_2$ la definimos lineal y llevando una base ortonormal en otra base ortonormal (podemos dado que tienen la misma dirección), sabemos que será ortogonal y biyectiva al llevar base en base. Si ahora ponemos una f afín cualquiera entre A_1 y A_2 tal que $\tilde{f} = \psi$, será isometría biyectiva. □

3.1. Movimientos

Definición 44. Un **movimiento** es una isometría de $\mathbb{A}^2_{\mathbb{R}}$ o $\mathbb{A}^3_{\mathbb{R}}$ en sí mismo, con la distancia usual.

Proposición 66 (Clasificación de movimientos en $\mathbb{A}^2_{\mathbb{R}}$). *Dado un movimiento f de $\mathbb{A}^2_{\mathbb{R}}$, sabemos que su aplicación asociada $\tilde{f}$ es ortogonal en $\mathbb{R}^2$, de modo que, elegida base $B = \{u_1, u_2\}$ ortonormal adecuada, tenemos las siguientes opciones:*

1. $\mathcal{M}_B(\tilde{f}) = Id_2$. *En este caso, $\tilde{f} = Id_{\mathbb{R}^2}$, con lo que se trata de una traslación, como ya sabemos.*

2. $\mathcal{M}_B(\tilde{f}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. *Este caso se caracteriza por* $\det f = -1$. *Podemos escribir:*

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

en el sistema $\mathcal{R} = \{P, u_1, u_2\}$, para algún P. Sabemos, además, que (a, b) es el vector $\overrightarrow{Pf(P)}$ en la base B. Este vector lo podemos separar en sus proyecciones sobre $< u_1 >$ y $< u_2 >$, obteniendo así:

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ b \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

Nos queda una traslación de vector $(a, 0)_B$ sumada con cierto movimiento, que resulta fácil de analizar, dado que tiene puntos fijos:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

luego:

$$\begin{pmatrix} 0 \\ -b \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

Es decir, los puntos fijos son la recta $y = \frac{b}{2}$ en este sistema de referencia. Tiene la dirección de $< u_1 >$.

En cuanto a la traslación de vector $(a, 0)_B = au_1$, es paralela al eje de simetría. Por ello, este movimiento es una ***simetría deslizante****. Nótese que $au_1 = P^{\perp}_{<u_1>}(\overrightarrow{Pf(P)}) = (u_1, \overrightarrow{Pf(P)})u_1$, según hemos visto. Esto servirá hallar los elementos geométricos cuando no tengamos el sistema de referencia R determinado.*

3. $\mathcal{M}_B(\tilde{f}) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ *Este caso queda caracterizado por* $\det f = 1$. *En el sistema $\mathcal{R} = \{P, u_1, u_2\}$ para algún P, tenemos que*

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

luego los puntos fijos son los que verifican:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

es decir:

$$\begin{pmatrix} -a \\ -b \end{pmatrix} = \begin{pmatrix} \cos\alpha - 1 & -\sin\alpha \\ \sin\alpha & \cos\alpha - 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

*Por tanto, si $\alpha \neq 0$ (caso degenerado de la identidad), la matriz anterior corresponde con una biyección, y solo hay un punto fijo $Q = (x, y)$, el **centro de giro**. Así, se trata de un **giro de ángulo** α respecto de Q. De hecho, en el sistema $\mathcal{R}' = \{Q, u_1, u_2\}$, se tiene:*

$$f\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

donde se aprecia este comportamiento. Al igual que en $\mathbb{R}^2$, se tiene esa matriz en cualquier base ortonormal.

Proposición 67 (Clasificación de movimientos en $\mathbb{A}^3_{\mathbb{R}}$)**.** *Dado un movimiento f de $\mathbb{A}^3_{\mathbb{R}}$, sabemos que su aplicación asociada $\tilde{f}$ es ortogonal en $\mathbb{R}^3$, de modo que, elegida base $B = \{u_1, u_2, u_3\}$ ortonormal adecuada, y en un sistema $R = \{P, B\}$ tenemos las siguientes opciones:*

1. *Si $M_B(\tilde{f}) = Id_3$, es una traslación.*

2. *Si $M_B(\tilde{f}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$, tenemos un movimiento con $\det\tilde{f} = 1$. Si separamos $\overrightarrow{Pf(P)}$ en su proyección sobre $< u_1 >$ (eje de giro) y su proyección sobre $< u_2, u_3 >$, vemos que, en R:*

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ b \\ c \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Es decir, se tiene un desplazamiento de vector au_1, y un giro cuyo eje tiene dirección $< u_1 >$ y es la recta de puntos fijos:

$$\begin{pmatrix} 0 \\ -b \\ -c \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \cos\alpha - 1 & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha - 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

*Se trata, pues, de un **movimiento helicoidal** cuyo ángulo se puede hallar como de costumbre ($tr(\tilde{f}) = 1 + 2\cos\alpha$ y $(f(u_2), u_3) = \sin\alpha$ si u_2, u_3 conforman junto al eje u_1 una base ortonormal orientada positivamente), su vector deslizante se calcula como $P^{\perp}_{<u_1>}(\overrightarrow{Pf(P)})$ y su eje de giro es la recta que se obtiene de resolver los puntos fijos de la función menos el vector deslizante (el giro propiamente dicho).*

3. *Si $M_B(\tilde{f}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, tenemos un movimiento con $\det\tilde{f} = -1$. Si separamos $\overrightarrow{Pf(P)}$ en su proyección sobre $< u_1, u_2 >$ (plano de simetría) y su proyección sobre $< u_3 >$, vemos que, en R:*

