PROGRAMA DE GEOMETRÍA AFÍN Y PROYECTIVA. MÓDULO TEÓRICO GRUPO B, CURSO 1999-2000

I. Geometría afín.

Concepto de espacio afín, subespacio afín y paralelismo.
Sistemas de referencia cartesianos.
Paralelismo, intersección e incidencia en coordenadas.
Aplicaciones afines y afinidades.

II. Geometría proyectiva

El espacio proyectivo y los subespacios proyectivos.
Relaciones de intersección e incidencia entre algunos subespacios proyectivos.
Coordenadas homogéneas en un espacio proyectivo.
Proyectividades.

III. Geometría afín métrica

Espacios vectoriales euclídeos
Definición de espacio afín euclídeo. Ángulos y distancias.
Isometrías entre espacios afines euclídeos.
Descomposición de una transformación ortogonal de un espacio vectorial euclídeo.
Descomposición de una isometría de un espacio afín euclídeo.

IV. Cónicas y cuádricas

Definición y ecuación de una hipercuádrica en el espacio afín euclídeo. Reducción y clasificación.

Particularización a las cónicas (espacio afín euclídeo de dimensión 2) y cuádricas (espacio afín euclídeo de dimensión 3).

GEOMETRÍA AFÍN

§1. Concepto de espacio afín, subespacio afín y paralelismo.

Para fijar ideas, en estos apuntes solo usaremos el cuerpo $\mathbb{R}$ de los números reales, aunque los enunciados y teoremas son válidos para cuerpos más generales (en particular para el cuerpo $\mathbb{C}$ de los números complejos). Los dibujos solo son ciertos para $\mathbb{R}$.

V denotará un espacio vectorial de dimensión finita sobre $\mathbb{R}$.

1.1.Definición. *Un conjunto E se dice que es un espacio afín asociado a un espacio vectorial V si existe una ley de composición externa $+$ de $E \times V$ en E*

$$E \times V \longrightarrow E \,/\, (P, v) \mapsto P + v,$$

verificando los axiomas

1.1.1. Para todo $(P, Q) \in E \times E$ existe un único $v \in V$ tal que $P + v = Q$. Se escribirá $v = \overrightarrow{PQ}$.

1.1.2. para cualesquiera $P \in E$ y $u, v \in V$ se verifica que

$$(P + u) + v = P + (u + v).$$

A los elementos de E se les llama puntos. De V se dice que es el espacio de direcciones de E, o también el espacio (vectorial) director de E.

1.2.Proposición. *Si $P, Q, R, P', Q' \in E$ y $v \in V$, entonces*

a) $P + v = P$ sii $v = 0$, siendo P un punto arbitrario de E.

b) $\overrightarrow{PQ} = -\overrightarrow{QP}$. (ver dibujo)

c) $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$. (ver dibujo)

d) $\overrightarrow{PQ} = \overrightarrow{P'Q'}$ implica que $\overrightarrow{PP'} = \overrightarrow{QQ'}$. (ver dibujo)

Demostración. a) De 1.1.1 resulta que, dado P, existe un único v tal que $P + v = P$. Usando esto y 1.1.2, tenemos $P + 0 = (P + v) + 0 = P + (v + 0) = P + v = P$, y, como el v es único, $v = 0$.

b) Por definición de $\overrightarrow{PQ}$, se tiene

$$P = Q + \overrightarrow{QP} = (P + \overrightarrow{PQ}) + \overrightarrow{QP} = P + (\overrightarrow{PQ} + \overrightarrow{QP}),$$

Typeset by $\mathcal{AMS}$-TEX

que, por a), implica $\overrightarrow{PQ}+\overrightarrow{QP}=0$, c.q.d.

c) Calculando como antes,

$$P+(\overrightarrow{PR})=R=Q+\overrightarrow{QR}=(P+\overrightarrow{PQ})+\overrightarrow{QR}=P+(\overrightarrow{PQ}+\overrightarrow{QR}),$$

y, por 1.1.1, $\overrightarrow{PR}=\overrightarrow{PQ}+\overrightarrow{QR}$, c.q.d.

d) Usando c), se tiene que

$$\overrightarrow{PQ'}=\overrightarrow{PQ}+\overrightarrow{QQ'} \quad \text{y} \quad \overrightarrow{PQ'}=\overrightarrow{PP'}+\overrightarrow{P'Q'}$$

y, usando la hipótesis, el resultado buscado sale de comparar estas dos igualdades. $\square$

1.2'.Ejercicios.

Comprobar o demostrar (y hacer dibujos siempre que sea posible) que

a) $\overrightarrow{P(P+v)}=v$ b) $P+\overrightarrow{PQ}=Q$

c) $\overrightarrow{P(Q+v)}=\overrightarrow{PQ}+v$ d) $\overrightarrow{(P+v)Q}=\overrightarrow{PQ}-v$

e) $\overrightarrow{(P+v)(Q+v)}=\overrightarrow{PQ}$ f) $P+v=Q$ sii $P=Q+(-v)$

g) $\overrightarrow{(P+v)(Q+w)}=-v+\overrightarrow{PQ}+w$ h) $R+\overrightarrow{PQ}=Q+\overrightarrow{PR}=P+(\overrightarrow{PQ}+\overrightarrow{PR})$.

1.2".Nota. Obsérvese que si E es un espacio afín asociado a un espacio vectorial V y $P\in E$, entonces $V=\{\overrightarrow{PQ}/Q\in E\}$, porque $\{\overrightarrow{PQ}/Q\in E\}\subset V$ por definición de $\overrightarrow{PQ}$ y, para todo $v\in V$, se tiene que $Q=P+v\in E$, luego $v=\overrightarrow{PQ}\in\{\overrightarrow{PQ}/Q\in E\}$, luego $V\subset\{\overrightarrow{PQ}/Q\in E\}$.

1.3.Ejemplo. *Si definimos $E=V$ y la operación $+$ como la suma de V, V es un espacio afín con espacio de direcciones V, y $\overrightarrow{vw}=w-v$.*

Recíprocamente, fijado un punto $P\in E$, el conjunto $T_PE=\{(P,Q)/Q\in E\}$ puede dotarse de la estructura de espacio vectorial que hace que la biyección

$$\varphi: T_PE\longrightarrow V/ \quad (P,Q)\mapsto\overrightarrow{PQ}$$

sea un isomorfismo de espacios vectoriales (es decir, la suma en T_pE se define por $(P,Q)+(P,R)=\varphi^{-1}(\overrightarrow{PQ}+\overrightarrow{PR})$ y el producto por un escalar $\lambda\in I\!R$ por $\lambda(P,Q)=\varphi^{-1}(\lambda\overrightarrow{PQ})$. (Que la aplicación anterior sea biyectiva es una consecuencia de 1.1.1). Resulta entonces que un espacio afín se puede considerar como un espacio vectorial en cuanto se fija un punto. Es decir, en cierto sentido, el espacio afín y el espacio vectorial son dos maneras distintas de mirar una misma cosa.

El conjunto T_PE, con la estructura de espacio vectorial indicada antes, se llama a veces *espacio vectorial tangente a E en P.*

1.4.Definición. *La dimensión del espacio afín E es la dimensión de su espacio de direcciones V.*

1.5.Definición. *Dados $P\in E$ y W s.e.v. de V, llamaremos subespacio afín (o variedad lineal afín) de E determinado por el punto P y el s.e.v. W (o pasando por P y con espacio director W) al subconjunto*

$$P+W:=\{P+w/w\in W\}=\{Q\in E/\overrightarrow{PQ}\in W\}.$$

Obviamente $P\in P+W$, pues $P=P+0$.

Si W es de dimensión 1, 2 ó $dim(E)-1$, respectivamente, se dice que $P+W$ es una recta, un plano o un hiperplano afín respectivamente.

Si $Q\in F=P+W$, se dice que el subespacio afín F pasa por el punto Q.

***1.6.Proposición.** *a) Si $F=P+W$ es un subespacio afín de E, entonces $W=\{\overrightarrow{PR}/R\in F\}=\mathcal{W}$, donde $\mathcal{W}=\{\overrightarrow{QR}/Q,R\in F\}$.*

b)$P+U=Q+W$ sii $U=W$ y $(P+U)\cap(Q+W)\neq\emptyset$

c) Dados $P,Q\in E$, y U, W s.e.v. de V, se tiene que $P+U=Q+W$ sii $U=W$ y $Q\in P+U$ ó $P\in Q+W$.

Demostración. a) Si $w\in W$, por la definición de $P+W$, $Q=P+w\in P+W$, luego $w=\overrightarrow{PQ}$, luego $W\subset\mathcal{W}$. Recíprocamente, si $Q,R\in F$, existen $w,u\in W$ tales que $Q=P+w$ y $R=P+u$, de donde, usando 1.1.2, $R+(-u)+w=P+w=Q$, luego $\overrightarrow{QR}=-w+u\in W$, de donde $\mathcal{W}\subset W$.

b) Es evidente que si $P+U=Q+W=F$, entonces, $(P+U)\cap(Q+W)\neq\emptyset$. Además, de a) se deduce que $U=\{\overrightarrow{QR}/Q,R\in F\}=W$. Recíprocamente, si se dan estas dos condiciones, sea $S\in(P+U)\cap(Q+W)$, se tiene que existe un $u\in U$ tal que $S=P+u$; además, para todo $w\in U$, $S+w=P+(u+w)\in P+U$, luego $S+U\subset P+U$. Por otro lado, $P=S+(-u)$ implica que para todo $P+w\in P+U$, $P+w=S+(-u+w)\in S+U$, luego $P+U\subset S+U$. Tenemos, por lo tanto, que $P+U=S+U$. De la misma manera se demuestra que $Q+U=S+U$, lo que acaba la demostración de b).

c) es consecuencia inmediata de b). $\square$

1.7.Corolario. *$P+W$ es un espacio afín con W como espacio de direcciones, con la ley de composición sobre $(P+W)\times W$ heredada de la ley de composición sobre $E\times V$.*

Demostración. Se verifican:

1.1.1 Para todo $P+u, P+w\in P+W$ existe un único $w-u\in W$ tal que $P+w=P+u+(w-u)$.

1.1.2 Para cualesquiera $P+u\in P+W$ y $w,t\in W$, se verifica

$$((P+u)+w)+t=P+(u+w+t)=(P+u)+(w+t).$$

□

1.8.Proposición. *Un subconjunto F de E es un subespacio afín sii es un espacio afín cuyo espacio de direcciones W es un subespacio vectorial de V y la ley externa $+$ es la restricción a $F\times W$ de la ley externa $+$ definida sobre $E\times V$. En este caso, para cualquier $P\in F$, $F=P+W$, i.e., F es el subespacio afín determinado por P y W.*

Demostración. Resulta de las dos últimas proposiciones que si F es un subespacio afín de E, entonces F verifica las propiedades del enunciado. Veamos el recíproco. De la propiedad 1.1.1 de la definición de espacio afín resulta que para cualquier $P\in F$, $F=P+W$, y esto prueba el recíproco. □

***1.9.Proposición.** *Sea $\{F_i\}_{i\in I}$ una familia de subespacios afines de E. Si la intersección de los F_i es no vacía, entonces es un subespacio afín cuyo espacio de direcciones es la intersección de los espacios de direcciones de cada F_i.*

Demostración. Para cada $i\in I$, sea V_i el espacio de direcciones de F_i. Sea $P\in\cap\{F_i/i\in I\}$. De 1.6 b) resulta que $F_i=P+V_i$. Se tiene que $Q\in\cap\{F_i/i\in I\}$ sii $Q\in P+V_i$ para todo $i\in I$ sii $\overrightarrow{PQ}\in V_i$ para todo $i\in I$ sii $\overrightarrow{PQ}\in\cap\{V_i/i\in I\}$ sii $Q\in P+(\cap\{V_i/i\in I\})$, de donde se tiene que $\cap\{F_i/i\in I\}=P+(\cap\{V_i/i\in I\})$, lo que acaba la demostración. □

***1.10.Definición.** *Si C es un subconjunto del espacio afín E, se llama subespacio afín $A(C)$ engendrado por C al menor (respecto de la inclusión) de los subespacios afines de E que contiene a C. Es decir, $A(C)$ es el subspacio afín de E tal que $C\subset A(C)$ y todo subespacio afín F de E tal que $C\subset F$ verifica $A(C)\subset F$.*

Esto es equivalente a decir que $A(C)=\cap_{i\in I}F_i$, siendo $I=\{i/F_i$es un subespacio afín de E y $C\subset F_i\}$. En efecto: Por 1.9 se tiene que $\cap_{i\in I}F_i$ es un subespacio afín que contiene a C por verificarse que $F_i\supset C$ para todo $i\in I$, luego $A(C)\subset\cap_{i\in I}F_i$. Como, además, $A(C)$ es un subespacio afín que contiene a C, existe un $j\in I$ tal que $A(C)=F_j$, luego $\cap_{i\in I}F_i\subset F_j=A(C)$. De estas dos inclusiones resulta la igualdad.

***1.11.Proposición.** *Sea $P\in C$, el subespacio afín $A(C)$ engendrado por un subconjunto C de E tiene como espacio de direcciones el subespacio vectorial de V engendrado por el conjunto $\{\overrightarrow{PQ}/\,Q\in C\}$.*

Demostración. Sea W el subespacio vectorial engendrado por $\{\overrightarrow{PQ}/\,Q\in C\}$. Como $P+W$ es un subespacio afín conteniendo a C, se tiene que $A(C)\subset P+W$. Para probar la inclusión contraria hay que probar que para todo subespacio afín $P+U$ de E conteniendo a C se tiene que $P+W\subset P+U$. Ahora bien, como $C\subset P+U$, se tiene que para todo $Q\in C$, $\overrightarrow{PQ}\in U$, luego $\{\overrightarrow{PQ}/\, Q\in C\}\subset U$, por lo tanto $W\subset U$, de donde se deduce $P+W\subset P+U$ $\square$

1.12.Corolario. *Dado un subconjunto $C\subset E$ y dados $P,Q\in C$, los espacios vectoriales generados por $\{\overrightarrow{PR}, R\in C\}$ y por $\{\overrightarrow{QR}, R\in C\}$ son iguales.*

Demostración. De la proposición anterior se deduce que $A(C)=P+<\{\overrightarrow{PR}, R\in C\}>=Q+<\{\overrightarrow{QR}, R\in C\}>$ y esto implica (por 6.1.(b)) que $<\{\overrightarrow{PR}, R\in C\}>=<\{\overrightarrow{QR}, R\in C\}>$.

1.13.Definición. *$h+1$ puntos $\{P_0,\, P_1,\, ...\,,\, P_h\}$ de un espacio afín E se dice que son independientes si el subespacio afín engendrado por ellos tiene dimensión h.*

1.14.Proposición. *$h+1$ puntos $\{P_0,\, P_1,\, ...\,,\, P_h\}$ de un espacio afín E son independientes sii el conjunto de vectores $\{\overrightarrow{P_0P_1},\, \overrightarrow{P_0P_2},\, ...,\, \overrightarrow{P_0P_h}\}$ es linealmente independiente.*

Demostración. De 1.11 se deduce que la dimensión del espacio afín generado por $\{P_0,\, P_1,\, ...\,,\, P_h\}$ es la dimensón del espacio vectorial generado por $\{\overrightarrow{P_0P_1},\, \overrightarrow{P_0P_2},\, ...,\, \overrightarrow{P_0P_h}\}$, y la dimensión de este espacio es h sii el conjunto de esos vectores es linealmente independiente. $\square$

***1.15.Definición.** *Dados M subespacios afines $F_1,\, ...\,,\, F_m$ de E, de intersección no vacía, se llama suma ($\sum_{i=1}^m F_i=F_1+\,...\,+F_m$) de estos subespacios al subespacio afín $A(\cup_{i=1}^m F_i)$*

***1.16.Proposición.** *Dados los subespacios afines $F_1=P_1+U_1,\, ...\,,\, F_m=P_m+U_m$ y dado $P\in\cap_{i=1}^m F_i$, se tiene que $\sum_{i=1}^m F_i=P+\sum_{i=1}^m U_i$*

Demostración. De la hipótesis $\cap_{i=1}^m F_i\neq\emptyset$ se deduce que existe $P\in\cap_{i=1}^m F_i$. De 1.11 resulta que el espacio director de $\sum_{i=1}^m F_i$ es el espacio vectorial engendrado por los vectores de la forma $\{\overrightarrow{PQ}/Q\in P+U_i,\, 1\le i\le m\}=\{v\in U_i,\, 1\le i\le m\}$ que es, por definición, el espacio $\sum_{i=1}^m U_i$, de donde se tiene lo que queríamos probar. $\square$

***1.17.Corolario.** *Si F_1, F_2 son dos subespacios afines de intersección no vacía, entonces*

$$\dim(F_1)+\dim(F_2)=\dim(F_1+F_2)+\dim(F_1\cap F_2).$$

Demostración. Sea $P\in F_1\cap F_2$ y escribamos $F_i=P+W_i$, $i=1,2$. De 1.16 resulta que $F_1+F_2=P+(W_1+W_2)$. De la definición de dimensión de un espacio afín y de la fórmula para las dimensiones en espacios vectoriales resulta que

$$\dim(F_1+F_2)=\dim(W_1+W_2)=\dim(W_1)+\dim(W_2)-\dim(W_1\cap W_2),$$

de donde se deduce la fórmula del enunciado, porque $F_1\cap F_2=P+(W_1\cap W_2)$. $\square$

1.18.Definición. *Dos subespacios afines $F = P + W$ y $G = Q + U$ de E se dice que son paralelos ($F||G$) sii $U \subset W$ ó $W \subset U$.*

Evidentemente, si $\dim(F) = \dim(G)$*, se tiene que $F||G$ sii $U = W$, porque dos espacios vectoriales de la misma dimensión tales que uno está contenido en el otro, son iguales.*

1.19.Proposición. *Si$F||G$, entonces $F \cap G = \emptyset$ ó $F \subset G$ ó $G \subset F$.*

Demostración. Si existe un $R \in F \cap G$, se puede escribir $F = R + W$ y $G = R + U$, de donde se deduce que $G \subset F$ o $F \subset G$. $\square$

1.20.Proposición. *Si F es un suabespacio afín de E y $Q \in E - F$, existe un único subespacio afín G de E de la misma dimensión que F tal que $Q \in G$ (G pasa por Q) y $F||G$. (i.e. el quinto postulado de Euclides siempre es cierto en Geometría Afín).*

Demostración. Si $F = P+W$ y $G = Q+U$, $F||G$ implica que $U \subset W$ ó $W \subset U$, y, como $\dim(U) = \dim(W)$, se tiene que $U = W$, luego $G = Q + W$, que está unívocamente determinado si se dan $F = P + U$ y Q. $\square$

1.21. Preguntas y ejercicios.

(a) Sea F un subespacio afín de E, U el espacio vectorial director de F y V el espacio vectorial asociado a E, y P un punto de F. Si $u \in U$, ¿qué es $P + u$ y donde está?. Si $v \in V - U$, ¿qué es $P + v$ y donde está?.

(b) Sea E un espacio afín de dimensión 4. Sean F y G subespacios afines de E de dimensión 2 y de intersección no vacía. ¿Cuáles son las posibles dimensiones de $F+G$?.

(c) Sea E un espacio afín con espacio vectorial director V, F un subespacio afín de E, $F = P + U$, siendo U un subespacio vectorial de V. Dados Q, $R \in F$, ¿donde está el vector que va de Q a R?.

(d) Sea $F = P + U$ un subespacio afín de E, $W = \{\overrightarrow{QR} \in V/\, Q, R \in F\}$. ¿Qué relación hay entre U y W?.

(e) Si F y G son subespacios afines de E, $\dim E = n$, $\dim F = 4$,$\dim G = m$, $n > m \geq 4$. ¿Cuáles son las posibles dimensiones de $F \cap G$?. ¿Es $F \cap G$ un subespacio afín?.

(f) Si $C \subset B \subset E$ y E es un espacio afín, ¿qué relación hay entre los espacios afines $A(C)$ y $A(B)$ generados por C y B respectivamente?. Si $\{P_0, ..., P_r\} \subset C$ y $\dim A(C) = 2$, ¿cuál es el valor máximo que puede tener r para que los puntos $P_0, ..., P_r$ sean independientes?.

(g) Si F, G son subespacios afines de E, $\dim F = n - 2$, $\dim G = n - 1$ y $\dim E = n$, ¿es posible que F y G sean paralelos?.

(h) Sea E_n un espacio afín de dimensión n. Sean F_m y G_k subespacios afines de dimensiones m y k respectivamente. En los siguientes casos, indicar todas las posibles posiciones relativas (son paralelos, se cortan en un subespacio afín de dimensión que hay que precisar, se cruzan -i.e. no se cortan ni son paralelos- conteniendo otrso subespacios afines paralelos de dimension a precisar, ...)

h1) Dos rectas F_1 y G_1 en E_2 c2) Dos rectas F_1 y G_1 en E_n con $n \geq 3$

h3) Una recta F_1 y un plano G_2 en E_3

h4) Una recta F_1 y un hiperplano G_{n-1} en E_n

h5) Dos planos F_2 y G_2 en E_3. h6) Dos planos F_2 y G_2 en E_4.

h7) Dos hiperplanos F_{n-1} y G_{n-1} en E_n

h8) F_2 y G_3 en E_5. h9)F_2 y G_3 en E_4.

h10) F_m y G_k en E_n, con $1 \leq m, k \leq n-1$ y $n = 5$.

§2. Sistemas de referencia cartesianos

2.1.Definición. *Si E es un espacio afín de dimensión n asociado a un espacio vectorial V, se llama referencia cartesiana del espacio E a un par $\{O, \{e_1, ..., e_n\}\}$ en el que $O \in E$ y $\{e_1, ...e_n\}$ es una base de V.*

Se llama i-ésimo eje coordenado de la referencia a la recta afín pasando por O y con espacio director el generado por e_i (se dice que es la recta pasando por O y con vector director e_i)

2.2.Definición. *Dada una referencia $\{O, e_1, ..., e_n\}$ de E, se llama coordenadas de un punto $P \in E$ a las componentes del vector $\overrightarrow{OP}$ de V en la base $\{e_1, ..., e_n\}$. Si $\overrightarrow{OP} = \lambda_1 e_1 + ... + \lambda_n e_n$, se escribirá $P = (\lambda_1, ..., \lambda_n)$.*

2.3. Proposición. *Si, en una referencia $\{O, e_1, ..., e_n\}$ de E, se tiene $P = (P^1, ..., P^n)$ y si $v = v^1 e_1 + ... + v^n e_n \in V$, entonces $P + v = (P^1 + v^1, ..., P^n + v^n)$. Si $Q = (Q^1, ..., Q^n)$, entonces $\overrightarrow{PQ} = (Q^1 - P^1)e_1 + ... + (Q^n - P^n)e_n$.*

Demostración. Las coordenadas de $P + v$ son, por definición, las componentes de $\overrightarrow{O(P+v)}$, pero $\overrightarrow{O(P+v)} = \overrightarrow{OP} + \overrightarrow{P(P+v)} = \overrightarrow{OP} + v$, de donde se tiene la primera parte de la proposición. La segunda se deduce de la primera tomando $v = \overrightarrow{PQ}$, pues se tiene entonces que $(Q^1, ..., Q^n) = P + \overrightarrow{PQ} = (P^1 + v^1, ..., P^n + v^n)$, y, de aquí, $v^i = Q^i - P^i$ para $1 \leq i \leq n$. $\square$

Vamos a ver ahora las ecuaciones de los puntos de un subespacio afín usando un sistema de coordenadas cartesiano

2.4.Proposición-Definición. *Sea $F = P + W$ un subespacio afín, $\{O, e_1, ..., e_n\}$ una referencia cartesiana de E, y $\{w_1, ..., w_k\}$ una base de W. Si $w_j = \sum_{i=1}^n w_j^i e_i$, $1 \leq j \leq k$, y $P = (P^1, ..., P^n)$, entonces F es el conjunto de puntos Q cuyas coordenadas $(Q^1, ..., Q^n)$ verifican*

$$Q^i = P^i + \sum_{j=1}^{k} \lambda^j w_j^i, \text{ para todo } \lambda^1, ..., \lambda^k \in I\!R, 1 \leq i \leq n, \tag{2.4.1}$$

o, en forma matricial,

$$\begin{pmatrix} Q^1 \\ \vdots \\ Q^n \end{pmatrix} = \begin{pmatrix} P^1 \\ \vdots \\ P^n \end{pmatrix} + \begin{pmatrix} w_1^1 & \dots & w_k^1 \\ \vdots & \ddots & \vdots \\ w_1^n & \dots & w_k^n \end{pmatrix} \begin{pmatrix} \lambda^1 \\ \vdots \\ \lambda^k \end{pmatrix}.$$

Las ecuaciones (2.4.1) se llaman ecuaciones paramétricas del subespacio afín F.

2.4'.Ejercicio. *Si (2.4.1) son las ecuaciones en paramétricas de un subespacio afín F de E en una referencia cartesiana $\{O, e_1, ..., e_n\}$ de E, dar las componentes en la base $\{e_1, ..., e_n\}$ de V de una base del espacio vectorial director de F.*

2.5. Cuando $\dim(F) = 1$, i.e., se trata de una recta, entonces (2.4.1) se transforma en $Q^i = P^i + \lambda w^i$, de donde resulta, despejando λ,

(2.5.1) $$\frac{Q^1 - P^1}{w^1} = ... = \frac{Q^n - P^n}{w^n}.$$

que son las ecuaciones en *forma contínua o implícita de una recta afín.*

Todo subespacio afín puede describirse, como acabamos de ver con las rectas, como el conjunto de puntos que verifican un sistema de ecuaciones lineales. Es lo que veremos en la proposición 2.7, para cuya demostración usaremos los siguientes resultados de álgebra lineal.

2.6^0. Lema de álgebra. *Sea V un espacio vectorial de dimensión n, V^* su espacio dual (i.e. el espacio vectorial de las aplicaciones lineales de V en $I\!R$), $\{e_1, ..., e_n\}$ una base de V y $\{e^{*1}, ..., e^{*n}\}$ la base de V^* dual de la de V (i.e., la base de V^* tal que $e^{*i}(e_j) = \delta_i^j$ para todo $i, j \in \{1, ..., n\}$).*

(a) Si $f \in V^*$ y $f = \sum_{i=1}^n f_i e^{*i}$, entonces $f_i = f(e_i)$.

(b) Si $x \in V$ y $x = \sum_{i=1}^n x^i e_i$, entonces $x^i = e^{*i}(x)$.

(c) $< \{e_1, ..., e_k\} >= \mathrm{Ker}\, e^{*k+1} \cap ... \cap \mathrm{Ker}\, e^{*n}$.

Demostración. : (a) $f(e_i) = \sum_{j=1}^n f_j e^{*j}(e_i) = \sum_{j=1}^n f_j \delta_i^j = f_i$.

(b) $e^{*i}(x) = e^{*i}(\sum_{j=1}^n x^j e_j) = \sum_{j=1}^n x^j e^{*i}(e_j) = \sum_{j=1}^n x^j \delta_j^i = x^i$.

(c) $< \{e_1, ..., e_k\} > = \{x^1 e_1 + ... + x^k e_k + 0\, e_{k+1} + ... + 0\, e_n / \, x^i \in I\!R, 1 \le i \le k\} = \{\sum_{i=1}^n x^i e_i / \, x^j = 0$ para $k+1 \le j \le n$ y $x^i \in I\!R\} = \{x \in V / \, e^{*j}(x) = 0$ para $k+1 \le j \le n\} = \mathrm{Ker}\, e^{*k+1} \cap ... \cap \mathrm{Ker}\, e^{*n}$. □

2.6. Lema. *W es un subespacio vectorial de V de dimensión k, si y solo si existen $n-k$ aplicaciones lineales (no únicas) $f^j : V \longrightarrow I\!R$ linealmente independientes (como elementos de V^*) tales que $W = \bigcap_{j=k+1}^n \mathrm{Ker}\, f^j$*

Demostración. Suponamos que W es un subespacio vectroial de dimensión k. Si $\{w_1, ..., w_k\}$ es una base de W, complétese hasta una base $\{w_1, ..., w_k, v_{k+1}, ..., v_n\}$ de V, con $v_j \in V - W$, $j = k+1, ..., n$. Sea $\{\phi^1, ..., \phi^k, f^{k+1}, ..., f^n\}$ la base de V^* dual de la anterior.

Por 2.6^0, todo $x \in V$ se puede escribir como $x = \sum_{i=1}^k \phi^i(x)w_i + \sum_{j=k+1}^n f^j(x)v_j$ y $x \in W$ sii $f^j(x) = 0$ para $k+1 \leq j \leq n$, i.e. sii $x \in \bigcap_{j=k+1}^n \operatorname{Ker} f^j$.

Reciprocamente, si $\{f^{k+1}, ...f^n\}$ son linealmente independientes en V^*, complétese a una base $\{\phi^1, ..., \phi^k, f^{k+1}, ..., f^n\}$ de V^*, y sea $\{w_1, ..., w_k, v_{k+1}, ..., v_n\}$ la base de V dual de la anterior. Por 2.6^0 se tiene que $\bigcap_{j=k+1}^n \operatorname{Ker} f^j = < \{w_1, ..., w_k\} >$ y, por tanto, es un espacio vectorial de dimensión k. $\square$

2.7. Proposición. *F es un subespacio afín de E de dimensión k sii es el conjunto de puntos $Q \in E$ cuyas coordenadas $(Q^1, ..., Q^n)$ verifican un sistema de $n-k$ ecuaciones lineales*

$$\begin{aligned} a_1^{k+1}Q^1 + a_2^{k+1}Q^2 + ... + a_n^{k+1}Q^n + a_0^{k+1} &= 0 \\ a_1^{k+2}Q^1 + a_2^{k+2}Q^2 + ... + a_n^{k+2}Q^n + a_0^{k+2} &= 0 \\ \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_1^{n}Q^1 + a_2^{n}Q^2 + ... + a_n^{n}Q^n + a_0^{n} &= 0 \end{aligned} \tag{2.7.1}$$

y la matriz

$$\begin{pmatrix} a_1^{k+1} & \dots & a_n^{k+1} \\ \vdots & \ddots & \vdots \\ a_1^n & \dots & a_n^n \end{pmatrix} \tag{2.7.2}$$

tiene rango $n-k$.

El sistema de ecuaciones (2.7.1) se llama ecuaciones en implícitas de F.

Demostración. Supongamos que $F = P + W$ es un subespacio afín de dimensión k. Sean $f^{k+1}, ..., f^n \in V^*$ aplicaciones lineales linealmente independientes tales que $W = \bigcap_{j=k+1}^n \operatorname{Ker} f^j$ y sean $a_i^j = f^j(e_i)$ para $i = 1, ..., n$, $j = k+1, ..., n$. Para cada j, $(a_1^j, ..., a_n^j)$ son las componentes de f^j en la base dual $\{e^{*1}, ..., e^{*n}\}$ de $\{e_1, ..., e_n\}$. Como $f^{k+1}, ..., f^n$ son linealmente independientes, la matriz (2.7.2) de las componentes de los f^j tiene rango $n-k$. Se tiene que $Q \in F = P + W$ sii $\overrightarrow{PQ} \in \bigcap_{j=k+1}^n \operatorname{Ker} f^j$ sii $f^j(\overrightarrow{PQ}) = 0$ para todo $j \in \{k+1, ..., n\}$ sii $0 = f^j(\sum_{i=1}^n (Q^i - P^i)e_i) = \sum_{i=1}^n (Q^i - P^i) f^j(e_i) = \sum_{i=1}^n (Q^i - P^i)a_i^j$ para todo $j \in \{k+1, ..., n\}$, lo que ocurre sii

$$\begin{array}{cccc} a_1^{k+1}(Q^1 - P^1) & + \dots & + a_n^{k+1}(Q^n - P^n) & = 0 \\ \dots & \dots & \dots & \\ a_1^{n}(Q^1 - P^1) & + \dots & + a_n^{n}(Q^n - P^n) & = 0 \end{array}$$

que es (2.7.1) si $-\sum_{i=1}^n a_i^j P^i = a_0^j$ para $j = k+1, ..., n$.

