6

Espacio afín

En este capítulo...

Definición de espació afín

En este tema $\mathbb{K} = \mathbb{R}$ ó $\mathbb{K} = \mathbb{C}$.

■ **Definición** Dado un espacio vectorial V sobre $\mathbb{K}$, decimos que un conjunto $\mathcal{A}$ es un espacio afín sobre V si existe una función

$$\begin{array}{rcl} + : \mathcal{A} \times V & \to & \mathcal{A} \\ (A, v) & \mapsto & A + v \end{array}$$

que satisface las siguientes propiedades:

A$_1$) para cualquier $A \in \mathcal{A}$ se tiene que $P + 0 = P$.

A$_2$) para cualquier $A \in \mathcal{A}$ y cualesquiera $u, v \in V$ se tiene que $(A + u) + v = A + (u + v)$.

A$_3$) para cualesquiera $A, B \in \mathcal{A}$ existe un *único* vector $v \in V$ tal que $A + v = B$, al cual denotaremos por $v = \overrightarrow{AB}$ y del que diremos que tiene como *origen* A y como *extremo* B (obsérvese que con esta notación se tiene la identidad natural $A + \overrightarrow{AB} = B$).

A los elementos de $\mathcal{A}$ se les llama *puntos*. Dado $A \in \mathcal{A}$ y $v \in V$ decimos que $A + v$ es el *trasladado* de A por v. La *dimensión* de $\mathcal{A}$ es por definición la del espacio vectorial V, es decir,

$$\dim(\mathcal{A}) = \dim(V).$$

Si además V es un *espacio vectorial Euclídeo*, es decir, tiene definido un producto escalar, entonces decimos que $\mathcal{A}$ es un espacio afín Euclídeo y definimos la distancia entre dos puntos A y B como

$$\mathrm{d}(A, B) = \|\overrightarrow{AB}\|.$$

Ejemplo El ejemplo que debemos tener en la cabeza es $\mathbb{R}^n$. Efectivamente, el conjunto de puntos $\mathbb{R}^n$ es un espacio afín sobre el espacio vectorial $(\mathbb{R}^n, +, \cdot)$ mediante la aplicación

$$\begin{array}{rcl} + : \mathbb{R}^n \times (\mathbb{R}^n, +, \cdot) & \to & \mathbb{R}^n \\ \big(\underbrace{(a_1, \ldots, a_n)}_{A}, \underbrace{(v_1, \ldots, v_n)}_{v}\big) & \mapsto & \underbrace{(a_1 + v_1, \ldots, a_n + v_n)}_{A+v} \end{array}$$

De hecho $\mathbb{R}^n$ es un espacio afín Euclídeo ya que en el espacio vectorial $(\mathbb{R}^n, +, \cdot)$ podemos considerar el producto escalar usual. En este caso, la distancia entre los dos puntos $A = (a_1, \ldots, a_n)$ y $B = (b_1, \ldots, b_n)$ viene dada por

$$\mathrm{d}(A, B) = \sqrt{(b_1 - a_1)^2 + \cdots + (b_n - a_n)^2}$$

Probemos a continuación algunas propiedades básicas de los espacios afines (las cuales resultan naturales si pensamos en $\mathbb{R}^n$).

■ **Proposición** *Para cualesquiera puntos A, B, C, D de un espacio afín $\mathcal{A}$ sobre V se tiene que*

1) $\overrightarrow{AB} = 0$ *si y solo si* $A = B$,

2) $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$,

3) $\overrightarrow{BA} = -\overrightarrow{AB}$,

4) $\overrightarrow{AB} = \overrightarrow{CD}$ *si y solo si* $\overrightarrow{AC} = \overrightarrow{BD}$.

Demostración. 1) Si $A = B$ queremos probar que $\overrightarrow{AB} = 0$, lo cual resulta obvio ya que por A_1 se tiene que $A + 0 = A$ y por tanto por la propia definición introducida en A_3 se tiene que $\overrightarrow{AA} = 0$. Si por el contrario se tiene que $\overrightarrow{AB} = 0$, entonces $B = A + \overrightarrow{AB} = A$, como queríamos probar.

2) Para probar que $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$ tan solo debemos ver que $A + (\overrightarrow{AB} + \overrightarrow{BC}) = C$, lo cual es cierto ya que por A_2 se tiene que

$$A + (\overrightarrow{AB} + \overrightarrow{BC}) = (A + \overrightarrow{AB}) + \overrightarrow{BC} = B + \overrightarrow{BC} = C.$$

3) Por la propiedad 2) anterior se tiene que $\overrightarrow{AB}+\overrightarrow{BA} = \overrightarrow{AA} = 0$ y por tanto $\overrightarrow{BA} = -\overrightarrow{AB}$.

4) Si $\overrightarrow{AB} = \overrightarrow{CD}$, por la propiedad 3) anterior tenemos que $\overrightarrow{AC} = \overrightarrow{AB}+\overrightarrow{BC} = \overrightarrow{CD}+\overrightarrow{BC} = \overrightarrow{BC} + \overrightarrow{CD} = \overrightarrow{BD}$. La otra dirección es parecida. □

■ **Proposición** *Para cualesquiera puntos A, B, C de un espacio afín euclídeo $\mathcal{A}$ se tiene que*

1) $d(A, B) = 0 \iff A = B$,

2) $d(A, B) = d(B, A)$,

3) $d(A, C) \leq d(A, B) + d(B, C)$.

Demostración. 1) $\mathrm{d}(A, B) = 0 \iff \|\overrightarrow{AB}\| = 0 \iff \overrightarrow{AB} = 0 \iff A = B$.

2) $\mathrm{d}(A, B) = \|\overrightarrow{AB}\| = \| - \overrightarrow{AB}\| = \|\overrightarrow{BA}\| = \mathrm{d}(B, A)$.

3) Por la desigualdad triangular que probamos en el Tema 4, sabemos que

$$\|\overrightarrow{AC}\| = \|\overrightarrow{AB} + \overrightarrow{BC}\| \leq \|\overrightarrow{AB}\| + \|\overrightarrow{BC}\|$$

y por tanto deducimos que $\mathrm{d}(A, C) \leq \mathrm{d}(A, B) + \mathrm{d}(B, C)$. □

A continuación, queremos definir lo que sería el equivalente a los subespacios vectoriales de un espacio vectorial. La idea es que van a ser "un punto más un subespacio vectorial".

■ **Definición** Sea $\mathcal{A}$ un espacio afín sobre el espacio vectorial V. Dado un punto $A \in \mathcal{A}$ y un subespacio vectorial W de V, llamamos *variedad afín que pasa por el punto A y tiene espacio de dirección W* al conjunto de puntos

$$A + W := \{A + w : w \in W\}.$$

O dicho de otra forma $A + W = \{B \in \mathcal{A} : \overrightarrow{AB} \in W\}$. Obsérvese que el punto A pertenece a $A + W$ ya que $A + 0 = A$.

Observación 6.1 1) Dada una variedad afín $A + W$ y dado un punto $B \in A + W$ es fácil comprobar que $A + W = B + W$, es decir, que podemos considerar cualquier punto de $A + W$ para definirla. En efecto, ya que $\overrightarrow{AB}$ es un vector del subespacio W, en particular $\overrightarrow{BA} = -\overrightarrow{AB}$ también lo es y se tiene que,

$$C \in A+W \Leftrightarrow \overrightarrow{AC} \in W \Leftrightarrow \overrightarrow{BA}+\overrightarrow{AC} \in W \Leftrightarrow \overrightarrow{BC} \in W \Leftrightarrow \overrightarrow{CB} = -\overrightarrow{BC} \in W \Leftrightarrow C \in B+W.$$

2)Usando la notación de la definición anterior, es fácil comprobar que

$$W = \{\overrightarrow{PQ} : P, Q \in A + W\}.$$

Veamos primero que $W \subset \{\overrightarrow{PQ} : P, Q \in A + W\}$. Dado $w \in W$, por definición A y $A + w$ son puntos de $A + W$ y el vector de origen A y extremo $A + w$ es precisamente w. Por tanto $w \in \{\overrightarrow{PQ} : P, Q \in A + W\}$, como queríamos probar. Veamos por último que $\{\overrightarrow{PQ} : P, Q \in A + W\} \subset W$. Dados dos puntos cualquiera $P, Q \in A + W$ se tiene que $\overrightarrow{AP} \in W$ y $\overrightarrow{AQ} \in W$ y en particular

$$\overrightarrow{PQ} = \overrightarrow{PA} + \overrightarrow{AQ} = -\overrightarrow{AP} + \overrightarrow{AQ} \in W.$$

3) En particular, si nos dan un subconjunto cualquiera de puntos L de $\mathcal{A}$, para comprobar que es una variedad afín tan solo tenemos que hacer dos cosas: primero, comprobar que el conjunto de vectores $\{\overrightarrow{PQ} : P, Q \in L\}$ es un subespacio vectorial de V, y segundo, en el caso de que lo sea, comprobar que

$$L = A + \{\overrightarrow{PQ} : P, Q \in L\}$$

para algún punto $A \in L$. De hecho, por la observación 1) anterior, da igual qué punto A escojamos.