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Que se trata de una traslación de vector $au_1 + bu_2$, paralelo al plano de simetría, más una simetría respecto al plano de puntos fijos dado por:

$$\begin{pmatrix} 0 \\ 0 \\ -c \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

*Es decir, el $z = \frac{c}{2}$ en estas coordenadas. Por tanto, es una **simetría deslizante** de vector $P^{\perp}_{<u_1>,<u_2>}(\overrightarrow{Pf(P)})$, con respecto al plano de puntos fijos de la simetría propiamente dicha, que se obtiene restando el vector deslizante a la aplicación.*

4. *Si $M_B(\tilde{f}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$, tenemos un movimiento con $\det \tilde{f} = -1$, que correspondería a una antirrotación en cuanto a aplicación lineal. Esta aplicación sabemos que no tiene vectores fijos salvo el 0, luego su espacio de puntos fijos es a lo sumo un solo punto. Si se tiene:*

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Es decir, la expresión en R como ya sabemos $\overrightarrow{Pf(P)} + \overrightarrow{f(P)f(X)}$, entonces el punto fijo vendría dado por:

$$\begin{pmatrix} -a \\ -b \\ -c \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & \cos\alpha - 1 & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha - 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

*Y como la matriz superior es biyección, (salvo el caso degenerado $\alpha = 0$, que es simetría deslizante) se obtiene la solución única $Q = (x, y, z)$, que es el **centro de la antirrotación**. El ángulo de giro se halla como de costumbre, y el plano de giro es el que tiene dirección del plano de giro de la aplicación $\tilde{f}$, pero pasando por Q.*

Cabe observar que el análisis se realiza en ese sistema de referencia para observar perfectamente las características de cada caso, pero si se mantienen los vectores y puntos de cada expresión, es válido en cualquier base (las operaciones no dependen de la base) y para cualquier centro del sistema de referencia (la elección de P ha sido arbitraria). La única diferencia será que la expresión de coordenadas y matrices no será tan clara.

La manera de identificar un movimiento será siempre estudiar su aplicación lineal asociada, y una vez identificada, clasificarlo según las pautas anteriores y obtener sus elementos geométricos en el espacio afín.

4. Formas cuadráticas y cónicas

Definición 45. Sea V un $\mathbb{R}$-espacio vectorial. Una **forma cuadrática** es una función $\phi : V \to \mathbb{R}$ que verifica:

1. $\forall \lambda \in \mathbb{R}$, $\phi(\lambda v) = \lambda^2 \phi(v)$.
2. La forma $\varphi_\phi : V \times V \to \mathbb{R}$ dada por $\varphi_\phi(u,v) = \frac{1}{2}(\varphi_\phi(u+v) - \varphi_\phi(u) - \varphi_\phi(v))$ es bilineal simétrica.

Observemos como se relacionan las cuadráticas y las bilineales:

Observación 20. En general, dada una forma bilineal $\varphi : V^2 \to \mathbb{R}$, la forma $\phi_\varphi(v) = \varphi(v,v)$ es cuadrática. Si φ ya es simétrica, coincide con la de la propiedad 2, y, por lo tanto, toda forma cuadrática se puede expresar como $\phi(u) = \varphi_\phi(u,u)$, es decir, toda forma cuadrática viene de una bilineal.

Razón. La propiedad 1 sigue de la bilinealidad, y la propiedad 2 se puede ver dado que si ψ es la función que se define en esa propiedad, entonces $\psi(u,v) = \frac{1}{2}(\varphi(u+v,u+v) - \varphi(u,u) - \varphi(v,v))$ que claramente es simétrica, y será bilineal por la bilinealidad de φ. Además sabemos que, en formas bilineales simétricas, se verifica $\varphi(u,v) = \frac{1}{2}(\varphi(u+v,u+v) - \varphi(u,u) - \varphi(v,v))$, de manera que si φ es simétrica, coincide con ψ.

No obstante, se pueden obtener desde dos formas bilineales distintas la misma cuadrática, sin embargo, cada cuadrática está caracterizada por una y solo una forma bilineal simétrica:

Proposición 68. *Fijada una base, la expresión de una forma cuadrática es* $\phi(x_1,...,x_n) = \sum_1^n \alpha_i x_i^2 + \sum\limits_{1 \le i < j \le n} \beta_{ij} x_i x_j$*. Solo hay una bilineal simétrica que genere esta expresión, y coincide con* φ_ϕ *según la observación anterior.*

Demostración. Se tiene, por la observación anterior, que:

$$\phi(x_1,...,x_n) = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} \alpha_1 & \gamma_{21} & \dots & \gamma_{n1} \\ \delta_{21} & \alpha_2 & \dots & \gamma_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{n1} & \delta_{n2} & \dots & \alpha_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_1^n \alpha_i x_i^2 + \sum_{1 \le i < j \le n} (\delta_{ji} + \gamma_{ji}) x_i x_j$$

Como queríamos ver. Así mismo, la única variante simétrica es aquella tal que $\delta_{ji} + \gamma_{ji} = \frac{\beta_{ij}}{2}$, y por tanto es única. $\square$

Definición 46. Se define la **matriz de la forma cuadrática ϕ en base** B, $\mathcal{M}_B(\phi)$, como la matriz de la forma bilineal simétrica φ_ϕ que la induce. (Sabemos que esta coincide con la que se obtiene de la propiedad 2).