Recíprocamente, si $F = \{Q = (Q^1, ..., Q^n)$ que verifican (2.7.1)$\}$, definamos $f^j : V \longrightarrow I\!R$ por $f^j(\sum_{i=1}^n x^i e_i) = \sum_{i=1}^n x^i a_i^j$ para $j \in \{k+1, ..., n\}$. Evidentemente $f^j \in V^*$, $f^j(e_i) = a_i^j$ y las componentes de f^j en la base dual de $\{e_1, ..., e_n\}$ son

$(a_1^j, ...a_n^j)$. Como la matriz (2.7.2) tiene rango $n-k$ entonces las f^j son linealmente independientes. Sea $(P^1, ..., P^n)$ una solución del sistema de ecuaciones (2.7.1) (tal solución existe, por ser un sistema de $n-k$ ecuaciones lineales con n incógnitas tal que la matriz de los coeficientes tiene rango $n-k$) y sea P el punto de F de coordenadas $(P^1, ..., P^n)$. Se tiene entonces que $Q \in F$ sii $0 = \sum_{i=1}^n a_i^j Q^i + a_0^j - (\sum_{i=1}^n a_i^j P^i + a_0^j) = \sum_{i=1}^n a_i^j (Q^i - P^i) = \sum_{i=1}^n f^j(e_i)(Q^i - P^i) = f^j(\sum_{i=1}^n (Q^i - P^i)e_i) = f^j(\overrightarrow{PQ})$ para todo $j = k+1, ..., n$ lo que ocurre sii $\overrightarrow{PQ} \in \bigcap_{j=k+1}^n \operatorname{Ker} f^j$, i.e. sii $Q \in P + \bigcap_{j=k+1}^n \operatorname{Ker} f^j$, que es un subespacio afín de dimensión k. $\square$

2.7'.Ejercicio.

Si un subespacio afín F de dimensión k viene descrito por el sistema de ecuaciones (2.7.1) en una referencia cartesiana $\{O, e_1, ..., e_n\}$ de E, dar las componentes en la base $\{e^{*1}, ..., e^{*n}\}$ de V^* dual de $\{e_1, ..., e_n\}$ de unas aplicaciones $f^{k+1}, ..., f^n : V \longrightarrow I\!R$ tales que el espacio vectorial director de F sea igual a $\bigcap_{j=k+1}^n \operatorname{Ker} f^j$.

2.8. Nota. Obsérvese que cuando F es un hiperplano afín $(k = n-1)$, la demostración de la proposición anterior dice que existe un $P \in F$ y una $f \in V^*$ tales que $F = P + \operatorname{Ker} f$. Cuando F es de dimensión k arbitraria, entonces existen $P \in F$ y $f^{k+1}, ..., f^n \in V^*$ tales que $F = P + \bigcap_{j=k+1}^n \operatorname{Ker} f^j = \bigcap_{j=k+1}^n (P + \operatorname{Ker} f^j)$. Como, por la observación anterior, cada $P + \operatorname{Ker} f^j$ representa un hiperplano, se tiene que: *todo subespacio afín de dimensión k considerarse como la intersección de $n-k$ hiperplanos.* En particular: *una recta viene dada por la intersección de $n-1$ hiperplanos.*

Si se considera la misma idea desde el punto de vista de las ecuaciones (2.7.1), se ve que cada subespacio afín de dimensión k puede expresarse por un sistema de $n-k$ ecuaciones lineales independientes, cada una de las cuales representa un hiperplano afín, llegándose a la misma conclusión que antes.

2.9.Proposición. *Sea E un espacio afín de dimensión n y $\{e_i\}$ una base de su espacio director V. Si se describe un subespacio afín $F = P+W \subset E$ de dimensión $k < n$ dando un punto $P \in F$ y una base $\{w_1, w_2, ..., w_k\}$ de su espacio director W, F es el conjunto* *de puntos Q de E cuyas coordenadas $(Q^1, ..., Q^n)$ verifican el sistema de ecuaciones*

$$\begin{vmatrix} Q^1 - P^1 & Q^2 - P^2 & \dots & Q^k - P^k & Q^j - P^j \\ w_1^1 & w_1^2 & \dots & w_1^k & w_1^j \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ w_k^1 & w_k^2 & \dots & w_k^k & w_k^j \end{vmatrix} = 0, \quad j = k+1, ..., n,$$

siendo $(P^1, ..., P^n)$ las coordenadas de P en la base $\{e_i\}$ y $w_i = \sum_{j=1}^n w_i^j e_j$, $i = 1, ..., k$, y habiendo ordenado la base $\{e_i\}$ de modo que $\det(w_i^j)_{1 \le i,j \le k} \neq 0$

Demostración. $x = \sum_{i=1}^n x^i e_i \in W$ sii x es combinación lineal de $w_1, ..., w_k$, y esto ocurre sii $\{x, w_1, ..., w_k\}$ es un conjunto de vectores linealmente dependientes, lo cual ocurre sii

$$\text{rango} \begin{pmatrix} x^1 & x^2 & \dots & x^n \\ w_1^1 & w_1^2 & \dots & w_1^n \\ \vdots & \vdots & \ddots & \vdots \\ w_k^1 & w_k^2 & \dots & w_k^n \end{pmatrix} = k,$$

lo cual ocurre, con la ordenación elegida de la base, sii las x^i verifican las ecuaciones escritas para las $Q^i - P^i$ del enunciado. De aquí, teniendo en cuenta que $Q \in F$ sii $\overrightarrow{PQ} \in W$, y que $\overrightarrow{PQ} = \sum_{i=1}^n (Q^i - P^i) e_i$, resulta la proposición. $\square$

2.10.NOTA: Comparando 2.7 con 2.4, se ve que las ecuaciones en paramétricas de un subespacio afín F de E aparecen de una manera natural cuando se describe el espacio vectorial director W dando una base del mismo, y las ecuaciones en implícitas cuando se describe W en la forma $\bigcap_{j=k+1}^n \operatorname{Ker} f^j$, siendo f^j elementos del espacio dual del espacio director de E. En la proposición 2.9 se ha partido de una base de W y se ha llegado, sin embargo, a un sistema de $n-k$ ecuaciones lineales (una descripción de F en implícitas). Puesto que los coeficientes w_i^j darían la ecuación en paramétricas $Q^j = P^j + \sum_{i=1}^k \lambda^i w_i^j$ de F, la proposición 2.9 da un procedimiento para pasar de una expresión en paramétricas a una expresión en implícitas. Este procedimiento consiste, esencialmente, en la eliminación de los parámetros $\lambda^1, ..., \lambda^k$ de las ecuaciones (2.4.1).

El paso inverso (de ecuaciones en implícitas a paramétricas) es aún más sencillo: basta con tomar k de las variables (las componentes de los puntos) como parámetros y despejar las otras $n-k$ variables en función de ellas usando las ecuaciones (2.7.1).

Como ejercicio: Dado el plano $F : \left.\begin{array}{ccc} x & +y & = 2 \\ z & +w & = -2 \end{array}\right\}$ de $E = I\!R^4$, obtener su expresión en paramétricas y, a partir de ellas, volver a obtener una expresión en implícitas.

Vamos a ver ahora las ecuaciones, en paramétricas y en implícitas, de un subespacio afín generado por $k+1$ puntos independientes

2.11. Proposición. *El subespacio afín F generado por (que pasa por) los $k+1$ puntos independientes $\{P_0, P_1, ..., P_k\}$ de coordenadas $P_l = (P_l^1, ..., P_l^n)$, $l = 0, 1, ..., k$*

es el conjunto de puntos $Q \in E$ cuyas coordenadas $(Q^1, ..., Q^n)$ verifican las ecuaciones

$$Q^i = P_0^i + \sum_{j=1}^{k} \lambda^j (P_j^i - P_0^i), \ \text{para todo } \lambda^1, ..., \lambda^k \in I\!R. \tag{2.1.1}$$

Cuando $k = 1$, las ecuaciones (2.11.1) pueden cambiarse por

$$\frac{Q^1 - P_0^1}{P_1^1 - P_0^1} = ... = \frac{Q^n - P_0^n}{P_1^n - P_0^n}. \tag{2.11.2}$$

Cuando k es arbitrario, las ecuaciones en implícitas que representan a F son

$$\begin{vmatrix} Q^1 - P_0^1 & Q^2 - P_0^2 & \dots & Q^k - P_0^k & Q^j - P_0^j \\ P_1^1 - P_0^1 & P_1^2 - P_0^2 & \dots & P_1^k - P_0^k & P_1^j - P_0^j \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ P_k^1 - P_0^1 & P_k^2 - P_0^2 & \dots & P_k^k - P_0^k & P_k^j - P_0^j \end{vmatrix} = 0, \qquad j = k+1, ..., n, \tag{2.11.3}$$

si la base $\{e_1, ..., e_n\}$ del sistema de referencia usado se ha ordenado de modo que $det(P_l^i - P_0^i)_{1 \le i,l \le k} \neq 0$.

En particular, cuando $k = n - 1$, se tiene

$$\begin{vmatrix} Q^1 - P_0^1 & Q^2 - P_0^2 & \dots & Q^n - P_0^n \\ P_1^1 - P_0^1 & P_1^2 - P_0^2 & \dots & P_1^n - P_0^n \\ \vdots & \vdots & \ddots & \vdots \\ P_{n-1}^1 - P_0^1 & P_{n-1}^2 - P_0^2 & \dots & P_{n-1}^n - P_0^n \end{vmatrix} = 0. \tag{2.11.4}$$

Demostración. Es consecuencia inmediata de que $F = P_0 + < \{\overrightarrow{P_0P_1}, ..., \overrightarrow{P_0P_k}\} >$ y de las proposiciones anteriores. $\square$

2.12. *Cambio de referencia cartesiana.* Sean $\{O, e_1, ..., e_n\}$ y $\{O', e_1', ...e_n'\}$ dos sistemas de referencia de un espacio afín E. Si las coordenadas de O' respecto de la primera referencia son $(O'^1, ..., O'^n)$ y $e_j' = \sum_{i=1}^n e_j'^i e_i$, dado un punto $P \in E$ con coordenadas $(P^1, ..., P^n)$ en el primer sistema de referencia y coordenadas $(P'^1, ..., P'^n)$ en el segundo sistema de referencia, de la expresión $\overrightarrow{OP} = \overrightarrow{OO'} + \overrightarrow{O'P}$ se deduce que

$$\sum_{i=1}^n P^i e_i = \sum_{i=1}^n O'^i e_i + \sum_{j=1}^n P'^j e_j' = \sum_{i=1}^n O'^i e_i + \sum_{j=1}^n P'^j \sum_{i=1}^n e_j'^i e_i = \sum_{i=1}^n (O'^i + \sum_{j=1}^n P'^j e_j'^i) e_i.$$

i.e.

$$P^i = O'^i + \sum_{j=1}^n P'^j e_j'^i, \tag{2.12.1}$$

o, en forma matricial

$$\begin{pmatrix} P^1 \\ \vdots \\ P^n \end{pmatrix} = \begin{pmatrix} O'^1 \\ \vdots \\ O'^n \end{pmatrix} + \begin{pmatrix} e_1'^1 & \dots & e_n'^1 \\ \vdots & & \vdots \\ e_1'^n & \dots & e_n'^n \end{pmatrix} \begin{pmatrix} P'^1 \\ \vdots \\ P'^n \end{pmatrix}, \tag{2.12.2}$$

que es la fórmula de cambio de coordenadas cuando se cambia de referencia cartesiana.

2.13. Definición. *Una referencia afín de un espacio afín E de dimensión n es un conjunto de $n+1$ puntos $\{P_0, P_1, ..., P_n\}$ de E independientes. Asociada a una tal referencia afín existe una referencia cartesiana $\{P_0, \overrightarrow{P_0P_1}, ..., \overrightarrow{P_0P_n}\}$.*

2.14. Preguntas y ejercicios.

(a) Sea F un subespacio afín de E de dimensión m. Sea n la dimensión de E. ¿Cuántas ecuaciones se usan para describir F en paramétricas?. ¿Cuántas para describir F en implícitas?.

(b) Sea E un espacio afín de dimensión 4 y sea F un subespacio afín de E cuyos puntos en una referencia cartesiana $\{O, e_1, e_2, e_3, e_4\}$ tienen unas coordenadas que satisfacen las ecuaciones $\left.\begin{array}{ll} x_1 + x_2 & = 1 \\ x_3 - x_4 & = 0 \end{array}\right\}$. ¿Qué dimensión tiene F?. ¿Qué relación hay entre las 4-adas $(1, 1, 0, 0)$, $(0, 0, 1, -1)$ y el espacio vectorial director de F?.

(c) Sea $\{O, e_1, ..., e_n\}$ un sistema de referencia cartesiano de un espacio afín E de dimensión n de espacio vectorial asociado V. Si $e_1, ..., e_n \in V$, ¿qué es O?, ¿qué relación hay entre este O y el vector $0 \in V$?.

(d) Sea E un espacio afín de dimensión $n > 4$ y sea F un subespacio afín de E cuyos puntos en una referencia cartesiana $\{O, e_1, ..., e_n\}$ tienen unas coordenadas que satisfacen las ecuaciones $\left.\begin{array}{ll} x_1 + x_2 + x_n & = 1 \\ x_3 - x_4 & = 0 \end{array}\right\}$. ¿Qué dimensión tiene F?. ¿Qué relación hay entre las n-adas $(1, 1, 0, ..., 0, 1)$, $(0, 0, 1, -1, 0, ..., 0)$ y el espacio vectorial director de F?.

§3. Paralelismo, intersección e incidencia en coordenadas

En todo este apartado usaremos una referencia cartesiana $\{O, e_1, ..., e_n\}$ del espacio afín E, y las coordenadas de los puntos serán siempre con respecto a este sistema de coordenadas.

Sean los subespacios afines $F = P + U$ y $G = Q + W$ de dimensiones respectivas m y k, $m \leq k$. $\{u_1, ..., u_m\}$ será una base de U y $\{w_1, ..., w_k\}$ una base de W. Además $u_j = \sum_{i=1}^n u_j^i e_i$ y $w_j = \sum_{i=1}^n w_j^i e_i$; $P = (P^1, ..., P^n)$ y $Q = (Q^1, ..., Q^n)$. En ecuaciones paramétricas, F y G vienen representados, respectivamente, por

$$F \equiv \qquad R^i = P^i + \sum_{j=1}^m \lambda^j u_j^i,$$

$$G \equiv \qquad S^i = Q^i + \sum_{j=1}^k \mu^j w_j^i.$$

Como $F||G$ sii $U \subset W$, se tiene que

3.1.Proposición. *F es paralelo a G sii*

$$\textit{rango}\begin{pmatrix} u_1^1 & \dots & u_1^n \\ \dots & \dots & \dots \\ u_m^1 & \dots & u_m^n \\ w_1^1 & \dots & w_1^n \\ \dots & \dots & \dots \\ w_k^1 & \dots & w_k^n \end{pmatrix} = k$$

o, equivalentemente,

$$\textit{rango}\begin{pmatrix} u_j^1 & \dots & u_j^n \\ w_1^1 & \dots & w_1^n \\ \dots & \dots & \dots \\ w_k^1 & \dots & w_k^n \end{pmatrix} = k \quad \textit{para todo} \quad j = 1, ..., m.$$

(ambas expresiones significan que, para $j = 1, ..., m$, u_j es combinación lineal de $w_1, ..., w_k$.

Es, por otra parte evidente que

3.2.Proposición. *a) $F \cap G \neq \emptyset$ sii existen $(\lambda^1, ..., \lambda^m, \mu^1, ..., \mu^k) \in I\!R^{m+k}$ tales que $P^i + \sum_{j=1}^m \lambda^j u_j^i = Q^i + \sum_{j=1}^k \mu^j w_j^i$. b) $F \subset G$ sii se cumple a) y las condiciones de la proposición 3.1. c) $F \cap G = \emptyset$ sii no se cumple la condición dada en a).*

Antes de ver como se estudian las posiciones relativas de F y G usando sus expresiones en implícitas, vamos a ver algunos resultados de álgebra lineal.

3.3.Definición. *Sea V un espacio vectorial de dimensión n, y W un subespacio de dimensión k. Se llama anulador W^N de W al subespacio de V^* definido por*

$$W^N = \{\beta \in V^* / \beta(w) = 0 \textit{ para todo } w \in W\}.$$

NOTA: Al no haber un producto escalar, el anulador juega, en geometría vectorial, o en geometría afín, el papel que juega el ortogonal en las geometrías vectorial o afín métricas.

3.4.Proposición. *Si $W = \bigcap_{j=k+1}^n \operatorname{Ker} f^j$, siendo $f^{k+1}, ..., f^n \in V^*$ linealmente independientes, entonces $W^N =< \{f^{k+1}, ..., f^n\} >$. Como consecuencia, $dim(W^N) = n - k$.*

Demostración. Completemos $\{f^{k+1}, ..., f^n\}$ a una base $\{\phi^1, ..., \phi^k, f^{k+1}, ..., f^n\}$ de V^*, y sea $\{w_1, ..., w_k, v_{k+1}, ..., v_n\}$ la base de V dual de la anterior. Si $\beta \in W^N$, entonces $\beta(w) = 0$ para todo $w \in W$, en particular, $\beta(w_i) = 0$, luego (usando 2.6^0) $\beta = \sum_{i=1}^k \beta(w_i)\phi^i + \sum_{j=k+1}^n \beta(v_j) f^j = \sum_{j=k+1}^n \beta(v_j) f^j \in < \{f^{k+1}, ..., f^n\} >$, lo que prueba la inclusión $W^N \subset < \{f^{k+1}, ..., f^n\} >$. Para probar la inclusión contraria, si $\varphi \in < \{f^{k+1}, ..., f^n\} >$, entonces $\varphi = \sum_{j=k+1}^n \phi_j f^j$ y para todo $w \in W = \bigcap_{j=k+1}^n \operatorname{Ker} f^j$ se tiene $\varphi(w) = \sum_{j=k+1}^n \phi_j f^j(w) = 0$. □ □

3.5.Corolario. *$W = \{w \in V/\, \beta(w) = 0$ para todo $\beta \in W^N\}$.*

Demostración. Si $w \in W$, entonces $\beta(w) = 0$ para todo $\beta \in W^N$, luego $W \subset \{w \in V/\, \beta(w) = 0$ para todo $\beta \in W^N\}$. Recíprocamente, si $\beta(w) = 0$ para todo $\beta \in W^N$, entonces, por 3.4, $f^j(w) = 0$ para todo $j = k+1, ..., n$, luego $w \in \bigcap_{j=k+1}^{n} \operatorname{Ker} f^j = W$. □ □

3.6.Proposición. *Sea U otro subespacio vectorial de V. Se tiene que $U \subset W$ si y solo si $U^N \supset W^N$.*

Demostración. Si $U \subset W$ y $\beta \in W^N$, entonces $\beta(u) = 0$ para todo $u \in U \subset W$, luego $\beta \in U^N$, luego $U \subset W$ implica que $U^N \supset W^N$.

Recíprocamente, si $U^N \supset W^N$, para todo $u \in U$ se tiene que si $\beta \in W^N \subset U^N$, entonces $\beta(u) = 0$ y, por 3.5, $u \in W$, luego $U^N \supset W^N$ implica que $U \subset W$. □ □

3.7. Vamos ahora a usar estos resultados para estudiar posiciones relativas de subespacios afines usando sus ecuaciones en implícitas. Sean F y G los subespacios afines representados, en implícitas, por las ecuaciones

$$(3.7.1) \qquad F \equiv \left.\begin{array}{ccccc} a_1^{m+1}R^1+ & \dots & +a_n^{m+1}R^n & +a_0^{m+1} & =0 \\ \dots & \dots & \dots & \dots & \dots \\ a_1^n R^1+ & \dots & +a_n^n R^n & +a_0^n & =0 \end{array}\right\}$$

y

$$(3.7.2) \qquad G \equiv \left.\begin{array}{ccccc} b_1^{k+1}S^1+ & \dots & +b_n^{k+1}S^n & +b_0^{k+1} & =0 \\ \dots & \dots & \dots & \dots & \dots \\ b_1^n S^1+ & \dots & +b_n^n S^n & +b_0^n & =0 \end{array}\right\}$$

De acuerdo con lo dicho en la demostración de 2.7 (ver también ejercicio 2.7), se tiene que si $U = \bigcap_{l=m+1}^{n} \operatorname{Ker} g^l$, $W = \bigcap_{r=k+1}^{n} \operatorname{Ker} f^r$, $F = P + U$, y $G = Q + W$, entonces, $g^l(e_i) = a_i^l$ y $f^r(e_j) = b_j^r$, $1 \leq i, j \leq n$, $m+1 \leq l \leq n$, $k+1 \leq r \leq n$. De ello se tiene que

3.8.Proposición. *a)$F||G$ sii*

$$(3.8.1) \qquad \text{rango} \begin{pmatrix} a_1^{m+1} & \dots & a_n^{m+1} \\ \dots & \dots & \dots \\ a_1^n & \dots & a_n^n \\ b_1^{k+1} & \dots & b_n^{k+1} \\ \dots & \dots & \dots \\ b_1^n & \dots & b_n^n \end{pmatrix} = n - m \qquad \textit{o, equivalentemente,}$$

$$(3.8.2) \qquad \text{rango} \begin{pmatrix} a_1^{m+1} & \dots & a_n^{m+1} \\ \dots & \dots & \dots \\ a_1^n & \dots & a_n^n \\ b_1^r & \dots & b_n^r \end{pmatrix} = n - m \textit{ para todo r=k+1,...,n.}$$

b) $F \cap G \neq \emptyset$ sii el sistema de ecuaciones formado por las ecuaciones (3.7.1) y (3.7.2) conjuntamente tiene solución, lo cual ocurre sii

$$\text{rango}\begin{pmatrix} a_1^{m+1} & \dots & a_n^{m+1} & a_0^{m+1} \\ \dots & \dots & \dots & \dots \\ a_1^n & \dots & a_n^n & a_0^n \\ b_1^{k+1} & \dots & b_n^{k+1} & b_0^{k+1} \\ \dots & \dots & \dots & \dots \\ b_1^n & \dots & b_n^n & b_0^n \end{pmatrix} = \text{rango}\begin{pmatrix} a_1^{m+1} & \dots & a_n^{m+1} \\ \dots & \dots & \dots \\ a_1^n & \dots & a_n^n \\ b_1^{k+1} & \dots & b_n^{k+1} \\ \dots & \dots & \dots \\ b_1^n & \dots & b_n^n \end{pmatrix}.$$

En este caso, la intersección es el conjunto de puntos cuyas coordenadas son solución de las ecuaciones (3.7.1) y (3.7.2) simultáneamente.

c)$F \subset G$ sii se dan a) y b) simultáneamente.

Demostración. a) Por la definición de paralelismo y por 3.4 y 3.6, $F||G$ sii $U \subset W$ sii $< \{g^{m+1}, ..., g^n\} > \supset < \{f^{k+1}, ..., f^n\} >$, lo que ocurre sii $\{g^{m+1}, ..., g^n, f^r\}$ es un conjunto linealmente dependiente para todo $r = k+1, ..., n$, lo que es equivalente a que la matriz de sus componentes tenga rango $n-m$, que es justamente la segunda de las condiciones (3.8.1), que, a su vez, es equivalente a la primera, por ser $\{g^{m+1}, ..., g^n\}$ linelamente independientes.

b) Es evidente que la intersección está formada por los puntos cuyas coordenadas verifican los sistemas de ecuaciones que definen cada uno de los subespacios afines, y la condición sobre los rangos de las matrices escrita en el enunciado es solo la condición necesaria y suficiente para que ese sistema de ecuaciones lineales tenga solución. □

Vamos a ver ahora algunos casos particulares.

Supongamos primero que F y G vienen dadas en ecuaciones paramétricas como las escritas antes de la proposición 3.1.

3.9.Corolario. *a) Cuando $k = m = 1$, $F||G$ sii*

$$\frac{u^1}{w^1} = \dots = \frac{u^n}{w^n}.$$

b)Si $m = 1$, $k = n-1$, $F||G$ sii

$$\begin{vmatrix} u^1 & \dots & u^n \\ w_1^1 & \dots & w_1^n \\ \dots & \dots & \dots \\ w_{n-1}^1 & \dots & w_{n-1}^n \end{vmatrix} = 0.$$

Si F y G vienen dadas en implícitas como (3.7.1) y (3.7.2), entonces

3.10.Corolario. *a) Cuando $k = m = n-1$, $F||G$ sii*

$$\frac{a_1}{b_1} = \frac{a_2}{b_2} = ... = \frac{a_n}{b_n}.$$

b) Cuando $m = 1$, $k = n-1$, $F||G$ sii

$$\begin{vmatrix} a_1^2 & \dots & a_n^2 \\ \dots & \dots & \dots \\ a_1^n & \dots & a_n^n \\ b_1 & \dots & b_n \end{vmatrix} = 0.$$

En esta situación se tiene, además,

?3.11.Proposición. *Cuando $k = m = 1$, $F||G$ sii*

$$\frac{\begin{vmatrix} a_1^2 & \dots & a_{n-1}^2 \\ \dots & \dots & \dots \\ a_1^n & \dots & a_{n-1}^n \end{vmatrix}}{\begin{vmatrix} b_1^2 & \dots & b_{n-1}^2 \\ \dots & \dots & \dots \\ b_1^n & \dots & b_{n-1}^n \end{vmatrix}} = \frac{\begin{vmatrix} a_1^2 & \dots & a_{i-1}^2 & a_{i+1}^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_1^n & \dots & a_{i-1}^n & a_{i+1}^n & \dots & a_n^n \end{vmatrix}}{\begin{vmatrix} b_1^2 & \dots & b_{i-1}^2 & b_{i+1}^2 & \dots & b_n^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ b_1^n & \dots & b_{i-1}^n & b_{i+1}^n & \dots & b_n^n \end{vmatrix}} = \frac{\begin{vmatrix} a_2^2 & \dots & a_n^2 \\ \dots & \dots & \dots \\ a_2^n & \dots & a_n^n \end{vmatrix}}{\begin{vmatrix} b_2^2 & \dots & b_n^2 \\ \dots & \dots & \dots \\ b_2^n & \dots & b_n^n \end{vmatrix}}.$$

para $i = 2, ..., n-1$.

Demostración. Esta condición para el paralelismo de rectas resulta de 3.9 y de que el vector

$$x = \left((-1)^n \begin{vmatrix} a_1^2 & \dots & a_{n-1}^2 \\ \dots & \dots & \dots \\ a_1^n & \dots & a_{n-1}^n \end{vmatrix}, ..., (-1)^i \begin{vmatrix} a_1^2 & \dots & a_{i-1}^2 & a_{i+1}^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_1^n & \dots & a_{i-1}^n & a_{i+1}^n & \dots & a_n^n \end{vmatrix}, \right.$$
$$\left. ..., (-1) \begin{vmatrix} a_2^2 & \dots & a_n^2 \\ \dots & \dots & \dots \\ a_2^n & \dots & a_n^n \end{vmatrix} \right).$$

genera el espacio vectorial director de la recta afín dada por las ecuaciones

$$\left.\begin{matrix} a_1^2 R^1 + & \dots & + a_n^2 R^n & + a_0^2 & = 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_1^n R^1 + & \dots & + a_n^n R^n & + a_0^n & = 0 \end{matrix}\right\}$$

En efecto, obsérvese que para todo $l = 2, ..., n$,

$$\sum_{i=1}^n a_i^l (-1)^i \begin{vmatrix} a_1^2 & \dots & a_{i-1}^2 & a_{i+1}^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_1^n & \dots & a_{i-1}^n & a_{i+1}^n & \dots & a_n^n \end{vmatrix} = - \begin{vmatrix} a_1^l & \dots & a_n^l \\ a_1^2 & \dots & a_n^2 \\ \dots & \dots & \dots \\ a_1^n & \dots & a_n^n \end{vmatrix} = 0$$

por haber dos filas repetidas, luego $x \in \operatorname{Ker} g^l$ para todo $l = 2, ..., n$, i.e., $x \in \cap_{l=2}^n \operatorname{Ker} g^l = U$, c.q.d. □

3.12. Preguntas y ejercicios.

(a) Sean E y F como en 2.14(d), y sea G el subespacio afín de E dado por $x_3 - x_4 = 1$. ¿Son F y G paralelos?.

§4 Aplicaciones afines y afinidades

4.1.Definición. *Un segmento es un par de puntos* $\{P, Q\}$ *de* E. *Un segmento orientado es un par ordenado* (P, Q) *de puntos distintos de* E.

En un espacio afín real se puede considerar el segmento $\{P, Q\}$ como el conjunto que contiene a $\{P, Q\}$ y a todos los puntos de la recta afín pasando por P y Q que están entre P y Q, i.e., al conjunto $\{P + \lambda\overrightarrow{PQ}/0 \leq \lambda \leq 1\}$. Esto solo tiene sentido si $I\!R$ es un cuerpo ordenado (por ejemplo $I\!R = I\!R$).

4.2.Definición. *Dos segmentos* $\{P, Q\}$ *y* $\{R, S\}$ *(orientados o no) se dice que son paralelos si lo son las rectas afines que los contienen, i.e., si existe un* $\lambda \in I\!R$ *tal que* $\overrightarrow{RS} = \lambda\overrightarrow{PQ}$.

4.3.Definición. *Sean* (P, Q) *y* (R, S) *dos segmentos orientados paralelos. La razón de* (R, S) *a* (P, Q) *es el escalar* $\rho \in I\!R$ *tal que* $\overrightarrow{RS} = \rho\overrightarrow{PQ}$. *Se denota este escalar* ρ *por*

$$\frac{(R, S)}{(P, Q)}$$

4.4.Definición. *Una aplicación* $f : E \longrightarrow E'$ *entre dos espacios afines* E *y* E' *de espacios vectoriales asociados* V *y* V' *respectivamente, se dice que es una aplicación afín si lleva rectas afines paralelas en rectas paralelas o en puntos y si conserva la razón de dos segmentos orientados paralelos, i.e., si se verifican las condiciones*

4.4.1. F *y* G *rectas de* E *y* $F||G$ *implica* $f(F)$ *y* $f(G)$ *son rectas paralelas o son puntos y*

4.4.2. $\frac{(f(R),f(S))}{(f(P),f(Q))} = \frac{(R,S)}{(P,Q)}$ *si* $f(R) \neq f(S)$. *(Obsérvese que esta condición implica que* $f(R) \neq f(S)$ *implica* $f(P) \neq f(Q)$, *pues si* $f(R) \neq f(S)$, *entonces* $\frac{(f(R),f(S))}{(f(P),f(Q))} = \frac{(R,S)}{(P,Q)} \in I\!R$*).*

4.5. *Dada una aplicación cualquiera* $f : E \longrightarrow E'$, *y dado* $P \in E$, f *define una aplicación* $f_P : V \longrightarrow V'$ *de la siguiente manera*

$$f_P(v) = \overrightarrow{f(P)f(P + v)}, \quad (\text{i.e.} \quad f_P(\overrightarrow{PQ}) = \overrightarrow{f(P)f(Q)}).$$

Con ayuda de esta aplicación, f *se escribe*

$$f(Q) = f(P) + f_P(\overrightarrow{PQ}). \tag{4.5.1}$$

***4.6.Proposición.** *Una aplicación* $f : E \longrightarrow E'$ *es afín sii* $f_P = f_Q$ *para todo* $P, Q \in E$ *(i.e.* f_P *no depende del* P *de* E *elegido) y* f_P *es lineal.*

Como f_P *no depende de* P *si* f *es afín, en adelante se escribirá* $\widetilde{f}$ *para denotar a* f_P, *y se la denominará aplicación lineal asociada a la aplicación afín* f. *A la expresión*

(4.5.1), que a partir de ahora se escribirá $f(Q) = f(P) + \widetilde{f}(\overrightarrow{PQ})$, se la llamará forma canónica de la aplicación afín f.