Observación 6.2 Esta observación solo es una formalidad, no es importante. En realidad, dado $\mathcal{A}$ un espacio afín sobre V, la forma natural de definir una subvariedad afín de $\mathcal{A}$ es como un subconjunto de puntos que *hereda* una estructura de espacio afín. Es decir, un subconjunto $L \subset \mathcal{A}$ es un *subespacio afín* si existe un subespacio vectorial W de V de forma que L es un espacio afín sobre W a través de la restricción

$$+|_{L\times W} : L \times W \to L.$$

No es difícil probar que ambas definiciones son equivalentes.

Ejercicio 6.3 Sean A_1+W_1 y A_2+W_2 dos variedades afines de un espacio afín $\mathcal{A}$. Demuestra que si $A_1 + W_1$ y $A_2 + W_2$ no son disjuntos, entonces

$$(A_1 + W_1) \cap (A_2 + W_2) = A + (W_1 \cap W_2)$$

donde A es un punto cualquiera de la intersección $(A_1 + W_1) \cap (A_2 + W_2)$.

Ejemplo 6.4 Sea $\mathcal{A}$ un espacio afín Euclídeo y sea $A+W$ una variedad afín de $\mathcal{A}$. Dado un punto cualquiera $X \in \mathcal{A}$, podemos considerar la variedad afín ortogonal a $A+W$ que pasa por X, es decir,

$$X + W^{\perp}.$$

Con la intuición que tenemos de $\mathbb{R}^n$, la intersección de ambas debería ser un solo punto. Comprobemos que es así (véase el dibujo de más abajo). Lo primero que debemos comprobar es que la intersección no es vacía. En efecto, sabemos que $V = W \oplus W^{\perp}$ y por tanto

$$\overrightarrow{AX} = w + v$$

para ciertos $w \in W$ y $v \in W^{\perp}$ (que además son únicos). Veamos que el punto

$$p_{A+W}(X) := A + w,$$

al cual llamamos *proyección ortogonal* de X sobre $A+W$, pertenece a $(A+W)\cap(X+W^{\perp})$. Evidentemente $A + w \in A + W$, así que solo hay que comprobar que pertenece a $X + W^{\perp}$. En efecto,

$$A + w = X + \overrightarrow{XA} + w = X - \overrightarrow{AX} + w = X - w - v + w = X - v \in X + W^{\perp}.$$

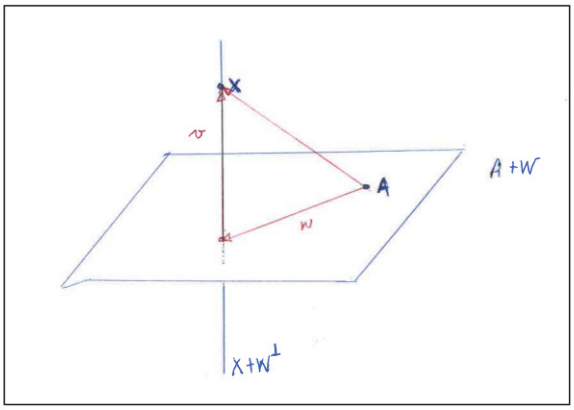
Finalmente, por el Ejercicio 6.3 tenemos que la intersección no solo no es vacía sino que es un solo punto ya que

$$(A + W) \cap (X + W^{\perp}) = p_{A+W}(X) + (W \cap W^{\perp}) = p_{A+W}(X) + \{0\}.$$

Obsérvese por último que

$$p_{A+W}(X) := A + p_W(\overrightarrow{AX}),$$

donde $p_W : V \to V$ es la proyección ortogonal de V sobre W.



Sistema de referencia y coordenadas

Nuestra objetivo es ahora el de definir en un espacio afín algo parecido al concepto de base y de coordenadas de un espacio vectorial. Sea $\mathcal{A}$ un espacio afín sobre un espacio vectorial V de dimensión n. Fijemos un punto O cualquiera de $\mathcal{A}$ y una base cualquiera $\mathcal{B}$ de V. Observamos que dado un punto $A \in \mathcal{A}$, el vector $\overrightarrow{OA}$ tiene unas ciertas coordenadas respecto de $\mathcal{B}$,

$$\overrightarrow{OA} = (a_1, \ldots, a_n)_{\mathcal{B}}.$$

Si nos dan otro punto $A' \in \mathcal{A}$ y las coordenadas de $\overrightarrow{OA}$ y $\overrightarrow{OA'}$ respecto de $\mathcal{B}$ coinciden entonces evidentemente $\overrightarrow{OA} = \overrightarrow{OA'}$ y en particular $A = A'$. Es decir, que esas coordenadas están describiendo unívocamente al punto A. Por tanto tiene sentido que digamos que

$$\mathcal{R} = \{O; \mathcal{B}\}$$

es un *sistema de referencia cartesiano* de $\mathcal{A}$ y que las coordenadas del punto A respecto de $\mathcal{R}$ son

$$A = (a_1, \ldots, a_n)_{\mathcal{R}}.$$

Además, si nos dan dos puntos $A, B \in \mathcal{A}$ cuyas coordenadas respecto de $\mathcal{R}$ son

$$A = (a_1, \ldots, a_n)_{\mathcal{R}} \qquad B = (a_1, \ldots, a_n)_{\mathcal{R}}$$

entonces tendremos que

$$\overrightarrow{AB} = \overrightarrow{AO} + \overrightarrow{OB} = -\overrightarrow{OA} + \overrightarrow{OB} = (b_1 - a_1, \ldots, b_n - a_n)_{\mathcal{R}}.$$

De lo anterior deducimos de forma inmediata que si $A = (a_1, \ldots, a_n)_{\mathcal{R}} \in \mathcal{A}$ y $v = (v_1, \ldots, v_n)_{\mathcal{B}} \in V$ entonces el trasladado B de A por v tiene coordenadas

$$B = (a_1 + v_1, \ldots, a_n + v_n)_{\mathcal{R}}.$$

Si el espacio vectorial V es de hecho un espacio vectorial Euclídeo y la base $\mathcal{B}$ es ortonormal, entonces diremos que $\mathcal{R}$ es un *sistema de referencia rectangular.* En este caso, dados dos puntos $A = (a_1, \ldots, a_n)_{\mathcal{R}}$ y $B = (a_1, \ldots, a_n)_{\mathcal{R}}$ se tiene que

$$\mathrm{d}(A, B) = \|\overrightarrow{AB}\| = \sqrt{(b_1 - a_1)^2 + \cdots + (b_n - a_n)^2}.$$

Ejemplo 6.5 En $\mathbb{R}^n$ como espacio afín euclídeo, tenemos el sistema de referencia rectangular canónico

$$\mathcal{R}_c = \{(0,\ldots,0) : \mathcal{B}_c\}$$

donde $\mathcal{B}_c$ es la base canónica de $\mathbb{R}^n$.

Y si nos dan otro sistema de referencia $\mathcal{R}' = \{O', \mathcal{B}'\}$, ¿cómo se realiza el cambio de coordenadas de $\mathcal{R}'$ a $\mathcal{R}$? Al igual que ocurría con las bases en los espacios vectoriales, necesitamos que $\mathcal{R}$ tenga información de $\mathcal{R}'$. En concreto, necesitamos dos cosas:

1) La matriz P de cambio de $\mathcal{B}'$ a $\mathcal{B}$, y

2) Las coordenadas $O' = (b_1, \ldots, b_n)_{\mathcal{R}}$ del punto O' respecto de $\mathcal{R}$, es decir, que
$$\overrightarrow{OO'} = (b_1, \ldots, b_n)_{\mathcal{B}}.$$

Veamos que con estos datos es suficiente. Consideremos un punto cualquiera $A \in \mathcal{A}$ del cual conocemos sus coordenadas respecto de $\mathcal{R}'$,

$$A = (x'_1, \ldots, x'_n)_{\mathcal{R}'}$$

es decir, que $\overrightarrow{O'A} = (x'_1, \ldots, x'_n)_{\mathcal{B}'}$. Queremos saber cuales son las coordenadas

$$A = (x_1, \ldots, x_n)_{\mathcal{R}}$$

de A respecto de $\mathcal{R}$. Por un lado tenemos que las coordenadas de $\overrightarrow{O'A}$ respecto de $\mathcal{B}$ son

$$P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

En particular, ya que $\overrightarrow{OA} = \overrightarrow{OO'} + \overrightarrow{O'A}$, obtenemos que

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} + P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

La fórmula anterior se suele escribir en forma reducida

$$X = O' + PX'$$

donde X denota la matriz columna de las coordenadas de A respecto de $\mathcal{R}$, X' denota la matriz columna de las coordenadas de A respecto de $\mathcal{R}'$ y O' denota la matriz columna de las coordenadas de O' respecto de $\mathcal{R}$.

De hecho, haciendo un pequeño artificio, incluso podemos capturar la fórmula anterior en forma matricial, lo cual siempre es más cómodo:

$$\left(\begin{array}{c} 1 \\ \hline X \end{array}\right) = \left(\begin{array}{c|c} 1 & 0 \\ \hline O' & P \end{array}\right) \left(\begin{array}{c} 1 \\ \hline X' \end{array}\right)$$

A la matriz cuadrada de orden $n+1$ que aparece en la expresión anterior

$$Q = \left(\begin{array}{c|c} 1 & 0 \\ \hline O' & P \end{array}\right)$$

la llamamos matriz del cambio de referencia de $\mathcal{R}'$ a $\mathcal{R}$. Al igual que pasaba en los cambios de base, la inversa de la matriz anterior es la matriz del cambio de referencia de $\mathcal{R}$ a $\mathcal{R}'$.