Proposición 69. *Sea V espacio euclídeo y $Q : V \to \mathbb{R}$ cuadrática. Entonces $\exists B' \subset V$ ortonormal tal que $\mathcal{M}_{B'}(Q)$ es diagonal, y por tanto la expresión en esa base se reduce a una suma de cuadrados con coeficientes.*

Demostración. Dada $B \subset V$ ortonormal, se tiene que $\mathcal{M}_B(Q)$ ha de ser simétrica como sabemos. Si la interpretamos ahora como una matriz de una aplicación de V en V, sabemos que diagonaliza (al ser autoadjunta en ese producto) en $\mathbb{R}$, en base ortonormal B'. Si hacemos el cambio de base P de B' a B, entendiendo la matriz como la de una aplicación, se tendría $\mathcal{M}_{B'}(Q) = P^{-1}\mathcal{M}_B(Q)P$, pero si entendemos P como aplicación que lleva base ortonormal en base ortonormal, entonces es ortogonal y $\mathcal{M}_{B'}(Q) = P^t \mathcal{M}_B(Q) P$, que es justo el cambio de base de forma bilineal, obteniendo así una matriz de la forma en base B' que es diagonal. $\square$

Como obtener los autovalores y autovectores no siempre es fácil, se tiene el siguiente método para diagonalizar la cuadrática (aunque este no garantiza la base ortonormal, ya que no se corresponde con los autovectores de la aplicación asociada como en el método anterior).

Observación 21 (Método de Gauss para la obtención de una forma canónica).

1. Completar cuadrados si es posible. Para ello se toma una de las variables elevada al cuadrado, y se completa correspondientemente. Por ejemplo: en $Q = x^2 + 3z^2 - 4xy - 4xz + 4yz$, podemos tomar la x, y agregar en un cuadrado todas la variables involucradas con la x: $Q = (x - 2y - 2z)^2 - z^2 - 4y^2 - 4yz$. Ahora se repite el método en la parte restante.

2. Si no hay cuadrados, se cambia de variable. Se toma un producto de 2 variables (por ejemplo, $4xy$) y se hace el cambio $x' = x + y$ y $y' = x - y$. Las demás variables involucradas se mantienen en este cambio. En el ejemplo anterior quedaría $4x^2 - 4y^2$.

3. Finalmente, se determina cuál es el cambio de base para que quede la forma canónica. Se obtendrá de combinar los cambios hechos en 2 con un último cambio haciendo que cada expresión al cuadrado en 1 sea una variable nueva.

Esto puede dar lugar a distintas formas canónicas, pero hay invariantes:

Proposición 70 (Ley de inercia)**.** *Sea Q una cuadrática en el $\mathbb{R}$-espacio V y sea $B = \{w_1, ..., w_n\}$ una base respecto de la que $Q(u) = \sum_1^p \mu_i y_i^2$ con y_i las coordenadas en B y $p \leq n$ para que $\mu_i \neq 0$. Si se tiene otra base $B' = \{v_1, ..., v_n\}$ en la que $Q(u) = \sum_1^k \lambda_i x_i^2$, con x_i las coordenadas y $\lambda_i \neq 0$ $(k \leq n)$, entonces:*

- *$p = k$, es decir, hay los mismos términos no nulos.*
- $|\{\lambda_i > 0 : 1 \leq i \leq k\}| = |\{\mu_i > 0 : 1 \leq i \leq p\}|$
- $|\{\lambda_i < 0 : 1 \leq i \leq k\}| = |\{\mu_i < 0 : 1 \leq i \leq p\}|$

*A estos dos últimos valores se les llama **índices de inercia positivo y negativo** y son invariantes en toda forma canónica de la misma cuadrática. Así mismo, la **signatura** de la cuadrática es el par $Sig(Q) = (I_p, I_n)$ de ambos índices.*

Demostración. Vamos a suponer sin perder en generalidad (si no, reorganizar la base) que los primeros q términos son positivos, es decir, $Q(u) = \sum_1^q \alpha_i y_i^2 + \sum_{q+1}^p (-\alpha_i) y_i^2$, con $|\mu_i| = \alpha_i > 0$, e igualmente, que Q(u) $= \sum_1^r \beta_i x_i^2 + \sum_{r+1}^k (-\beta_i) x_i^2$, con $|\lambda_i| = \beta_i > 0$.

Ahora definimos $V_1 =< w_1, ..., w_q >$ y $V_2 =< v_{r+1}, ..., v_n >$. Supongamos que $q > r$, entonces como $\dim V_1 + \dim V_2 = q + n - r > n$, luego intersecan, así que existe un $u = \sum_1^q a_i w_i = \sum_{r+1}^n b_i v_i$, luego $Q(u) = \sum_1^q \alpha_i a_i^2 = -\sum_{r+1}^n (\beta_i) b_i$, contradicción dado que uno es positivo y el otro negativo, de modo que $q \leq r$. Intercambiando V_1 y V_2 correspondientemente, se puede mostrar que $r \leq q$ del mismo modo, luego $r = q$. Si en lugar de esto se toma V_1 con los vectores cuyas coordenadas corresponden a coeficientes negativos, se puede probar mediante dos similares pasos que $p - q = k - r$, luego $p = k$. $\square$

Definición 47. La forma **normal** de una cuadrática corresponde a la canónica con coeficientes de módulo 1 (o nulos). Es única por la proposición anterior y, dada $Q(u) = \sum_1^p \mu_i y_i^2$, se obtiene con un último cambio $z_i = \sqrt{|\mu_i|} y_i$.