Demostración. Supongamos primero que f es una aplicación afín. Sean $P, Q \in E$, $v \in V$. Como $\overrightarrow{P(P+v)} = \overrightarrow{Q(Q+v)} = v$, se tiene que $(P, P+v)$ y $(Q, Q+v)$ son segmentos paralelos de razón 1.

Si $f(P) = f(P+v)$, de la propiedad 4.4.2 resulta que si f es afín, entonces $f(Q) = f(Q+v)$, y $f_P(v) = 0 = f_Q(v)$ (pues si $f(Q) \neq f(Q+v)$, entonces $f(P) \neq f(P+v)$, como se observó en 4.4.2).

Si $f(P) \neq f(P+v)$, como f es afín, de la definición de razón de dos segmento resulta que

$$\overrightarrow{f(P)f(P+v)} = \frac{(f(P), f(P+v))}{(f(Q), f(Q+v))}\overrightarrow{f(Q)f(Q+v)} \tag{4.6.1}$$

y de la propiedad 4.4.2 resulta que

$$\frac{(f(P), f(P+v))}{(f(Q), f(Q+v))} = \frac{(P, P+v)}{(Q, Q+v)} = 1,$$

que, sustituido en (4.6.1), da

$$f_P(v) = \overrightarrow{f(P)f(P+v)} = \overrightarrow{f(Q)f(Q+v)} = f_Q(v).$$

Luego está bien definida $\widetilde{f} = f_P$. Veamos que $\widetilde{f}$ es lineal. Sean $v, w \in V$ y $\lambda \in I\!R$, se tiene que

$$\begin{aligned}\widetilde{f}(v+w) = \overrightarrow{f(P)f(P+v+w)} &= \overrightarrow{f(P)f(P+v)} + \overrightarrow{f(P+v)f(P+v+w)} \\ &= f_P(v) + f_{P+v}(w) = \widetilde{f}(v) + \widetilde{f}(w)\end{aligned}$$

y, usando que $(P, P+v)$ y $(P, P+\lambda v)$ son segmentos orientados paralelos de razón λ y volviendo a aplicar las propiedades 4.4.1 y 4.4.2,

$$\begin{aligned}\widetilde{f}(\lambda v) = \overrightarrow{f(P)f(P+\lambda v)} &= \frac{(f(P), f(P+\lambda v))}{(f(P), f(P+v))}\overrightarrow{f(P)f(P+v)} \\ &= \frac{(P, P+\lambda v)}{(P, P+v)}\overrightarrow{f(P)f(P+v)} = \lambda\widetilde{f}(v).\end{aligned}$$

Supongamos ahora que $\widetilde{f} = f_P$ no depende de P y que es lineal. Si $u \in V$ y $P \in E$,

$$f(P+\lambda u) = f(P) + \widetilde{f}(\lambda u) = f(P) + \lambda\widetilde{f}(u),$$

luego, si $F = P + <\{u\}>$ es una recta de E, entonces

$$f(F) = \{f(P) + \lambda\widetilde{f}(u) / \lambda \in I\!R\} = f(P) + \widetilde{f}(<\{u\}>).$$

Por tanto, $f(F)$ es un subespacio afín de espacio vectorial director $\widetilde{f}(<\{u\}>)$, que es una recta si $\widetilde{f}(u) \neq 0$ y un punto si $\widetilde{f}(u) = 0$. Por tanto, si $F||G$ y $f(F)$ es una recta, entonces $f(G)$ es otra recta que tiene el mismo espacio director que $f(F)$, y si $f(F)$ es un punto, entonces $f(G)$ es un punto. Además, si (P,Q) y (R,S) son segmentos paralelos, entonces

$$\overrightarrow{f(R)f(S)} = \widetilde{f}(\overrightarrow{RS}) = \widetilde{f}(\frac{(R,S)}{(P,Q)}\overrightarrow{PQ}) = \frac{(R,S)}{(P,Q)}\widetilde{f}(\overrightarrow{PQ}) = \frac{(R,S)}{(P,Q)}\overrightarrow{f(P)f(Q)},$$

de donde se tiene, usando la definición de razón de segmentos orientados paralelos, que

$$\frac{(R,S)}{(P,Q)} = \frac{(f(R),f(S))}{(f(P),f(Q))}.$$

□

4.7.Proposición. *a) Una aplicación afín $f : E \longrightarrow E'$ lleva un subespacio afín $F = P+U$ de E en el subespacio afín $f(F) = f(P)+\widetilde{f}(U)$ de E' de dimensión igual al rango de $\widetilde{f}$*

b) Si $G = Q+W$ es un subespacio afín de E' y $P \in f^{-1}(G) \neq \emptyset$, entonces $f^{-1}(G)$ es un subespacio afín de E que se puede escribir de la forma $P + \widetilde{f}^{-1}(W)$. En particular, $f^{-1}(Q)$ es un subespacio afín de E de dimensión igual a $\dim \operatorname{Ker} \widetilde{f}$.

Demostración. Basta con escribir las definiciones:

a)$f(F) = \{f(P+u)/u \in U\} = \{f(P)+\widetilde{f}(u)/u \in U\} = f(P)+\widetilde{f}(U)$.

b)$f^{-1}(G) = \{P+v/f(P+v) = f(P)+\widetilde{f}(v) \in G = f(P)+W\} = \{P+v/\widetilde{f}(v) \in W\} = P+\widetilde{f}^{-1}(W)$. Y la afirmación sobre $\dim f^{-1}(Q)$ resulta de la igualdad anterior cuando $G = \{Q\}$ y, por tanto, $W = \{0\}$. □

La siguiente proposición dice que una aplicación afín viene unívocamente determinada por la imagen de una referencia afín (recordar def. 2.15).

***4.8.Proposición.** *Sean $P_0, P_1, ..., P_n$ $n+1$ puntos independientes de un espacio afín E de dimensión n, y $P'_0, P'_1, ..., P'_n$ $n+1$ puntos arbitrarios (pueden incluso repetirse) de un espacio afín E'. Existe una única aplicación afín $f : E \longrightarrow E'$ tal que $f(P_i) = P'_i$ para cada $i = 0,1,...,n$.*

Demostración. Si f existe, por 4.7, se puede escribir de la forma $f(P) = f(P_0)+\widetilde{f}(\overrightarrow{P_0P})$. Como $P_0, P_1, ..., P_n$ son independientes, se escribe $\overrightarrow{P_0P} = \sum_{i=1}^n \lambda^i \overrightarrow{P_0P_i}$, de donde resulta que

$$(4.8.1) \quad f(P) = f(P_0) + \widetilde{f}(\sum_{i=1}^n \lambda^i \overrightarrow{P_0P_i}) = P'_0 + \sum_{i=1}^n \lambda^i \widetilde{f}(\overrightarrow{P_0P_i})$$
$$= P'_0 + \sum_{i=1}^n \lambda^i \overrightarrow{f(P_0)f(P_i)} = P'_0 + \sum_{i=1}^n \lambda^i \overrightarrow{P'_0P'_i}.$$

Luego, si f existe, es la aplicación dada por (4.8.1). Por otro lado, para ver que f existe, basta con definirla por la expresión (4.8.1) y comprobar que cumple las condiciones del teorema. $\square$

Vamos a ver ahora algunos casos especiales (o ejemplos) de aplicaciones afines

4.9. Definición. *Se llama traslación T_v, de vector $v \in V$ en un espacio afín E a una aplicación afín $T_v : E \longrightarrow E$ de la forma $T_v(Q) = Q + v$.*

T_v es realmente una aplicación afín, de aplicación lineal asociada la identidad. En efecto, si se fija un $P \in E$, se tiene

$$T_v(Q) = Q + v = P + \overrightarrow{PQ} + v = P + v + \overrightarrow{PQ} = T_v(P) + Id(\overrightarrow{PQ}).$$

4.10.Definición. *Una homotecia o dilatación es una aplicación afín de un espacio afín en sí mismo que aplica cada recta en una recta paralela*

Veamos cual la forma canónica de una homotecia. Si f es una dilatación, y $\widetilde{f}$ su aplicación lineal asociada, fijado $P \in E$, para cada $Q \in E$, la recta $P+ < \{\overrightarrow{PQ}\} >$ ha de ser paralela a $f(P)+ < \{\widetilde{f}(\overrightarrow{PQ})\} >$, luego existe un $\lambda(Q) \in I\!R$ tal que $\widetilde{f}(\overrightarrow{PQ}) = \lambda(Q)\overrightarrow{PQ}$. Se define así una función $\lambda : E \longrightarrow I\!R$ que asigna a cada $Q \in E$ el elemento $\lambda(Q) \in I\!R$ tal que $\widetilde{f}(\overrightarrow{PQ}) = \lambda(Q)\overrightarrow{PQ}$. Ahora bien, como $\widetilde{f}$ ha de ser lineal,

$$\lambda(P + v + w)(v + w) = \widetilde{f}(\overrightarrow{P(P + v + w)}) = \widetilde{f}(v) + \widetilde{f}(w) = \lambda(P + v)v + \lambda(P + w)w$$

lo que implica que, si v y w son l.i., entonces

$$\lambda(P + v + w) = \lambda(P + v) = \lambda(P + w). \tag{4.10.1}$$

Por otra parte,

$$\lambda(P + \mu v)(\mu v) = \widetilde{f}(\overrightarrow{P(P + \mu v)}) = \mu\widetilde{f}(v) = \mu\lambda(P + v)v,$$

lo que implica que

$$\lambda(P + \mu v) = \mu\,\lambda(P + v). \tag{4.10.2}$$

De (4.10.1) y (4.10.2) resulta que $\lambda(Q) = \lambda(R)$ para cualesquiera $Q, R \in E$. Luego toda homotecia se puede escribir de la forma

$$f(Q) = f(P) + \rho\overrightarrow{PQ}, \quad \rho \in I\!R. \tag{4.10.3}$$

El escalar $\rho \in I\!R$ se llama razón de la homotecia. Evidentemente $\widetilde{f} = \rho\, Id$.

4.11. Corolario. *Una traslación es una homotecia de razón 1.*

Antes de dar el siguiente ejemplo-definición (las proyecciones paralelas), recordemos que si tenemos una descomposición de un espacio vectorial V en suma directa de dos subespacios U y W ($V = U \oplus W$), se definen las proyecciones $\pi_U : V \longrightarrow U$ y $\pi_W : V \longrightarrow W$ de la siguiente manera: dado un $v \in V$, existen vectores únicos $u \in U$ y $w \in W$ tales que $v = u + w$, se definen entonces $\pi_U(v) = u$ y $\pi_W(v) = w$. Se comprueba que π_U y π_W son aplicaciones lineales.

4.12.Definición. *Sean $F = P + U$ y $G = P + W$ subespacios afines de E tales que $F \cap G = \{P\}$ y $\dim F + \dim G = \dim E$. Se llama proyección de E sobre G paralelamente a F a la aplicación*

$$\Pi_G : E \longrightarrow G / \Pi_G(P + v) = P + \pi_W(v), \text{ i.e. } \Pi_G(Q) = P + \pi_W(\overrightarrow{PQ}).$$

Obsérvese que la definición anterior es correcta, porque $F \cap G = \{P\}$ implica que $U \cap W = \{0\}$ y $\dim F + \dim G = \dim E$ implica que $\dim U + \dim W = \dim V$, luego $V = U \oplus W$, y la aplicación π_W está bien definida.

De la misma definición 4.12 se deduce que Π_G es una aplicación afín de aplicación lineal asociada π_W.

Vamos a ver ahora un teorema que da una definición geométrica de de proyección paralela equivalente a 4.12.

?4.13.Proposición. *Sean E, F y G como en 4.12. Si $Q \notin G$, existe una única recta paralela a F, pasando por Q y que corta a G, y $\Pi_G(Q)$ es la intersección de esta recta con G. Si $Q \in G$, $\Pi_G(Q) = Q$.*

Demostración. . Supongamos $Q \notin G$, lo que implica $Q \neq P$. Sea $Q + Z$ una recta pasando por Q y paralela a F (por tanto $Z \subset U$). Como $G = P + W$, se tiene que $(Q + Z) \cap G \neq \emptyset$ sii existen $z \in Z - \{0\}$ y $w \in W - \{0\}$ tales que $Q + z = P + w$, i.e., tales que $\overrightarrow{PQ} = -z + w \in U \oplus W$, lo que es equivalente a que $-z = \pi_U(\overrightarrow{PQ})$ y $w = \pi_W(\overrightarrow{PQ})$. La primera de esta igualdades implica que la recta $P + Z$ de intersección no vacía con G está unívocamente determinada por $Z =< \{z\} >$, y la segunda que $\Pi_G(Q) = P + \pi_W(\overrightarrow{PQ}) = P + w$ es el punto de intersección de la recta $Q + Z$ con el subespacio afín G. Si $Q \in G$, entonces $\overrightarrow{PQ} \in W$, y $\Pi_G(Q) = P + \pi_W(\overrightarrow{PQ}) = P + \overrightarrow{PQ} = Q$. $\square$

Vamos ahora a ver la relación entre la composición de aplicaciones afines o la inversa de una aplicación afín y la composición o la inversa de sus correspondientes aplicaciones lineales asociadas.

***4.14.Proposición.** *Sean $f : E \longrightarrow E'$ y $g : E' \longrightarrow E"$ aplicaciones afines de aplicaciones lineales asociadas $\widetilde{f} : V \longrightarrow V'$ y $\widetilde{g} : V' \longrightarrow V"$. Para cada $P, Q \in E$, se tiene que*

a) $(g \circ f)(Q) = (g \circ f)(P) + (\widetilde{g} \circ \widetilde{f})(\overrightarrow{PQ})$. Por tanto $g \circ f$ es una aplicación afín y $\widetilde{g \circ f} = \widetilde{g} \circ \widetilde{f}$

b) f es suprayectiva sii $\widetilde{f}$ es suprayectiva.

c) f es inyectiva sii $\widetilde{f}$ es inyectiva.

d) f es biyectiva sii $\widetilde{f}$ es un isomorfismo. Entonces f^{-1} es una aplicación afín y $\widetilde{f^{-1}} = \widetilde{f}^{-1}$.

Demostración. a) $(g \circ f)(Q) = g(f(Q)) = g(f(P) + \widetilde{f}(\overrightarrow{PQ})) = g(f(P)) + \widetilde{g}(\overrightarrow{f(P)(f(P) + \widetilde{f}(\overrightarrow{PQ}))})$ $= g(f(P)) + \widetilde{g}(\widetilde{f}(\overrightarrow{PQ}))$.

b) f es suprayectiva sii para todo $Q' = f(P) + v' \in E'$, existe $Q = P + v \in E$ tal que $f(Q) = Q'$, lo cual ocurre sii para todo $v' \in V'$, existe un $v \in E$ tal que $f(P+v) = f(P) + v'$ y, como $f(P+v) = f(P) + \widetilde{f}(v)$, esto es equivalente a que para todo $v' \in V'$ exista un $v \in V$ tal que $\widetilde{f}(v) = v'$, que es la definición de que $\widetilde{f}$ sea suprayectiva.

c)f es inyectiva sii $f(P) + \widetilde{f}(v) = f(P+v) = f(P+w) = f(P) + \widetilde{f}(w)$ implica que $P + v = P + w$, i.e., $v = w$, y esto es equivalente a que $\widetilde{f}(v) = \widetilde{f}(w)$ sii $v = w$, que es la definición de que $\widetilde{f}$ sea inyectiva.

d) De b) y c) resulta que f es biyectiva sii $\widetilde{f}$ es un isomorfismo. Se tiene que, si $Q' = f(P) + v'$, entonces

$$f^{-1}(Q') = f^{-1}(f(P) + v') = P + v \text{ tal que } f(P) + v' = f(P+v) = f(P) + \widetilde{f}(v),$$

de donde se deduce que

$$v' = \widetilde{f}(v) \text{ i.e. } v = \widetilde{f}^{-1}(v').$$

De todo ello se deduce que

$$f^{-1}(Q') = P + \widetilde{f}^{-1}(v') = f^{-1}(f(P)) + \widetilde{f}^{-1}(\overrightarrow{f(P)Q'}).$$

Luego f^{-1} es una aplicación afín de aplicación lineal asociada $\widetilde{f^{-1}} = \widetilde{f}^{-1}$. □

4.15. Definición. *Una aplicación afín biyectiva se dice que es un isomorfismo afín, o una afinidad o una aplicación equiafín.*

4.16. Proposición-Definición. *El conjunto $GA(E)$ de las afinidades de un espacio afín E en sí mismo es un grupo con respecto a la ley de composición de aplicaciones. Se le llama grupo afín de E.*

Demostración. Por 4.14.a), $\circ$ es una ley de composición interna, es asociativa por serlo la composición de aplicaciones, el elemento neutro es la aplicación identidad, y 4.14.d) demuestra que existe la inversa. □

Obsérvese que el conjunto de las traslaciones $\mathcal{T}(E)$ y el de las homotecias $\mathcal{H}(E)$ de E son subgrupos del grupo $GA(E)$.

?4.17.Proposición. *Sea $GL(V)$ el grupo de los isomorfismos de V en V. La aplicación*

$$\Psi : GA(E) \longrightarrow GL(V)/\Psi(f) = \widetilde{f}$$

es un epimorfismo de grupos con $\operatorname{Ker}\Psi = \mathcal{T}(E)$ *y, por tanto,* $GA(E)/\mathcal{T}(E) \cong GL(V)$.

Demostración. Por 4.14.a) se tiene que $\Psi(f \circ g) = \widetilde{f \circ g} = \widetilde{f} \circ \widetilde{g} = \Psi(f) \circ \Psi(g)$, luego Ψ es un homomorfismo. Por otro lado, dada $\phi \in GL(V)$, si definimos $f \in GA(E)$ por $f(Q) = P + \phi(\overrightarrow{PQ})$, se tiene que $\Psi(f) = \phi$, luego Ψ es sobre. Por otro lado, de la definición 4.9 resulta que $\operatorname{Ker}\Psi = \{f/\widetilde{f} = Id\} = \mathcal{T}(E)$. La parte final de la proposición sale de todo esto aplicando el primer teorema de isomorfismo. $\square$

***4.18.Proposición.** *Dado $P \in E$, el conjunto $GA_P(E)$ de las afinidades de E que dejan P fijo (i.e., tales que $f(P) = P$) es un subgrupo de $GA(E)$ isomorfo a $GL(V)$.*

Demostración. La proposición es consecuencia de que la aplicación

$$\chi : GL(V) \longrightarrow GA(E)/\phi \mapsto f/f(Q) = P + \phi(\overrightarrow{PQ})$$

es un homomorfismo inyectivo cuya imagen es $GA_P(E)$. Vamos a demostrar esto: si $\chi(\phi) = f$ y $\chi(\psi) = g$, entonces

$$\chi(\psi \circ \phi)(Q) = P + (\psi \circ \phi)(\overrightarrow{PQ}) = P + \psi(\phi(\overrightarrow{PQ}))$$

y

$$g \circ f(Q) = g(f(Q)) = g(P + \phi(\overrightarrow{PQ})) = P + \psi(\phi(\overrightarrow{PQ})),$$

luego $\chi(\psi \circ \phi) = g \circ f = \chi(\psi) \circ \chi(\phi)$, es decir, χ es un homomorfismo de grupos.

Si $\chi(\psi) = \chi(\phi)$, entonces, para todo $v \in V$,

$$P + \psi(v) = \chi(\psi)(P + v) = \chi(\phi)(P + v) = P + \phi(v),$$

lo que implica que $\psi(v) = \phi(v)$ para todo $v \in V$, luego $\psi = \phi$, por lo tanto χ es inyectiva.

Sea $f \in GA_P(E)$, entonces f se puede escribir en la forma $f(Q) = f(P) + \widetilde{f}(\overrightarrow{PQ}) = P + \widetilde{f}(\overrightarrow{PQ})$. Como f es biyectiva, $\widetilde{f} \in GL(V)$ (por 4.14.d)), y

$$\chi(\widetilde{f})(Q) = P + \widetilde{f}(\overrightarrow{PQ}) = f(Q),$$

luego $\chi(\widetilde{f}) = f$, y χ es sobre. $\square$

***4.19.Proposición.** *Fijado $P \in E$, todo $f \in GA(E)$ se puede escribir de manera única como la composición de un elemento de $GA_P(E)$ y una traslación. En concreto:*

$$f = T_{\overrightarrow{Pf(P)}} \circ g, \quad \textit{con} \quad g \in GA_P(E) \ \textit{definida por} \ g(Q) = P + \widetilde{f}(\overrightarrow{PQ}).$$

Demostración. $f(Q) = f(P) + \widetilde{f}(\overrightarrow{PQ}) = P + \overrightarrow{Pf(P)} + \widetilde{f}(\overrightarrow{PQ}) = P + \widetilde{f}(\overrightarrow{PQ}) + \overrightarrow{Pf(P)} = T_{\overrightarrow{Pf(P)}}(g(Q))$. Por otro lado, si $v \in V$ y $h \in GA_P(E)$ son tales que $f = T_v \circ h$, entonces

$$P + \widetilde{h}(\overrightarrow{PQ}) + v = T_v \circ h(Q) = f(Q) = P + \widetilde{f}(\overrightarrow{PQ}) + \overrightarrow{Pf(P)}.$$

Cuando $Q = P$, esta igualdad implica $v = \overrightarrow{Pf(P)}$, y sustituyendo esto de nuevo en la igualdad anterior, resulta $\widetilde{f} = \widetilde{h}$, lo que prueba que $\widetilde{g} = \widetilde{h}$, y, por tanto, $g = h$. $\square$

Vamos ahora a estudiar la expresión de una aplicación afín en coordenadas, i.e., la expresión matricial de una aplicación afín.

A partir de ahora E será un espacio afín de dimensión n, $\mathcal{R} = \{O, e_1, ..., e_n\}$ una referencia cartesiana de E, E' un espacio afín de dimensión m, y $\mathcal{R}' = \{O', e'_1, ..., e'_m\}$ una referencia cartesiana de E'. $f : E \longrightarrow E'$ será una aplicación afín. Si S es un punto de un espacio afín E en el que se ha dado una referencia cartesiana $\mathcal{R}$, escribiremos $S = (S^1, ..., S^n)$ para indicar que $(S^1, ..., S^n)$ son las coordenadas del punto S en la referencia $\mathcal{R}$.

4.20.Proposición. *Si $\widetilde{f}(e_i) = \sum_{j=1}^m f_i^j e'_j$, $P = (P^1, ..., P^n) \in E$, $Q = f(P) = (Q^1, ...Q^m) \in E'$ y $f(O) = (\theta^1, ..., \theta^m) \in E'$, entonces*

$$Q^j = \theta^j + \sum_{i=1}^n f_i^j P^i, \quad \text{para } j = 1, ..., m, \tag{4.20.1}$$

o, en forma matricial

$$\begin{pmatrix} Q^1 \\ \vdots \\ Q^m \end{pmatrix} = \begin{pmatrix} f_1^1 & \dots & f_n^1 \\ \vdots & \ddots & \vdots \\ f_1^m & \dots & f_n^m \end{pmatrix} \begin{pmatrix} P^1 \\ \vdots \\ P^n \end{pmatrix} + \begin{pmatrix} \theta^1 \\ \vdots \\ \theta^m \end{pmatrix},$$

o, equivalentemente,

$$\begin{pmatrix} Q^1 \\ \vdots \\ Q^m \\ 1 \end{pmatrix} = \begin{pmatrix} f_1^1 & \dots & f_n^1 & \theta^1 \\ \vdots & \ddots & \vdots & \vdots \\ f_1^m & \dots & f_n^m & \theta^m \\ 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} P^1 \\ \vdots \\ P^n \\ 1 \end{pmatrix}.$$

Las ecuaciones 4.20.1 se llaman ecuaciones de f en las referencia $\mathcal{R}$ y $\mathcal{R}'$, y la matriz

$$\begin{pmatrix} f_1^1 & \dots & f_n^1 & \theta^1 \\ \vdots & \ddots & \vdots & \vdots \\ f_1^m & \dots & f_n^m & \theta^m \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

se llama matriz de f en las referencias $\mathcal{R}$ y $\mathcal{R}'$.

Demostración. De la definición de coordenadas de un punto se tiene que $\overrightarrow{OP} = \sum_{i=1}^n P^i e_i$ y $f(O) = O' + \sum_{j=1}^m \theta^j e'_j$. Por lo tanto

$$\begin{aligned} f(P) &= f(O) + \widetilde{f}(\overrightarrow{OP}) = f(O) + \widetilde{f}(\sum_{i=1}^n P^i e_i) \\ &= O' + \sum_{j=1}^m \theta^j e'_j + \sum_{i=1}^n P^i \sum_{j=1}^m f_i^j e'_j = O' + \sum_{j=1}^m (\theta^j + \sum_{i=1}^n P^i f_i^j) e'_j. \end{aligned}$$

De donde resultan las ecuaciones (4.20.1). □

Recíprocamente

4.21.Proposición. *Sean $M = (a_i^j)_{1 \le i \le n, 1 \le j \le m}$ una matriz $m \times n$ y $b = (b^j)_{1 \le j \le m}$ una matriz $m \times 1$. Existe una única aplicación afín $f : E \longrightarrow E'$ que, en los sistemas de referencia $\mathcal{R}$ y $\mathcal{R}'$, tiene como matriz*

$$\begin{pmatrix} a_1^1 & \dots & a_n^1 & b^1 \\ \vdots & \ddots & \vdots & \vdots \\ a_1^m & \dots & a_n^m & b^m \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

Demostración. Para ver la existencia basta con definir

$$f(P) = (Q^1, ..., Q^m)/ \quad Q^j = b^j + \sum_{i=1}^n a_i^j P^i, \tag{4.21.1}$$

siendo $P = (P^1, ..., P^n)$ y $Q = (Q^1, ..., Q^m)$. Entonces

$$f(P) = f(O) + \widetilde{f}(\overrightarrow{OP}), \text{ siendo } f(O) = O' + \sum_{j=1}^m b^j e'_j \text{ y } \widetilde{f}(e_i) = \sum_{j=1}^m a_i^j e'_j.$$

La unicidad resulta de que si g es una aplicación afín con esa matriz, y se hace actuar sobre un punto $(P^1, ..., P^n)$, el resultado viene dado por las ecuaciones (4.21.1), que son las que definen f, luego $g = f$. □

Veamos ahora cual es la matriz de una composición de aplicaciones afines.

4.22.Proposición. *Sea E'' un espacio afín de dimensión r y $\mathcal{R}'' = \{O'', e''_1, ..., e''_r\}$ una referencia cartesiana de E''. Si $f : E \longrightarrow E'$ y $g : E' \longrightarrow E''$ son aplicaciones afines con matrices F y G respectivamente, en las referencias $\mathcal{R}$, $\mathcal{R}'$, $\mathcal{R}'$, entonces la matriz de $g \circ f$ es $G \circ F$.*

Demostración. Si

$$\widetilde{g}(e'_j) = \sum_{k=1}^{r} g_j^k e''_k \quad \text{y} \quad g(O') = O'' + \sum_{k=1}^{r} \varphi^k e''_k,$$

entonces, usando 4.14.a),

$$\begin{aligned}(g \circ f)(P) &= (g \circ f)(O) + (\widetilde{g} \circ \widetilde{f})(\overrightarrow{OP}) = g(O' + \sum_{j=1}^{m} \theta^j e'_j) + \widetilde{g}(\sum_{j=1}^{m}\sum_{i=1}^{n} P^i f_i^j e'_j) \\ &= g(O') + \widetilde{g}(\sum_{j=1}^{m} \theta^j e'_j) + \sum_{j=1}^{m}\sum_{i=1}^{n} P^i f_i^j \widetilde{g}(e'_j) \\ &= O'' + \sum_{k=1}^{r} \varphi^k e''_k + \sum_{k=1}^{r}\sum_{j=1}^{m} \theta^j g_j^k e''_k + \sum_{k=1}^{r}\sum_{j=1}^{m}\sum_{i=1}^{n} P^i f_i^j g_j^k e''_k) \\ &= O'' + \sum_{k=1}^{r}(\sum_{i=1}^{n}\sum_{j=1}^{m} g_j^k f_i^j P^i + \sum_{j=1}^{m} g_j^k \theta^j + \varphi^k) e''_k).\end{aligned}$$

o, en forma matricial,

$$\begin{pmatrix} Q^1 \\ \vdots \\ Q^r \end{pmatrix} = \begin{pmatrix} g_1^1 & \dots & g_m^1 \\ \vdots & \ddots & \vdots \\ g_1^r & \dots & g_m^r \end{pmatrix} \begin{pmatrix} f_1^1 & \dots & f_n^1 \\ \vdots & \ddots & \vdots \\ f_1^m & \dots & f_n^m \end{pmatrix} \begin{pmatrix} P^1 \\ \vdots \\ P^n \end{pmatrix} + \begin{pmatrix} g_1^1 & \dots & g_m^1 \\ \vdots & \ddots & \vdots \\ g_1^r & \dots & g_m^r \end{pmatrix} \begin{pmatrix} \theta^1 \\ \vdots \\ \theta^m \end{pmatrix} + \begin{pmatrix} \phi^1 \\ \vdots \\ \phi^r \end{pmatrix}$$

o bien

$$\begin{pmatrix} Q^1 \\ \vdots \\ Q^r \\ 1 \end{pmatrix} = \begin{pmatrix} g_1^1 & \dots & g_m^1 & \phi^1 \\ \vdots & \ddots & \vdots & \vdots \\ g_1^r & \dots & g_m^r & \phi^r \\ 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} f_1^1 & \dots & f_n^1 & \theta^1 \\ \vdots & \ddots & \vdots & \vdots \\ f_1^m & \dots & f_n^m & \theta^m \\ 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} P^1 \\ \vdots \\ P^n \\ 1 \end{pmatrix}.$$

□

4.23.Corolario. *Sea $f : E \longrightarrow E'$ un isomorfismo afín y $f^{-1} : E' \longrightarrow E$ su inversa. Sean F y $\widetilde{F}$ sus respectivas matrices en las referencias $\mathcal{R}$ y $\mathcal{R}'$. Se tiene entonces que $\widetilde{F} = F^{-1}$.*

Demostración. De 4.22, como $f \circ f^{-1} = Id = f^{-1} \circ f$, y como la matriz de la aplicación identidad en cualquier referencia es la matriz identidad I, se tiene que $F \circ \widetilde{F} = I = \widetilde{F} \circ F$, lo que da $\widetilde{F} = F^{-1}$, c.q.d. □

4.24. Preguntas y ejercicios.

a) Sean los puntos de E_3 que, en una referencia cartesiana $\{O, e_1, e_2, e_3\}$ tienen las coordenadas

$$P_0 = (1,0,1), \quad P_1 = (0,1,1), \quad P_2 = (1,0,0), \quad P_3 = (-1,1,2),$$

$$Q_0 = (1,2,0), \quad Q_1 = (-1,2,0), \quad Q_2 = (0,2,0), \quad Q_3 = (1,-2,0).$$

Probar que el conjunto $\{P_0, P_1, P_2, P_3\}$ consta de 4 puntos independientes y encontrar la matriz de la aplicación afín $f : E \longrightarrow E$ tal que $f(P_i) = Q_i$, $0 \leq i \leq 3$.