Para terminar la sección, veamos que dado un sistema de referencia $\mathcal{R} = \{O, \mathcal{B}\}$ de $\mathcal{A}$, es posible describir una variedad afín $A + W$ de $\mathcal{A}$ mediante ecuaciones paramétricas e implícitas. En efecto, si $A = (a_1, \ldots, a_n)$ y $W = L[w_1, \ldots, w_\ell]$ donde los vectores $w_1, \ldots, w_\ell$ son una base de W cuyas coordenadas respecto de $\mathcal{B}$ son

$$\begin{cases} w_1 = (w_{11}, \ldots, w_{\ell 1})_{\mathcal{B}} \\ \vdots \\ w_\ell = (w_{1\ell}, \ldots, w_{\ell\ell})_{\mathcal{B}} \end{cases}$$

entonces tendremos que un punto $X = (x_1, \ldots, x_n)_{\mathcal{R}}$ pertenece a $A+W$ si y solo si $\overrightarrow{AX} \in W$, es decir, si existen parámetros $\lambda_1, \ldots, \lambda_\ell \in \mathbb{R}$ tales que

$$\begin{cases} x_1 = a_1 + \lambda_1 w_{11} + \ldots + \lambda_\ell w_{1\ell} \\ \vdots \\ x_n = a_n + \lambda_1 w_{\ell 1} + \ldots + \lambda_\ell w_{\ell\ell} \end{cases}$$

las cuales son las *ecuaciones paramétricas* de $A + W$ en el sistema de referencia $\mathcal{R}$. Si eliminamos los parámetros de estas ecuaciones obtendremos unas *ecuaciones implícitas* de $A + W$.

Coordenadas Baricéntricas

Estudiemos otra forma de referencia en los espacios afines. Primero debemos probar lo siguiente:

■ **Proposición** (Combinación afín) *Sea $\mathcal{A}$ un espacio afín sobre el $\mathbb{K}$-espacio vectorial V de dimensión n, y sean $A_0, \ldots, A_n \in \mathcal{A}$. Probar que las siguientes condiciones son equivalentes:*

1) Los vectores $\{\overrightarrow{A_0A_1}, \ldots, \overrightarrow{A_0A_n}\}$ son linealmente independientes.

2) Fijado cualquier índice $j \in \{0, 1, \ldots, n\}$, los vectores

$$\{\overrightarrow{A_jA_i} : i \neq j\}$$

son linealmente independientes.

3) Dado cualquier punto $X \in \mathcal{A}$, si existen $\alpha_0, \ldots, \alpha_n \in \mathbb{K}$ tales que $\alpha_0 + \cdots + \alpha_n = 0$ y

$$\alpha_0 \overrightarrow{XA_0} + \cdots + \alpha_n \overrightarrow{XA_n} = 0$$

entonces $\alpha_0 = \cdots = \alpha_n = 0$.

Demostración. Veamos que 1) implica 2). Tenemos que probar que, fijado $j \in \{1, \ldots, n\}$, los vectores

$$\{\overrightarrow{A_jA_i} : i \neq j\}$$

son linealmente independientes. Supongamos que existen $\beta_i \in \mathbb{K}$ para $i \neq j$ tales que

$$\sum_{i \neq j} \beta_i \overrightarrow{A_jA_i} = 0.$$

En particular,

$$\sum_{i \neq j} \beta_i (\overrightarrow{A_jA_0} + \overrightarrow{A_0A_i}) = 0.$$

y por tanto, como $\overrightarrow{A_0A_0} = 0$ obtenemos que

$$\beta_1\overrightarrow{A_0A_1} + \cdots + \beta_{j-1}\overrightarrow{A_0A_{j-1}} - (\sum_{i\neq j}\beta_i)\overrightarrow{A_0A_j} + \beta_{j+1}\overrightarrow{A_0A_{j+1}} + \cdots + \beta_n\overrightarrow{A_0A_n} = 0.$$

Por (1) entonces $\beta_1, \ldots, \beta_{j-1}, \ldots, \beta_{j+1}, \ldots, \beta_n = 0$ y $\sum_{i\neq j}\beta_i = 0$, de lo que deducimos que también $\beta_0 = 0$, como queríamos probar.

Veamos que 2) implica 3). Sea $X \in \mathcal{A}$ y $\alpha_0, \ldots, \alpha_n \in \mathbb{K}$ tales que $\alpha_0 + \cdots + \alpha_n = 0$ y

$$\alpha_0\overrightarrow{XA_0} + \cdots + \alpha_n\overrightarrow{XA_n} = 0.$$

En particular,

$$\alpha_0(\overrightarrow{XA_0} + \overrightarrow{A_0A_0}) + \cdots + \alpha_n(\overrightarrow{XA_0} + \overrightarrow{A_0A_n}) = 0$$

y por tanto

$$(\alpha_0 + \cdots + \alpha_n)\overrightarrow{XA_0} + \alpha_1\overrightarrow{A_0A_1} + \cdots + \alpha_n\overrightarrow{A_0A_n} = 0.$$

Como $\alpha_0 + \cdots + \alpha_n = 0$ deducimos que

$$\alpha_1\overrightarrow{A_0A_1} + \cdots + \alpha_n\overrightarrow{A_0A_n} = 0$$

y por (2) se tiene que $\alpha_1 = \cdots = \alpha_n = 0$. En particular,

$$\alpha_0 = -\alpha_1 - \cdots - \alpha_n = 0,$$

como queríamos probar.

Por último, veamos que (3) implica (1). Sea $\alpha_1, \ldots, \alpha_n \in \mathcal{A}$ tales que

$$\alpha_1\overrightarrow{A_0A_1} + \cdots + \alpha_n\overrightarrow{A_0A_n} = 0.$$

Definamos $\alpha_0 := -\alpha_1 - \cdots - \alpha_n$. Puesto que $\overrightarrow{A_0A_0} = 0$ se tiene que

$$\alpha_0\overrightarrow{A_0A_0} + \alpha_1\overrightarrow{A_0A_1} + \cdots + \alpha_n\overrightarrow{A_0A_n} = 0,$$

y dado que $\alpha_0+\alpha_1+\cdots+\alpha_n = 0$ deducimos de (3) que $\alpha_0, \ldots, \alpha_n = 0$, como deseábamos. □

■ **Definición** Dado un espacio afín $\mathcal{A}$ de dimensión n, decimos que los puntos $A_0, \ldots, A_n \in \mathcal{A}$ son *afínmente independientes* si alguna de las condiciones (o equivalentemente, todas) de la proposición anterior se satisface.

Sea $\mathcal{A}$ un espacio afín sobre el $\mathbb{K}$-espacio vectorial V de dimensión n, y sean $A_0, \ldots, A_n \in \mathcal{A}$ puntos afínmente independientes. Dado un punto cualquier $X \in \mathcal{A}$, dado que V tiene dimensión n, los vectores $\overrightarrow{XA_0}, \ldots, \overrightarrow{XA_n}$ deben ser linealmente dependientes, es decir, existen $\alpha_0, \ldots, \alpha_n \in \mathbb{K}$ no todos nulos tales que

$$\alpha_0\overrightarrow{XA_0} + \cdots + \alpha_n\overrightarrow{XA_n} = 0.$$

Observamos que si $\alpha_0 + \cdots + \alpha_n = 0$ entonces por el ejercicio anterior tendríamos que $\alpha_0 = \cdots = \alpha_n = 0$, lo cual es una contradicción. Por tanto $\alpha_0 + \cdots + \alpha_n \neq 0$ y en particular si definimos

$$x_0 := \frac{\alpha_0}{\alpha_0 + \cdots + \alpha_n} \quad \& \quad \cdots \quad \& \quad x_n := \frac{\alpha_n}{\alpha_0 + \cdots + \alpha_n}$$

se tiene que

$$x_0\overrightarrow{XA_0} + \cdots + x_n\overrightarrow{XA_n} = \frac{1}{\alpha_0 + \cdots + \alpha_n}\big(\alpha_0\overrightarrow{XA_0} + \cdots + \alpha_n\overrightarrow{XA_n}\big) = 0,$$

y

$$x_0 + \cdots + x_n = \frac{\alpha_0 + \cdots + \alpha_n}{\alpha_0 + \cdots + \alpha_n} = 1.$$

■ **Lema** *Sea $\mathcal{A}$ un espacio afín de dimensión n. Sean $A_0, \dots, A_n \in \mathcal{A}$ puntos afínmente independientes. Entonces existen unos únicos $x_0, \dots, x_n \in \mathbb{K}$ tales que*

$$x_0\overrightarrow{XA_0} + \cdots + x_n\overrightarrow{XA_n} = 0 \quad \& \quad x_0 + \cdots + x_n = 1.$$

Es más, si $Y \in \mathcal{A}$ es un punto tal que $x_0\overrightarrow{YA_0} + \cdots + x_n\overrightarrow{YA_n} = 0$ entonces $X = Y$.