Definición 48. La forma Q es **definida positiva** si $Q(u) \geq 0$ con igualdad si y solo si $u = 0$. Si no se da esto último, es **semidefinida positiva**. Es **definida negativa** si $Q(u) \leq 0$ con igualdad si y solo si $u = 0$. Si no se da esto último, es **semidefinida negativa**. Si no se da nada de lo anterior, es **indefinida**.

Dada una curva en forma canónica, es muy fácil determinar su definición. Si $\dim V = n$, entonces $Sig(Q) = (n, 0)$ es evidentemente definida positiva, $Sig(Q) = (0, n)$ es definida negativa, $Sig(Q) = (p, 0)$ con $p < n$ es semidefinida positiva y $Sig(Q) = (0, p)$ con $p < n$ es semidefinida negativa.

Observación 22 (Criterio de Sylvester para definidas negativas). Como Q es definida negativa $\iff$ $-Q$ es definida positiva, y los determinantes menores de $-Q$ verifican $\det_i(-Q) = (-1)^i \det(Q)$, entonces Q es definida negativa si y solo si $\det_i(Q) < 0$ en i impar y $\det_i(Q) > 0$ en i par, fijada una base.

4.1. Curvas cónicas

Las curvas cónicas con las que resultan de seccionar un cono (como el de ecuación $px^2+qy^2-rz^2=0$) con un plano. Dependiendo de la posición del plano respecto del cono, se obtendrá una curva u otra. En general, si el plano tiene ecuación $ax+by+cz+d=0$, entonces $z=\frac{-ax-by-d}{c}$ y por lo tanto la curva resultante tiene la forma $px^2+qy^2-r(\frac{-ax-by-d}{c})^2=0$, que, tras simplificaciones, acabará siendo de la forma:

$$\alpha x^2+\beta y^2+\gamma xy+\delta x+\epsilon y+\omega=0.$$

O lo que es lo mismo:

$$\begin{pmatrix} x & y \end{pmatrix}\cdot A\cdot\begin{pmatrix} x \\ y \end{pmatrix}+B\cdot\begin{pmatrix} x \\ y \end{pmatrix}+\gamma=0.$$

Con las A y B correspondientes. Es decir tiene una parte cuadrática y una parte lineal. Veremos posteriormente que si se toma la A de la cuadrática simétrica, se puede diagonalizar mediante un cambio ortogonal, simplificando su ecuación sin alterar sus características métricas ni dimensiones.

4.1.1. Elipse, hipérbola y parábola

Vamos a definir las tres curvas cónicas (no degeneradas) que se pueden obtener:

Definición 49. Una **elipse** en el plano afín $\mathbb{A}^2_{\mathbb{R}}$, fijados dos puntos F_1 y F_2 llamados **focos** y una constante positiva a (llamada **semieje mayor**) es el conjunto de puntos $X\in\mathbb{A}^2_{\mathbb{R}}$ tales que $d(A,F_1)+d(A,F_2)=2a$.

La distancia de cada foco al centro se denomina **distancia focal** y se denota por c. Es claro que debe darse $a\geq c$ para que exista la elipse, dado que como $d(A,F_1)+d(A,F_2)\geq 2c$ por la desigualdad triangular, si $c>a$ entonces $d(A,F_1)+d(A,F_2)>2a$ y no existiría.

Se conoce como **centro de la elipse** al punto medio de los focos $\frac{1}{2}(F_1+F_2)$.

El **eje mayor** de la elipse es la recta que une a los focos, $F_1+<\overrightarrow{F_1F_2}>$.

El **eje menor o secundario** de la elipse es la recta ortogonal al eje mayor que pasa por el centro.

Los **vértices** de la elipse A_1 y A_2 son los puntos del eje mayor que pertenecen a la elipse.

La **excentricidad de la elipse** es la cantidad $\epsilon=\frac{c}{a}$. En este caso $\epsilon<1$.

El **semieje menor** es la magnitud $b=\sqrt{a^2-c^2}=a\sqrt{1-\epsilon^2}$. En la siguiente proposición veremos que coincide con la magnitud del segmento entre el centro y la intersección de la elipse con el eje menor.

A continuación veremos cual es la ecuación de una elipse centrada en el sistema de referencia. Cuando no lo esté, emplearemos movimientos rígidos para alinearla con los ejes y obtener así esta expresión:

Proposición 71 (Ecuación reducida de la elipse)**.** *Dada una elipse en un sistema de referencia ortonormal, centrada en el* $(0,0)$*, cuyo eje mayor es el eje* X*, semieje mayor* a *y focos* $F_1=(-c,0)$ *y* $F_2=(c,0)$ $(c<a)$*, entonces su ecuación es:*

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

Si está alineada con el eje Y*, será, evidentemente (tras un cambio ortonormal que permute* x *e* y*, no alterando sus medidas),* $\frac{x^2}{b^2}+\frac{y^2}{a^2}=1$*.*

Con esto demostramos también que la distancia del centro a la intersección de la elipse con el semieje menor es b*, dado que esta intersección se da en los puntos donde* $x=0$ *(pasa por el centro y es ortogonal a* $y=0$*), luego se da donde* $y^2=b^2$*, o lo que es lo mismo, en* $(0,b)$ *y* $(0,-b)$*.*