References

[Burgos] J. de Burgos, *Curso de Algebra y Geometria*, Alhambra, Madrid, 1977.

[García y López] J. García García y M. López Pellicer, *álgebra lineal y geometría*, marfil, Alcoy, 1977.

[Mulet] Josep Mulet, *Espais afins i Aplicacions afins*, Burjassot, 1993-94.

[Snapper y Troyer] Ernst Snapper and Robert J. Troyer, *Metric Affine Geomstry*, Academic Press, New York and London, 1971.

GEOMETRÍA PROYECTIVA

En todo este capítulo $I\!\!K$ designará el cuerpo de los números reales $I\!\!R$ o el de los complejos $\mathbb{C}$ indistintamente, salvo cuando se indique expresamente de que cuerpo se trata. Todos los espacios vectoriales considerados serán de dimensión finita.

§1. El espacio proyectivo y los subespacios proyectivos.

1.1. Motivación heurística.

Comentarios a la fig. 1. En el capítulo de Geometría Afín hemos definido las proyecciones paralelas. Otra forma clásica de proyectar es la proyección cónica. Esta forma parte de nuestro universo sensible. Corresponde a la visión monocular del mundo físico, si se admite la propagación rectílinea de la luz, de modo que la sensación de relieve se adquiere por comparación de dos proyecciones cónicas.

Una proyección cónica sobre un hiperplano F de un espacio afín E desde un punto P de E no situado en F es la que a cada punto Q de E (excepto los situados en el hiperplano G paralelo a F pasando por P) le hace corresponder el punto de intersección con F de la recta de E que pasa por P y Q.

Cada recta que pasa por P (excepto las contenidas en G) proyecta sobre un único punto de F. Como límite de las imágenes de rectas próximas, las rectas de G pasando por P proyectarían sobre puntos del infinito que no están en F. Si ampliamos F con esos puntos del infinto obtendríamos un espacio que estaría en biyección con el conjunto de rectas de E pasando por P, y la proyección cónica sería una aplicación bien definida del espacio afín E en ese nuevo espacio.

Comentarios a la fig. 2. Ahora F es un plano afín de un espacio afín E de dimensión 3, y O es un punto de E que no está en F. L y M son dos rectas paralelas contenidas en F. A cada punto del plano F corresponde una recta de E (no contenida en el plano G que pasa por O y es paralelo a F), la que une ese punto con O, y viceversa. La posición límite de las rectas r_m, s_m que se corresponden con los puntos R_m y S_m de las rectas L y M respectivamente, es la recta ℓ paralela a L y a M y que pasa por O. Su correspondiente en F sería el punto del infinito en que se cortarían las rectas paralelas L y M. Se ve así como el conjunto de las rectas que pasan por O está en biyección con el conjunto formado por F más los puntos del infinito en que se cortan las rectas paralelas.

Resulta por tanto natural, para ampliar el espacio afín de dimensión n a un nuevo espacio de la misma dimensión de modo que contenga los "puntos de corte de las rectas

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

paralelas” o tal que permita definir la proyección cónica sobre él, definir ese nuevo espacio como el conjunto de las rectas de un espacio afín de dimensión $n+1$ que pasan por un punto fijo, conjunto que está en biyección con las rectas vectoriales de un espacio vectorial de dimensión $n+1$, por ello vamos a definir:

1.2.Definición. *Sea V un espacio vectorial sobre $\mathbb{K}$. El espacio proyectivo $P(V)$ asociado a V es el conjunto de las rectas vectoriales (subespacios vectoriales de dimensión 1) de V.*

Se define la dimensión de $P(V)$ por $\dim(P(V)) = \dim(V) - 1$.

Si $V = \mathbb{K}^{n+1}$, se escribe $\mathbb{K}P^n$ para designar a $P(\mathbb{K}^{n+1})$, el espacio proyectivo de dimensión n sobre $\mathbb{K}$.

Sea $V' = V - \{0\}$. Para cada $v \in V'$, denotaremos $[v] =< \{v\} >$, i.e., $[v]$ denotará la recta vectorial de V conteniendo a v.. Por definición de $P(V)$, $[v] \in P(V)$.

1.3. Definición. *La aplicación*

$$\pi : V' \longrightarrow P(V) \text{ tal que } \pi(v) = [v],$$

se llama proyección canónica de V' sobre $P(V)$.

1.4.Definición. *Un subconjunto $A \subset P(V)$ se dice que es un subespacio proyectivo o una variedad lineal proyectiva de $P(V)$ sii el conjunto $\hat{A} = \pi^{-1}(A) \cup \{0\}$ es un subespacio vectorial de V. La dimensión de A es $\dim(\hat{A}) - 1$.*

1.5.Nota. *Un subespacio proyectivo A es, él mismo, un espacio proyectivo de dimensión igual a la dimensión de A. En efecto, $P(\hat{A}) = \{[v]/v \in \hat{A} - \{0\}\} = \{[v]/v \in \pi^{-1}(A)\} = \{[v]/\pi(v) \in A\}$.*

Una recta proyectiva R es un (sub)espacio proyectivo de dimensión 1 ($\dim(\hat{R}) = 2$).

Un plano proyectivo Π es un (sub)espacio proyectivo de dimensión 2 ($\dim(\hat{\Pi}) = 3$).

Un hiperplano proyectivo H es un subespacio proyectivo de dimensión $n-1$ ($\dim(\hat{H}) = n$), siendo n la dimensión del espacio proyectivo ambiente.

Un punto $p = \pi(v)$ es un (sub)espacio proyectivo de dimensión 0 ($\dim < \{v\} >= 1$).

***1.6.Proposición.** *Si $\{X_i\}_{i\in I}$ es una familia de subespacios proyectivos de $P(V)$, su intersección es el subespacio proyectivo $P(\cap_{i\in I}\hat{X}_i) = \pi(\cap_{i\in I}\hat{X}_i - \{0\})$.*

Demostración. $[v] \in \cap_{i\in I}X_i$ sii $[v] \in X_i$ para todo $i \in I$ sii $\{\lambda v/\lambda \in I\!\!K\} \subset \hat{X}_i$ para todo $i \in I$ sii $\{\lambda v/\lambda \in I\!\!K\} \subset \cap_{i\in I}\hat{X}_i$ sii $[v] = \pi(v) \in \pi(\cap_{i\in I}\hat{X}_i - \{0\}) = P(\cap_{i\in I}\hat{X}_i)$.

***1.7.Definición.** *Si $A \subset P(V)$, la variedad lineal proyectiva $V_P(A)$ engendrada por A es el menor (respecto de la inclusión) subespacio proyectivo que contiene a A. Es decir, $V_P(A)$ es un subespacio proyectivo de $P(V)$ tal que $A \subset V_P(A)$ y para todo subespacio proyectivo $H \supset A$ se verifica que $H \supset V_P(A)$.*

***1.8.Proposición.** *$V_P(A) = P(< \{\pi^{-1}(A)\} >)$, es decir, $\widehat{V_P(A)} =< \{\pi^{-1}(A)\} >$.*

Demostración. Como $V_P(A) \supset A$, se tiene que $\pi^{-1}(V_P(A)) \supset \pi^{-1}(A)$. Usando esto y la definición de subespacio proyectivo, resulta que $\pi^{-1}(V_p(A)) \cup \{0\}$ es un espacio vectorial conteniendo a $\pi^{-1}(A)$, luego $\pi^{-1}(V_P(A)) \cup \{0\} \supset< \{\pi^{-1}(A)\} >$, de donde $\pi^{-1}(V_P(A)) \supset< \{\pi^{-1}(A)\} > -\{0\}$ y, tomando la imagen por π, $V_P(A) \supset P(< \{\pi^{-1}(A)\} >) \supset A$ (porque $P(< \{\pi^{-1}(A)\} >) = \pi(< \{\pi^{-1}(A)\} > -\{0\}) \supset \pi(\pi^{-1}(A)))$ y, como $V_P(A)$ es el menor de los subespacios proyectivos conteniendo a A, se tiene $V_P(A) = P(< \{\pi^{-1}(A)\} >)$, c.q.d.

1.9.Definición. *$h + 1$ puntos $\{p_1, ..., p_{h+1}\}$ de un espacio proyectivo $P(V)$ se dice que son independientes si la variedad lineal proyectiva engendrada por ellos tiene dimensión h.*

1.10.Proposición. *$h + 1$ puntos $\{p_1, ..., p_{h+1}\}$ de un espacio proyectivo $P(V)$ son independientes sii existen vectores $v_1, ..., v_{h+1}$ linealmente independientes de V tales que $p_i = \pi(v_i)$ para $1 \leq i \leq h + 1$. Si esto ocurre para unos vectores $v_1, ..., v_{h+1}$, también ocurre para cualesquiera otros $w_1, ..., w_{h+1}$ verificando $p_i = \pi(w_i)$ para $1 \leq i \leq h + 1$.*

Demostración. Sea $A = \{p_1, ..., p_{h+1}\}$. Aplicando la proposición 1.8, se tiene que $V_P(A) = P(< \{v_1, ..., v_{h+1}\} >)$, que tiene dimensión h sii $v_1, ..., v_{h+1}$ son linealmente independientes.

1.11.Corolario. *Por dos puntos distintos pasa una única recta proyectiva.*

Demostración. Si $p = [x] \neq [y] = q$, no existe ningún $\lambda \in I\!\!K - \{0\}$ tal que $y = \lambda x$, luego x, y son linealmente independientes y, por 1.10, p y q generan un subespacio proyectivo de dimensión 1.

?1.12.Proposición-Definición. *Sea $\{X_i\}_{i=1}^r$ una familia finita de subespacios proyectivos, se tiene*

$$V_P(\cup_{i=1}^r X_i) = P(\hat{X}_1 + ... + \hat{X}_r).$$

A este subespacio se le llama suma de los subespacios proyectivos $X_1, ..., X_r$.

Demostración. Aplicando 1.8, se tiene que $V_P(\cup_{i=1}^r X_i) = P(< \{\pi^{-1}(\cup_{i=1}^r X_i)\} >) = P(< \{\cup_{i=1}^r \pi^{-1}(X_i)\} >) = P(< \{\cup_{i=1}^r \hat{X}_i\} >) = P(\sum_{i=1}^r \hat{X}_i)$.

1.13.Preguntas.

a) ¿Qué relación hay entre $\hat{A}$ y $\pi^{-1}(A)$?. Si A es un subespacio proyectivo de $P(V)$, ¿qué es $\hat{A}$?.

b) Si R es una recta proyectiva de $P(V)$, ¿qué es su antiimagen en V?.

§2. Relaciones de intersección e incidencia entre algunos subespacios proyectivos.

En todo este apartado, V será un espacio vectorial de dimensión $n+1$. En los siguientes teoremas se ve que, para muchos valores de las dimensiones de los subespacios proyectivos, dichos subespacios necesariamente se cortan, i.e., al contrario que en la geometría afín, no existen (para esas dimensiones) subespacios proyectivos paralelos.

***2.1.Proposición.** *Sean H y R un hiperplano y una recta proyectiva, respectivamente, en el espacio proyectivo $P(V)$. Necesariamente ha de ocurrir que: o bien $R \subset H$ o bien $R \cap H$ es un punto y solo un punto.*

Demostración. $\dim(\hat{H}) = n$, $\dim \hat{R} = 2$, luego $\dim(\hat{H} \cap \hat{R}) = n + 2 - \dim(\hat{H} + \hat{R}) \geq n + 2 - (n+1) = 1$, luego $\dim(\hat{H} \cap \hat{R})$ es 1 o 2. Resulta de aquí que $\hat{H} \cap \hat{R}$ es $\hat{R}$ o una recta vectorial. Como vimos en 1.6, $H \cap R = \pi(\hat{H} \cap \hat{R} - \{0\})$. Esto, junto con la observación anterior, dice que $H \cap R$ es R o un punto, lo que acaba la demostración.

2.2.Corolario. a) *Dos rectas proyectivas distintas en un espacio proyectivo de dimensión 2 tienen siempre un punto en común y uno solo.*

b) En un espacio proyectivo de dimensión 3, toda recta no contenida en un plano corta a ese plano en un punto y solo uno.

? 2.3.Proposición. *Todo plano proyectivo Q no contenido en un hiperplano proyectivo H lo corta en una recta proyectiva. En particular, dos planos proyectivos distintos en un espacio proyectivo de dimensión 3 se cortan en una recta proyectiva.*

Demostración. Es análoga a la dada en 2.1. $\dim(\hat{Q} \cap \hat{H}) = 3 + n - \dim(\hat{Q} + \hat{H}) \geq 3 + n - (n+1) = 2$, luego $\dim(\hat{Q} \cap \hat{H})$ es 2 o 3. Si es 3, $\hat{Q} \subset \hat{H}$ y, por tanto $Q \subset H$, luego, si Q no está contenido en H, entonces $\dim(Q \cap H) = \dim(\hat{Q} \cap \hat{H}) - 1 = 2 - 1 = 1$, i.e. $Q \cap H$ es una recta proyectiva, como queríamos demostrar.

2.4.Preguntas y ejercicios.

a) Sean A y B subespacios proyectivos de $P(V)$. ¿Qué relación hay entre $\pi^{-1}(A) \cap \pi^{-1}(B)$ y $A \cap B$?.

b) Decir (razonándolo) para que dimensiones son ciertas las siguientes afirmaciones:

b1) La intersección de un hiperplano proyectivo y de un plano proyectivo no contenido en él es una recta proyectiva.

b2) La intersección de dos planos proyectivos distintos es una recta proyectiva.

b3) La intersección de dos rectas proyectivas distintas es un punto.

§3. Coordenadas homogéneas en un espacio proyectivo.

3.1.Definición. *Sea $\{e_0, e_1, ..., e_n\}$ una base de V. Se llama coordenadas homogéneas de un punto $p = [x] \in P(V)$ a cualquiera de las $(n+1)$-tuplas que dan las coordenadas $(x^0, x^1, ..., x^n)$ de un vector $x \in \pi^{-1}(p) \subset V$ en la base $\{e_0, e_1, ..., e_n\}$. Es decir, $(x^0, x^1, ..., x^n)$ son las coordenadas homogéneas de $p \in P(V)$ sii $x = x^0 e_0 + x^1 e_1 + ... + x^n e_n \in \pi^{-1}(p)$.*

Como $0 \notin \pi^{-1}(p)$, $(x^0, x^1, ..., x^n) \neq (0, 0, ..., 0)$.

De la definición resulta que si $(x^0, x^1, ..., x^n)$ son coordenadas homogéneas de p, también lo son $(\lambda x^0, \lambda x^1, ..., \lambda x^n)$ para todo $\lambda \in I\!\!K - \{0\}$, y si $(x^0, x^1, ..., x^n)$ y $(y^0, y^1, ..., y^n)$ son coordenadas homogéneas de un mismo punto $p \in P(V)$, entonces existe un $\lambda \in I\!\!K - \{0\}$ tal que $y^j = \lambda x^j$ para $0 \leq j \leq n$.

3.2. El hecho de que un mismo punto de $P(V)$ venga representado, en coordenadas homogéneas, por muchas $(n+1)$-tuplas distintas es realmente un inconveniente para calcular. En muchos casos ese inconveniente se puede evitar de la siguiente manera.

Observemos primero que si $(x^0, x^1, ..., x^n)$ son coordenadas homogéneas de un punto $p \in P(V)$ y $x^j \neq 0$ (resp $x^j = 0$), entonces $y^j \neq 0$ (resp. $y^j = 0$) para cualesquiera otras coordenadas homogéneas $(y^0, ..., y^n)$ del mismo punto p.

Si $(x^0, x^1, ..., x^n)$ son coordenadas homogéneas de un punto $p \in P(V)$, como $(x^0, x^1, ..., x^n) \neq (0, ..., 0)$, existe un $j \in \{0, 1, ..., n\}$ tal que $x^j \neq 0$. La n-tupla $(\frac{x^0}{x^j}, ..., \frac{x^{j-1}}{x^j}, 1, \frac{x^{j+1}}{x^j}, ..., \frac{x^n}{x^j})$ representa el mismo punto p, y es la única de las n-tuplas representando p cuya coordenada j-ésima es 1. Esto da una manera única de representar los puntos p cuya coordenada j-ésima sea distinta de 0 para alguna $(n+1)$-tupla que sea coordenadas homogéneas de p.

Obsérvese que si consideramos este tipo de coordenadas unívocas sobre los puntos con $x^n \neq 0$ (por ejemplo), entonces el procedimiento anterior (dividir por x^n) aplicado a los puntos con $x^n = 0$ da ∞ para sus coordenadas. Estos son los puntos del infinito para este sistema de coordenadas. En el dibujo de $I\!\!RP^2$ como una semiesfera con identificaciones en los puntos del ecuador (ver apéndice), y para la base canónica de $I\!\!R^3$, estos puntos del infinito son los puntos del ecuador. Se corresponden con los puntos que, en la introducción heurística, se añaden a un plano paralelo al plano XY para contener los puntos que serían los de intersección de rectas paralelas.

Vamos a ver ahora cuales serían las ecuaciones, en paramétricas y en implícitas, de una variedad lineal proyectiva. A partir de ahora $\{e_0, e_1, ..., e_n\}$ será una base fijada de

V, y las coordenadas homogéneas de los puntos de $P(V)$ lo serán con respecto a esa base.

3.3.Proposición-Definición. *Sea S un subespacio proyectivo de dimensión d y sea $\{v_0, ..., v_d\}$ una base de $\hat{S}$ tal que $v_i = \sum_{j=0}^n v_i^j e_j$, para $0 \leq i \leq d$, entonces S es el conjunto de puntos $x \in P(V)$ cuyas coordenadas homogéneas $(x^0, ..., x^n)$ verifican*

$$\rho x^j = \sum_{i=0}^d \lambda^i v_i^j, \text{ para todo } \rho, \lambda^0, ..., \lambda^d \in I\!K, \tag{3.3.1}$$

o, en forma matricial

$$\begin{pmatrix} \rho x^0 \\ \vdots \\ \rho x^n \end{pmatrix} = \begin{pmatrix} v_0^0 & \dots & v_d^0 \\ \vdots & \ddots & \vdots \\ v_0^n & \dots & v_d^n \end{pmatrix} \begin{pmatrix} \lambda^0 \\ \vdots \\ \lambda^d \end{pmatrix}.$$

Las ecuaciones (3.3.1) se llaman ecuaciones paramétricas del subespacio proyectivo S.

Demostración. $(x^0, ..., x^n)$ son las coordenadas homogéneas de un punto $[x] \in S$ sii $x = x^0 e_0 + ... + x^n e_n \in \hat{S}$ sii x es combinación lineal de $\{v_0, ..., v_d\}$ sii existen $\lambda^0, ... \lambda^d \in I\!K$ tales que $x = \lambda^0 v_0 + ... + \lambda^d v_d$ sii $x^0 e_0 + ... + x^n e_n = \lambda^0 \sum_{j=0}^n v_0^j e_j + ... + \lambda^d \sum_{j=0}^n v_d^j e_j = \sum_{i=0}^d \lambda^i v_i^0 e_0 + ... + \sum_{i=0}^d \lambda^i v_i^n e_n$, de donde se deduce (3.3.1).

3.4.Corolario. *Sea S el subespacio proyectivo de dimensión d generado por los $d+1$ puntos independientes $\{p_0, ..., p_d\}$y sean $(p_i^0, ..., p_i^n)$ coordenadas homogéneas de p_i, $0 \leq i \leq d$. Entonces S es el conjunto de puntos $x \in P(V)$ cuyas coordenadas homogéneas $(x^0, ..., x^n)$ verifican*

$$\rho x^j = \sum_{i=0}^d \lambda^i p_i^j, \text{ para todo } \rho, \lambda^0, ..., \lambda^d \in I\!K, \tag{3.4.1}$$

o, en forma matricial

$$\begin{pmatrix} \rho x^0 \\ \vdots \\ \rho x^n \end{pmatrix} = \begin{pmatrix} p_0^0 & \dots & p_d^0 \\ \vdots & \ddots & \vdots \\ p_0^n & \dots & p_d^n \end{pmatrix} \begin{pmatrix} \lambda^0 \\ \vdots \\ \lambda^d \end{pmatrix}.$$

Demostración. Basta recordar que $\{p_0, ..., p_d\}$ son independientes sii tomando $0 \neq v_i \in \pi^{-1}(p_i), 1 \leq i \leq d$ (por ejemplo $v_i = \sum_{j=0}^n p_i^j e_j$), $\{v_0, ..., v_d\}$ son linealmente independientes y una base de $\hat{S}$, y aplicar después las ecuaciones (3.3.1).

3.5.Proposición-Definición. *Sea S un subespacio proyectivo de dimensión d. Sean $f^{d+1},...,f^n \in V^*$ linealmente independientes tales que $\hat{S} = \operatorname{Ker} f^{d+1} \cap ... \cap \operatorname{Ker} f^n$ (cfr. apuntes de Geometría Afín, 2.6 y demostración de 2.7, y Práctica 0). Si $f^j(e_i) = a_i^j$, $0 \leq i \leq n$ y $d+1 \leq j \leq n$, entonces S es el conjunto de puntos $x \in P(V)$ cuyas coordenadas homogéneas $(x^0,...,x^n)$ verifican*

$$\left.\begin{array}{ll} a_0^{d+1}x^0 + ... + a_n^{d+1}x^n & = 0 \\ \cdots\cdots\cdots\cdots & \cdots \\ a_0^n x^0 + ... + a_n^n x^n & = 0 \end{array}\right\} \tag{3.5.1}$$

Las ecuaciones (3.5.1) se llaman ecuaciones en forma implícita del subespacio proyectivo S. Obsérvese que coinciden con las ecuaciones en forma implícita de $\hat{S}$.

3.6. Pregunta.. *Sea V un espacio vectorial, $v \in V$, $\{e_1,...,e_n\}$ una base de V, $p \in P(V)$ tal que $v \in \pi^{-1}(p)$, siendo $\pi : V - \{0\}$ la proyección canónica. ¿Qué relación y diferencia hay entre p, $[v]$ y $\pi^{-1}(p)$?. Si $v = \sum_{i=1}^n v^i e_i$, ¿qué relación y diferencia hay entre $(v^1,...,v^n)$ y las coordenadas homogéneas de p?.*

§4 Proyectividades

En este apartado V y W serán dos espacios vectoriales de dimensiones n y m respectivamente, sobre un mismo cuerpo $I\!\!K$.

***4.1.Definición.** *Sea S un subespacio proyectivo de $P(V)$. Se llama aplicación proyectiva (o proyectividad) de $P(V)$ en $P(W)$ de centro S a toda aplicación $f : P(V) - S \longrightarrow P(W)$ tal que existe una aplicación lineal $\widetilde{f} : V \longrightarrow W$ tal que $\operatorname{Ker} \widetilde{f} = \widehat{S}$ y $f(\pi(v)) = \pi(\widetilde{f}(v))$. Se dice que $\widetilde{f}$ es una aplicación lineal asociada a la proyectividad f.*

Quizás se entienda mejor 4.1 si se indica que lo que dice es que exite una $\widetilde{f}$ tal que el diagrama

$$\begin{array}{ccc} V - \widehat{S} & \xrightarrow{\widetilde{f}} & W - \{0\} \\ {\scriptstyle \pi}\Big\downarrow & & {\scriptstyle \pi}\Big\downarrow \\ P(V) - S & \xrightarrow{f} & P(W) \end{array}$$

es conmutativo.

4.2.Nota. *Obsérvese que para cualquier aplicación lineal $\widetilde{f} : V \longrightarrow W$, tomando $S = \pi(\operatorname{Ker} \widetilde{f} - \{0\})$, se tiene que $f : P(V) - S \longrightarrow P(W)$ dada por $f(\pi(v)) = \pi(\widetilde{f}(v))$ es una aplicación bien definida y, por tanto, una aplicación proyectiva. Es decir, a cada aplicación lineal de V en W corresponde una proyectividad de $P(V)$ en $P(W)$.*

En efecto. Dado un $\lambda v \in \pi^{-1}(\pi(v))$ arbitrario, se tiene que $\pi(\widetilde{f}(\lambda v)) = \pi(\lambda \widetilde{f}(v)) = \pi(\widetilde{f}(v))$, luego $f(v)$ no depende del elemento de $[v]$ elegido.

***4.2'.Nota.** *Si $\widetilde{f}$ es una aplicación lineal asociada a f también lo es $\lambda\widetilde{f}$. Si φ es otra aplicación lineal asociada a f, entonces existe un $\lambda \in I\!K - \{0\}$ tal que $\varphi = \lambda\widetilde{f}$.*

En efecto. $\pi(\widetilde{f}(v)) = f(\pi(v)) = \pi(\varphi(v))$, luego existe $\lambda_v \in I\!K$ tal que $\varphi(v) = \lambda_v\widetilde{f}(v)$. Consideremos dos casos:

- Supongamos primero que v y w son dos elementos de V tales que $\widetilde{f}(v)$ y $\widetilde{f}(w)$ son l.i.. Se tiene:

$$\lambda_{v+w}(\widetilde{f}(v) + \widetilde{f}(w)) = \lambda_{v+w}\widetilde{f}(v+w) = \varphi(v+w) = \varphi(v) + \varphi(w) = \lambda_v\widetilde{f}(v) + \lambda_w\widetilde{f}(w),$$

que, por la independencia lineal de $\widetilde{f}(v)$ y $\widetilde{f}(w)$, implican $\lambda_v = \lambda_{v+w} = \lambda_w$.

-Supongamos ahora que v y w son dos elementos de V tales que $\widetilde{f}(v)$ y $\widetilde{f}(w)$ son l.d.. Entonces existe un $\mu \in I\!K$ tal que $\widetilde{f}(w) = \mu\widetilde{f}(v)$, i.e., $w - \mu v \in \operatorname{Ker}\widetilde{f} = \widehat{S} = \operatorname{Ker}\varphi$, luego $\varphi(w) = \mu\varphi(v)$ y

$$\mu\lambda_v\widetilde{f}(v) = \mu\varphi(v) = \varphi(w) = \lambda_w\widetilde{f}(w) = \lambda_w\mu\widetilde{f}(v),$$

lo que implica $\lambda_v = \lambda_w$.

Luego, para cualesquiera v, $w \in V$ se tiene que $\lambda_v = \lambda_w$, i.e., existe una $\lambda \in I\!K$ independiente de $v \in I\!K$ tal que $\varphi(v) = \lambda\widetilde{f}(v)$.

***4.3.Proposición.** *Sean f una proyectividad de $P(V)$ en $P(W)$, y $\widetilde{f} : V \longrightarrow W$ su aplicación lineal asociada. Se verifican:*

a) Im$(f) = \pi(\mathrm{Im}(\widetilde{f}) - \{0\})$.

b) f es inyectiva sii $\widetilde{f}$ es inyectiva (entoces f es una proyectividad de centro vacío).

c) f es suprayectiva sii $\widetilde{f}$ es suprayectiva.

Demostración. a) Si S es el centro de f, $\mathrm{Im}f = \{f([v])/[v] \in P(V) - S\} = \{f(\pi(v))/v \in V - \operatorname{Ker}\widetilde{f}\} = \{\pi(\widetilde{f}(v))/v \in V - \operatorname{Ker}\widetilde{f}\} = \{\pi(\widetilde{f}(v))/v \in V, \widetilde{f}(v) \neq 0\} = \pi(\mathrm{Im}\widetilde{f} - \{0\})$.

b) Supongamos primero que $\widetilde{f}$ es inyectiva. $f(\pi(x)) = f(\pi(y))$ sii $\pi(\widetilde{f}(x)) = \pi(\widetilde{f}(y))$ sii existe $\lambda \in I\!K$ tal que $\widetilde{f}(y) = \lambda\widetilde{f}(x) = \widetilde{f}(\lambda x)$, lo que implica (por ser $\widetilde{f}$ inyectiva) que $y = \lambda x$ y $\pi(x) = \pi(y)$.

Si suponemos que es f inyectiva, sean $x, y \in V - \operatorname{Ker}\widetilde{f}$. $\widetilde{f}(x) = \widetilde{f}(y)$ implica $\pi(\widetilde{f}(x)) = \pi(\widetilde{f}(y))$, de donde $f(\pi(x)) = f(\pi(y))$ y (como f es inyectiva) $\pi(x) = \pi(y)$ y, por tanto, existe un $\lambda \in I\!K - \{0\}$ tal que $y = \lambda x$ que, sustituido en la igualdad inicial, da $\widetilde{f}(x) = \widetilde{f}(\lambda x) = \lambda\widetilde{f}(x)$, de donde $(\lambda - 1)\widetilde{f}(x) = 0$, lo que implica $\lambda = 1$, ya que $x \notin \operatorname{Ker}\widetilde{f}$. Tenemos por tanto que $\widetilde{f}$ es inyectiva sobre $V - \operatorname{Ker}\widetilde{f}$. Supongamos que existe $a \in \operatorname{Ker}\widetilde{f}$, $a \neq 0$. Sea $x \in V - \operatorname{Ker}\widetilde{f}$, entonces $x + a$ y $x - a$ no pertenecen a $\operatorname{Ker}\widetilde{f}$ (porque $\widetilde{f}(x+a) = \widetilde{f}(x) + \widetilde{f}(a) = \widetilde{f}(x) \neq 0$ y $\widetilde{f}(x-a) = \widetilde{f}(x) - \widetilde{f}(a) = \widetilde{f}(x) \neq 0$) y son distintos, luego, por lo que acabamos de probar, $\widetilde{f}(x+a) \neq \widetilde{f}(x-a)$. Por otra parte $\widetilde{f}(x+a) - \widetilde{f}(x-a) = \widetilde{f}(2a) = 0$, lo que está en contradicción con la desigualdad anterior, luego $\operatorname{Ker}\widetilde{f} = \{0\}$ y $\widetilde{f}$ es inyectiva.

c) Si $\widetilde{f}$ es suprayectiva, $\mathrm{Im}(\widetilde{f}) = W$, y usando a), $\mathrm{Im}(f) = \pi(\mathrm{Im}(\widetilde{f}) - \{0\}) = \pi(W - \{0\}) = P(W)$, luego f es suprayectiva.

Si f es suprayectiva, $\mathrm{Im}(f) = P(W)$ y, usando a), $\pi(\mathrm{Im}(\widetilde{f}) - \{0\}) = P(W)$, de donde resulta que $\mathrm{Im}(\widetilde{f}) = P(\hat{W}) = W$, y $\widetilde{f}$ es suprayectiva.

4.3'.Proposición. *Si f es una proyectividad de $P(V)$ en $P(W)$ de centro S y A es un subespacio proyectivo de $P(V)$, entonces $f|_A$ es una proyectividad de A en $P(W)$ de centro $A \cap S$.*

Demostración. La aplicación $\widetilde{f}|_{\hat{A}} : \hat{A} \longrightarrow W$ es una aplicación lineal cuyo núcleo es $\hat{A} \cap \hat{S}$. Además $f|_A(\pi(a)) = \pi(\widetilde{f}|_{\hat{A}}(a))$ para todo $a \in \hat{A}$, luego $f|_A$ es una aplicación proyectiva de aplicación lineal asociada $\widetilde{f}|_{\hat{A}}$ y centro $S \cap A$.