Demostración. Justo antes del enunciado de este lema hemos probado que existen $x_0, \dots, x_n \in \mathbb{K}$ tales que

$$x_0\overrightarrow{XA_0} + \cdots + x_n\overrightarrow{XA_n} = 0 \quad \& \quad x_0 + \cdots + x_n = 1.$$

Sean $x_0', \dots, x_n' \in \mathbb{K}$ tales que

$$x_0'\overrightarrow{XA_0} + \cdots + x_n'\overrightarrow{XA_n} = 0 \quad \& \quad x_0' + \cdots + x_n' = 1,$$

y probemos que $x_0' = x_0, \dots, x_n' = x_n$. En efecto, si definimos $\alpha_i := x_i - x_i'$ para $i = 0, \dots, n$, se cumple que

$$\alpha_0\overrightarrow{XA_0} + \cdots + \alpha_n\overrightarrow{XA_n} = 0 \quad \& \quad \alpha_0 + \cdots + \alpha_n = 0$$

y por la Proposición de la Combinación de afín deducimos que $\alpha_i = x_i - x_i' = 0$ para todo $i = 0, \dots, n$, como queríamos probar.

Por último, sea $Y \in \mathcal{A}$ un punto tal que $x_0\overrightarrow{YA_0} + \cdots + x_n\overrightarrow{YA_n} = 0$. En particular,

$$x_0\overrightarrow{A_0Y} + \cdots + x_n\overrightarrow{A_nY} = -x_0\overrightarrow{YA_0} - \cdots - x_n\overrightarrow{YA_n} = 0$$

y por tanto, como también $x_0\overrightarrow{XA_0} + \cdots + x_n\overrightarrow{XA_n} = 0$, deducimos que

$$x_0(\overrightarrow{XA_0} + \overrightarrow{A_0Y}) + \cdots + x_n(\overrightarrow{XA_n} + \overrightarrow{A_nY}) = 0.$$

Así pues, dado que $x_0 + \cdots, x_n = 1$, se tiene que

$$0 = x_0\overrightarrow{XY} + \cdots + x_n\overrightarrow{XY} = (x_0 + \cdots + \cdots + x_n)\overrightarrow{XY} = \overrightarrow{XY}$$

de lo que deducimos que $X = Y$. □

■ **Definición** Sea $\mathcal{A}$ un espacio afín de dimensión n. Sean $A_0, \dots, A_n \in \mathcal{A}$ puntos afínmente independientes. Entonces decimos que

$$\mathcal{R} := \{A_0, \dots, A_n\}$$

es un *sistema de referencia baricéntrico* de $\mathcal{A}$. Dado un punto $X \in \mathcal{A}$, si $x_0, \dots, x_n \in \mathbb{K}$ son los únicos escalares tales que

$$x_0\overrightarrow{XA_0} + \cdots + x_n\overrightarrow{XA_n} = 0 \quad \& \quad x_0 + \cdots + x_n = 1,$$

entonces denotamos $X = (x_0, \dots, x_n)_{\mathcal{R}}$ y decimos que son sus *coordenadas baricentricas* respecto de $\mathcal{R}$.

Ejercicio Sea $\mathbb{R}^2$ como espacio afín. Considérense los puntos $A_0 = (1, 0)$, $A_1 = (0, 1)$ y $A_2 = (-1, 0)$. Probar que son afínmente independientes. Calcular las coordenadas baricentricas del punto $(0, 0)$ respecto del sistema de referencia baricéntrico $\mathcal{R} := \{A_0, A_1, A_2\}$ de $\mathbb{R}^2$. Calcular las coordenadas del punto A_1 respecto de $\mathcal{R}$. Dibujar el conjunto de puntos cuyas coordenadas baricentricas son de la forma $(0, x_1, x_2)_{\mathcal{R}}$. Dibujar el conjunto de puntos cuyas coordenadas baricentricas son de la forma $(0, x_1, x_2)_{\mathcal{R}}$ con $x_1, x_2 \geq 0$. Dibujar el conjunto de puntos cuyas coordenadas baricentricas son de la forma $(x_0, x_1, x_2)_{\mathcal{R}}$ con $x_0, x_1, x_2 \geq 0$.

Aplicaciones afines

Siguiendo el espíritu de las secciones anteriores, queremos definir una aplicación afín como "un punto más una aplicación lineal".

■ **Definición** Sean $\mathcal{A}$ y $\mathcal{A}'$ espacios afines sobre los espacios vectoriales V y V' respectivamente. Dada una aplicación lineal

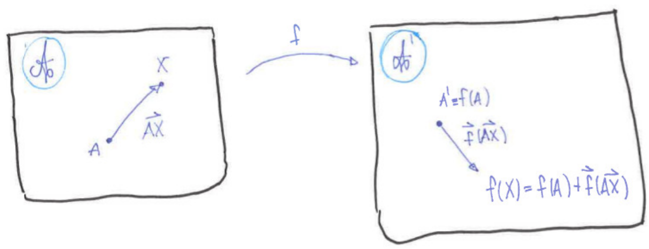
$$\vec{f} : V \to V'$$

y fijados dos puntos $A \in \mathcal{A}$ y $A' \in \mathcal{A}'$, decimos que la aplicación

$$\begin{array}{rcl} f : \mathcal{A} & \to & \mathcal{A}' \\ X & \mapsto & A' + \vec{f}(\overrightarrow{AX}) \end{array}$$

es una *aplicación afín* con aplicación lineal asociada $\vec{f}$.

Observación 6.6 1) Obsérvese que por definición $f(A) = A' + \vec{f}(\overrightarrow{AA}) = A'$. Así que podemos pensar que lo que hemos hecho al fin y al cabo es fijar un punto A de $\mathcal{A}$, a continuación hemos elegido la imagen $A' = f(A)$ de ese punto en $\mathcal{A}'$, para después establecer que la imagen de un punto cualquiera X de $\mathcal{A}$ depende linealmente del vector $\overrightarrow{AX}$ por medio de $\vec{f}$, es decir, $f(X) = f(A) + \vec{f}(\overrightarrow{AX})$.



2) Si fijamos un punto $B \in \mathcal{A}$ cualquiera, entonces podemos reescribir f de la siguiente manera:

$$f(X) = f(B) + \vec{f}(\overrightarrow{BX}).$$

En efecto,

$$\begin{aligned} f(X) &= A' + \vec{f}(\overrightarrow{AX}) = A' + \vec{f}(\overrightarrow{AB} + \overrightarrow{BX}) = \\ &= A' + \vec{f}(\overrightarrow{AB}) + \vec{f}(\overrightarrow{BX}) = f(B) + \vec{f}(\overrightarrow{BX}). \end{aligned}$$

Esta igualdad nos está asegurando que podemos usar un punto cualquiera para definir la aplicación.

3) La identidad $I : \mathcal{A} \mapsto \mathcal{A} : X \mapsto X$ es una aplicación afín cuya aplicación lineal asociada es también la identidad: en efecto, si fijamos un punto cualquiera $A \in \mathcal{A}$, podemos escribir

$$I(X) = X = A + \overrightarrow{AX}.$$

4) Si sabemos que $f : \mathcal{A} \to \mathcal{A}'$ es una aplicación afín, podemos recuperar la aplicación lineal asociada $\vec{f}$ de la siguiente forma:

$$\begin{array}{rcl} \vec{f} : V & \to & V' \\ \overrightarrow{XY} & \mapsto & \overrightarrow{f(X)f(Y)} \end{array}$$

5) Si nos dan una aplicación

$$f : \mathcal{A} \to \mathcal{A}'$$

¿cómo podemos saber si es una aplicación afín? Fácil, tan solo tenemos que fijar un punto cualquiera $A \in \mathcal{A}$ y comprobar que la correspondencia

$$\begin{array}{rcl} \vec{f} : V & \to & V' \\ \overrightarrow{AX} & \mapsto & \vec{f}(\overrightarrow{AX}) := \overrightarrow{f(A)f(X)} \end{array}$$

es una aplicación lineal (la cual será evidentemente la aplicación lineal asociada a f). En caso de que lo sea, tendremos que

$$\begin{array}{rcl} f : \mathcal{A} & \to & \mathcal{A}' \\ X & \mapsto & f(A) + \vec{f}(\overrightarrow{AX}) \end{array}$$

es una aplicación afín.

Ejercicio 6.7 Demostrar que si $f : \mathcal{A} \to \mathcal{A}'$ y $g : \mathcal{A}' \to \mathcal{A}''$ son aplicaciones afines entonces $g \circ f : \mathcal{A} \to \mathcal{A}''$ es una aplicación afín. Prueba que $\overrightarrow{g \circ f} = \vec{g} \circ \vec{f}$.

Ejercicio 6.8 Demuestra que una aplicación afín $f : \mathcal{A} \to \mathcal{A}'$ es una biyección si y solo si $\vec{f}$ es un isomorfismo. En este caso, decimos que f es una *isomorfismo afín.*

Ejemplo 6.9 La aplicación

$$\begin{array}{rcl} f : \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (x, y, z) & \mapsto & (x + y + 1, x - z, x + y + z - 1) \end{array}$$

es una aplicación afín.