Demostración. Los puntos de la elipse son los (x, y) tales que:

$$\sqrt{(x-c)^2+y^2}+\sqrt{(x+c)^2+y^2}=2a.$$

Desarrollando como es habitual para eliminar las raíces, queda:

$$a^4+c^2x^2+2a^2cx=a^2x^2+a^2c^2+2a^2cx+a^2y^2.$$

Es decir:

$$a^2(a^2-c^2)=x^2(a^2-c^2)+y^2a^2.$$

Con la definición de b:

$$a^2b^2=x^2b^2+y^2a^2.$$

Y dividiendo se llega a lo que se quería. □

Definición 50. Una **hipérbola** en el plano afín $\mathbb{A}^2_{\mathbb{R}}$, fijados dos puntos F_1 y F_2 llamados **focos** y una constante positiva a (llamada **semieje mayor**) es el conjunto de puntos $X \in \mathbb{A}^2_{\mathbb{R}}$ tales que $|d(A, F_1) - d(A, F_2)| = 2a$.

La distancia de cada foco al centro se denomina **distancia focal** y se denota por c. Es claro que debe darse $a \leq c$ para que exista la elipse, dado que como $|d(A, F_1) - d(A, F_2)| \leq 2c$ por la desigualdad triangular (en su variante inversa), si $c < a$ entonces $|d(A, F_1) - d(A, F_2)| < 2a$ y no existiría.

Se conoce como **centro de la hipérbola** al punto medio de los focos $\frac{1}{2}(F_1 + F_2)$.

El **eje mayor** de la hipérbola es la recta que une a los focos, $F_1 + < \overrightarrow{F_1F_2} >$.

El **eje menor o secundario** de la elipse es la recta ortogonal al eje mayor que pasa por el centro.

Los **vértices** de la elipse A_1 y A_2 son los puntos del eje mayor que pertenecen a la hipérbola.

La **excentricidad de la hipérbola** es la cantidad $\epsilon = \frac{c}{a}$. En este caso $\epsilon > 1$.

El **semieje menor** es la magnitud $b = \sqrt{c^2 - a^2} = a\sqrt{\epsilon^2 - 1}$.

Las **asíntotas** de la hipérbola son las rectas que pasan por el centro y por los puntos (a, b) y $(-a, b)$ con respecto al sistema afín ortonormal que forman el centro, el eje mayor y el eje menor. Es decir, con respecto a ese sistema, las rectas $y' = \frac{\pm b}{a}x'$.

También tiene una ecuación determinada cuando está alineada:

Proposición 72 (Ecuación reducida de la hipérbola)**.** *Dada una hipérbola en un sistema de referencia ortonormal, centrada en el* $(0, 0)$*, cuyo eje mayor es el eje* X*, semieje mayor* a *y focos* $F_1 = (-c, 0)$ *y* $F_2 = (c, 0)$ $(c > a)$*, entonces su ecuación es:*

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1.$$

Si está alineada con el eje Y*, será, evidentemente (tras un cambio ortonormal que permute* x *e* y*, no alterando sus medidas),* $-\frac{x^2}{b^2}+\frac{y^2}{a^2}=1$.

Demostración. Al igual que en el caso de la elipse, un punto (x, y) estará en la hipérbola descrita si y solo si:

$$|\sqrt{(x-c)^2+y^2}-\sqrt{(x+c)^2+y^2}|=2a.$$

Elevando al cuadrado las veces necesarias para eliminar las raíces:

$$a^4 + c^2x^2 + 2a^2cx = a^2x^2 + a^2c^2 + 2a^2cx + a^2y^2.$$

Es decir:

$$a^2(a^2 - c^2) = x^2(a^2 - c^2) + y^2a^2.$$

O lo que es lo mismo:

$$-a^2b^2 = -b^2x^2 + a^2y^2.$$

Y dividiendo se obtiene lo que se quería. □

Definición 51. Una **parábola** en el plano afín $\mathbb{A}^2_{\mathbb{R}}$, fijado un punto F llamado **foco** y una recta r llamada **directriz**, es el conjunto de los puntos $X \in \mathbb{A}^2_{\mathbb{R}}$ tales que $d(X, F) = d(X, r)$.

El **vértice de la parábola** es el punto más cercano a la directriz, o lo que es lo mismo, el punto medio del segmento que une F con r ortogonalmente.

El **eje mayor o de simetría** de la parábola es la recta ortogonal a r que pasa por F.

El **eje menor o tangente en el vértice** de la parábola es la recta ortogonal al eje mayor (paralela a la directriz) que pasa por el vértice.

Proposición 73 (Ecuación reducida de la parábola)**.** *Dada una parábola en un sistema de referencia ortonormal, centrada (su vértice) en el* $(0, 0)$*, cuyo eje mayor es el eje* X*, su foco es el* $F = (\frac{p}{2}, 0)$ *y su directriz es, por tanto, la recta* $x = -\frac{p}{2}$*, entonces su ecuación es:*

$$y^2 = 2px$$

Si está alineada con el eje Y*, será, evidentemente (tras un cambio ortonormal que permute* x *e* y*, no alterando sus medidas),* $x^2 = 2py$*.*

Demostración. El punto (x, y) está en la parábola si y solo si:

$$x + \frac{p}{2} = \sqrt{(x - \frac{p}{2})^2 + y^2}$$

Elevando al cuadrado:

$$x^2 + px + \frac{p^2}{4} = x^2 - px + \frac{p^2}{4} + y^2$$

Simplificando:

$$y^2 = 2px$$

Y se tiene lo que se quería. □

El objetivo es, dada una cónica cualquiera, desplazarla y girarla mediante movimientos ortogonales para que quede alineada como se describe anteriormente, de forma que se obtenga su ecuación reducida, lo que permite obtener sus elementos geométricos y referirlos al sistema estándar deshaciendo el movimiento.