4.4.Corolario. *La imagen de un subespacio proyectivo A por una proyectividad de centro S tal que $S \cap A = \emptyset$ es un subespacio proyectivo.*

Demostración. De 4.3' y 4.3.a) se deduce que $f(A) = \pi(\widetilde{f}(\hat{A}) - \{0\})$ y, como $\widetilde{f}$ es lineal, $\widetilde{f}(\hat{A})$ es un subespacio vectorial de W, y $\pi(\widetilde{f}(\hat{A}) - \{0\}) = P(\widetilde{f}(\hat{A}))$.

***4.5.Proposición.** *Sea f una proyectividad de $P(V)$ en $P(W)$ de centro S y g una proyectividad de $P(W)$ en $P(U)$ de centro R. Si $Im(f) \subset P(W) - R$, entonces $g \circ f$ es una proyectividad de $P(V)$ en $P(U)$ de centro S con aplicación lineal asociada $\widetilde{g} \circ \widetilde{f}$.*

Demostración. Como $\widetilde{f}$ y $\widetilde{g}$ son lineales, $\widetilde{g} \circ \widetilde{f}$ es lineal, y se tiene que

$$(4.5.1) \qquad \mathrm{Ker}(\widetilde{g} \circ \widetilde{f}) = \{v \in V / \widetilde{g}(\widetilde{f}(v)) = 0\} = \{v \in V / \widetilde{f}(v) \in \mathrm{Ker}\, \widetilde{g}\} = \widetilde{f}^{-1}(\mathrm{Ker}\, \widetilde{g}),$$

y, como $Im(f) = \pi(\widetilde{f}(V) - \{0\}) \subset P(W) - R$, se tiene que $Im(\widetilde{f}) \cap \hat{R} = \{0\}$, de donde resulta, como $\hat{R} = \mathrm{Ker}\, \widetilde{g}$, que $\widetilde{f}^{-1}(0) = \widetilde{f}^{-1}(\mathrm{Ker}\, \widetilde{g})$, lo que, sustituido en (4.5.1), da $\mathrm{Ker}(\widetilde{g} \circ \widetilde{f}) = \mathrm{Ker}\, \widetilde{f} = \hat{S}$.

Además, como $\widetilde{f}$ y $\widetilde{g}$ son las aplicaciones lineales asociadas de f y g respectivamente, entonces

$$g(f(\pi(v))) = g(\pi(\widetilde{f}(v))) = \pi(\widetilde{g}(\widetilde{f}(v))) \text{ para todo } \pi(v) \in P(V) - S,$$

lo que acaba la demostración de que $g \circ f$ es una proyectividad de centro S y aplicación lineal asociada $\widetilde{g} \circ \widetilde{f}$.

4.6.Definición. *Una proyectividad biyectiva se llama una homografía.*

***4.7.Proposición.** *La inversa f^{-1} de una homografía f es una homografía y $\widetilde{f^{-1}} = \widetilde{f}^{-1}$.*

Demostración. Sea g la proyectividad definida por $\widetilde{f}^{-1}$ (i.e. $g : P(W) \longrightarrow P(V)$ tal que $g(\pi(w)) = \pi(\widetilde{f}^{-1}(w))$. Aplicando 4.5 resulta que $g \circ f(\pi(v)) = g(\pi(\widetilde{f}(v))) = \pi(\widetilde{f}^{-1}(\widetilde{f}(v))) = \pi(v)$, y $f \circ g(\pi(w)) = f(\pi(\widetilde{f}^{-1}(w))) = \pi(\widetilde{f}(\widetilde{f}^{-1}(w))) = \pi(w)$, luego $g = f^{-1}$, luegof^{-1} es una homografía con isomorfismo asociado $\widetilde{f}^{-1}$.

***4.8.Proposición.** *a) El conjunto de homografías de $P(V)$ en si mismo, con la ley de composición de aplicaciones, es un grupo, lo denotaremos por $PGL(V)$. b)La aplicación $\chi : GL(V) \longrightarrow PGL(V)$ definida por $\chi(\widetilde{f})(\pi(v)) = \pi(\widetilde{f}(v))$ es un epimorfimo (homomorfismo suprayectivo) de grupos cuyo núcleo es el conjunto aplicaciones que son múltiplo de la identidad.*

Demostración. a) Que $PGL(V)$ es un grupo es consecuencia de 4.5 y 4.7. El elemento neutro es la aplicación identidad.

b) Veamos que χ es suprayectiva: Dado $f \in PGL(V)$, sea $\widetilde{f} : V \longrightarrow V$ tal que $f(\pi(v)) = \pi(\widetilde{f}(v))$, se tiene que $\chi(\widetilde{f})(\pi(v)) = \pi(\widetilde{f}(v)) = f(\pi(v))$, i.e. $\chi(\widetilde{f}) = f$, luego χ es suprayectiva. Que es un homomorfismo de grupos, $\chi(\widetilde{g} \circ \widetilde{f}) = \chi(\widetilde{g}) \circ \chi(\widetilde{f})$, es consecuencia de la proposición 4.5.

Que el núcleo de χ es el conjunto de las aplicaciones que son múltiplo de la identidad es consecuencia de lo que dijimos en la nota 4.2': puesto que la identidad Id en V es una aplicación lineal asociada a la identidad en $P(V)$, cualquier otra aplicación lineal asociada a la misma homografía es de la forma λId.

Vamos a ver ahora cuales son las ecuaciones (o la matriz) de una proyectividad:

4.9.Proposición. *Sean V y W espacios vectoriales de dimensiones $n+1$ y $m+1$ respectivamente. Sean $\{e_0, ..., e_n\}$ una base de V y $\{w_0, ..., w_m\}$ una base de W. Sea f una proyectividad de $P(V)$ en $P(W)$ de centro S, y $\widetilde{f} : V \longrightarrow W$ una aplicación lineal asociada a f. Sean $a_i^j \in I\!\!K$, $0 \leq i \leq n$, $0 \leq j \leq m$ los escalares definidos por $\widetilde{f}(e_i) = \sum_{j=0}^m a_i^j w_j$. Si $(x^0, ..., x^n)$ son coordenadas proyectivas de $x \in P(V)$ y $(y^0, ..., y^m)$ de $f(x)$, entonces*

$$\text{existe } \rho \in I\!\!K - \{0\} \text{ tal que } \rho\, y^j = \sum_{i=0}^n a_i^j x^i.$$

Dicho de otra manera: La matriz de f es, salvo el producto por un escalar $\rho \in I\!\!K - \{0\}$, la matriz de $\widetilde{f}$.

Demostración. $f(x) = f(\pi(\sum_{i=0}^n x^i e_i)) = \pi(\widetilde{f}(\sum_{i=0}^n x^i e_i)) = \pi(\sum_{i=0}^n x^i \widetilde{f}(e_i)) = \pi(\sum_{i=0}^n x^i \sum_{j=0}^m a_i^j w_j) = \pi(\sum_{j=0}^m(\sum_{i=0}^n x^i a_i^j) w_j)$, luego $(\sum_{i=0}^n a_i^0 x^i, ..., \sum_{i=0}^n a_i^m x^i)$ son coordenadas proyectivas de $f(x)$.

De otra manera: $f([x]) = [\widetilde{f}(x)]$, luego las coordenadas homogéneas de $f[x])$ son las componentes de $\widetilde{f}(x)$, salvo el producto por un escalar, luego la matriz de f es la de $\widetilde{f}$ excepto el producto por un escalar.

4.10.Definición. *En las condiciones de la proposición 4.9, la matriz*

$$\lambda \begin{pmatrix} a_0^0 & \dots & a_n^0 \\ \vdots & \ddots & \vdots \\ a_0^m & \dots & a_n^m \end{pmatrix}$$

se llama matriz de la proyectividad f respecto de las bases $\{e_0, ...e_n\}$ y $\{w_0, ..., w_m\}$. Obsérvese que, salvo el escalar arbitrario no nulo λ, coincide con la matriz de $\widetilde{f}$.

***4.11.Definición.** *Se llama referencia proyectiva de $P(V)$ a un conjunto de $n+2$ puntos $\{p_0, ..., p_{n+1}\}$ tal que cualesquiera $n+1$ puntos de ese conjunto sean independientes.*

***4.12.Proposición.** *Si $\{p_0, ..., p_{n+1}\}$ es una referencia proyectiva de $P(V)$, existe una base $\{e_0, ..., e_n\}$ de V única salvo el producto por un escalar, tal que $\pi(e_i) = p_i$ para $0 \leq i \leq n$ y $\pi(e_0 + ... + e_n) = p_{n+1}$.*

Demostración. Elijamos $v_i \in \pi^{-1}(p_i)$, $0 \leq i \leq n+1$. Por ser $\{p_0, ..., p_{n+1}\}$ una referencia proyectiva, se tiene que $\{v_0, ..., v_n\}$ es una base de V y, por lo tanto, se puede escribir $v_{n+1} = \sum_{i=0}^{n} \mu^i v_i$. Definamos $e_i = \mu^i v_i$ para $0 \leq i \leq n$. Se tiene que $\pi(e_i) = \pi(v_i) = p_i$ y $\pi(e_0 + ... + e_n) = \pi(\mu^0 v_0 + ... + \mu^n v_n) = \pi(v_{n+1}) = p_{n+1}$, luego la base $\{e_0, ..., e_n\}$ de V verifica las condiciones deseadas. Veamos que es única salvo el producto por un escalar. Si $\{w_0, ..., w_n\}$ es otra base verificando las mismas condiciones, entonces $\pi(w_i) = \pi(e_i)$, lo que implica que existe un ν^i tal que $w_i = \nu^i e_i$ para cada $i \in \{0, ..., n\}$. Además, se ha de verificar $\pi(e_0 + ... + e_n) = p_{n+1} = \pi(w_0 + ... + w_n) = \pi(\nu^0 e_0 + ... + \nu^n e_n)$, lo que implica que existe un λ tal que $\nu^0 e_0 + ... + \nu^n e_n = \lambda(e_0 + ... + e_n)$, de donde $(\nu^0 - \lambda)e_0 + ... + (\nu^n - \lambda)e_n = 0$, lo que implica $\nu^i = \lambda$ para todo $i \in \{0, ..., n\}$, i.e. $w_i = \lambda e_i$ para todo $i \in \{0, ..., n\}$, c.q.d.

? 4.13.Definición. *Se llama coordenadas homogéneas de $p \in P(E)$ en la referencia proyectiva $\{p_0, ..., p_{n+1}\}$ a las coordenadas homogéneas de p con respecto a una de las bases $\{e_0, ..., e_n\}$ de V definidas por la referencia proyectiva del modo que se vió en 4.12.*

***4.14.Proposición.** *Si $m = n$, dadas dos referencias proyectivas $\{p_0, ..., p_{n+1}\}$ de $P(V)$ y $\{q_0, ..., q_{n+1}\}$ de $P(W)$, existe una única homografía $f : P(V) \longrightarrow P(W)$ tal que $f(p_i) = q_i$ para $0 \leq i \leq n+1$.*

Demostración. Sean $\{e_0, ..., e_n\}$ y $\{w_0, ..., w_n\}$ bases de V y W respectivamente, definidas a partir de las referencias proyectivas dadas como en 4.12. Definamos $\widetilde{f} : V \longrightarrow W$ por $\widetilde{f}(\sum_{i=0}^{n} \mu^i e_i) = \sum_{i=0}^{n} \mu^i w_i$, y $f : P(V) \longrightarrow P(W)$ por $f(\pi(v)) = \pi(\widetilde{f}(v))$. De la definición de f, por ser $\widetilde{f}$ un isomorfismo, resulta que f es una homografía de aplicación lineal asociada $\widetilde{f}$. Además , para $0 \leq i \leq n$, se tiene $f(p_i) = f(\pi(e_i)) = \pi(\widetilde{f}(e_i)) = \pi(w_i) = q_i$, y $f(p_{n+1}) = f(\pi(e_0 + ... + e_n)) = \pi(\widetilde{f}(e_0 + ... + e_n)) = \pi(w_0 + ... + w_n) = q_{n+1}$.

Si g es otra homografía verificando $g(p_i) = q_i$, entonces, para cualquier $\widetilde{g} : V \longrightarrow W$ aplicación lineal asociada a g, se tiene

$$q_i = g(p_i) = g(\pi(e_i)) = \pi(\widetilde{g}(e_i)) \text{ para } 0 \leq i \leq n,$$

$$q_{n+1} = g(p_{n+1}) = g(\pi(e_0 + ... + e_n)) = \pi(\widetilde{g}(e_0) + ... + \widetilde{g}(e_n)),$$

de donde, por 4.12, resulta que $\widetilde{g}(e_i) = \lambda w_i = \lambda \widetilde{f}(e_i)$ para $0 \leq i \leq n+1$ y, por lo tanto, $\widetilde{g} = \lambda \widetilde{f}$, y $\widetilde{g}$ es una aplicación lineal asociada a f, luego $f = g$, luego f es única.

APÉNDICE:
El espacio proyectivo como espacio cociente. Visualización de $\mathbb{R}P^2$.

A continuación vamos a dar una definición (equivalente) más manejable de espacio proyectivo. Para ello necesitamos primero algunos conceptos de teoría de conjuntos.

A.1.Definición. *Dado un conjunto X, una relación binaria R sobre X es un subconjunto $\mathcal{R} \subset X \times X$, se escribe*

$$x\,R\,y \text{ si y solo si } (x, y) \in \mathcal{R}.$$

Si no es cierto que $x\,R\,y$, escribiremos $x \not R\, y$

A.2.Definición. *Una relación binaria de equivalencia (R.B.E.) sobre un conjunto X es una relación binaria R sobre X que cumple las propiedades*
-Reflexiva: $x\,R\,x$.
-Simétrica: $x\,R\,y$ sii $y\,R\,x$.
-Transitiva: $x\,R\,y$ y $y\,R\,z$ implican $x\,R\,z$.

A.3.Definición. *Una partición de un conjunto X es una familia $\{X_i\}_{i\in I}$ de subconjuntos de X disjuntos y tales que $X = \cup_{i\in I} X_i$.*

A.4.Definición. *Si R es una relación binaria de equivalencia sobre X, y $x \in X$, se llama clase de equivalencia de x al conjunto $[x] = \{y \in X/\, x\,R\,y\}$.*

A.5.Proposición-Definición. *Dada una una relación binaria de equivalencia R sobre X, las clases de equivalencia de los elementos de X definen una partición de X. Al conjunto cuyos elementos son las clases de equivalencia se le llama conjunto cociente de X por la relación de equivalencia R, y se denota por X/R.*

Demostración. Evidentemente $X = \cup_{x\in X}[x]$, luego, si vemos que

$$[x] \neq [y] \text{ sii } [x] \cap [y] = \emptyset, \tag{A.5.1}$$

las clases de equivalencia serán disjuntas y la familia de las clases de equivalencia será una partición de X. Vamos, pues, a probar (A.5.1). Si $[x]\cap[y] \neq \emptyset$, existe un $z \in [x]\cap[y]$, (i.e., existe un z tal que $x\,R\,z$ y $y\,R\,z$) y, aplicando las propiedades simétrica y transitiva, $x\,R\,y$. Entonces, para cualquier $w \in [y]$, se tiene que $x\,R\,y$ y $y\,R\,w$, luego $x\,R\,w$ y $w \in [x]$, luego $[y] \subset [x]$. De la misma manera se ve que $[x] \subset [y]$. Hemos visto, por tanto, que $[x] \cap [y] \neq \emptyset$ implica $[x] = [y]$, luego $[x] \neq [y]$ implica $[x] \cap [y] = \emptyset$. Recíprocamente, es evidente que $[x] = [y]$ implica $[x]\cap[y] \neq \emptyset$, luego $[x]\cap[y] = \emptyset$ implica $[x] \neq [y]$, y (A.5.1) queda probado.

Volvemos ahora a usar la notación de la definición 1.2.

A.6.Proposición. *Sea $\sim$ la relación binaria sobre $V' = V - \{0\}$ definida por*

$$x \sim y \text{ sii existe un } \lambda \in I\!K - \{0\} \text{ tal que } y = \lambda x.$$

$\sim$ es una relación binaria de equivalencia

Demostración.
- $x \sim x$ porque $x = 1\,x$.
-$x \sim y$ implica que existe $\lambda \in I\!K - \{0\}$ tal que $y = \lambda x$, luego existe $\lambda^{-1} \in I\!K - \{0\}$ tal que $x = \lambda^{-1} y$, luego $y \sim x$.
- $x \sim y$ y $y \sim z$ implican que existen λ y μ en $I\!K - \{0\}$ tales que $y = \lambda x$ y $z = \mu y$, de donde resulta que existe $\lambda\mu \in I\!K - \{0\}$ tal que $z = \lambda\mu y$ y, por tanto, $x \sim z$.

A.7.Proposición. *$P(V)$ está en biyección con $V'/\sim$.*
Por ello, de ahora en adelante se identificarán ambos conjuntos.

Demostración. Definamos la aplicación

$$\phi : V'/\sim \longrightarrow P(V) \text{ tal que } \phi([v]) = \{\lambda v : \lambda \in I\!K\}.$$

ϕ está bien definida, pues si $w \sim v$, existe $\mu \in I\!K$ tal que $v = \mu w$, y $\{\lambda v : \lambda \in I\!K\} = \{\lambda\mu w : \lambda \in I\!K\} = \{\nu w : \nu \in I\!K\}$. Luego ϕ no depende el elemento de $[v]$ elegido para definir la recta vectorial imagen por ϕ.

Definamos ahora

$$\psi : P(V) \longrightarrow V'/\sim \ \text{ tal que } \psi(U) = [u], \text{ siendo } u \in U - \{0\}.$$

Esta aplicación también está bien definida, es decir, no depende del elemento u de U elegido. En efecto, si $w \in U - \{0\}$, como $\dim(U) = 1$, existe un $\lambda \in I\!K - \{0\}$ tal que $w = \lambda u$, luego $u \sim w$ y $[u] = [w]$. Vamos a ver ahora que ϕ y ψ son aplicaciones inversa una de la otra, lo que prueba que son biyecciones y que, por lo tanto, la proposición es cierta. Componiendo estas aplicaciones, tenemos:

$$\psi \circ \phi([v]) = \psi(\{\lambda v : \lambda \in I\!K\}) = [v] \text{ y,}$$
$$\text{si } u \in U - \{0\},\ \phi\psi(U) = \phi([u]) = \{\lambda u : \lambda \in I\!K\} = U.$$

Para el caso $I\!K = I\!R$, vamos a ver otro modelo equivalente del espacio proyectivo.

Sea $|\ |$ la norma canónica de $I\!R^{n+1}$, es decir, dado $x = (x^1, ..., x^{n+1}) \in I\!R^{n+1}$, $|x|^2 = (x^1)^2 + ... + (x^{n+1})^2$. La esfera de radio 1 en $I\!R^{n+1}$ es el conjunto $S^n = \{x \in I\!R^{n+1} : |x| = 1\}$.

La relación de equivalencia $\sim$ definida sobre $I\!R^{n+1} - \{0\}$ como en A.6 se puede restringir a S^n dando una nueva relación de equivalencia, que seguiremos denotando por $\sim$. Obsérvese que:

$$\text{Si } x,\, y \in S^n, \quad x \sim y \text{ sii } y = x \text{ ó } y = -x,$$

puesto que $x \sim y$ sii $y = \lambda x$, y $y \in S^n$ implica $1 = |\lambda x| = |\lambda||x| = |\lambda|$.

A.8.Proposición. *$I\!\!RP^n$ está en biyección con $S^n/\sim$.*

Demostración. La aplicación

$$\xi : (I\!\!R^{n+1} - \{0\})/\sim \longrightarrow S^n/\sim \text{ dada por } \xi([x]) = \left[\frac{x}{|x|}\right]$$

está bien definida y es una biyección. En efecto: si $y \in [x]$, $y = \lambda x$, luego

$$\frac{y}{|y|} = \frac{\lambda x}{|\lambda||x|} = \frac{\lambda}{|\lambda|}\frac{x}{|x|} \text{ lo que implica } \frac{x}{|x|} \sim \frac{y}{|y|},$$

luego ξ no depende del elemento de $[x]$ elegido. Además se comprueba fácilmente que la aplicación de $S^n/\sim$ en $(I\!\!R^{n+1} - \{0\})/\sim$ definida por $[u] \mapsto [u]$ es su inversa, luego ξ es una biyección, y esto prueba la proposición.

La última proposición da la imagen de $I\!\!RP^n$ como una semiesfera de S^n en la que se han identificado los puntos opuestos del ecuador. En los casos $n = 1, 2$ tenemos los siguientes dibujos

En el caso n=1 se ve que $I\!\!RP^1$ está en biyección con S^1.

Denotaremos por $P(V)$ tanto el conjunto de la definición 1.2 como $V'/\sim$. Análogamente, $I\!\!RP^n$ denotará indistintamente $P(I\!\!R^{n+1})$, $(I\!\!R^{n+1} - \{0\})/\sim$ y $S^n/\sim$.

En el dibujo de $I\!\!RP^2$ a continuación de A.8, se puede observar que “las rectas proyectivas (como se definieron en 1.5) no son rectas, sino circunferencias”, y que (en $I\!\!RP^2$) siempre se cortan.

References

[Burgos] J. de Burgos, *Curso de Algebra y Geometria*, Alhambra, Madrid, 1977.
[Donnedu] A. Donnedu, *Compléments de Géométrie Algébrique*, Dunod, Paris, 1968.
[Frenkel] J. Frenkel, *Géométrie pour l'élève-professeur*, Hermann, Paris, 1977.
[Jennings] George A. Jennings, *Modern Geometry with Applications*, Springer-Verlag. Universitext, New York, Berlin,..., Barcelona, 1994.
[Tisseron] C. Tisseron, *Géométries affine, projective et euclidienne*, Hermann, Paris, 1983.

GEOMETRÍA AFÍN MÉTRICA

En todo este capítulo V designará un espacio vectorial de dimensión n sobre el cuerpo $\mathbb{R}$ de los números reales.

§0. Algunas cuestiones de espacios vectoriales euclídeos

Para estudiar la geometría afín métrica necesitamos conocer unas cuestiones de geometría vectorial euclídea que estudiamos en este epígrafe.

0.1.Definición. *Una forma bilineal sobre V es una aplicación $g : V \times V \longrightarrow \mathbb{R}$ que es lineal en cada componente, es decir*

$$g(\lambda v + \mu w,\, x) = \lambda g(v,\, x) + \mu g(w,\, x) \text{ para todo } \lambda, \mu \in \mathbb{R} \text{ y todo } v, w, x \in V, \quad \text{y}$$

$$g(v,\, \lambda x + \mu y) = \lambda g(v,\, x) + \mu g(v,\, y) \text{ para todo } \lambda, \mu \in \mathbb{R} \text{ y todo } v, x, y \in V.$$

Obsérvese que para toda forma bilineal g se tiene $g(0, y) = g(x, 0) = 0$ para cualesquiera $x,\, y \in V$.

0.2.Definiciones. *Una forma bilineal g sobre V es simétrica si $g(x, y) = g(y, x)$ para todo $x,\, y \in V$. Se dice que g es definida positiva si $g(x, x) > 0$ para todo $x \in V - \{0\}$.*

Obsérvese que se deduce de la definición que para una forma bilineal simétrica definida positiva $g(x, x) = 0$ sii $x = 0$.

0.3.Definición. *Se llama producto escalar euclídeo, o métrica euclídea, o producto euclídeo, o, simplemente, producto escalar sobre V a una forma bilineal simétrica definida positiva (que denotaremos por $\langle\, ,\, \rangle$) definida sobre V. Al espacio V con el producto escalar $\langle\, ,\, \rangle$ lo llamaremos espacio vectorial euclídeo o métrico.*

0.3'.Ejemplo. *1) La aplicación bilineal definida sobre $\mathbb{R}^n$ por*

$$\langle x, y \rangle = \sum_{i=1}^{n} x^i y^i, \quad \text{si} \quad x = (x^1, ..., x^n) \text{ e } y = (y^1, ..., y^n)$$

es un producto escalar, y $\mathbb{R}^n$, con este producto escalar, es un espacio vectorial euclídeo.

2) Si $(V; \langle, \rangle)$ es un espacio vectorial euclídeo y U es un subespacio vectorial de V, si $\langle, \rangle_U = \langle, \rangle|_{U \times U}$, entonces $(U; \langle, \rangle_U)$ es un espacio vectorial euclídeo.

A partir de ahora V denotará un espacio vectorial euclídeo de dimensión n.

0.4.Definición-Proposición. *Si U es un subespacio vectorial de V, de dimensión $m < n$, se llama subespacio ortogonal a U o complementario ortogonal de U al subespacio vectorial $U^{\perp}$ definido por*

$$U^{\perp} = \{w \in V / \langle w, u \rangle = 0 \text{ para todo } u \in U\}.$$

$U^{\perp}$ es, realmente, un subespacio vectorial de V.

Demostración. Que $U^{\perp}$ es un subespacio vectorial de V es consecuencia de que, si $x, y \in U^{\perp}$ y $\lambda, \mu \in I\!R$, entonces, para cualquier $u \in U$, se tiene $\langle \lambda x + \mu y, u \rangle = \lambda \langle x, u \rangle + \mu \langle y, u \rangle = 0$.

0.5.Definición. *Dos vectores u y v se dice que son ortogonales sii $\langle u, v \rangle = 0$. Dos subespacios vectoriales U y W de V se dice que son ortogonales, y se escribe $U \perp W$, sii para cualesquiera $u \in U$ y $w \in W$ se tiene que u y w son ortogonales.*

En particular, U y $U^{\perp}$ son ortogonales.

0.6.Definición-Proposición. *Se define el módulo o la norma $|u|$ de un vector $u \in U$ como el número real positivo*

$$|u| = (\langle u, u \rangle)^{\frac{1}{2}}.$$

Se verifica la siguiente desigualdad de Schwarz

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

La desigualdad de Schwarz permite dar la siguiente definición de ángulo

0.7.Definición. *Dados dos vectores $u, v \in V$, se define el ángulo $\theta = \measuredangle(u, v)$ formado por u y v como el número real $\theta \in [0, \pi]$ tal que $\langle u, v \rangle = |u||v| \cos \theta$.*

Que la definición anterior es posible resulta de que $|u/|u|| = 1$, $|v/|v|| = 1$, y, por la desigualdad de Schwarz,

$$|\langle \frac{u}{|u|}, \frac{v}{|v|} \rangle| \leq 1 \text{ i.e. } -1 \leq \frac{1}{|u||v|} \langle u, v \rangle \leq 1.$$

0.8.Definición. *Un conjunto de vectores no nulos $\{v_1, ..., v_m\}$ de dice que es ortogonal sii $i \neq j$ implica $\langle v_i, v_j \rangle = 0$. Se dice que ese conjunto es ortonormal sii es ortogonal y todos sus vectores son unitarios ($|v_i| = 1$), i.e. sii $\langle v_i, v_j \rangle = \delta_{ij}$.*

Si m vectores son ortogonales, entonces son linealmente independientes (pues $\sum_{i=1}^{m} \lambda_i v_i = 0$ implica $0 = \langle \sum_{i=1}^{m} \lambda_i v_i, v_j \rangle = \lambda_j \langle v_j, v_j \rangle$ y $\lambda_j = 0$ para todo $j \in \{1, ..., m\}$, ya que $\langle v_j, v_j \rangle = 0$ implica $v_j = 0$, pero suponemos -cfr. def 0.8- que $v_j \neq 0$), lo que implica $m \leq n$. Cuando $m = n$, dichos vectores forman una base, que se llama *base ortogonal* si los vectores son ortogonales y *base ortonormal* si los vectores son ortonormales.

Para el espacio vectorial euclídeo $I\!R^n$ del ejemplo 0.3', la base canónica $\{(1, 0, ..., 0), (0, 1, 0, ..., 0), ... , (0, ..., 0, 1)\}$ es una base ortonormal.

***0.9.Proposición.** *Sobre un espacio vectorial euclídeo siempre existe una base ortonormal. A partir de cualquier base $\{v_1, ..., v_n\}$ de V se puede construir una base ortonormal $\{e_1, ..., e_n\}$ de la siguiente manera:*

$$e_1 = \frac{v_1}{|v_1|}$$
$$e_2 = \frac{v_2 - \langle v_2, e_1\rangle e_1}{|v_2 - \langle v_2, e_1\rangle e_1|}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$e_n = \frac{v_n - \sum_{i=1}^{n-1}\langle v_n, e_i\rangle e_i}{|v_n - \sum_{i=1}^{n-1}\langle v_n, e_i\rangle e_i|}$$

Este método de obtener una base ortonormal a partir de una base dada se llama método de ortonormalización de Gram-Schmidt.

Demostración. Que cada vector es de norma unidad es inmediato, pues

$$|e_j| = \left|\frac{v_j - \sum_{i=1}^{j-1}\langle v_j, e_i\rangle e_i}{|v_j - \sum_{i=1}^{j-1}\langle v_j, e_i\rangle e_i|}\right| = 1$$

Para ver la ortogonalidad, demostrémoslo por inducción:

Veamos primero que $\langle e_1, e_2\rangle = 0$. En efecto:

$$\langle e_1, e_2\rangle = \langle e_1, \frac{v_2 - \langle v_2, e_1\rangle e_1}{|v_2 - \langle v_2, e_1\rangle e_1|}\rangle = \frac{1}{|v_2 - \langle v_2, e_1\rangle e_1|}\langle e_1, v_2 - \langle v_2, e_1\rangle e_1\rangle$$
$$= \frac{1}{|v_2 - \langle v_2, e_1\rangle e_1|}(\langle e_1, v_2\rangle - \langle v_2, e_1\rangle\langle e_1, e_1\rangle) = 0.$$

Supongamos ahora que, para un cierto $k \geq 3$ ($k \leq n$) se tiene

(0.9.1) $$\langle e_i, e_j\rangle = 0 \text{ para } i \neq j, \quad 1 \leq i, j \leq k-1,$$

y vamos a ver que esto implica $\langle e_i, e_j\rangle = 0$ para $i \neq j$ $1 \leq i, j \leq k$, (lo que, por inducción, probará (0.9.1) para $1 \leq i, j \leq n$). Por la hipótesis de inducción, lo único que falta por probar es que $\langle e_i, e_k\rangle = 0$ para todo $i \leq k-1$. Ahora bien, usando de nuevo que $\langle e_i, e_j\rangle = \delta_{ij}$ para $1 \leq i, j \leq k-1$, se tiene

$$\langle e_i, e_k\rangle = \langle e_i, \frac{v_k - \sum_{j=1}^{k-1}\langle v_k, e_j\rangle e_j}{|v_k - \sum_{j=1}^{k-1}\langle v_k, e_j\rangle e_j|}\rangle = \frac{1}{|v_k - \sum_{j=1}^{k-1}\langle v_k, e_j\rangle e_j|}\langle e_i, v_k - \sum_{j=1}^{k-1}\langle v_k, e_j\rangle e_j\rangle$$
$$= \frac{1}{|v_k - \sum_{j=1}^{k-1}\langle v_k, e_j\rangle e_j|}(\langle e_i, v_k\rangle - \sum_{j=1}^{k-1}\langle v_k, e_j\rangle\langle e_i, e_j\rangle)$$
$$= \frac{1}{|v_k - \sum_{j=1}^{k-1}\langle v_k, e_j\rangle e_j|}(\langle e_i, v_k\rangle - \sum_{j=1}^{k-1}\langle v_k, e_j\rangle\delta_{ij})$$
$$= \frac{1}{|v_k - \sum_{j=1}^{k-1}\langle v_k, e_j\rangle e_j|}(\langle e_i, v_k\rangle - \langle v_k, e_i\rangle) = 0.$$

0.10.Proposición. *En una base ortonormal* $\{e_1, ..., e_n\}$ *de* V*, todo vector* v *se escribe de la forma* $v = \sum_{i=1}^n \langle v, e_i\rangle e_i$*, es decir, si* $v = \sum_{i=1}^n v^i e_i$*, entonces* $v^i = \langle v, e_i\rangle$*.*

En efecto. Si $v = \sum_{i=1}^n v^i e_i$, entonces $\langle v, e_i\rangle = \langle \sum_{j=1}^n v^j e_j, e_i\rangle = \sum_{j=1}^n v^j \delta_{ji} = v^i$.