Ejemplo 6.10 Sea $\mathcal{A}$ un espacio afín sobre V y sea $v \in V$. La *traslación* de vector v es la aplicación

$$\begin{array}{rcl} t_v : \mathcal{A} & \to & \mathcal{A} \\ X & \mapsto & X + v \end{array}$$

la cual, si fijamos un punto A, podemos escribir como

$$t_v(X) = (A + v) + \overrightarrow{AX}$$

y por tanto es evidentemente una aplicación afín cuya aplicación lineal asociada es la identidad. Por ejemplo, la traslación en $\mathbb{R}^3$ de vector $v = (1, 1, 2)$ es la aplicación

$$\begin{array}{rcl} t_v : \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (x, y, z) & \mapsto & (x, y, z) + (1, 1, 2) = (x + 1, y + 1, z + 2) \end{array}$$

Ejemplo 6.11 Sea $A + W$ una variedad afín de un espacio afín $\mathcal{A}$ sobre V. Recuérdese que en el Ejemplo 6.4 vimos que dado un punto $X \in \mathcal{A}$ podemos considerar la proyección ortogonal de X sobre $A + W$. Evidentemente, esto define una aplicación afín

$$\begin{array}{rcl} p_{A+W} : \mathcal{A} & \to & \mathcal{A} \\ X & \mapsto & p_{A+W}(X) = A + p_W(\overrightarrow{AX}) \end{array}$$

que llamamos *proyección ortogonal* de $\mathcal{A}$ sobre $A + W$.

Ejemplo 6.12 Con la notación del ejemplo anterior, podemos definir también la simetría de $\mathcal{A}$ respecto de $A + W$, la cual viene dada por

$$\begin{array}{rcl} s_{A+W} : \mathcal{A} & \to & \mathcal{A} \\ X & \mapsto & A + s_W(\overrightarrow{AX}) \end{array}$$

donde s_w denota la simetría ortogonal de V respecto de W. Es fácil probar (ejercicio) que $s_{A+W}(X) = X - 2p_{A+W}(X)$ usando el hecho de que $s_W = I - 2p_W$.

Al igual que hicimos con la aplicaciones lineales, nos gustaría trabajar con las aplicaciones afines de forma matricial. Sean $\mathcal{A}$ y $\mathcal{A}'$ espacios afines sobre V y V' respectivamente. Sean $\mathcal{R} = \{O; \mathcal{B}\}$ y $\mathcal{R}' = \{O'; \mathcal{B}'\}$ sistemas de referencia de $\mathcal{A}$ y $\mathcal{A}'$ respectivamente. Consideremos una aplicación afín

$$f : \mathcal{A} \to \mathcal{A}$$

con aplicación lineal asociada $\vec{f} : V \to V'$. Sabemos entonces que para cualquier punto $X \in \mathcal{A}$ se tiene que

$$f(X) = f(O) + \vec{f}(\overrightarrow{OX}).$$

Denotemos por $X = (x_1, \ldots, x_n)_{\mathcal{R}}$ y $f(X) = (y_1, \ldots, y_m)_{\mathcal{R}'}$ a las coordenadas de X y $f(X)$ respecto de $\mathcal{R}$ y $\mathcal{R}'$. Entonces,

1) si M es la matriz de $\vec{f}$ respecto de $\mathcal{B}$ y $\mathcal{B}'$, y

2) $f(O) = (c_1, \ldots, c_m)_{\mathcal{R}'}$,

se tiene que

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} + M \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Esta expresión podemos reescribirla de la siguiente forma

$$\begin{pmatrix} 1 \\ \hline y_1 \\ \vdots \\ y_m \end{pmatrix} = \underbrace{\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline c_1 & & & \\ \vdots & & M & \\ c_m & & & \end{array}\right)}_{N} \begin{pmatrix} 1 \\ \hline x_1 \\ \vdots \\ x_n \end{pmatrix}$$

donde a la matriz N la denominamos *matriz de f asociada a los sistemas de referencia $\mathcal{R}$ y $\mathcal{R}'$.*

Ejemplo 6.13 La aplicación

$$\begin{array}{rcl} f : \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (x, y, z) & \mapsto & (x + y + 1, x - z, x + y + z - 1) \end{array}$$

es una aplicación afín cuya matriz respecto del sistema de referencia canónico $\mathcal{R}_c$ es

$$N = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 1 & 1 & 1 \end{array}\right)$$

La traslación de vector $v = (1, 1, 2)$,

$$\begin{array}{rcl} t_v : \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (x, y, z) & \mapsto & (x, y, z) + (1, 1, 2) = (x + 1, y + 1, z + 2) \end{array}$$

tiene matriz asociada respecto del sistema de referencia canónico $\mathcal{R}_c$

$$N' = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{array}\right)$$

En particular, la matriz de $f \circ t_v$ respecto del sistema de referencia canónico $\mathcal{R}_c$ es

$$N \cdot N' = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 1 & 1 & 1 \end{array}\right) \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 3 & 1 & 1 & 0 \\ -1 & 1 & 0 & -1 \\ 3 & 1 & 1 & 1 \end{array}\right)$$

Ejercicio 6.14 Sea $f : \mathcal{A} \to \mathcal{A}'$ una aplicación afín. Demostrar que

$$\mathrm{Im}(f) = \{f(A) : A \in \mathcal{A}\}$$

es una variedad afín de $\mathcal{A}'$. Dado $B \in \mathcal{A}'$, demostrar que $f^{-1}(B)$ es una variedad afín de $\mathcal{A}$.

El conjunto de los puntos fijos de una aplicación afín va a ser muy importante en las siguientes secciones:

■ **Proposición** (de los puntos fijos) *Sea $f : \mathcal{A} \to \mathcal{A}$ una aplicación afín. Entonces el conjunto $\mathcal{F}_f = \{A \in \mathcal{A} : f(A) = A\}$ de puntos fijos de f, si no es un conjunto vacío, es una variedad afín de $\mathcal{A}$. De hecho, si A es un punto fijo cualquiera de f, se tiene que*

$$\mathcal{F}_f = A + Ker(\vec{f} - I). \qquad (*)$$

Demostración. En efecto, si existe un punto $A \in \mathcal{A}$ que es fijo entonces $f(A) = A$ y por tanto se tiene que

$$X = f(X) = f(A) + \vec{f}(\overrightarrow{AX}) = A + \vec{f}(\overrightarrow{AX}) \Longleftrightarrow \overrightarrow{AX} = \vec{f}(\overrightarrow{AX}) \Longleftrightarrow$$

$$\Longleftrightarrow \overrightarrow{AX} \in \mathrm{Ker}(\vec{f} - I) \Longleftrightarrow X \in A + \mathrm{Ker}(\vec{f} - I).$$

Es decir, los puntos fijos de f son $A + \mathrm{Ker}(\vec{f} - I)$. □

Movimientos

En esta sección $\mathcal{A}$ denotará siempre un espacio afín euclídeo, es decir, que su espacio vectorial asociado V es un espacio vectorial Euclídeo y tiene por tanto asociado un cierto producto escalar $\langle -, - \rangle$. Dado que tenemos un concepto de distancia en $\mathcal{A}$, resulta natural estudiar las aplicaciones afines que respetan dicha distancia.

■ **Definición** Decimos que una aplicación afín $f : \mathcal{A} \to \mathcal{A}$ es un *movimiento* si conserva la distancia entre puntos, es decir,

$$\mathrm{d}(A, B) = \mathrm{d}(f(A), f(B)) \qquad \text{para cualesquiera } A, B \in \mathcal{A}.$$

Ejercicio 6.15 Demostrar que la composición de dos movimientos es un movimiento.

■ **Proposición** *Una aplicación afín $f : \mathcal{A} \to \mathcal{A}$ es un movimiento si y solo si su aplicación lineal asociada es un endomorfismo ortogonal.*

Demostración. Para cualesquiera $A, B \in \mathcal{A}$ se tiene que

$$\mathrm{d}(A, B) = \mathrm{d}(f(A), f(B)) \iff \|\overrightarrow{AB}\| = \|\overrightarrow{f(A)f(B)}\| = \|\vec{f}(\overrightarrow{AB})\|$$

es decir, que f respeta la distancia si y solo si $\vec{f}$ respeta la norma. □

■ **Corolario** *Sea $\mathcal{R} = \{0; \mathcal{B}\}$ un sistema de referencia rectangular de $\mathcal{A}$. Sea $f : \mathcal{A} \to \mathcal{A}$ una aplicación afín cuya matriz respecto de $\mathcal{R}$ es*

$$\left(\begin{array}{c|c} 1 & 0 \\ \hline C & M \end{array}\right).$$

Entonces f es un movimiento si y solo si M es ortogonal.

Demostración. Es inmediato, ya que la matriz M es la matriz de $\vec{f}$ respecto de la base ortonormal $\mathcal{B}$, la cual es ortogonal si y solo si $\vec{f}$ es un endomorfismo orotogonal. □

Vemos por tanto que las *traslaciones son movimientos* ya que la matriz de su aplicación lineal asociada es la identidad, la cual es obviamente ortogonal. Las traslaciones no tienen puntos fijos (recuérdese la definición en la Proposición de los puntos fijos), a no ser que traslademos por el vector 0. Por otro lado, el siguiente resultado nos asegura que siempre podemos trasladar un movimiento para conseguir que tenga puntos fijos (lo cual será muy importante para su clasificación):

■ **Teorema** (Descomposición de movimientos) *Sea $f : \mathcal{A} \to \mathcal{A}$ un movimiento* sin *puntos fijos. Entonces existe un vector no nulo $u \in Ker(\vec{f} - id)$ y un movimiento $g : \mathcal{A} \to \mathcal{A}$* con *puntos fijos tales que*

$$f = t_u \circ g.$$

En particular, se tiene que $\vec{g} = \vec{f}$ y que la variedad de puntos fijos $\mathcal{F}_g$ es una variedad invariante de f, es decir, $f(\mathcal{F}_g) = \mathcal{F}_g$.