4.1.2. Secciones del cono

Antes, vamos a comprobar algunos detalles geométricos de estas curvas, como por ejemplo, que efectivamente son las que se obtienen al intersecar el cono con un plano. Para ello, unos resultados previos:

Lema 7. *Sean A, B puntos del plano. El punto P pertenece a la circunferencia de diámetro $\overline{AB}$ $\iff$ $\overrightarrow{AP} \perp \overrightarrow{BP}$*

Demostración. Se tiene, si r es el radio de la circunferencia y P está en ella, que: $r^2 = \|CP\|^2 = (\overrightarrow{CA}+\overrightarrow{AP}, \overrightarrow{CB}+\overrightarrow{BP}) = (\overrightarrow{CA}, \overrightarrow{CB})+(\overrightarrow{AP}, \overrightarrow{CB})+(\overrightarrow{AP}, \overrightarrow{BP})+(\overrightarrow{CA}, \overrightarrow{BP}) = -r^2-(\overrightarrow{AP}, \overrightarrow{CA})+(\overrightarrow{CA}, \overrightarrow{BP})+(\overrightarrow{AP}, \overrightarrow{BP}) = -r^2 + (\overrightarrow{AP}, \overrightarrow{BP}) + (\overrightarrow{CA}, \overrightarrow{BA}) = -r^2 + 2r^2 + (\overrightarrow{AP}, \overrightarrow{BP})$, luego se tiene $(\overrightarrow{AP}, \overrightarrow{BP}) = 0$. Si P no está en ella, entonces se tiene $(\overrightarrow{AP}, \overrightarrow{BP}) \neq 0$. □

Lema 8. *Dada una circunferencia (o una esfera) de centro C y un punto exterior A, si P, P' son los puntos de tangencia de dos rectas tangentes a la circunferencia (o esfera) que pasan por A, entonces $\left\|\overrightarrow{AP}\right\| = \left\|\overrightarrow{AP'}\right\|$.*

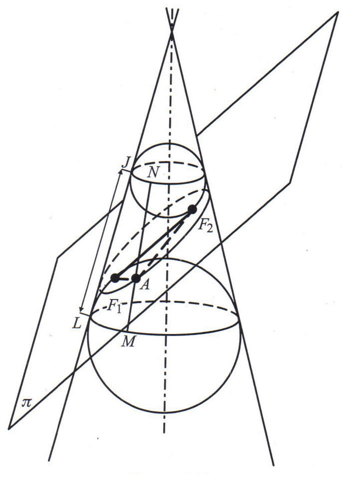
Demostración. Sea r el radio de la circunferencia (o de la esfera), y u un vector unitario en dirección de $\overrightarrow{AP}$, tal que $\overrightarrow{AP} = \lambda u$, es decir, $\|AP\| = \lambda$. Como $r^2 = \|CP\|^2 = \left\|\overrightarrow{CA} + \overrightarrow{AP}\right\|^2 = \|CA\|^2 + 2\lambda(\overrightarrow{CA}, u) + \lambda^2$.

Como P es un punto de tangencia, no puede haber dos valores de λ distintos que satisfagan esta ecuación, dado que en ese caso habría dos puntos P distintos posibles alineados con A en la circunferencia (o esfera), con lo que no serían de tangencia. Por tanto, el discriminante de esa ecuación debe ser 0, es decir: $\overrightarrow{CA}, u = \left\|\overrightarrow{CA}^2\right\| - r^2$, y entonces $\lambda = -(\overrightarrow{CA}, u)$. Combinando ambas, se tiene que $\lambda^2 = \left\|\overrightarrow{AP}\right\|^2 = \|CA\|^2 - r^2$.

Como vemos, el razonamiento anterior da lugar a una expresión de esa norma que no depende de P, mientras que sea de tangencia. Así, un razonamiento análogo con P', que también es de tangencia, da lugar a $\left\|\overrightarrow{AP'}\right\|^2 = \|CA\|^2 - r^2 = \left\|\overrightarrow{AP}\right\|^2$ □

Proposición 74. *La curva resultante de seccionar un cono con un plano cuyo ángulo con la horizontal es menor que el que forma la directriz del cono, es una elipse. Si el ángulo es mayor, se trata de una hipérbola.*

Demostración. Para la elipse, observemos la siguiente figura:



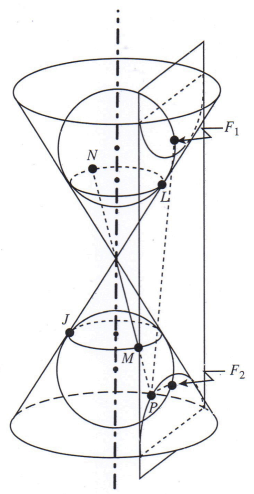
Las esferas son tangentes al plano y al cono. Los puntos de tangencia F_1 y F_2 de cada esfera con el plano constituirán los focos de la curva intersección. Para ver que efectivamente es una elipse, veamos que si A pertenece a esa curva, entonces, por el lema previo, $\left\|\overrightarrow{AF_1}\right\| = \left\|\overrightarrow{AM}\right\|$, y $\left\|\overrightarrow{AF_2}\right\| = \left\|\overrightarrow{AN}\right\|$. Por tanto, $\left\|\overrightarrow{AF_1}\right\| + \left\|\overrightarrow{F_2}\right\| = \left\|\overrightarrow{AM}\right\| + \left\|\overrightarrow{AN}\right\| = \left\|\overrightarrow{MN}\right\| = \left\|\overrightarrow{JL}\right\|$, al estar alineados. No obstante, $\overrightarrow{JL}$ es constante luego es una elipse.