***0.11.Proposición.** $V = U \oplus U^\perp$*. Como consecuencia, se tiene que* $\dim(U^\perp) = n - m$*.*

Demostración. Sea $\{v_1, ...v_m\}$ una base de U. Completemosla a una base $\{v_1, ..., v_m, v_{m+1}, ..., v_n\}$ de V. Aplicando el proceso de Gram-Schmidt, obtenemos una base ortonormal $\{e_1, ..., e_n\}$ de V tal que los primeros m vectores $\{e_1, ..., e_m\}$ son una base ortonormal de U (pues son independientes y son combinación lineal de $v_1, ..., v_m$). Por definición de base ortonormal, los vectores $e_{m+1}, ..., e_n$ son ortogonales a U, luego pertenecen a $U^\perp$, y, si $v \in V$, por 0.10, $v = \sum_{i=1}^m \langle v, e_i\rangle e_i + \sum_{j=m+1}^n \langle v, e_j\rangle e_j = u + w$, con $u = \sum_{i=1}^m \langle v, e_i\rangle e_i \in U$ y $w = \sum_{j=m+1}^n \langle v, e_j\rangle e_j \in U^\perp$, luego $V = U + U^\perp$. Para ver que la suma es directa, sea $v \in U \cap U^\perp$, se tiene que $\langle v, v\rangle = 0$ y, por lo tanto, $v = 0$.

0.12.Proposición-Definición. *Sea* $V = U \oplus U^\perp$*,* $\{e_1, ..., e_m\}$ *es una base ortonormal de* U *y* $\{e_{m+1}, ..., e_n\}$ *es una base ortonormal de* $U^\perp$*, las aplicaciones proyección* π_U *y* $\pi_{U^\perp}$ *asociadas a la descomposición* $V = U \oplus U^\perp$ *vienen dadas por*

$$(0.12.1) \qquad \pi_U(v) = \sum_{i=1}^m \langle v, e_i\rangle e_i \quad y \quad \pi_{U^\perp}(v) = \sum_{i=m+1}^n \langle v, e_i\rangle e_i.$$

Se define la proyección ortogonal $\pi_U^\perp : V \longrightarrow U$ *como igual a la proyección* $\pi_U : V \longrightarrow U$ *definida por la descomposición* $V = U \oplus U^\perp$*.*

0.12'. Nota Obsérvese que $(U^\perp)^\perp = U$.

En efecto, por definición de ortogonal es evidente que $U \subset (U^\perp)^\perp$ y, como U tiene dimensión m y $(U^\perp)^\perp$ tiene dimensión $n - (n - m) = m$, estos subespacios vectoriales son iguales.

Por lo tanto, la descomposición de V definida por $U^\perp$ es $V = U^\perp \oplus U = U \oplus U^\perp$, la misma que la definida por U, luego $\pi_{U^\perp}^\perp$, que se define como la proyección $\pi_{U^\perp} : V \longrightarrow U^\perp$ definida por la descomposición anterior, es la misma que la $\pi_{U^\perp}$ que aparecía en la definición de $\pi_U^\perp$, luego

$$v = \pi_U^\perp(v) + \pi_{U^\perp}^\perp(v),$$

lo que permite calcular

$$\pi_U^\perp(v) = v - \pi_{U^\perp}^\perp(v),$$

y, si $\{e_{m+1}, ..., e_n\}$ es una base ortonormal de $U^\perp$, entonces

$$(0.12'.1) \qquad \pi_U^\perp(v) = v - \sum_{i=m+1}^n \langle v, e_i\rangle e_i.$$

0.12". Nota. Cuando quede claro por el contexto a que nos referimos (casi siempre), denotaremos $\pi_U^\perp$ por π_U.

0.13.Proposición. $(U+W)^{\perp} = U^{\perp} \cap W^{\perp}$.

Demostración. $x \in (U+W)^{\perp}$ sii para todo $u \in U$ y todo $w \in W$ se cumple $\langle x, u+w\rangle = 0$, en particular, tomando $w=0$, $\langle x, u\rangle = 0$ para todo $u \in U$ y, tomando $u=0$, $\langle x, w\rangle = 0$ para todo $w \in W$, luego $x \in U^{\perp} \cap W^{\perp}$, de donde $(U+W)^{\perp} \subset U^{\perp} \cap W^{\perp}$. Reciprocamente, si $x \in U^{\perp} \cap W^{\perp}$, entonces, para todo $u \in U$ y todo $w \in W$, se tiene $\langle x, u+w\rangle = \langle x, u\rangle + \langle x, w\rangle = 0$, luego $x \in (U+W)^{\perp}$, y $U^{\perp} \cap W^{\perp} \subset (U+W)^{\perp}$.

?0.14.Proposición. *El producto escalar define los isomorfismos (musicales):*

$$\sharp : V^* \longrightarrow V/\langle \alpha^{\sharp}, y\rangle = \alpha(y) \text{ para todo } y \in V,$$

$$\flat : V \longrightarrow V^*/x^{\flat}(y) = \langle x, y\rangle \text{ para todo } y \in V,$$

que son uno la aplicación inversa del otro.

Que $\sharp$ es una aplicación bien definida es consecuencia de que un vector $x \in V$ está determinado por su producto escalar por todos los vectores de V, como se deduce de 0.10.

El próximo corolario relaciona los papeles del núcleo de un elemento del dual y del ortogonal a un vector en el estudio de los subespacios vectoriales.

?0.14'.Corolario. *Si $f \in V^*$ y $v \in V$, entonces* $\operatorname{Ker} f =< \{f^{\sharp}\} >^{\perp}$ *y* $< \{v\} >^{\perp} = \operatorname{Ker} v^{\flat}$.

En efecto. $\operatorname{Ker} f = \{x \in V/f(x) = 0\} = \{x \in V/\langle f^{\sharp}, x\rangle = 0\} =< \{f^{\sharp}\} >^{\perp}$ y $< \{v\} >^{\perp} = \{x \in V/\langle v, x\rangle = 0\} = \{x \in V/v^{\flat}(x) = 0\} = \operatorname{Ker} v^{\flat}$.

?0.15.Proposición. *Los isomorfismos musicales llevan una base en su dual sii la correspondiente base de V es ortonormal.*

0.16.Definición. *Una aplicación lineal $f : U \longrightarrow V$ entre espacios vectoriales euclídeos de la misma dimensión se dice que es una isometría o transformación ortogonal sii $\langle f(x), f(y)\rangle = \langle x, y\rangle$ para todo $x, y \in U$. El nombre de transformación ortogonal lo reservaremos para el caso $U = V$.*

0.17.Proposición. *Una isometría $f : U \longrightarrow V$ entre espacios vectoriales euclídeos es un isomorfismo de espacios vectoriales.*

Demostración. Bastará con ver que $\operatorname{Ker} f = \{0\}$, pues entonces será una aplicación lineal inyectiva entre espacios vectoriales de la misma dimensión y, por tanto, un isomorfismo. Ahora bien, si $f(x) = 0$, se tiene $\langle x, x\rangle = \langle f(x), f(x)\rangle = 0$, lo que implica $x = 0$.

0.18.Lema. *Si $\{e_i\}_{i=1}^{n}$ es una base ortonormal de un e.v. euclídeo V y $x, y \in V$, entonces*

$$\langle x, y\rangle = \sum_{i=1}^{n} \langle x, e_i\rangle \langle y, e_i\rangle.$$

Demostración. Sustituyendo las expresiones de x e y dadas en 0.10 se tiene $\langle x,y\rangle = \langle \sum_{i=1}^n \langle x,e_i\rangle e_i, \sum_{j=1}^n \langle y,e_j\rangle e_j\rangle = \sum_{i=1}^n \sum_{j=1}^n \langle x,e_i\rangle\langle y,e_j\rangle\langle e_i,e_j\rangle = \sum_{i=1}^n \sum_{j=1}^n \langle x,e_i\rangle\langle y,e_j\rangle\delta_{ij} = \sum_{i=1}^n \langle x,e_i\rangle\langle y,e_i\rangle$.

0.19.Proposición. *Una aplicación lineal $f : U \longrightarrow V$ entre espacios vectoriales euclídeos es una isometría sii la imagen de una base ortonormal de U es una base ortonormal de V.*

Demostración. Si f es una isometría, es un isomorfismo, luego la imagen de una base ortonormal $\{e_i\}_{i=1}^n$ de U es una base $\{f(e_i)\}_{i=1}^n$ de V que, además, por ser f isometría, verifica $\langle f(e_i), f(e_j)\rangle = \langle e_i, e_j\rangle = \delta_{ij}$ y es una b.o.n. de V.

Recíprocamente, si la imagen de una b.o.n. $\{e_i\}_{i=1}^n$ de U es una base ortonormal $\{f(e_i)\}_{i=1}^n$ de V, para todo vector $x, y \in U$, se tiene

$$\begin{aligned}\langle f(x), f(y)\rangle &= \langle f(\sum_{i=1}^n \langle x,e_i\rangle e_i), f(\sum_{j=1}^n \langle y,e_j\rangle e_j)\rangle \\ &= \langle \sum_{i=1}^n \langle x,e_i\rangle f(e_i), \sum_{j=1}^n \langle y,e_j\rangle f(e_j)\rangle = \sum_{i=1}^n \langle x,e_i\rangle\langle y,e_i\rangle = \langle x,y\rangle\end{aligned}$$

.

0.20.Definición-Proposición. *Una matriz A, $n \times n$, se dice que es ortogonal sii su inversa es igual a su transpuesta, $A^{-1} = A^t$ (implicitamente se está suponiendo que A es regular).*

Si A es una matriz ortogonal, $\det A = \pm 1$.

***0.21.Proposición.** *Una aplicación lineal entre espacios vectoriales euclídeos de la misma dimensión es una isometría sii su matriz en una base ortonormal de cada espacio euclídeo es una matriz ortogonal.*

Demostración. Sea $f : U \longrightarrow V$ la aplicación lineal, $\{e_1, ..., e_n\}$ una base ortonormal de U y $\{v_1, ..., v_n\}$ una base ortonormal de V. La matriz F de f es la matriz

$$\begin{pmatrix} f_1^1 & \dots & \dots f_n^1 \\ \vdots & \ddots & \vdots \\ f_1^n & \dots & \dots f_n^n \end{pmatrix}$$

cuyas componentes vienen definidas por $f(e_i) = \sum_{j=1}^n f_i^j v_j$. Se verifica entonces que

$$\begin{aligned} F^t\, F &= \begin{pmatrix} f_1^1 & \dots & f_1^n \\ \vdots & \ddots & \vdots \\ f_n^1 & \dots & f_n^n \end{pmatrix} \begin{pmatrix} f_1^1 & \dots & f_n^1 \\ \vdots & \ddots & \vdots \\ f_1^n & \dots & f_n^n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n f_1^i f_1^i & \dots & \sum_{i=1}^n f_1^i f_n^i \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n f_n^i f_1^i & \dots & \sum_{i=1}^n f_n^i f_n^i \end{pmatrix} \\ &= \begin{pmatrix} \langle f(e_1), f(e_1)\rangle & \dots & \langle f(e_1), f(e_n)\rangle \\ \vdots & \ddots & \vdots \\ \langle f(e_n), f(e_1)\rangle & \dots & \langle f(e_n), f(e_n)\rangle \end{pmatrix} \end{aligned}$$

y la última matriz es la identidad sii $\langle f(e_i), f(e_j)\rangle = \delta_{ij}$, lo que equivale, por 0.19, a que f es una isometría. Esto prueba que f es una isometría sii $F^t\,F = Id$. Luego F ortogonal implica que f es una isometría y, para ver que f isometría implica que F es ortogonal, solo falta ver que $F^t\,F = Id$ implica $F\,F^t = Id$. Veámoslo: Como f es una isometría y, por tanto, un isomorfismo, de donde resulta que su matriz asociada F es regular, luego existe F^{-1}. Multiplicando $F^t\,F = Id$ a la izquierda por F, se tiene $F\,F^t\,F = F$ y, multiplicando a la derecha por F^{-1}, resulta $F\,F^t = Id$.

0.22.Proposición-Definición. *Sea U un subespacio vectorial de V. La aplicación*

$$s_U : V \longrightarrow V / s_U(x) = \pi_U(x) - \pi_{U^\perp}(x)$$

es una isometría y se llama simetría respecto de U.

Demostración. s_U es lineal por serlo $\pi_U(x)$ y $\pi_{U^\perp}$. Además $\langle s_U(x), s_U(y)\rangle = \langle \pi_U(x) - \pi_{U^\perp}(x), \pi_U(y) - \pi_{U^\perp}(y)\rangle = \langle \pi_U(x), \pi_U(y)\rangle + \langle \pi_{U^\perp}(x), \pi_{U^\perp}(y)\rangle = \pi_U(x) + \pi_{U^\perp}(x), \pi_U(y) + \pi_{U^\perp}(y)\rangle = \langle x, y\rangle$.

0.23.Definición. *Una simetría respecto de un hiperplano vectorial (s.e.v. de dimensión $n-1$) se llama una reflexión.*

Si U es un hiperplano de V, sea a un vector unitario generador de $U^\perp$, se tiene

$$s_U(x) = \pi_U(x) - \pi_{U^\perp}(x) = x - \langle x, a\rangle a - (\langle x, a\rangle a) = x - 2\langle x, a\rangle a.$$

0.24.Proposición-Definición. *El conjunto de las isometrías de un espacio vectorial euclídeo V en sí mismo, con respecto a la ley de composición de aplicaciones, es un grupo, que denotaremos por $O(V)$ y se llama grupo ortogonal de V. En particular, $O(\mathbb{R}^n)$ se denota $O(n, \mathbb{R})$.*

Demostración. Puesto que la aplicación identidad es, evidentemente, una isometría, bastará con demostrar que la composición de dos isometrías es una isometría y que la inversa de una isometría es una isometría, lo que es consecuencia de que si f y g son isometrías, entonces, para todo $x, y \in V$,

$$\langle g \circ f(x), g \circ f(y)\rangle = \langle g(f(x)), g(f(y))\rangle = \langle f(x), f(y)\rangle = \langle x, y\rangle$$

$$\text{y} \quad \langle f^{-1}(x), f^{-1}(y)\rangle = \langle f(f^{-1}(x)), f(f^{-1}(y))\rangle = \langle x, y\rangle.$$

0.25.Proposición. *Si $f \in O(V)$, y U es un subespacio vectorial de V, entonces $f(U^\perp) = f(U)^\perp$.*

Demostración. Si $x \in U^\perp$ y $u \in U$, entonces

$$\langle f(x), f(u)\rangle = \langle x, u\rangle = 0,$$

luego $f(U^\perp) \subset f(U)^\perp$, pero como f es un isomorfismo, $f(U^\perp)$ y $f(U)^\perp$ son espacios de la misma dimensión, luego son iguales.

0.26. Preguntas

a) Si $x = \sum_{i=1}^n x^i e_i$ y $y = \sum_{i=1}^n y^i e_i$, donde $\{e_i\}_{i=1}^n$ es una base ortonormal de V, ¿Conoces una expresión de $\langle x, y\rangle$?.

b) ¿Qué relación hay entre $(U+W)^\perp$ y $U^\perp \cap W^\perp$?.

c) Si f es una aplicación lineal entre espacios vectoriales euclídeos que lleva una base ortonormal en otra, ¿qué se puede decir de f?.

§1. Definición de espacio afín euclídeo. Angulos y distancias.

1.1.Definición. *Un espacio afín euclídeo es un espacio afín real E asociado a un espacio vectorial V dotado con un producto escalar $\langle\ ,\ \rangle$.*

1.2.Definición. *En un espacio afín euclídeo se define la distancia como la aplicación*

$$d : E \times E \longrightarrow I\!R^+ / \quad d(P,Q) = |\overrightarrow{PQ}|,$$

donde la norma o módulo $|\overrightarrow{PQ}|$ de un vector se define como en 0.6.

***1.3.Proposición.** *La distancia d definida en 1.2 es una distancia en el sentido de la topología, es decir, se verifican:*

(1.3.1) $d(P,Q) \geq 0$ para todo $P, Q \in E$ y $d(P,Q) = 0$ sii $P = Q$.

(1.3.2) $d(P,Q) = d(Q,P)$ para todo $P, Q \in E$.

(1.3.3) $d(P,Q) \leq d(P,R) + d(R,Q)$ para todo $P, Q, R \in E$.

Demostración. (1.3.1) Que $d(P,Q) \geq 0$ es consecuencia de que la distancia se ha definido como una norma. La proposición "$d(P,Q) = 0$ sii $P = Q$" es consecuencia de que el producto escalar es una forma bilineal definida positiva, pues, entonces, $d(P,Q) = 0$ sii $|\overrightarrow{PQ}| = 0$ sii $\langle \overrightarrow{PQ}, \overrightarrow{PQ}\rangle = 0$ sii $\overrightarrow{PQ} = 0$.

(1.3.2) resulta de que $d(P,Q) = |\overrightarrow{PQ}| = |-\overrightarrow{PQ}| = |\overrightarrow{QP}| = d(Q,P)$.

(1.3.3) se deduce usando la bilinealidad de $\langle\ ,\ \rangle$ y la desigualdad de Schwarz de la siguiente manera:

$$\begin{aligned} d(P,Q)^2 &= |\overrightarrow{PQ}|^2 = |\overrightarrow{PR} + \overrightarrow{RQ}|^2 = \langle \overrightarrow{PR} + \overrightarrow{RQ}, \overrightarrow{PR} + \overrightarrow{RQ}\rangle \\ &= |\overrightarrow{PR}|^2 + |\overrightarrow{RQ}|^2 + 2\,\langle \overrightarrow{PR}, \overrightarrow{RQ}\rangle \leq |\overrightarrow{PR}|^2 + |\overrightarrow{RQ}|^2 + 2\,|\langle \overrightarrow{PR}, \overrightarrow{RQ}\rangle| \\ &\leq |\overrightarrow{PR}|^2 + |\overrightarrow{RQ}|^2 + 2\,|\overrightarrow{PR}||\overrightarrow{RQ}| = (|\overrightarrow{PR}| + |\overrightarrow{RQ}|)^2 = (d(P,R) + d(R,Q))^2. \end{aligned}$$

1.3'.Ejemplo. *Sea V un espacio vectorial euclídeo. Al estudiar el espacio afín, vimos (ejemplo 1.3) que $E = V$, con la operación $+$ la suma de V, es un espacio afín con espacio vectorial asociado el propio V. Entonces, si V es euclídeo, $E = V$ es un espacio afín euclídeo.*

En particular, $I\!R^n$, como espacio afín asociado al espacio vectorial euclídeo $I\!R^n$ definido en el ejemplo 0.3' es un espacio afín euclídeo. De la definición de distancia en 1.2, el hecho (visto en Afin 1.3) de que, en un espacio vectorial, $\overrightarrow{xy} = y - x$, y la definición de producto escalar en $I\!R^n$ vista en 0.3', resulta que la distancia entre dos puntos $x = (x^1, ..., x^n)$, $y = (y^1, ..., y^n) \in I\!R^n$ viene dada por

$$d(x, y) = |y - x| = \sqrt{(y^1 - x^1)^2 + ... + (y^n - x^n)^2}.$$

1.4.Definición. *Sean $F = P+ < \{u\} >$ y $G = P+ < \{w\} >$ dos rectas de E que se cortan en un punto P. Se llama ángulo formado por F y G al menor de los ángulos $\measuredangle(u, w)$ y $\measuredangle(u, -w)$ definidos como en 0.7.*

Obsérvese que de la definición 0.7 se deduce que

$$\measuredangle(u, w) + \measuredangle(u, -w) = \pi,$$

porque $|u||w| \cos \measuredangle(u, -w) = \langle u, -w\rangle = -\langle u, w\rangle = -|u||w| \cos \measuredangle(u, w)$, de donde se deduce que $\cos \measuredangle(u, -w) = -\cos \measuredangle(u, w)$ y, por tanto $\measuredangle(u, -w) + \measuredangle(u, w) = \pi$. De aquí se deduce que el ángulo formado por las dos rectas es el que sea menor o igual que $\pi/2$.

1.5.Teorema de Pitágoras. *Sean P, Q, R tres puntos de un espacio afín métrico tal que la recta pasando por P y R forma un ángulo de $\pi/2$ con la que pasa por Q y R, entonces*

$$d(P, R)^2 + d(R, Q)^2 = d(P, Q)^2.$$

Este teorema es un corolario de 1.3, su demostración es la de (1.3.3) cuando $\overrightarrow{PR}$ es perpendicular a $\overrightarrow{RQ}$.

1.6.Definición. *Sea F un subespacio afín de E de dimensión m y espacio vectorial director U. Se llama subespacio afín pasando por $P \in F$ ortogonal a F al subespacio afín de dimensión $n - m$ definido por $P + U^\perp$.*

1.7.Definición. *Sean $F = P + U$ y $G = P + W$ dos hiperplanos de E que se cortan (en un subespacio afín de dimensión $n - 2$). Se llama ángulo formado por F y G al ángulo que forman dos rectas $P + U^\perp$ y $P + W^\perp$ perpendiculares a F y G respectivamente y pasando por un punto $P \in F \cap G$.*

De las definiciones 1.4 y 1.7 resulta que el ángulo formado por dos hiperplanos F y G como en 1.7 es

$$\measuredangle(F, G) = \measuredangle(P+U^\perp, P+W^\perp) = \min\{\measuredangle(u, w), \measuredangle(u, -w)\}, \quad \text{con} \quad 0 \neq u \in U^\perp \text{ y } 0 \neq w \in W^\perp.$$

Hemos definido el ángulo entre dos subespacios afines para el caso de rectas e hiperplanos. La perpendicularidad (o ángulo de $\pi/2$) la vamos a definir para subespacios afines de dimensión arbitraria, pero, de modo semejante a como hicimos en la definición de de los ángulos, vamos a tener que distinguir entre los casos $\dim F + \dim G \leq n$ y $\dim F + \dim G > n$.

1.8.Definición. *Sean $F = P + U$ y $G = P + W$ dos subespacios afines de E que se cortan y tales que* $\dim F + \dim G \leq n$. *Se dice que F es perpendicular u ortogonal a G sii $U \perp W$ (o, equivalentemente, $U \subset W^{\perp}$ ó $W \subset U^{\perp}$).*

1.9.Definición. *Sean $F = P + U$ y $G = P + W$ dos subespacios afines de E que se cortan y tales que* $\dim F + \dim G > n$. *Se dice que F es perpendicular u ortogonal a G sii $U^{\perp} \perp W^{\perp}$ (o, equivalentemente, $U \supset W^{\perp}$ ó $W \supset U^{\perp}$).*

Vamos ahora a definir y ver como se calcula la distancia entre dos subespacios afines.

1.10.Definición. *Dadas dos variedades lineales afines F y G de E, se define la distancia entre ellas como se hace con la distancia entre dos subconjuntos de un espacio métrico, i.e.*

$$d(F, G) = \inf\{d(P, Q)/P \in F \text{ y } Q \in G\}.$$

Obsérvese que, en la definición anterior, el ínfimo existe, pues la distancia es siempre mayor o igual que 0. El objetivo de las próximas proposiciones es dar un método efectivo del cálculo de la distancia. Al mismo tiempo, mostrará que el ínfimo de la definición de distancia es, en realidad, un mínimo.

***1.11.Lema.** *Sea $v \in V$, U un subespacio vectorial de V. Entonces $\pi_{U^{\perp}}(v)$ es el único vector de la forma $w = v + x$, con $x \in U$ tal que $|w| = \min\{|v + x|/x \in U\}$ (y, por tanto, $\min\{|v + x|/x \in U\} = |\pi_{U^{\perp}}(v)|$).*

Demostración.

$$\begin{aligned}|v + x|^2 = \langle v + x, v + x\rangle = |v|^2 + 2\langle v, x\rangle + |x|^2 \\ = |\pi_U(v) + \pi_{U^{\perp}}(v)|^2 + 2\langle \pi_U(v) + \pi_{U^{\perp}}(v), x\rangle + |x|^2,\end{aligned}$$

pero

$$\begin{aligned}&|\pi_U(v) + \pi_{U^{\perp}}(v)|^2 = \langle \pi_U(v) + \pi_{U^{\perp}}(v), \pi_U(v) + \pi_{U^{\perp}}(v)\rangle \\ &= \langle \pi_U(v), \pi_U(v)\rangle + \langle \pi_{U^{\perp}}(v), +\pi_{U^{\perp}}(v)\rangle + 2\langle \pi_U(v), +\pi_{U^{\perp}}(v)\rangle = |\pi_U(v)|^2 + |\pi_{U^{\perp}}(v)|^2,\end{aligned}$$

y $\langle \pi_{U^{\perp}}(v), x\rangle = 0$, luego

$$|v+x|^2 = |\pi_U(v)|^2 + |\pi_{U^{\perp}}(v)|^2 + 2\langle \pi_U(v), x\rangle + |x|^2 = |\pi_{U^{\perp}}(v)|^2 + |\pi_U(v) + x|^2 \geq |\pi_{U^{\perp}}(v)|^2,$$

y se tiene la igualdad en la última desigualdad (por tanto el valor mínimo) sii $|\pi_U(v) + x|^2 = 0$, es decir, sii $x = -\pi_U(v)$, lo que es equivalente a que $w = v + x = v - \pi_U(v) = \pi_U(v) + \pi_{U^{\perp}}(v) - \pi_U(v) = \pi_{U^{\perp}}(v)$, c.q.d..

***1.12.Proposición.** *Sean F y G dos variedades lineales afines con espacios vectoriales directores U y W respectivamente. Existen $P \in F$ y $Q \in G$ (no necesariamente únicos) tales que $\overrightarrow{PQ} \in (U + W)^{\perp}$ y, para esos puntos, $d(F, G) = d(P, Q)$.*

Demostración. Sabemos que $V = (U + W) \oplus (U + W)^{\perp}$. Sean $R \in F$ y $S \in G$, erntonces $\overrightarrow{RS} \in V$ y, por la descomposición anterior, existen $u \in U$, $w \in W$ y $x \in (U + W)^{\perp}$ $(u + w$

y x únicos, dependiendo solo de $\overrightarrow{RS}$) tales que $\overrightarrow{RS} = u + w + x$. Tomemos $P = R + u$, $Q = S - w$. Se tiene entonces que

$$\begin{aligned}\overrightarrow{PQ} = \overrightarrow{(R+u)(S-w)} = \overrightarrow{(R+u)R} + \overrightarrow{RS} + \overrightarrow{S(S-w)} \\ = -u + \overrightarrow{RS} - w = \overrightarrow{RS} - (u+w) = x \in (U+W)^{\perp},\end{aligned}$$

lo que prueba la primera parte del enunciado.

Vamos a probar ahora que $d(F,G) = d(P,Q)$. Sean $R' \in F$ y $S' \in G$, aplicando el lema al caso en que el subespacio de V es $U + W$ y el vector v es $\overrightarrow{R'Q}$, se tiene que, como $\overrightarrow{QS'} \in W \subset U + W$,

$$d(R', S') = |\overrightarrow{R'S'}| = |\overrightarrow{R'Q} + \overrightarrow{QS'}| \geq |\pi_{(U+W)^{\perp}}(\overrightarrow{R'Q})|,$$

pero $\overrightarrow{R'Q} = \overrightarrow{R'P} + \overrightarrow{PQ}$, con $\overrightarrow{R'P} \in U \subset U+W$ y $\overrightarrow{PQ} \in (U+W)^{\perp}$, luego $\pi_{(U+W)^{\perp}}(\overrightarrow{R'Q}) = \overrightarrow{PQ}$, luego

$$d(R', S') \geq |\overrightarrow{PQ}| = d(P,Q) \text{ para todo } R' \in F \text{ y } S' \in G,$$

luego $d(F,G) = d(P,Q)$.

1.13.Nota. *La demostración de la proposición anterior da también un método práctico de calcular la distancia entre dos subespacios afines F y G: se eligen $R \in F$ y $S \in G$, se calcula $x = \pi_{(U+W)^{\perp}}(\overrightarrow{RS})$, y $d(F,G) = |x| = |\pi_{(U+W)^{\perp}}(\overrightarrow{RS})|$.*

1.14.Modo de calcular la distancia entre dos subespacios afines, cuyas ecuaciones vienen dadas en paramétricas, usando 1.13.

En una referencia cartesiana ortonormal $\{O, e_1, ..., e_n\}$ (i.e. una referencia cartesiana de E en la que $\{e_1, ..., e_n\}$ es una base ortonormal de V), sean

$$F \equiv \quad Q^i = R^i + \sum_{j=1}^{m} \lambda^j v_j^i, \quad 1 \leq i \leq n \quad \text{y}$$

$$G \equiv \quad T^i = S^i + \sum_{j=1}^{r} \lambda^j w_j^i, \quad 1 \leq i \leq n.$$

Entonces podemos tomar $\overrightarrow{RS} = (S^1 - R^1, ..., S^n - R^n)$. Para calcular su proyección sobre el ortogonal a $U + W$, partimos de los vectores generadores de este espacio, $v_j = \sum_{i=1}^{n} v_j^i e_i$ y $w_k = \sum_{i=1}^{n} w_k^i e_i$, $1 \leq j \leq m$ y $1 \leq k \leq r$. Aplicando el proceso de ortonormalización de Gram-Schmidt a estos vectores, obtenemos una base ortonormal $\{a_1, ..., a_s\}$ de $U + W$ (a pesar de que el conjunto de vectores inicial puede no ser una base de $U + W$). Entonces

$$d(F,G) = |\pi_{(U+W)^{\perp}}(\overrightarrow{RS})| = |\overrightarrow{RS} - \sum_{i=1}^{s} \langle \overrightarrow{RS}, a_i \rangle a_i|.$$

En algunos casos es fácil calcular una base ortonormal $\{b_1, ..., b_h\}$ de $(U+W)^\perp$. En ese caso

$$d(F,G) = |\pi_{(U+W)^\perp}(\overrightarrow{RS})| = |\sum_{i=1}^{h} \langle \overrightarrow{RS}, b_i \rangle b_i|.$$

1.15.Ejemplos.

1.15.1) Sea E un espacio afín euclídeo de dimensión 3. En una referencia ortonormal $\{O, e_1, e_2, e_3\}$ de E, el punto P y el plano H tienen las siguientes expresiones:

$$P = (0,0,3), \quad H \equiv \left.\begin{array}{r} x = \mu \\ y = \lambda + \mu \\ z = \lambda \end{array}\right\}$$

Calcular la distancia de P a H. Para ello observar que H es el plano que pasa por $(0,0,0)$ con espacio vectorial director generado por $e_2 + e_3$ y $e_1 + e_2$.