Demostración. Veamos primero que

$$V = \mathrm{Ker}(\vec{f} - id) \oplus \mathrm{Im}(\vec{f} - id).$$

Ya sabemos que $\dim(V) = \dim\big(\mathrm{Ker}(\vec{f} - id)\big) + \dim\big(\mathrm{Im}(\vec{f} - id)\big)$, así que basta probar que su intersección es nula. En efecto, en el ejercicio 2 de la hoja 5 probamos incluso que $\mathrm{Ker}(\vec{f} - id)^{\perp} = \mathrm{Im}(\vec{f} - id)$.

Dado un punto O cualquiera de $\mathcal{A}$, sabemos que podemos descomponer de forma única el vector $\overrightarrow{Of(O)}$ como suma de un vector de $\mathrm{Ker}(\vec{f} - id)$ y otro de $\mathrm{Im}(\vec{f} - id)$, es decir, que

$$\overrightarrow{Of(O)} = u + (\vec{f}(v) - v))$$

para cierto $v \in V$ y para cierto $u \in V$ que satisface $\vec{f}(u) = u$. Definimos entonces $g = t_{-u} \circ f$. Se tiene que g es un movimiento por ser composición de dos movimientos (Ejercicio 6.15), y que $f = (t_{-u})^{-1} \circ g = t_u \circ g$. Obsérvese que $\vec{f} = \vec{t_u} \circ \vec{g} = id \circ \vec{g} = \vec{g}$.

Comprobemos que g tiene puntos fijos. De hecho, mostremos que el único punto A de $\mathcal{A}$ que satisface que $\overrightarrow{AO} = v$ es un punto fijo de g. En efecto, observamos que

$$\overrightarrow{Af(A)} = \overrightarrow{AO} + \overrightarrow{Of(O)} + \overrightarrow{f(O)f(A)} = v + u + \vec{f}(v) - v + \vec{f}(-v) = u,$$

y por tanto $g(A) = t_{-u} \circ f(A) = f(A) - u = A$. En particular se tiene que el conjunto de puntos fijos $\mathcal{F}_g$ de g es

$$A + \mathrm{Ker}(\vec{g} - id) = A + \mathrm{Ker}(\vec{f} - id)$$

y dado que $u \in \mathrm{Ker}(\vec{f} - id)$, el vector u pertenece al espacio de dirección de $\mathcal{F}_g$. Obsérvese también que u no puede ser el vector nulo, en caso contrario tendríamos que $f = t_0 \circ g = g$ y por tanto f tendría puntos fijos, lo cual es una contradicción.

Comprobemos por último que $\mathcal{F}_g$ es una variedad invariante de f, es decir, que dado $A \in \mathcal{F}_g$ se tiene que $f(A) \in \mathcal{F}_g$. En efecto, se tiene que $f(A) = t_u \circ g(A) = t_u(A) = A + u \in \mathcal{F}_g$ ya que u es un vector en el espacio de dirección de $\mathcal{F}_g$. De hecho hemos probado que la aplicación $f|_{\mathcal{F}_g} : \mathcal{F}_g \to \mathcal{F}_g$ es la traslación por medio del vector u. □

Usando este resultado, y gracias a que ya sabemos cómo son los endomorfismos ortogonales de los espacios vectoriales euclídeos $\mathbb{R}^2$ y $\mathbb{R}^3$, podemos clasificar los movimientos de $\mathbb{R}^2$ y $\mathbb{R}^3$ como espacios afines euclídeos. La mejor forma de hacerlo será a través de sus puntos fijos. Recuérdese que en la Proposición de los puntos fijos vimos que si $f : \mathcal{A} \to \mathcal{A}$ es un movimiento y A es un punto fijo, entonces los puntos fijos de f son

$$\mathcal{F}_f = A + \operatorname{Ker}(\vec{f} - I). \qquad (*)$$

Los puntos fijos de f nos darán información de los vectores fijos de $\vec{f}$, el cual ya sabemos que es un endomorfismo ortogonal. De hecho, dado que ya tenemos una clasificación de los endomorfismos orotogonales, seremos capaces de llevar a cabo una clasificación de los movimientos.

Clasificación de los movimientos del plano

A continuación clasificamos los movimientos $f : \mathbb{R}^2 \to \mathbb{R}^2$ en función de su variedad $\mathcal{F}_f$ de puntos fijos:

- Si $\dim(\mathcal{F}_f) = 2$, es decir, si todo $\mathbb{R}^2$ queda fijo por medio de f, entonces se trata claramente de la *identidad*.
- Si $\dim(\mathcal{F}_f) = 1$, es decir, si f tiene una recta de puntos fijos, entonces por $(*)$ deducimos que $\lambda = 1$ es un autovalor de $\vec{f}$ y la dimensión de su subespacio propio $E_1(1) = \operatorname{Ker}(\vec{f} - I)$ es 1. Por tanto $\vec{f}$ debe ser la simetría axial $s_{E_1(1)}$ respecto de $E_1(1)$. De hecho, si tomamos un punto fijo concreto A entonces podemos escribir

  $$f(X) = A + s_{E_1(1)}(\overrightarrow{AX}),$$

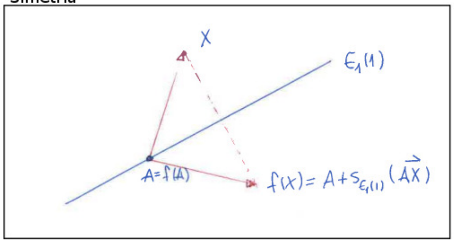
  es decir, que f es la *simetría ortogonal* respecto de la recta afín de puntos fijos $\mathcal{F}_f = A + E_1(1)$, véase el Ejemplo 6.12.
- Si $\dim(\mathcal{F}_f) = 0$, es decir, si f tiene solo un punto fijo $A \in \mathbb{R}^2$, de nuevo por $(*)$ se tiene que $\operatorname{Ker}(\vec{f} - I) = \{0\}$ y por tanto $\vec{f}$ debe ser un giro r de ángulo α (lo cual incluye a la simetría central, es decir, el giro de ángulo π). En particular, si tomamos el punto fijo A de f, entonces observamos que $f(X) = A + r(\overrightarrow{AX})$ es simplemente un *giro* de ángulo α alrededor del punto A.
- Si $\mathcal{F}_f = \emptyset$, es decir, si f no tiene puntos fijos, entonces sabemos por el Teorema de descomposición de movimientos que existe un vector no nulo $u \in \operatorname{Ker}(\vec{f} - id)$ tal que $f = t_u \circ g$ donde g es un movimiento con puntos fijos (recuérdese que en particular $\vec{f} = \vec{t_u} \circ \vec{g} = id \circ \vec{g} = \vec{g}$). Así que, como ya sabemos cómo son los movimientos con puntos fijos, tenemos los siquientes casos:
  i) Si $\dim(\mathcal{F}_g) = \dim(\operatorname{Ker}(\vec{g} - id)) = \dim(\operatorname{Ker}(\vec{f} - id)) = 2$ entonces g es la identidad y por tanto f es la traslación t_u.
  ii) Si $\dim(\mathcal{F}_g) = \dim(\operatorname{Ker}(\vec{g} - id)) = \dim(\operatorname{Ker}(\vec{f} - id)) = 1$ entonces g es la simetría ortogonal respecto de la recta $\mathcal{F}_g$. Como u es un vector en $\operatorname{Ker}(\vec{g} - id)$ deducimos que f es una *simetría deslizante*.

  Obsérvese que el caso $\dim(\mathcal{F}_g) = \dim(\operatorname{Ker}(\vec{g} - id)) = \dim(\operatorname{Ker}(\vec{f} - id)) = 0$ no debemos considerarlo ya que entonces tendríamos que el vector de la traslación sería $u = 0$, lo cual es una contradicción.