Ahora, queda ver que esa suma es igual a $2a$, como habíamos definido. Esto es, no obstante, una propiedad general de las elipses, así que se podría prescindir de este detalle en la definición (y pedir, simplemente, que la suma sea constante). Para verlo, observemos que si A_1, A_2 son los vértices de la elipse, se tiene $\left\|\overrightarrow{A_1F_1}\right\| + \|\dim A_1F_2\| = \left\|\overrightarrow{A_2F_1}\right\| + \|\dim A_2F_2\|$. Reescribiendo ambos lados de la igualdad:

$$\left\|\overrightarrow{A_1F_1}\right\| + \left\|\overrightarrow{F_1F_2}\right\| + \left\|\overrightarrow{A_1F_1}\right\| = \left\|\overrightarrow{A_2F_2}\right\| + \left\|\overrightarrow{F_1F_2}\right\| + \left\|\overrightarrow{A_2F_2}\right\|$$

Con lo que $\left\|\overrightarrow{A_1F_1}\right\| = \left\|\overrightarrow{A_2F_2}\right\|$. Ahora, si A es genérico en la elipse, $\left\|\overrightarrow{AF_1}\right\| + \left\|\overrightarrow{AF_2}\right\| = \left\|\overrightarrow{A_1F_1}\right\| + \left\|\overrightarrow{A_1F_2}\right\| = \left\|\overrightarrow{A_2F_2}\right\| + \left\|\overrightarrow{A_1F_2}\right\| = \|A_1A_2\| = 2a$

Para la hipérbola, vale un razonamiento completamente análogo al anterior, que se puede visualizar en la figura siguiente:

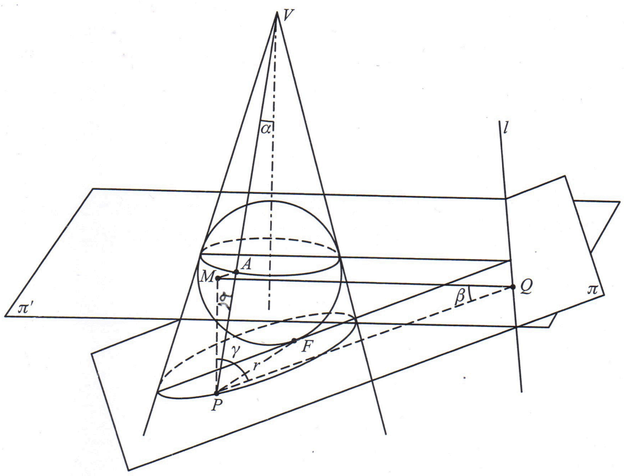


□

Lo último que veremos es que la definición foco-directriz es válida para elipses e hipérbolas también:

Proposición 75. *Se pueden definir las elipses, hipérbolas y parábolas como el lugar geométrico de puntos X tales que, fijado un foco F y una directriz l, se tiene $\frac{d(X,F)}{d(X,l)} = \epsilon$. Si $\epsilon = 1$, como sabemos, es una parábola. Con $\epsilon < 1$ es una elipse, y con $\epsilon > 1$ es una hipérbola.*

Demostración. Utilizaremos las definiciones de elipse e hipérbola que acabamos de demostrar (como secciones cónicas). Es evidente que para la parábola la proposición es cierta dado que es su definición original. Observemos la siguiente figura:

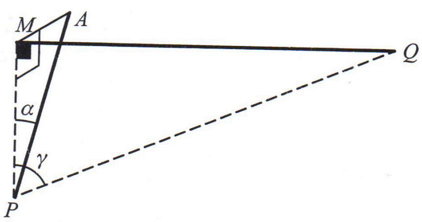


Se interseca al cono con el plano π para obtener la cónica, como hemos demostrado. Ahora, construimos la esfera tangente al cono y al plano. Vale cualquiera de las dos (superior o inferior), y será tangente al plano en el punto F que constituirá el foco. Ahora, construimos el plano π' horizontal y pasando por la circunferencia de tangencia de la esfera con el cono. La intersección entre π' y π será la directriz l.

Ahora, dado P en la cónica, ponemos M su proyección ortogonal en π', Q su proyección sobre l en dirección del eje mayor de la elipse, y A la intersección de $\overrightarrow{VP}$, con V el vértice del cono, con la circunferencia de tangencia de la esfera con el cono. (Todo aparece en la figura anterior). Si α es el ángulo de la directriz con la vertical, y γ el de π con la vertical, se tiene, finalmente:

$$\frac{\left\|\overrightarrow{PF}\right\|}{\left\|\overrightarrow{PQ}\right\|} = \frac{\|PA\|}{\left\|\overrightarrow{PQ}\right\|} = \frac{\frac{\left\|\overrightarrow{PM}\right\|}{\cos\alpha}}{\frac{\left\|\overrightarrow{PM}\right\|}{\cos\gamma}} = \frac{\cos\gamma}{\cos\alpha}$$

Para este razonamiento puede ser de ayuda el siguiente detalle de la figura anterior:



Con $\epsilon = \frac{\cos\gamma}{\cos\alpha}$ se tiene lo que se quería. Está claro que si el ángulo del plano es el mismo que el de la directriz, sale una parábola. Si es mayor el del plano (con la vertical), se obtiene la elipse (como vimos en las demostraciones anteriores) y si es menor el del plano se obtiene la elipse. □

Las figuras de las dos últimas demostraciones provienen del libro 'Álgebra lineal y Geometría' de Eugenio Hernández, María Jesús Vázquez Gallo y María Ángeles Zurro.