1.15.2) Sea E un espacio afín euclídeo de dimensión 3. En una referencia ortonormal $\{O, e_1, e_2, e_3\}$ de E, las rectas F y G tienen las siguientes expresiones:

$$F \equiv \left.\begin{array}{r} x = \lambda \\ y = \lambda \\ z = \lambda \end{array}\right\}, \quad G \equiv \left.\begin{array}{r} x = 3 + \lambda \\ y = 2 \\ z = 0 \end{array}\right\}$$

Calcular la distancia de F a G. Para ello observar que F es la recta que pasa por $(0,0,0)$ con espacio vectorial director generado por $e_1 + e_2 + e_3$ y G es la recta que pasa por $(3,2,0)$ con espacio vectorial director generado por e_1.

Vamos a ver ahora como calcular la distancia cuando los subespacios afines vienen dados en implícitas. Para ello necesitamos la siguiente versión euclídea de los apartados 2.12 y 3.13 de la Geometría Afín.

1.16.Proposición. *Si F es un subespacio afín de E de dimensión m, con espacio vectorial director U, F es el conjunto de puntos Q de E cuyas coordenadas, en una referencia ortonormal $\{O, e_1, ..., e_n\}$, verifican*

$$\sum_{j=1}^{n} Q^j a_i^j + a_i^0 = 0, \quad m+1 \leq i \leq n, \tag{1.16.1}$$

donde los vectores $\{w_i\}_{i=m+1}^n$ definidos por $w_i = \sum_{j=1}^n a_i^j e_j$ son una base de $U^\perp$.

Demostración. Ya vimos en geometría afín que un subespacio afín de dimensión m se podía describir por $n-m$ ecuaciones como la (1.16.1). Lo que falta ver es que los vectores w_i definidos como en el enunciado de esta proposición son una base de $U^\perp$. Para ello, observemos primero que el número de vectores es $n-m$, es decir, la dimensión de $U^\perp$. Además son linealmente independientes, porque sus componentes son los coeficientes de las ecuaciones (1.16.1), y la matriz formada por estos coeficientes tiene rango $n-m$ (Proposición 2.12 de Geometría Afín). Por lo tanto solo falta probar que $w_i \in U^\perp$ para $m+1 \leq i \leq n$. Ahora bien, todos los vectores de U son de la forma $\overrightarrow{PQ}$, donde $P, Q \in F$, luego todo lo que hay que ver es que $\langle \overrightarrow{PQ}, w_i \rangle = 0$ para todo $P, Q \in F$ y $m+1 \leq i \leq n$. Si $P = (P^1, ..., P^n)$ y $Q = (Q^1, ..., Q^n)$, las coordenadas P^j y Q^j verifican (1.16.1), y, usando 0.18,

$$\langle \overrightarrow{PQ}, w_i \rangle = \sum_{j=1}^{n} (Q^j - P^j) a_i^j = \sum_{j=1}^{n} Q^j a_i^j - \sum_{j=1}^{n} P^j a_i^j = -a_i^0 + a_i^0 = 0.$$

1.17.Modo de calcular la distancia entre dos subespacios afines, cuyas ecuaciones vienen dadas en implícitas, usando 1.13.

En una referencia cartesiana ortonormal $\{O, e_1, ..., e_n\}$ de E, sean

$$(1.17.1) \qquad F = P + U \equiv \sum_{j=1}^{n} Q^j a_i^j + a_i^0 = 0, \quad m+1 \leq i \leq n,$$

$$(1.17.2) \qquad G = T + W \equiv \sum_{j=1}^{n} Q^j b_i^j + b_i^0 = 0, \quad k+1 \leq i \leq n.$$

Entonces podemos tomar $R = (R^1, ..., R^n)$ verificando las ecuaciones (1.17.1) y $S = (S^1, ..., S^n)$ verificando las ecuaciones (1.17.2), así $\overrightarrow{RS} = (S^1 - R^1, ..., S^n - R^n)$. Para calcular su proyección sobre el ortogonal a $U + W$, partimos de los vectores $u_i = \sum_{j=1}^{n} a_i^j e_j$ $(m+1 \leq i \leq n)$ generadores de $U^\perp$ y $w_i = \sum_{j=1}^{n} b_i^j e_j$ $(k+1 \leq i \leq n)$ generadores de $W^\perp$. A partir de estas dos bases se obtiene una base ortonormal $\{v_1, ..., v_h\}$ de $(U+W)^\perp = U^\perp \cap W^\perp$ y

$$d(F, G) = |\pi_{(U+W)^\perp}(\overrightarrow{RS})| = |\sum_{i=1}^{h} \langle \overrightarrow{RS}, v_i \rangle v_i|.$$

1.18.Ejemplos.

1.18.1) Sea E un espacio afín euclídeo de dimensión 3. En una referencia ortonormal $\{O, e_1, e_2, e_3\}$ de E, las rectas F y el plano G tienen las siguientes expresiones:

$$F \equiv \left. \begin{array}{r} x+y+z=0 \\ x+y-z=1 \end{array} \right\}, \quad G \equiv \left. \begin{array}{r} x=z \\ x=y \end{array} \right\}$$

Calcular la distancia de F a G.

1.18.2) Sea E un espacio afin euclídeo de dimensión 3. En una referencia ortonormal $\{O, e_1, e_2, e_3\}$ de E, el punto P y el plano H tienen las siguientes expresiones:

$$P = (0, 0, 3), \quad H \equiv\ x + y + z = 0.$$

Calcular la distancia de P a H.

1.18.3) Sea E un espacio afin euclídeo de dimensión n. En una referencia ortonormal $\{O, e_1, ..., e_n\}$ de E, el punto P y el plano H tienen las siguientes expresiones:

$$P = (P^1, ..., P^n), \quad H \equiv\ a^1Q^1 + ... + a^nQ^n + a^0 = 0.$$

Calcular la distancia de P a H.

Vamos a acabar este epígrafe viendo como se define la proyección ortogonal sobre un subespacio afín.

1.19. Proyección ortogonal. *Sea F un subespacio afín de E, de espacio vectorial director U. Se define la proyección ortogonal π_F de E sobre F como la aplicación $\pi_F : E \longrightarrow F$ definida por $\pi_F(Q) = P + \pi_U(\overrightarrow{PQ})$, donde $P \in F$ y π_U es la proyección ortogonal sobre U definida en 0.12.*

La proyección ortogonal π_F no depende del punto $P \in F$ elegido. En efecto: Si R es otro punto de F, $\overrightarrow{PR} \in U$, y

$$\begin{aligned} R + \pi_U(\overrightarrow{RQ}) &= P + \overrightarrow{PR} + \pi_U(\overrightarrow{RQ}) = P + \pi_U(\overrightarrow{PR}) + \pi_U(\overrightarrow{RQ}) \\ &= P + \pi_U(\overrightarrow{PR} + \overrightarrow{RQ}) = P + \pi_U(\overrightarrow{PQ}). \end{aligned}$$

? Para cualquier $P \in F$, sea G el subespacio ortogonal a F pasando por P. Entonces la proyección ortogonal sobre F coincide con la proyección sobre F paralelamente a G tal y como se definió en Geometría Afín 4.12, como se deduce de las definiciones 1.19, 0.12 y Geometría Afín 4.12.

1.20. Preguntas.

a) El ángulo entre dos hiperplanos de un espacio métrico es θ. ¿Cuál es el ángulo que forman dos rectas perpendiculares a esos hiperplanos y que se cortan en un punto?.

b) Un plano P y un hiperlano H de un espacio afín métrico E no son paralelos. ¿Cuál es la distancia entre ellos?. ¿Por qué?.

§2. Isometrías entre espacios afines euclídeos

2.1.Definición. *Sean E y F dos espacios afines euclídeos de espacios vectoriales asociados V y U respectivamente. Una aplicación afín $f : E \longrightarrow F$ diremos que es una isometría si su aplicación lineal asociada $\widetilde{f} : V \longrightarrow U$ es una isometría de espacios vectoriales euclídeos.*

Como una isometría de espacios vectoriales euclídeos es un isomorfismo, f es un isomorfismo afín y E y F son espacios afines de la misma dimensión.

2.2.Proposición. *Todo espacio afín métrico E de dimensión n es isométrico al espacio afín métrico $\mathbb{R}^n$ definido en el ejemplo 1.3'. La isometría no es canónica, sino que depende de la elección de una referencia ortonormal en E.*

Demostración. Elijamos $\{O, e_1, ..., e_n\}$ una referencia ortonormal en E. Definamos $f : E \longrightarrow \mathbb{R}^n$ por

$$f(O + \lambda^1 e_1 + ... + \lambda^n e_n) = (\lambda^1, ..., \lambda^n).$$

Esta aplicación verifica $f(O) = (0, ..., 0)$ y su aplicación lineal asociada $\widetilde{f}$ lleva la base ortonormal $\{e_1, ..., e_n\}$ en la base ortonormal $\{(1, 0, ..., 0), (0, 1, 0, ..., 0), ... , (0, ..., 0, 1)\}$, luego, por 0.19, $\widetilde{f}$ es una isometría de espacios vectoriales, y F es también una isometría.

2.3.Proposición. *La composición de dos isometrías entre espacios afines es una isometría.*

Demostración. En la demostración de 0.24 se demuestra que la composición entre dos isometrías de espacios vectoriales euclídeos es una isometría. Si f y g son isometrías entre espacios afines, de aplicaciones lineales asociadas $\widetilde{f}$ y $\widetilde{g}$ respectivamente, sabemos (G.A.4.14) que $\widetilde{g} \circ \widetilde{f}$ es la aplicación lineal asociada a $g \circ f$. Como $\widetilde{f}$ y $\widetilde{g}$ son isometrías, $\widetilde{g} \circ \widetilde{f}$ es una isometría y, por lo tanto, $g \circ f$ es una isometría.

2.4.Corolario. *Dos espacios afines euclídeos de la misma dimensión son isométricos.*

Demostración. Es consecuencia de 2.2 y 2.3.

***2.5.Teorema.** *Una aplicación $f : E \longrightarrow F$ entre espacios afines es una isometría sii $d(f(P), f(Q)) = d(P, Q)$ para todo $P, Q \in E$.*

Demostración. Denotaremos por V y U los espacios vectoriales asociados a E y F respectivamente. Si f es una isometría, como $\widetilde{f}(\overrightarrow{PQ}) = \overrightarrow{f(P)f(Q)}$ (cfr. 4.5 del capítulo de Geometría Afín), se tiene

$$d(f(P), f(Q)) = |\overrightarrow{f(P)f(Q)}| = |\widetilde{f}(\overrightarrow{PQ})| = |\overrightarrow{PQ}| = d(P, Q).$$

Recíprocamente, si $d(f(P), f(Q)) = d(P, Q)$, entonces (buscamos $\widetilde{f}$ tal que f se escriba $f(Q) = f(P) + \widetilde{f}(\overrightarrow{PQ})$, i.e. $\widetilde{f}(\overrightarrow{PQ}) = \overrightarrow{f(P)f(Q)}$, luego) definimos, para $P \in E$ fijado,

$$\widetilde{f} : V \longrightarrow U / \quad \widetilde{f}(v) = \overrightarrow{f(P)f(P + v)}.$$

Veamos que $\widetilde{f}$ conserva el producto escalar

$$(2.5.1) \quad d(P + v, P + w)^2 = |\overrightarrow{(P + v)(P + w)}|^2 = \langle w - v, w - v\rangle = |w|^2 + |v|^2 - 2\langle w, v\rangle,$$

$$\begin{aligned}
(2.5.2) \quad d(f(P + v), f(P + w))^2 &= |\overrightarrow{f(P + v)f(P + w)}|^2 \\
&= \langle \overrightarrow{f(P + v)f(P)} + \overrightarrow{f(P)f(P + w)}, \overrightarrow{f(P + v)f(P)} + \overrightarrow{f(P)f(P + w)}\rangle \\
&= |\overrightarrow{f(P + v)f(P)}|^2 + |\overrightarrow{f(P)f(P + w)}|^2 + 2\langle \overrightarrow{f(P + v)f(P)}, \overrightarrow{f(P)f(P + w)}\rangle \\
&= | - \widetilde{f}(v)|^2 + |\widetilde{f}(w)|^2 - 2\langle \widetilde{f}(w), \widetilde{f}(v)\rangle,
\end{aligned}$$

pero

$$|v|^2 = d(P, P+v)^2 = d(f(P), f(P+v))^2 = |\overrightarrow{f(P)f(P+v)}|^2 = |\widetilde{f}(v)|^2 \quad \text{y}$$

$$|w|^2 = d(P, P+w)^2 = d(f(P), f(P+w))^2 = |\overrightarrow{f(P)f(P+w)}|^2 = |\widetilde{f}(w)|^2,$$

luego, de (2.5.1) y (2.5.2) resulta

$$\langle v, w\rangle = \langle \widetilde{f}(v), \widetilde{f}(w)\rangle.$$

Vamos a ver ahora que si $\widetilde{f}$ conserva el producto escalar, entonces es lineal (lo que probará que $\widetilde{f}$ es una isometría). Veamos:

$$|\widetilde{f}(v+w) - \widetilde{f}(v) - \widetilde{f}(w)|^2 = \langle \widetilde{f}(v+w) - \widetilde{f}(v) - \widetilde{f}(w), \widetilde{f}(v+w) - \widetilde{f}(v) - \widetilde{f}(w)\rangle$$
$$= |\widetilde{f}(v+w)|^2 + |\widetilde{f}(v)|^2 + |\widetilde{f}(w)|^2 - 2\langle \widetilde{f}(v+w), \widetilde{f}(v)\rangle - 2\langle \widetilde{f}(v+w), \widetilde{f}(w)\rangle + 2\langle \widetilde{f}(v), \widetilde{f}(w)\rangle$$

$= |$como $\widetilde{f}$ conserva el producto escalar$| =$

$$= |v+w|^2 + |v|^2 + |w|^2 - 2\langle v+w, v\rangle - 2\langle v+w, w\rangle + 2\langle v, w\rangle$$
$$= \langle v+w-v-w, v+w-v-w\rangle = 0.$$

Luego, como $\langle x, x\rangle = 0$ implica $x = 0$, se tiene

$$\widetilde{f}(v+w) = \widetilde{f}(v) + \widetilde{f}(w).$$

Además, usando de nuevo que $\widetilde{f}$ conserva el producto escalar,

$$|\widetilde{f}(\lambda v) - \lambda\widetilde{f}(v)|^2 = \langle \widetilde{f}(\lambda v) - \lambda\widetilde{f}(v), \widetilde{f}(\lambda v) - \lambda\widetilde{f}(v)\rangle$$
$$= |\widetilde{f}(\lambda v)|^2 - 2\lambda\langle \widetilde{f}(\lambda v), \widetilde{f}(v)\rangle + \lambda^2\langle \widetilde{f}(v), \widetilde{f}(v)\rangle = |\lambda v|^2 - 2\lambda\langle \lambda v, v\rangle + \lambda^2\langle v, v\rangle = 0,$$

y, de nuevo, esto implica

$$\widetilde{f}(\lambda v) = \lambda\widetilde{f}(v).$$

Luego $\widetilde{f}$ es lineal y, por tanto, como conserva el producto escalar, una isometría.

De la definición de $\widetilde{f}$ resulta que $f(Q) = f(P) + \widetilde{f}(\overrightarrow{PQ})$, luego f es una aplicación afín de aplicación lineal asociada $\widetilde{f}$. Como $\widetilde{f}$ es una isometría, f también lo es.

Obsérvese que si $f : E \longrightarrow F$ es una isometría, f^{-1} también lo es, porque $\widetilde{f^{-1}} = \widetilde{f}^{-1}$, como se vió en G.A.4.14. De esta observación y de 2.3, resulta:

2.6.Proposición. *Las isometrías de un espacio afín métrico E en si mismo forman un grupo con respecto a la composición de aplicaciones. Lo denotaremos por $Is(E)$, se llama también grupo de movimientos de E (es decir, una isometría de E en sí mismo se llama un movimiento de E).*

2.7.Ejemplos.

a) Una traslación de E en E es una isometría, pues su aplicación lineal asociada es la identidad.

b) La *simetría* s_F *respecto del subespacio afín* F de E se define por

$$s_F : E \longrightarrow E/ \quad s_F(Q) = P + s_U(\overrightarrow{PQ}) \text{ para todo } Q \in E,$$

siendo P un punto de F y U el espacio vectorial director de F.

Esta aplicación no depende del $P \in F$ elegido. En efecto, si $R \in F$,

$$R + s_U(\overrightarrow{RQ}) = R + s_U(\overrightarrow{RP} + \overrightarrow{PQ}) = R + s_U(\overrightarrow{RP}) + s_U(\overrightarrow{PQ}),$$

y, como $R, P \in F$, entonces $\overrightarrow{RP} \in U$, luego $s_U(\overrightarrow{RP}) = \pi_U(\overrightarrow{RP}) - \pi_{U^\perp}(\overrightarrow{RP}) = \overrightarrow{RP}$, y

$$R + s_U(\overrightarrow{RQ}) = R + \overrightarrow{RP} + s_U(\overrightarrow{PQ}) = P + s_U(\overrightarrow{PQ}).$$

De la definición de s_F se deduce que su aplicación lineal asociada es s_U y, como esta es una isometría de espacios vectoriales, s_F es una isometría de espacios afines euclídeos.

Las proposiciones G.A. 4.18 y 4.19 se traducen aquí en el teorema que damos a continuación, cuya demostración es la misma de G.A.4.18,19, teniendo en cuenta que $f \in Is(E)$ sii $\widetilde{f} \in O(V)$.

2.8.Proposición. *a) Dado $P \in E$, el conjunto $Is_P(E)$ de los movimientos (isometrías) de E que dejan P fijo (i.e. $f(P) = P$) es un subgrupo de $Is(E)$ que es isomorfo al grupo ortogonal $O(V)$ del espacio vectorial euclídeo V asociado al espacio afín E.*

b)Fijado $P \in E$, toda $f \in Is(E)$ se puede escribir de manera única como la composición $T_{Pf(P)} \circ g$ de un elemento $g \in Is_P(E)$ definido por $g(Q) = P + \widetilde{f}(\overrightarrow{PQ})$ y una traslación $T_{\overrightarrow{Pf(P)}}$.

2.9.Matriz de una isometría.

Sea $f : E \longrightarrow E'$ una isometría, $\{O, e_1, ..., e_n\}$ una referencia ortonormal de E y $\{O', e'_1, ..., e'_n\}$ una referencia ortonormal de E'. Por G.A. 4.20, sabemos que la matriz F de f en estas referencias es

$$F = \begin{pmatrix} \widetilde{F} & \theta \\ 0 & 1 \end{pmatrix},$$

siendo $\widetilde{F}$ la matriz de la aplicación lineal asociada $\widetilde{f}$ y $\theta = (\theta^1, ..., \theta^n)^t$ definida por $O'f(O) = \sum_{i=1}^n \theta^i e'_i$. Como $\widetilde{f}$ es una transformación ortogonal, $\widetilde{F}$ es una matriz ortogonal. Si $E = E'$ y $\{O, e_1, ..., e_n\} = \{O', e'_1, ..., e'_n\}$, entonces $\sum_{i=1}^n \theta^i e'_i$ es precisamente el vector que define la parte traslación de f en 2.8.b) cuando se toma $P = O$.

2.10. Ejercicios y pregunta.

a) La descomposición de 2.8.b) no es conmutativa, i.e., $T_{\overrightarrow{Pf(P)}} \circ g \neq g \circ T_{\overrightarrow{Pf(P)}}$. Probar que la relación de conmutación correcta es $g \circ T_v = T_{\vec{g}(v)} \circ g$.

b) Sean $g, h \in Is_P(E)$. Dadas $f = T_v \circ g$ y $\varphi = T_w \circ h$, encontrar la descomposición de $f \circ \varphi$ y $\varphi \circ f$ dada por 2.8.b).

c) Sea $g \in Is_P(E)$. Dada $f = T_v \circ g$, encontrar la descomposición de f^{-1} dada por 2.8.b).

d) Sea $I\!R^n$ con su estructura canónica de espacio afín euclídeo. En cada uno de los casos siguientes, decídase si f es una isometría de $I\!R^3$. De ser así, encontrar su componente traslación y su componente transformación ortogonal.

(1) $f(p) = -p$.

(2) $f(p) = < p, a > \ a$, donde $|a| = 1$.

(3) $f(p) = (p_3 - 1, p_2 - 2, p_1 - 3)$.

(4) $f(p) = (p_1, p_2, 1)$.

e) La matriz de una isometría de un espacio afín métrico en si mismo tiene la forma $\begin{pmatrix} M & \theta \\ 0 & 1 \end{pmatrix}$. ¿Qué puedes decir de la matriz M?.

§3. Orientación de espacios vectoriales y descomposición de una transformación ortogonal.

Intuitivamente, en el espacio ordinario, el concepto de orientación corresponde a la posibilidad de distinguir entre mano derecha e izquierda. Son “intrínsecamente iguales” y, sin embargo, es imposible hacer coincidir exactamente una con la otra. La idea para “formalizar” ese concepto, en una teoría lineal, es sustituir las manos por sistemas de referencia. Así, en el plano ($I\!R^2$), los sistemas de referencia de la fig. 1 hacen el papel de la mano derecha y la mano izquierda y, además, solo hay dos manos (sistemas de referencia) : la derecha y la izquierda (ver 3.2 para una demostración de este hecho en dimensión finita arbitraria).

sistema de referencia derecho

sistema de referencia izquierdo

Figura 1

?3.1. Definición. *Dado un $I\!R$-espacio vectorial V de dimensión finita n, decimos que dos bases ordenadas $e = \{e_1, ..., e_n\}$, $f = \{f_1, ..., f_n\}$ tienen la misma orientación si la matriz de cambio de base tiene determinante positivo (i.e.: si $f_j = \sum_{i=1}^n f_j^i e_i$, entonces $\det(f_j^i) > 0$).*

?3.2. *Si en el conjunto $\mathcal{B}$ de las bases ordenadas de V definimos la relación*

$$e \sim f \quad \textit{sii} \quad e \;\; \textit{y} \;\; f \textit{ tienen la misma orientación,}$$

se ve fácilmente que $\sim$ es una relación de equivalencia en $\mathcal{B}$. Como el determinante de una matriz de cambio de base es positivo o negativo (no puede ser cero), hay dos clases de equivalencia en $\mathcal{B}$.

En efecto: $\mathcal{B}/\sim$ tiene al menos dos elementos: si $e = \{e_1, e_2, ..., e_n\} \in \mathcal{B}$, entonces $e' = \{e_2, e_1, ..., e_n\}$ no es equivalente a e.

Sean $[e]$, $[e']$ las clases de quivalencia de e y e' respectivamente. Entonces $\mathcal{B}/\sim = \{[e], [e']\}$. En efecto: sea $g \in \mathcal{B}$, si $g \sim e$, $[g] = [e]$; si g no es equivalente a e, sean G y F las matrices tales que $g = Ge$, $e = Fe'$. Entonces $g = GFe'$, pero g no equivalente a e, y e no equivalente a e' implican $\det G < 0$ y $\det F < 0$ respectivamente, luego $\det(GF) > 0$ y $g \sim e'$.

?3.3.Definición. *Se llama orientación de un espacio vectorial V a la elección de un elemento de $\mathcal{B}/\sim$.*

Se llama espacio vectorial orientado $(V, \mathbf{or})$ a un espacio vectorial V en el que se ha elegido una orientación $\mathbf{or} \in \mathcal{B}/\sim$.

Diremos que una base ordenada e de $(V, \mathbf{or})$ está orientada positivamente si $e \in \mathbf{or}$, y diremos que está orientada negativamente si pertenece a la otra orientación $-\mathbf{or}$ de V.

De acuerdo con estas definiciones, si $e \in \mathbf{or}$, otra base ordenada f está orientada positivamente si la matriz de cambio de base F de e a f tiene determinante positivo, y negativamente orientada si $\det F < 0$.

?3.4. *En el caso $V = I\!R^n$, como hay una base canónica, llamaremos orientación canónica (**can**) de $I\!R^n$ a aquella clase de equivalencia a la que pertenece la base canónica; denotaremos por $-$**can** la otra orientación. Si no decimos lo contrario, consideraremos a $I\!R^n$ siempre con la orientación **can**.*

?3.5.Definición. *Si una transformación ortogonal transforma una base ordenada en otra con la misma orientación, se dice que conserva la orientación.*

De la definición de orientación se deduce que una transformación ortogonal A conserva la orientación sii $\det A > 0$, lo cual, de acuerdo con las propiedades de las matrices ortogonales, equivale a $\det A = 1$.

?3.6. El grupo de las transformaciones ortogonales que conservan la orientación se llama *grupo especial ortogonal* $SO(n, I\!R)$, y los elementos de ese grupo se llaman *rotaciones* .

3.7.Definición. *Sea V un conjunto, U un subconjunto de V, $f : U \longrightarrow U$ una aplicación. Se dice que $g : V \longrightarrow V$ es una extensión de f o que f se extiende a g si $g|_U = f$.*

3.8.Proposición. *Sea U un subespacio vectorial de V. Toda isometría $f : U \longrightarrow U$ se puede extender a una isometría $\bar{f} : V \longrightarrow V$.*

Demostración. Consideremos la descomposición $V = U \oplus U^\perp$ y definamos $\bar{f} : V \longrightarrow V$ por

$$\bar{f}(u + x) = f(u) + x \text{ para todo } u \in U \text{ y todo } x \in U^\perp.$$

$\bar{f}$ es una isometría, porque, para todo $u, v \in U$ y todo $x, y \in U^\perp$, como f es una isometría,

$$\begin{aligned}\langle \bar{f}(u + x), \bar{f}(v + y)\rangle &= \langle f(u) + x, f(v) + y\rangle \\ &= \langle f(u), f(v)\rangle + \langle x, y\rangle = \langle u, v\rangle + \langle x, y\rangle = \langle u + x, v + y\rangle.\end{aligned}$$

3.9.Definición. *Sean U, V y f como en 3.8. Llamaremos extensión de f, y la denotaremos por $\bar{f}$, a la extensión $\bar{f}$ de f definida en la demostración de 3.8.*

3.10.Proposición. *Sean U, V y f como en 3.8. Sea g otra isometría de U en U. Se verifica que*

$$\overline{g \circ f} = \bar{g} \circ \bar{f}.$$

Demostración. Calculemos: para $u \in U$ y $x \in U^\perp$,

$$\overline{g \circ f}(u + x) = g \circ f(u) + x = \bar{g}(f(u) + x) = \bar{g}(\bar{f}(u + x)).$$

3.11.Lema. *Sea U un subespacio vectorial de V, y W un hiperplano vectorial de $U^\perp$ (obsérvese que, entonces, $U \oplus W$ es un hiperplano vectorial de V). Se tiene que la extensión de la reflexión $s_W : U^\perp \longrightarrow U^\perp$ es $s_{U\oplus W} : V \longrightarrow V$ ($\overline{s_W} = s_{U\oplus W}$).*

Demostración. Consideremos la descomposición $V = U \oplus U^\perp = U \oplus W \oplus W^\perp$, donde $W^\perp$ denota el ortogonal a W en $U^\perp$. Para cualesquiera $u \in U$, $w \in W$, $x \in W^\perp$, se tiene

$$\overline{s_W}(u + w + x) = u + s_W(w + x) = u + w - x = s_{U\oplus W}(u + w + x).$$

3.12.Lema. *Para toda simetría $s_U : V \longrightarrow V$ se verifica que $s_U^2 = Id$.*

Demostración. Calculemos: para todo $x \in V$,

$$\begin{aligned}s_U^2(x) &= s_U(\pi_U(x) - \pi_{U^\perp}(x)) \\ &= \pi_U(\pi_U(x) - \pi_{U^\perp}(x)) - \pi_{U^\perp}(\pi_U(x) - \pi_{U^\perp}(x)) = \pi_U(x) + \pi_{U^\perp}(x) = x.\end{aligned}$$

***3.13.Teorema de Cartan-Dieudonné (versión débil).** *Toda transformación ortogonal de V es el producto de m reflexiones, siendo m un número entero entre* 0 *y* n, *y donde convenimos en que la aplicación identidad es el producto de* 0 *reflexiones.*

Demostración. La haremos por inducción sobre la dimensión de V. Sea $f \in O(V)$. Si $n = 1$, entonces, para todo $x \in V$, $f(x) = \lambda x$ y $|x| = |f(x)| = |\lambda x|$, luego $|\lambda| = 1$ y $f = \pm Id$, de modo que $-Id = s_{\{0\}}$ es la única reflexión de V. Como $Id = (-Id)^0$, el teorema es trivial para $n = 1$. Supongamos que el teorema es cierto para $\dim(V) < n$. Distinguiremos dos casos:

i) f deja fijo un vector $x \neq 0$ de V, entonces $V =< \{x\} > \oplus < \{x\} >^\perp$ y, si denotamos por f' la restricción $f|_{<\{x\}>^\perp}$ (que está bien definida porque,por 0.25, $f(< \{x\} >^\perp) = f(< \{x\} >)^\perp = \{x\}^\perp$ por ser $f(x) = x$), $f(\lambda x + y) = \lambda x + f'(y)$ para todo $\lambda \in I\!\!R$ y todo $y \in< \{x\} >^\perp$. Como $\dim(< \{x\} >^\perp) = n - 1$, por la hipótesis de inducción, existen hiperplanos vectoriales W_1, ... ,W_m $(m \leq n-1)$ de $< \{x\} >^\perp$ tales que $f' = s_{W_1} \circ ... \circ s_{W_m}$. De la definición de f' se sigue que $f = \overline{f'}$ y, usando 3.10 y 3.11, se tiene

$$f = \overline{f'} = \overline{s_{W_1}} \circ ... \circ \overline{s_{W_m}} = s_{<\{x\}>\oplus W_1} \circ ... \circ s_{<\{x\}>\oplus W_m},$$

i.e., f es el producto de m reflexiones, con $m \leq n - 1$.

ii) f no tiene ningún punto fijo salvo el 0. Sea $x \in V - \{0\}$, entonces $f(x) - x \neq 0$ y

$$\begin{aligned} s_{<\{f(x)-x\}>^\perp}(f(x)) &= f(x) - 2\langle f(x), f(x) - x\rangle \frac{f(x)-x}{|f(x)-x|^2} \\ &= f(x) - 2\frac{1}{2}\langle f(x) + x + f(x) - x, f(x) - x\rangle \frac{f(x)-x}{|f(x)-x|^2} \\ &= f(x) - (\langle f(x) + x, f(x) - x\rangle - \langle f(x) - x, f(x) - x\rangle)\frac{f(x)-x}{|f(x)-x|^2} \\ &= f(x) - (f(x) - x) = x, \end{aligned}$$

luego $s_{<\{f(x)-x\}>^\perp} \circ f$ tiene x como vector fijo, estamos pues en las hipótesis i), y existen reflexiones $s_1, ..., s_m$ $(m \leq n - 1)$ de V tales que $s_{<\{f(x)-x\}>^\perp} \circ f = s_1 \circ ... \circ s_m$, y, componiendo a la izquierda son $s_{<\{f(x)-x\}>^\perp}$ y aplicando 3.12, resulta $f = s_{<\{f(x)-x\}>^\perp} \circ s_1 \circ ... \circ s_m$, y f es el producto de $m + 1 \leq n$ reflexiones.

3.14.Lema. *Sea $f \in O(V)$. El conjunto $\mathcal{F}$ de los puntos fijos de f es un subespacio vectorial de V.*

Demostración. Si $x, y \in \mathcal{F}$, entonces $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) = \lambda x + \mu y$, luego $\lambda x + \mu y \in \mathcal{F}$, y $\mathcal{F}$ es un subespacio vectorial de V.