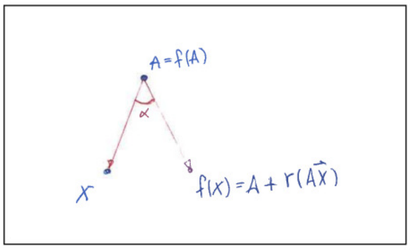
Clasificación de los movimientos del espacio

A continuación clasificamos los movimientos $f : \mathbb{R}^3 \to \mathbb{R}^3$ en función de su variedad $\mathcal{F}_f$ de puntos fijos:

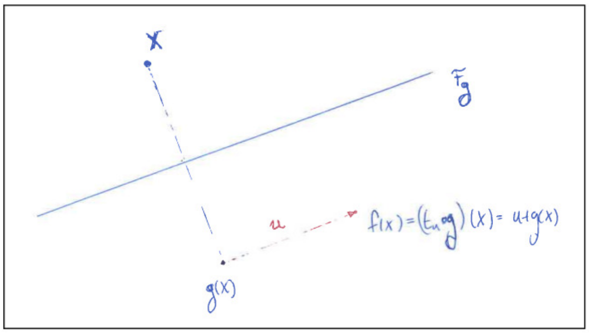
Simetría



Giro



Simetría deslizante



- Si $\dim(\mathcal{F}_f) = 3$, es decir, si todo $\mathbb{R}^3$ queda fijo por medio de f, entonces se trata claramente de la *identidad*.
- Si $\dim(\mathcal{F}_f) = 2$, es decir, si f tiene un plano de puntos fijos, entonces por $(*)$ deducimos que $\lambda = 1$ es un autovalor de $\vec{f}$ y la dimensión de su subespacio propio $E_1(1) = \operatorname{Ker}(\vec{f} - I)$ es 2. Por tanto $\vec{f}$ debe ser la simetría especular $s_{E_1(1)}$ respecto del plano $E_1(1)$. De hecho, si tomamos un punto fijo concreto A entonces podemos escribir
$$f(X) = A + s_{E_1(1)}(\overrightarrow{AX}),$$
es decir, que f es la *simetría ortogonal* respecto del plano afín de puntos fijos $\mathcal{F}_f = A + E_1(1)$, véase el Ejemplo 6.12.
- Si $\dim(\mathcal{F}_f) = 1$, es decir, si f tiene una recta de puntos fijos, entonces por $(*)$ deducimos que $\lambda = 1$ es un autovalor de $\vec{f}$ y la dimensión de su subespacio propio $E_1(1) = \operatorname{Ker}(\vec{f} - I)$ es 1. Por tanto $\vec{f}$ debe ser una rotación axial $r_{E_1(1)}$ de ángulo α alrededor del eje $E_1(1)$. De hecho, si tomamos un punto fijo concreto A entonces podemos escribir
$$f(X) = A + r_{E_1(1)}(\overrightarrow{AX}),$$
es decir, que f es la *rotación axial* de ángulo α respecto de la recta de puntos fijos. Obsérvese que si $\alpha = \pi$ entonces se trata de la simetría axial respecto de la recta de puntos fijos.
- Si $\dim(\mathcal{F}_f) = 0$, es decir, si f tiene un único punto fijo, entonces por $(*)$ deducimos que $\lambda = 1$ es un autovalor de $\vec{f}$ y la dimensión de su subespacio propio $E_1(1) = \operatorname{Ker}(\vec{f} - I)$ es 0. Deducimos entonces que $\vec{f}$ debe ser la composición de una rotación axial $r_{E_1(-1)}$ de ángulo α alrededor del eje $E_1(-1)$ con la simetría especular $s_{E_1(-1)^\perp}$ respecto del plano $E_1(-1)^\perp$. De hecho, si tomamos el único punto fijo A de f entonces podemos escribir
$$f(X) = A + s_{E_1(-1)^\perp} \circ r_{E_1(-1)}(\overrightarrow{AX}),$$

Así pues, f es la composición de la rotación de ángulo α alrededor de la recta $A + E_1(-1)$ con la simetría especular respecto al plano $A + E_1(-1)^{\perp}$.

- Si $\mathcal{F}_f = \emptyset$, es decir, si f no tiene puntos fijos, entonces sabemos por el Teorema de descomposición de movimientos que existe un vector no nulo $u \in \operatorname{Ker}(\vec{f} - id)$ tal que $f = t_u \circ g$ donde g es un movimiento con puntos fijos (recuérdese que en particular $\vec{f} = \vec{t_u} \circ \vec{g} = id \circ \vec{g} = \vec{g}$). Así que, como ya sabemos cómo son los movimientos con puntos fijos, tenemos los siguientes casos:

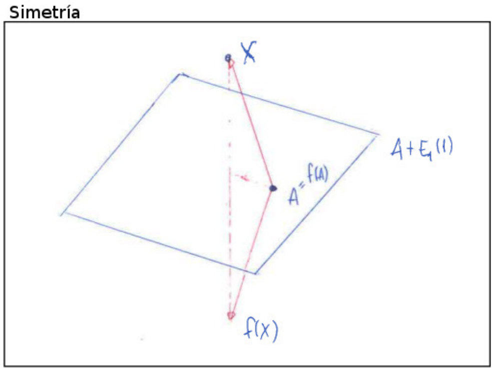
  i) Si $\dim(\mathcal{F}_g) = \dim(\operatorname{Ker}(\vec{g} - id)) = \dim(\operatorname{Ker}(\vec{f} - id)) = 3$ entonces g es la identidad y por tanto f es la traslación t_u.

  ii) Si $\dim(\mathcal{F}_g) = \dim(\operatorname{Ker}(\vec{g} - id)) = \dim(\operatorname{Ker}(\vec{f} - id)) = 2$ entonces g es la simetría ortogonal respecto del plano $\mathcal{F}_g$. Como u es un vector en $\operatorname{Ker}(\vec{g} - id)$ deducimos que f es una *simetría deslizante*.

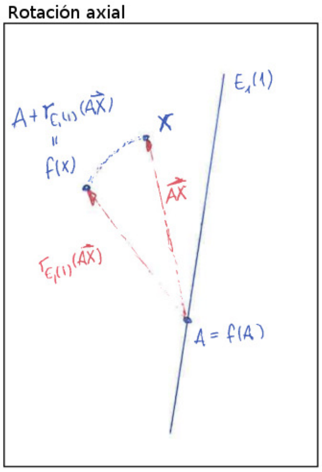
  iii) Si $\dim(\mathcal{F}_g) = \dim(\operatorname{Ker}(\vec{g} - id)) = \dim(\operatorname{Ker}(\vec{f} - id)) = 1$ entonces g es una rotación axial respecto a la recta $\mathcal{F}_g$. Como u es un vector en $\operatorname{Ker}(\vec{g} - id)$ deducimos que f es una *movimiento helicodial*.

  Obsérvese que el caso de que $\dim(\mathcal{F}_g) = \dim(\operatorname{Ker}(\vec{g} - id)) = \dim(\operatorname{Ker}(\vec{f} - id)) = 0$ no debemos considerarlo ya que entonces tendríamos que el vector de la traslación sería $u = 0$, lo cual es una contradicción.

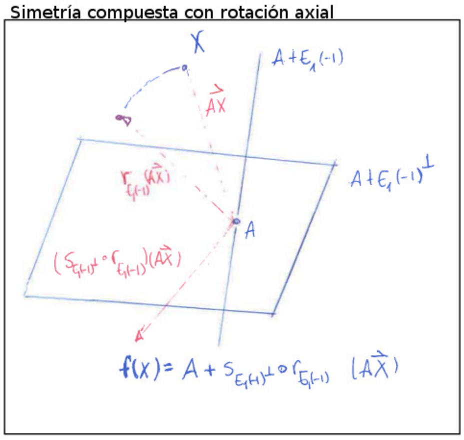
Simetría



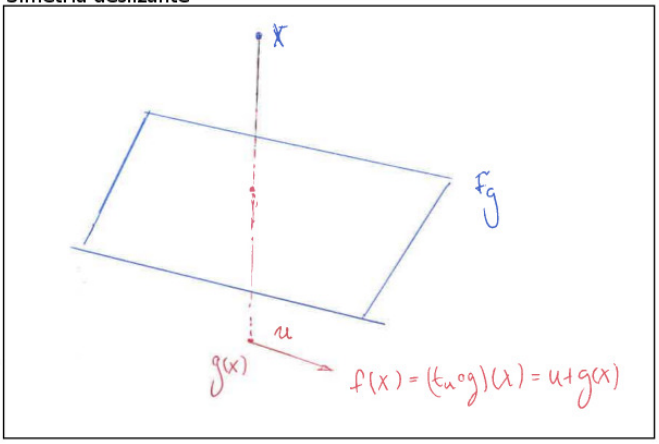
Rotación axial



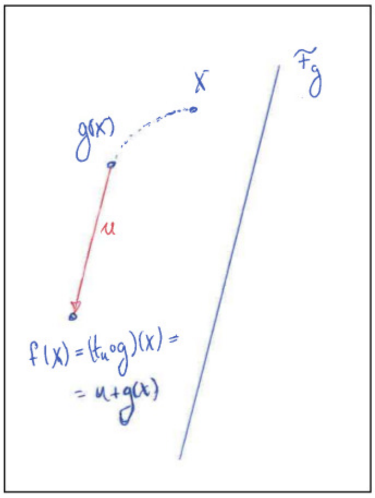
Simetría compuesta con rotación axial



Simetría deslizante



Movimiento helicoidal



Cónicas

Recordemos brevemente las definiciones de las *cónicas no degeneradas*:

Elipse Dados dos puntos F_1 y F_2 de $\mathbb{R}^2$ y un número positivo real a, la *elipse de focos F_1 y F_2 y semieje a* es el lugar geométrico de los puntos X de $\mathbb{R}^2$ que verifican

$$\mathrm{d}(X, F_1) + \mathrm{d}(X, F_2) = 2a.$$

Al punto medio de F_1 y F_2 lo llamamos *centro* de la elipse. A la recta generada por F_1 y F_2, y a la recta perpendicular a ésta que pasa por el centro, las denominamos *ejes*.

En el caso concreto en el que $F_1 = (-d, 0)$ y $F_2 = (d, 0)$ para cierto $d \in \mathbb{R}$ tal que $0 < d < a$, es fácil comprobar (ejercicio) que un punto $X = (x, y) \in \mathbb{R}^2$ pertenece a la elipse si y solo si

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

donde $b = \sqrt{a^2 - d^2}$.