4.1.3. Obtener elementos geométricos de una cónica

Si volvemos ahora a la definición de cónica como expresión de la forma $X^tAX + BX + c = 0$ con $X = (x, y) \in \mathbb{A}^2_{\mathbb{R}}$, vamos a ver como modificarla mediante un movimiento rígido (isometría) para alinearla con los ejes como describimos anteriormente, obteniendo la ecuación canónica:

Teorema 9. *Cualquier ecuación de cónica de la forma $X^tAX+BX+c=0$, en un sistema de referencia ortonormal, puede transformarse en una ecuación reducida (Con A diagonal y $B=0$ si A no tiene autovalores nulos, o únicamente un coeficiente de los de B o c no nulo en caso contrario) mediante un cambio de sistema de referencia $X''=C+QX$, con $\det Q=1$ y Q ortogonal.*

Es decir, a través de una rotación y un giro de $A^2_{\mathbb{R}}$, se obtiene la ecuación reducida.

Demostración. Puesto que A es simétrica, sabemos que es diagonalizable en base ortonormal. Es decir, existe D diagonal y P ortogonal (con columnas la base ortonormal de autovectores, ly lleva base ortonormal en ortonormal) tales que:

$$D=P^{-1}AP=P^tAP.$$

Se puede seleccionar con $\det P=1$, si no, basta con permutar las columnas. Por tanto, el cambio dado por $X=PX'$, es un giro del plano que convierte la ecuación de la cónica en:

$$(PX')^tA(PX')+B(PX')+c=0.$$

Es decir:

$$X'^tDX'+BPX'+c=0.$$

Ahora, la expresión tiene la forma:

$$\alpha x'^2+\beta y'^2+ax'+by'+c=0.$$

Lo único que hay que hacer ahora es completar cuadrados en x y en y, siempre que se pueda (si α o β son nulos, no se podrá en esa variable). Una vez hecho esto, quedará una expresión de uno de los siguientes tipos:

1. Si $\beta,\alpha\neq 0$

   $$(x'+b_1)^2+(y'+b_2)^2+c-b_1^2-b_2^2=0.$$

   En cuyo caso la traslación $x''=x'+b_1$, $y''=y'+b_2$ da lugar a la forma canónica.

2. Si $\beta=0$ (análogo si $\alpha=0$):

   $$(x'+b_1)^2+b(y'+\frac{c-b_1^2}{b})=0.$$

   En cuyo caso la traslación $x''=x'+b_1$, $y''=y'+\frac{c-b_1^2}{b}$ da lugar a la forma canónica.

3. Si en el caso anterior $b=0$:

   $$(x'+b_1)^2+c-b_1^2=0.$$

   En cuyo caso la traslación $x''=x'+b_1$, $y''=y'$ da lugar a la forma canónica.

Con lo cual, tenemos dos cambios ($X''=X'+C$ y $X=PX'$), que, al componerlos, se tiene: $X''=P^tX+C$, como queríamos. Este nuevo sistema de referencia, donde está expresado X'', es el de vectores las columnas de P y centro el correspondiente a $X''=0$ (es decir, $-PC$) □

Observación 23. Una vez en forma canónica, tenemos las siguientes posibilidades:

1. **Elipse real.** $\alpha^2x^2+\beta^2y^2=c^2$. Dividiendo por c^2, llegamos a la expresión que conocemos, pudiendo obtener los elementos geométricos en el sistema de referencia final, lo que permite revertirlos al estándar con las ecuaciones ya deducidas.
2. **Elipse imaginaria.** $\alpha^2x^2+\beta^2y^2=-c^2$. No tiene soluciones reales.
3. **Hipérbola.** $\alpha^2x^2-\beta^2y^2=c^2$. Dividiendo por c^2, llegamos a la expresión que conocemos, pudiendo obtener los elementos geométricos en el sistema de referencia final.
4. **Parábola.** $y^2=2px$. Esta es la expresión que conocemos, pudiendo obtener los elementos geométricos en el sistema de referencia final.
5. **Un punto.** $\alpha x^2+\beta y^2=0$
6. **Dos rectas que se cortan.** $\alpha x^2-\beta y^2=0$. En el sistema final, son las rectas $\alpha x=\beta y$ y $\alpha x=-\beta y$.
7. **Recta doble.** $x^2=0$. Es la recta $x=0$ en el sistema final.
8. **Dos rectas paralelas.** $x^2=c^2$. Son las rectas $x=c$ y $x=-c$
9. **Dos rectas imaginarias paralelas.** $x^2=-c^2$. No tiene soluciones reales.

Por supuesto, es posible que curva final quede alineada sobre el eje Y, lo que intercambiaría x e y en las expresiones anteriores, pero el razonamiento es análogo.