Hipérbola Dados dos puntos F_1 y F_2 de $\mathbb{R}^2$ y un número positivo real a, la *hipérbola de focos F_1 y F_2 y semieje a* es el lugar geométrico de los puntos X de $\mathbb{R}^2$ que verifican

$$|\mathrm{d}(X, F_1) - \mathrm{d}(X, F_2)| = 2a.$$

Al punto medio de F_1 y F_2 lo llamamos *centro* de la elipse. A la recta generada por F_1 y F_2, y a la recta perpendicular a ésta que pasa por el centro, las denominamos *ejes*.

En el caso concreto en el que $F_1 = (-d, 0)$ y $F_2 = (d, 0)$ para cierto $d \in \mathbb{R}$ tal que $0 < d < a$, es fácil comprobar (ejercicio) que un punto $X = (x, y) \in \mathbb{R}^2$ pertenece a la hipérbola si y solo si

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

donde $b = \sqrt{a^2 - d^2}$.

Parábola Dado un punto F de $\mathbb{R}^2$ y una recta r de $\mathbb{R}^2$, la *parábola de foco F y directriz r* es el lugar geométrico de los puntos X de $\mathbb{R}^2$ que satisfacen

$$\mathrm{d}(X, F) = \mathrm{d}(X, r).$$

Al punto de la parábola más cercano a la recta r lo denominamos *vértice*.

En el caso concreto en el que $F = (\frac{p}{2}, 0)$ para cierto $p \in \mathbb{R}$ positivo, y la directriz es la recta de ecuación $x = -\frac{p}{2}$, es fácil comprobar (ejercicio) que un punto $X = (x, y) \in \mathbb{R}^2$ pertenece a la hipérbola si y solo si

$$y^2 = 2px.$$

El nombre de *cónicas* viene de que los lugares geométricos antes descritos se obtienen al intersecar un cono de revolución con un plano. Aunque en realidad, hay otros tres lugares geométricos llamados *cónicas degeneradas* que se obtiene al considerar dichas intersecciones: un punto, una recta, dos rectas que se cortan en un punto, y dos rectas paralelas (véase la hoja de SAGE colgada en el Campus Virtual).

Antes hemos descrito las ecuaciones de las cónicas no degeneradas cuando sus elementos geométricos están "bien colocados" respecto del sistema de referencia canónico. A estas ecuaciones las llamamos las *ecuaciones reducidas*. Pero evidentemente, dichos elementos geométricos no tendrían por qué estar tan bien colocados con respecto del sistema de referencia canónico. Por ejemplo, podriamos considerar una elipse cuyos focos son $F_1 = (2, 3)$ y $F_2 = (-10, 4)$. Sin embargo, es fácil imaginar que respecto de un cierto sistema de referencia rectangular $\mathcal{R}'$ de $\mathbb{R}^2$ los elementos geométricos sí están bien colocados y por tanto la ecuación respecto de $\mathcal{R}'$ es reducida. En el ejemplo de la elipse, podemos imaginar la recta que une los focos F_1 y F_2 y el punto medio de ambos en dicha recta; acto seguido podemos centrar nuestra mirada en dicho punto medio y girar nuestra cabeza para que tenga la inclinación de la recta que une los focos. Es decir, que lo que hemos hecho es centrar nuestra mirada en el centro de la elipse y hemos inclinado la cabeza para alinearlas con los ejes de la elipse. Si hacemos esto obviamente ahora sí nos parecerá que la elipse está "bien colocada". Es decir, que podemos encontrar un sistema de referencia rectangular $\mathcal{R}'$ tal que la elipse en cooordenadas $(x', y')_{\mathcal{R}'}$ respecto de $\mathcal{R}'$ tiene ecuación:

$$\frac{[x']^2}{a^2} + \frac{[y']^2}{b^2} = 1.$$

O dicho de otra forma, una cónica "mal colocada" la obtenemos girando y trasladando una cónica "bien colocada". ¿Cómo será la ecuación de esta cónica "mal colocada" en coordenadas respecto de $\mathcal{R}_c$? Las ecuaciones de paso de $\mathcal{R}_c$ a $\mathcal{R}'$ serán de la forma

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \alpha_2 & \alpha_3 \\ \beta_2 & \beta_3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

y por tanto la ecuación de la cónica usando las coordenadas respecto de $\mathcal{R}$ será

$$\frac{(\alpha_1 x + \alpha_2 y + \alpha_3)^2}{a^2} + \frac{(\beta_1 x + \beta_2 y + \beta_3)^2}{b^2} = 1.$$

Si desarrollamos la ecuación anterior, seguro que nos quedará algo de la forma:

$$a_{11}x^2 + a_{22}y^2 + 2a_{21}xy + a_{01}x + a_{02}y + a_{00} = 0. \qquad (*)$$

para ciertos $a_{11}, a_{22}, a_{21}, a_{01}, a_{02}, a_{00} \in \mathbb{R}$.

Así que es natural preguntarse lo siguiente: si nos dan una ecuación del tipo $(*)$, ¿cómo podemos saber de qué cónica se trata? ¿cómo podemos calcular un sistema de referencia respecto del cual la cónica tiene una ecuación reducida?

Dada una ecuación del tipo $(*)$, la podemos escribir de la forma

$$X^t AX + BX + a_{00} = 0$$

donde

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \qquad B = \begin{pmatrix} a_{01} & a_{02} \end{pmatrix} \qquad X = \begin{pmatrix} x \\ y \end{pmatrix}$$

Nuestro objetivo es encontrar un sistema de referencia rectangular $\mathcal{R}'$ de forma que cuando escribimos la ecuación anterior respecto de las coordenadas respecto de $\mathcal{R}'$ nos quede reducida. Para ello debemos seguir, a grandes rasgos, dos pasos:

1) calcular (usando el teorema espectral) una matriz P ortogonal tal que $P^{-1}AP$ es una matriz diagonal y $\det(P) = 1$, y

2) completar cuadrados (teniendo cuidado de que en el cambio de coordenadas en este paso no se multipliquen las variables por coeficientes, tan solo se sumen números).

En el primer paso lo que estamos haciendo a fin de cuentas es aplicar un giro, y en el segundo paso una traslación. Lo mostramos con un par de ejemplos.

Ejemplo 6.16 Consideremos la ecuación

$$x^2 + 6x + 5y + 14 = 0$$

la cual podemos escribir como

$$X^t AX + BX + 14 = 0$$

donde

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 6 & 5 \end{pmatrix} \qquad X = \begin{pmatrix} x \\ y \end{pmatrix}$$

La matriz A ya es diagonal, así que tan solo queda completar cuadrados:

$$0 \quad = \quad x^2 + 6x + 5y + 14 = (x+3)^2 - 9 + 5y + 14 = (x+3)^2 + 5y + 5 = (x+3)^2 + 5(y+1)$$

Así que si hacemos la sustitución

$$\left\{ \begin{array}{l} x' = x + 3 \\ y' = y + 1 \end{array} \right.$$

entonces la ecuación quedaría

$$[x']^2 + 5y' = 0$$

es decir, es una parábola. O dicho de otra manera, la ecuación con la que empezamos, respecto del sistema de referencia $\mathcal{R}' = \{(-3,-1); (1,0), (0,1)\}$ es $[x']^2 + 5y' = 0$.

Ejemplo 6.17 Consideremos la ecuación

$$2x^2 - 4xy - y^2 - 4x - 8y + 14 = 0$$

la cual podemos escribir como

$$X^t AX + BX + 14 = 0$$

donde

$$A = \begin{pmatrix} 2 & -2 \\ -2 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} -4 & -8 \end{pmatrix} \qquad X = \begin{pmatrix} x \\ y \end{pmatrix}$$

Haciendo los cálculos necesarios, podemos diagonalizar la matriz A de forma que

$$\begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}^{-1} \begin{pmatrix} 2 & -2 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix}$$

En particular, si hacemos el cambio de variable

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

entonces la ecuación quedará de la forma

$$[X']^t P^t A P X' + B P X' + 14 = 0$$

es decir,

$$-2[x']^2 + 3[y']^2 - \frac{20}{\sqrt{5}} x' + 14 = 0.$$

Ahora debemos completar cuadrados, pero primero, debemos sacar factor común para asegurarnos que el cambio que nos quede después de haber completado cuadrados corresponda a una traslación:

$$-2\big([x']^2 + \frac{10}{\sqrt{5}} x'\big) + 3[y']^2 + 14 = 0.$$

Completando cuadrados queda

$$-2[x' + \frac{5}{\sqrt{5}}]^2 + 3[y']^2 + 24 = 0.$$

Así que si hacemos la sustitución

$$\begin{cases} x'' = x' + \frac{5}{\sqrt{5}} \\ y'' = y' \end{cases}$$

la ecuación queda

$$\frac{[x'']^2}{12} - \frac{[y'']^2}{4} = 1$$

es decir, es una hipérbola. Recapitulando, el cambio de variable que hemos realizado es

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} \frac{5}{\sqrt{5}} \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix} \quad \leftrightsquigarrow \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \end{pmatrix} + \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} x'' \\ y'' \end{pmatrix}$$

Es decir, que la hipérbola tiene ecuación reducida respecto del sistema de referencia rectangular:

$$\mathcal{R}'' = \{(-1,-2); (\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}), (-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}})\}$$