***3.15.Teorema de Scherk (versión débil).** *Sea $f \in O(V)$ y sea U el conjunto de los puntos fijos de f (por el lema 3.14, U será un subespacio vectorial de V). Sea $r = \dim(U^\perp)$. Entonces f es el producto de r y no menos de r reflexiones.*

Demostración. . Por el teorema 3.13 sabemos que f se puede escribir como composición de $m \leq n$ reflexiones.

Vamos a ver primero que $m \geq r$. En efecto, sea $f = s_1 \circ ... \circ s_m$, donde s_i es la reflexión respecto del hiperplano W_i de V. Como los puntos de W_i son fijos para s_i (pues si $x \in W_i$, $s_i(x) = \pi_{W_i}(x) - \pi_{W_i^\perp}(x) = x$), f deja fijos los puntos de $W_1 \cap ... \cap W_m$, luego $W_1 \cap ... \cap W_m \subset U$, luego

$$\dim(W_1 \cap ... \cap W_m) \leq \dim(U) = n - r, \tag{3.15.1}$$

pero como los W_i son hiperplanos vectoriales, se tiene:

$$n = \dim(V) \geq \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2) = 2n - 2 - \dim(W_1 \cap W_2),$$

de donde

$$\dim(W_1 \cap W_2) \geq n - 2$$

y, usando esto, se obtiene

$$n = \dim(V) \geq \dim(W_1 \cap W_2) + \dim(W_3) - \dim(W_1 \cap W_2 \cap W_3) = n-2+n-1-\dim(W_1 \cap W_2 \cap W_3),$$

$$\text{y} \quad \dim(W_1 \cap W_2 \cap W_3) \geq n - 2 + n - 1 - n = n - 3,$$

y, si $\dim(W_1 \cap ... \cap W_{m-1}) \geq n - (m-1)$, entonces

$$n = \dim(V) \geq \dim(W_1 \cap ... \cap W_{m-1}) + \dim(W_m) - \dim(W_1 \cap ... \cap W_m), \quad \text{y}$$

$$\dim(W_1 \cap ... \cap W_m) \geq n - (m-1) + n - 1 - n = n - m. \tag{3.15.2}$$

De (3.15.1) y (3.15.2) resulta que $n - r \geq n - m$, luego $m \geq r$.

Vamos a ver ahora que f puede ponerse como el producto de r reflexiones. Como $f(U) = U$, entonces $f(U^\perp) = U^\perp$ (por 0.25), y $f|_{U^\perp} : U^\perp \longrightarrow U^\perp$ es una isometría entre espacios vectoriales de dimensión r. Aplicando el teorema 3.13, $f|_{U^\perp}$ se puede escribir como composición de $m \leq r$ simetrías, $f|_{U^\perp} = s_{W_1} \circ ... \circ s_{W_m}$, siendo W_i hiperplanos de $U^\perp$ y, tomando las extensiones, y aplicando los lemas 3.10 y 3.11, $f = \overline{f|_{U^\perp}} = \overline{s_{W_1}} \circ ... \circ \overline{s_{W_m}} = s_{U \oplus W_1} \circ ... \circ s_{U \oplus W_m}$, luego f se puede escribir como producto de $m \leq r$ simetrías. Como hemos visto antes que $m \geq r$, se tiene que $m = r$ y el teorema está demostrado.

3.16. Caso de $\dim(V) = 2$**.**

De acuerdo con el teorema de Scherk, una transformación ortogonal f puede:

i) Tener un subespacio vectorial de dimensión 2 de puntos fijos. Entonces $f = Id$.

ii) Tener una recta vectorial U de puntos fijos. Entonces $f = s_W$, siendo W una recta vectorial. Ahora bien, $s_W(x) = x$ es equivalente a que $2\pi_W(x) - x = x$, i.e., $x = \pi_W(x)$, lo que equivale a $x \in W$. Hemos visto, por lo tanto que $s_W(x) = x$ sii $x \in W$. Luego $W = U$, i.e., f es una reflexión respecto de U.

Obsérvese que, si $\{e_1, e_2\}$ es una base ortonormal de V tal que $e_2 \in U$, entonces la imagen por f de esa base es $\{-e_1, e_2\}$, que tiene orientación opuesta a la base inicial, luego f cambia la orientación.

iii) Tener el 0 como único vector fijo. Entonces, por 3.15, $f = s_L \circ s_M$, siendo L y M rectas vectoriales de V distintas. Elijamos una base ortonormal $\{e_1, e_2\}$ de V de modo que $e_1 \in M$. Sea l un vector unitario $|l| = 1$ de L de modo que $\measuredangle(M, L) = \measuredangle(e_1, l)$. Entonces

$$l = \langle l, e_1\rangle e_1 + \langle l, e_2\rangle e_2 = \cos\theta e_1 + \operatorname{sen}\theta e_2,$$

siendo θ el ángulo que forman l y e_1 (que es el mismo que forman L y M). Sea $x \in V$. Se puede escribir

$$x = \langle x, e_1\rangle e_1 + \langle x, e_2\rangle e_2 = |x|\{\cos\phi\, e_1 + \operatorname{sen}\phi\, e_2\},$$

siendo ϕ el ángulo que forman x y e_1. Calculemos $f(x)$,

$$f(x) = |x|\{\cos\phi f(e_1) + \operatorname{sen}\phi f(e_2)\},$$

ahora bien,

$$\begin{aligned} (3.16.1)\quad f(e_1) = s_L \circ s_M(e_1) = s_L(e_1) = 2\langle e_1, l\rangle l - e_1 &= 2\cos\theta(\cos\theta e_1 + \operatorname{sen}\theta e_2) - e_1 \\ &= (2\cos^2\theta - 1)e_1 + 2\cos\theta\operatorname{sen}\theta e_2 = \cos(2\theta)e_1 + \operatorname{sen}(2\theta)e_2, \end{aligned}$$

$$\begin{aligned} f(e_2) = s_L \circ s_M(e_2) = -s_L(e_2) = -(2\langle e_2, l\rangle l - e_2) &= -2\operatorname{sen}\theta(\cos\theta e_1 + \operatorname{sen}\theta e_2) + e_2 \\ &= -2\cos\theta\operatorname{sen}\theta e_1 + (-2\operatorname{sen}^2\theta + 1)e_2 = -\operatorname{sen}(2\theta)e_1 + \cos(2\theta)e_2, \end{aligned}$$

de donde

$$\begin{aligned} (3.16.2)\quad f(x) &= |x|\{\cos\phi(\cos(2\theta)e_1 + \operatorname{sen}(2\theta)e_2) + \operatorname{sen}\phi(-\operatorname{sen}(2\theta)e_1 + \cos(2\theta)e_2)\} \\ &= |x|\{(\cos\phi\cos(2\theta) - \operatorname{sen}\phi\operatorname{sen}(2\theta))e_1 + (\cos\phi\operatorname{sen}(2\theta) + \operatorname{sen}\phi\cos(2\theta))e_2\} \\ &= |x|\{\cos(\phi + 2\theta)e_1 + \operatorname{sen}(\phi + 2\theta)e_2\}. \end{aligned}$$

Vemos pues que, para todo $x \in V$, $f(x)$ es un vector que forma un ángulo 2θ con x. Por tanto, f es lo que en geometría elemental llamábamos una rotación de ángulo 2θ.

Obsérvese que la imagen por f de la base $\{e_1, e_2\}$ es la base $\{\cos(2\theta)e_1 + \operatorname{sen}(2\theta)e_2, -\operatorname{sen}(2\theta)e_1 + \cos(2\theta)e_2\}$, que tiene la misma orientación que $\{e_1, e_2\}$, luego f conserva la orientación. Esto justifica el que se llame rotaciones a las transformaciones ortogonales que conservan la orientación.

De (3.16.1) se deduce que la matriz de f en la base $\{e_1, e_2\}$ es

$$\begin{pmatrix} \cos(2\theta) & -\operatorname{sen}(2\theta) \\ \operatorname{sen}(2\theta) & \cos(2\theta) \end{pmatrix}.$$

3.17. Caso de $\dim(V) = 3$**.**

De acuerdo con el teorema de Scherk, una transformación ortogonal f puede:

i) Ser la identidad.

ii) Tener un un subespacio vectorial U de puntos fijos de dimensión 2. Entonces es una simetría respecto de U (demostración como la de 3.16.ii). Si se toma una base ortonormal $\{e_1, e_2, e_3\}$ de V tal que $e_1, e_2 \in U$, entonces su imagen por f es $\{e_1, e_2, -e_3\}$, que tiene orientación opuesta a la inicial y, por tanto, cambia la orientación.

iii) Tener un subespacio vectorial U de puntos fijos de dimensión 1. Entonces se puede escribir como composición de dos reflexiones. Resulta entonces que $f|_{U^\perp}$ es una rotacion (en el sentido de 3.16) sobre $U^\perp$. La actuación de f sobre V es entonces

$$f(x) = \pi_U(x) + f|_{U^\perp}(\pi_{U^\perp}(x)),$$

lo que se llama en geometría elemental una rotación de eje U.

Si se elige una base ortonormal $\{e_1, e_2, e_3\}$ de V tal que $e_1 \in U$, entonces la matriz de f en esa base tiene la forma

$$F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\operatorname{sen}\alpha \\ 0 & \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}.$$

Obsérvese que $\cos\alpha = \frac{1}{2}(\operatorname{tr} F - 1)$.

iv) Tener 0 como único punto fijo. Entonces se puede escribir como composición de tres reflexiones y su determinante (el determinante de la matriz que representa a f en una base elegida) ha de ser -1. En efecto, por ser la matriz de f ortogonal, su determinante puede ser ± 1, pero como cada reflexión cambia la orientación, el determinante de cada reflexión es -1, y el determinante del producto de tres reflexiones es -1.

Consideremos entonces la aplicación $-f$. Su determinante será 1 y, por lo tanto, solo puede ser la identidad o composición de dos reflexiones. Se tiene entonces que, o bien $f = -Id$, o $-f$ es una transformación ortogonal del tipo considerado en iii), y tendrá un subespacio vectorial invariante U de dimensión 1. Entonces $-f(U^\perp) = U^\perp$ y $f(U^\perp) = -Id(-f(U^\perp)) = -Id(U^\perp) = U^\perp$. Resulta entonces que f actúa de la siguiente forma

$$\begin{aligned} f(x) &= -(-f(x)) = -(\pi_U(x) + (-f|_{U^\perp})(\pi_{U^\perp}(x))) \\ &= -\pi_U(x) + f|_{U^\perp}(\pi_{U^\perp}(x)) = s_{U^\perp}(\pi_U(x) + f|_{U^\perp}(\pi_{U^\perp}(x))). \end{aligned}$$

Es, por tanto, la composición de una rotación $\overline{f|_{U^\perp}}$ de eje U y una reflexión $s_{U^\perp}$.

Si se elige una base ortonormal $\{e_1, e_2, e_3\}$ de V tal que $e_1 \in U$, entonces la matriz de f en esa base tiene la forma

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\alpha & -\operatorname{sen}\alpha \\ 0 & \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}.$$

Obsérvese que $\cos\alpha = \frac{1}{2}(\operatorname{tr} F + 1)$.

3.18.Ejercicio. *Sea V un espacio vectorial euclídeo de dimensión 3. Sea $\{e_1, e_2, e_3\}$ una base ortonormal de V. Sea f la isometría tal que $f(e_1) = e_2$, $f(e_2) = e_3$ y $f(e_3) = -e_1$. Describir f como la composición de una reflexión o la identidad y una rotación (decir cuales son el eje y el ángulo de rotación).*

§4 Descomposición de una isometría en un espacio afín euclídeo

En este apartado E será un espacio afín métrico de dimensión n, con espacio vectorial euclídeo asociado V.

4.1.Lema. *El subconjunto F de puntos fijos de una aplicación afín $f : E \longrightarrow E$ es un subespacio afín.*

Demostración. Sea $P \in F$. Si $U = \{u \in V/P + u \in F\}$ se tiene que, $\widetilde{f}(u) = u$ sii $\overrightarrow{f(P)f(P+u)} = \overrightarrow{P(P+u)}$ sii $\overrightarrow{Pf(P+u)} = \overrightarrow{P(P+u)}$ sii $f(P+u) = P+u$ sii $P+u \in F$, luego U es el conjunto de vectores fijos de $\widetilde{f}$, que es un subespacio vectorial (por lema 3.14) de V, luego $F = P + U$ es un subespacio afín de E.

4.2.Lema. *Sea $\varphi : V \longrightarrow V$ una isometría. Sea U el subespacio vectorial de vectores fijos de φ. Se tiene que*

$$U^\perp = (\varphi - Id)(V).$$

Demostración. Observemos primero que $\mathrm{Ker}(\varphi - Id) = \{x \in V/\varphi(x) - x = 0\} = U$, por lo tanto, $\dim((\varphi - Id)(V)) = n - \dim(U) = \dim(U^\perp)$. Ademas, si $x = \varphi(v) - v \in (\varphi - Id)(V)$, entonces, para $u \in U$,

$$\begin{aligned}\langle x, u\rangle = \langle \varphi(v) - v, u\rangle &= \langle \varphi(v), u\rangle - \langle v, u\rangle \\ &= \langle \varphi(v), \varphi(u)\rangle - \langle v, u\rangle = \langle v, u\rangle - \langle v, u\rangle = 0.\end{aligned}$$

Luego $(\varphi - Id)(V) \subset U^\perp$ y, por las dimensiones, se tiene la igualdad.

***4.3.Teorema.** *Sea $f \in Is(E)$, entonces existen una traslación T_u y una isometría $g \in Is(E)$ únicos tales que g tiene un subespacio afín no vacío F de puntos fijos, con espacio vectorial director U, $u \in U$, y $f = T_u \circ g$. Se verifica además que $f = g \circ T_u$, que U es el subespacio vectorial de vectores fijos de $\widetilde{f}$ y que $\widetilde{f} = \widetilde{g}$.*

Demostración. Vamos a ver primero la existencia de T_u y de g. Sea U el subespacio vectorial de puntos fijos de $\widetilde{f}$. Definamos $u \in U$ por $u = \pi_U(\overrightarrow{Pf(P)})$, siendo $P \in E$. Este u no depende de la elección hecha de P. En efecto, sea $R \in E$, se tiene

$$\pi_U(\overrightarrow{Rf(R)}) = \pi_U(\overrightarrow{RP} + \overrightarrow{Pf(P)} + \overrightarrow{f(P)f(R)}) = \pi_U(-\overrightarrow{PR} + \widetilde{f}(\overrightarrow{PR}) + \overrightarrow{Pf(P)}),$$

pero, por lema 4.2, $-\overrightarrow{PR} + \widetilde{f}(\overrightarrow{PR}) \in U^\perp$, luego

$$\pi_U(\overrightarrow{Rf(R)}) = \pi_U(\overrightarrow{Pf(P)}).$$

Usando de nuevo 4.2, se tiene

$$\overrightarrow{Pf(P)} = \pi_U(\overrightarrow{Pf(P)}) + \pi_{U^\perp}(\overrightarrow{Pf(P)})$$
$$= u + \pi_{(\widetilde{f}-Id)(V)}(\overrightarrow{Pf(P)}) = u + \widetilde{f}(b) - b$$

para algún $b \in V$. Sea ahora $A = P - b$, se tiene

$$\overrightarrow{Af(A)} = \overrightarrow{AP} + \overrightarrow{Pf(P)} + \widetilde{f}(\overrightarrow{PA}) = b + u + \widetilde{f}(b) - b + \widetilde{f}(-b) = u,$$

luego si definimos g por $g(Q) = f(Q) - u$, se tiene que

$$g(A) = f(A) - u = A + \overrightarrow{Af(A)} - u = A, \text{ y } \widetilde{g} = \widetilde{f},$$

y, para todo $x \in V$,

$$g(A + x) = A + \widetilde{g}(x) = A + \widetilde{f}(x) = A + x \text{ sii } x \in U,$$

es decir, $A + U$ es el conjunto de puntos fijos de g. Hemos visto, pues, que existen g y T_u verificando las condiciones del teorema.

Veamos ahora la unicidad. Sean T_y y h verificando las condiciones del teorema. Entonces $\widetilde{h} = \widetilde{f}$. Sea $H = B + W$ el espacio de los puntos fijos de h. Se verifica que $w \in W$ sii $\widetilde{f}(w) = \widetilde{h}(w) = \overrightarrow{Bh(B+w)} = \overrightarrow{B(B+w)} = w$, luego $W = U$. Por otro lado,

$$\overrightarrow{Bf(B)} = \overrightarrow{B(h(B)+y)} = \overrightarrow{B(B+y)} = y.$$

Como $\overrightarrow{Af(A)} = u$, se tiene, aplicando el lema 4.2,

$$y - u = \overrightarrow{Bf(B)} - \overrightarrow{Af(A)} = \overrightarrow{BA} + \overrightarrow{Af(B)} - \overrightarrow{Af(B)} - \overrightarrow{f(B)f(A)}$$
$$= \overrightarrow{BA} - \widetilde{f}(\overrightarrow{BA}) = (Id - \widetilde{f})(\overrightarrow{BA}) \in U^\perp,$$

luego $y = u$, ya que $y - u \in U$.

4.4. Caso $n = 2$.

Estudiemos como se puede descomponer $f \in Is(E_2)$ si $f \neq Id$. Sea U el espacio vectorial de vectores fijos de $\widetilde{f}$. Seguiremos los casos posibles con el esquema de 3.16:

i) $\dim(U) = 2$. En este caso $\widetilde{f}$ es la identidad y f es una traslación. No tiene ningún punto fijo.

ii) $\dim(U) = 1$. Entonces $\widetilde{f}$ es una simetría s_U respecto de una recta vectorial.

Si $T_u = T_0$, entonces $f = g$ tiene una recta afín $F = P + U$ de puntos fijos, y $f(Q) = P + \widetilde{f}(\overrightarrow{PQ}) = P + s_U(\overrightarrow{PQ}) = s_F(Q)$.

Si $T_u \neq T_0$, entonces f es la composición de una simetría respecto de una recta afín F y una traslación no nula paralela a F. En este caso f no tiene puntos fijos.

iii) $\dim(U) = 0$. Entonces el vector de la parte traslación es nulo, y $f = g$ tiene un punto fijo P. Además $\widetilde{f}$ es una rotación (de ángulo α). Para todo punto $Q \in E$, elegida una b.o.n. e_1, e_2 de V, se puede escribir $Q = P + \cos\phi e_1 + \operatorname{sen}\phi e_2$, de modo que

$$f(Q) = P + \widetilde{f}(\overrightarrow{PQ}) = P + \widetilde{f}(\cos\phi e_1 + \operatorname{sen}\phi e_2)$$
$$= P + \cos(\phi + \alpha)e_1 + \operatorname{sen}(\phi + \alpha)e_2,$$

es decir, f es una rotación (de ángulo α) alrededor de su punto fijo P.

4.5. Caso $n = 3$.

Estudiemos como se puede descomponer $f \in Is(E_3)$ si $f \neq Id$. Sea U el espacio vectorial de vectores fijos de $\widetilde{f}$. Seguiremos los casos posibles con el esquema de 3.17:

i) $\dim(U) = 3$. Entonces $\widetilde{f}$ es la identidad y f es una traslación. No tiene ningún punto fijo.

ii) $\dim(U) = 2$. Entonces $\widetilde{f}$ es una simetría s_U respecto de un plano vectorial.

Si $T_u = T_0$, entonces $f = g$ tiene un plano afín $F = P + U$ de puntos fijos, y $f(Q) = P + \widetilde{f}(\overrightarrow{PQ}) = P + s_U(\overrightarrow{PQ}) = s_F(Q)$.

Si $T_u \neq T_0$, entonces f es la composición de una simetría respecto de un plano afín F y una traslación no nula paralela a F. En este caso f no tiene puntos fijos.

iii) $\dim(U) = 1$. Entonces $\widetilde{f}$ es una rotación alrededor de U.

Si $T_u = T_0$, entonces $f = g$ tiene una recta afín $F = P + U$ de puntos fijos, y $f(Q) = P + \widetilde{f}(\overrightarrow{PQ})$ es una rotación alrededor de $P + U$. Este nombre se justifica porque, para todo $Q \in E$, existe un único $R = \pi_F(Q) \in F$ tal que $\overrightarrow{QR} \perp U$ y se puede escribir como $Q = R + \cos\phi e_2 + \operatorname{sen}\phi e_3$, siendo $\{e_2, e_3\}$ una b.o.n. de $U^\perp$. Si $\widetilde{f}$ es una rotación de ángulo α alrededor de U, entonces

$$f(Q) = R + \widetilde{f}(\overrightarrow{RQ}) = R + \widetilde{f}(\cos\phi e_2 + \operatorname{sen}\phi e_3) = R + \cos(\phi+\alpha)e_2 + \operatorname{sen}(\phi+\alpha)e_3,$$

que es lo que se conoce como una rotación de ángulo α alrededor de F en geometría elemental.

Si $T_u \neq T_0$, entonces f es la composición de una rotación alrededor de una recta afín F y una traslación no nula paralela a F. En este caso f no tiene puntos fijos.

iv) $\dim(U) = 0$. Entonces (cfr. 3.17.iv) existe una recta vectorial W de V tal que $\widetilde{f}$ es la composición de una rotación ψ alrededor de W y una aplicación φ de la forma $\varphi(w + x) = -w + x$ para $w \in W$ y $x \in W^\perp$.

En este caso, como el vector u de la traslación ha de verificar $u \in U$, $u = 0$ y entonces $f = g$ tiene un punto fijo, y $f(Q) = P + \widetilde{f}(\overrightarrow{PQ})$ es una rotación alrededor de $P + W$ seguida de una simetría $s_{P+W^\perp}$. En efecto: sea $Q = P + w + x$, con $w \in W$ y $x \in W^\perp$, entonces $f(Q) = P + \widetilde{f}(\overrightarrow{PQ}) = P + \varphi(\psi(\overrightarrow{PQ})) = P + \varphi(w + \psi(x)) = P - w + \psi(x)$ $= P + s_{W^\perp}(w + \psi(x)) = s_{P+W^\perp}(P + w + \psi(x))$, y la aplicación $P + w + x \mapsto P + w + \psi(x)$ es una rotación alrededor de $P + W$.

4.6.Ejercicio. *Sea E un espacio afín euclídeo de dimensión 3, y V su espacio vectorial euclídeo asociado. Sea $\{O, e_1, e_2, e_3\}$ una referencia ortonormal de E. Sea f la isometría que lleva el punto O en el punto de coordenadas $(0, 0, 1)$ y cuya aplicación lineal asociada es la dada en el ejercicio 3.18. Describir f como una composición de (posiblemente) una rotación, una reflexión y una traslación, siendo la traslación la dada por el teorema 4.3.*

CÓNICAS Y CUÁDRICAS

En todo este capítulo V designará un espacio vectorial euclídeo de dimensión n sobre el cuerpo $\mathbb{R}$ de los números reales.

$\{O, e_1, ..., e_n\}$ denotará una referencia ortonormal de un espacio afín euclídeo E de espacio vectorial asociado V. Cuando no se indique otra cosa, las coordenadas de un punto se considerarn respecto de esta referencia.

§1. Hipercuádricas y su clasificación

1.1.Definición. *Una hipercuádrica Q de E es el conjunto de puntos de E cuyas coordenadas x^i, $1 \leq i \leq n$, verifican una ecuación polinómina de segundo grado del tipo:*

$$\sum_{i,j=1}^{n} a_{ij} x^i x^j + 2 \sum_{i=1}^{n} b_i x^i + c = 0, \tag{1.1.1}$$

donde $a_{ij} = a_{ji}$.

Expresada en forma matricial, la ecuación (1.1.1) se puede escribir de las formas siguientes:

$$X^t \, A \, X + 2 \, B \, X + c = 0, \tag{1.1.2}$$

donde

$$X = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} b_1 & \dots & b_n \end{pmatrix},$$

(obsérvese que $a_{ij} = a_{ji}$ es equivalente a decir que la matriz A es simétrica) o bien

$$\widetilde{X}^t \, \widetilde{A} \, \widetilde{X} = 0, \tag{1.1.3}$$

donde

$$\widetilde{X} = \begin{pmatrix} X \\ \\ 1 \end{pmatrix} \quad \text{y} \quad \widetilde{A} = \begin{pmatrix} A & B^t \\ \\ B & c \end{pmatrix}.$$

Typeset by $\mathcal{AMS}$-TEX

1.2.Teorema. *Dada una hipercuádrica Q definida por (1.1.1), existe una referencia ortonormal $\{P, f_1, ..., f_n\}$ en la que la hipercuádrica tiene una de las formas siguientes:*

(i) $\sum_{i=1}^{r} \lambda_i\, (x^i)^2 + \delta = 0, \quad 1 \leq r \leq n, \quad r = rango(A).$

(ii) $\sum_{i=1}^{r} \lambda_i\, (x^i)^2 - 2\,\beta x_n = 0, \quad 1 \leq r \leq n, \quad r = rango(A), \quad \beta > 0.$

Estas expresiones (i) ó (ii) se llaman forma canónica o ecuación reducida de la hipercuádrica.

Si $r = 0$, la hipercuádrica se reduce a un hiperplano, y no consideramos ese caso.

Demostración. Como A es una matriz simétrica, se puede diagonalizar, y existe una base ortonormal $\{e'_1, ..., e'_n\}$ de V tal que, si

$$\overrightarrow{Ox} = \sum_{i=1}^{n} x^i e_i = \sum_{j=1}^{n} y^j e'_j, \tag{1.2.1}$$

entonces

$$\sum_{i,j=1}^{n} a_{ij} x^i x^j = \sum_{i=1}^{r} \lambda_i (y^i)^2,$$

donde λ_1, ... λ_r son los elementos no nulos de la diagonal de la matriz diagonalizada de A, y, además, $r = \text{rango}(A)$.

Haciendo el cambio de referencia de $\{O, e_1, ..., e_n\}$ a $\{O, e'_1, ..., e'_n\}$, teniendo en cuenta (1.2.1), si

$$e'_i = \sum_{j=1}^{n} E_i^j e_j, \ \text{ o, en forma matricial } \ e' = Ee$$

entonces $\sum_{j=1}^{n} x^j e_j = \sum_{i=1}^{n} y^i e'_i = \sum_{i=1}^{n} \sum_{j=1}^{n} y^i e'^j_i e_j$, se deduce

$$x^j = \sum_{i=1}^{n} y^i e'^j_i, \ \text{ o, en forma matricial } \ x = E^t y,$$

de donde, teniendo en cuenta que E es una matriz de cambio de base ortonormal y, por lo tanto, $E^{-1} = E^t$, resulta que

$$y^i = \sum_{j=1}^{n} e'^j_i x^j, \ \text{ o, en forma matricial } \ y = Ex,$$

Con este cambio de coordenadas resulta que

$$\sum_{i=1}^{n} b_i x^i = \sum_{i=1}^{n} \sum_{j=1}^{n} b_i e'^i_j y^j,$$

de modo que si definimos d_i por

$$d_j = \sum_{i=1}^{n} b_i e'^i_j \ \text{ o, en forma matricial, } \ D = BE^t,$$

tenemos que el conjuntos de puntos de E cuyas coordenadas x^i verifican la ecuación (1.1.1) en la referencia $\{O, e_1, ..., e_n\}$ coincide con el conjunto de puntos cuyas coordenadas y^i en la referencia $\{O, e'_1, ..., e'_n\}$ verifican la ecuación

$$\sum_{i=1}^{r} \lambda_i (y^i)^2 + 2\sum_{i=1}^{n} d_i y^i + c = 0.$$

Vamos a hacer ahora dos nuevos cambios de coordenadas con el objetivo final de eliminar el mayor número posible de d_i. Para ello comenzamos completando cuadrados en los dos primeros sumandos de (1.1.2) de la manera siguiente:

$$\lambda_i (y^i)^2 + 2d_i y^i = \lambda_i (y^i + \frac{d_i}{\lambda_i})^2 - \frac{d_i^2}{\lambda_i},$$

y, para que en (1.1.2) aparezcan los sumandos agrupados de esta manera, tomamos un nuevo origen del sistema de coordenadas (sin cambiar la base e')

$$O' = O - \sum_{i=1}^{r} \frac{d_i}{\lambda_i} e'_i,$$

de modo que, si $O' + \sum_{i=1}^n z^i e'_i = O' + \overrightarrow{O'x} = O + \overrightarrow{Ox} = O + \sum_{i=1}^n y^i e'_i$, entonces

$$z^i = y^i + \frac{d_i}{\lambda_i} \text{ para } 1 \leq i \leq r \text{ y } z^i = y^i \text{ para } r+1 \leq i \leq n.$$

Ahora, si llamamos $c' = c - \sum_{i=1}^r \frac{d_i^2}{\lambda_i}$, se tiene que

$$\sum_{i=1}^{r} \lambda_i (y^i)^2 + 2\sum_{i=1}^{n} d_i y^i + c = \sum_{i=1}^{r} \lambda_i (z^i)^2 + 2\sum_{i=r+1}^{n} d_i z^i + c'.$$

Si $d_i = 0$ para todo i variando entre $r+1$ y n, entonces la hipercuádrica es el conjunto de puntos cuyas coordenadas en la referencia $\{O, e'_1, ..., e'_n\}$ satisfacen la ecuación

$$\sum_{i=1}^{r} \lambda_i (z^i)^2 + c' = 0,$$

que es la ecuación de la forma (i) que buscábamos.

Si existe un $d_i \neq 0$, entonces necesariamente $r < n$, y el vector $d = \sum_{i=r+1}^n d_i e'_i \neq 0$ y pertenece al espacio ortogonal al generado por $e'_1, ..., e'_r$. Sea $\{w_{r+1}, ..., w_{n-1}, w_n = -\frac{d}{|d|}\}$ una base ortonormal de este último espacio. Si u^i son las nuevas coordenadas del punto

x en la referencia $\{O', e'_1, ..., e'_r, w_{r+1}, ..., w_n\}$, se tiene que $\sum_{i=r+1}^{n} d_i z^i = \langle d, \overrightarrow{O'x}\rangle = \langle -|d|w_n\, , \sum_{i=1}^{r} u^i e'_i + \sum_{j=r+1}^{n} u^j w_j\rangle = -|d|u^n$, de donde

$$\sum_{i=1}^{r} \lambda_i (z^i)^2 + 2 \sum_{i=r+1}^{n} d_i z^i + c' = \sum_{i=1}^{r} \lambda_i (u^i)^2 - 2|d|u^n + c'.$$

Si tomamos ahora $O'' = O' + \frac{1}{2}c'\frac{d}{|d|^2} = O' - \frac{1}{2}c'|d|^{-1}w_n$, las coordenadas v^i del punto x en la referencia $\{O'', e'_1, ..., e'_r, w_{r+1}, ..., w_n\}$ son las del vector $\overrightarrow{O''x} = \overrightarrow{O''O'} + \overrightarrow{O'x} = -\frac{1}{2}c'|d|^{-1}w_n + u^1 e'_1 + ... + u^r e'_r + u^{r+1}w_{r+1} + ... + u^n w_n$, de donde $v^i = u^i$ para $1 \leq i \leq n-1$ y $v^n = u^n - \frac{1}{2}c'|d|^{-1}$. Con este último cambio de coordenadas, se tiene que $-2|d|u^n + c' = -2|d|(v^n + \frac{1}{2}c'|d|^{-1}) + c' = -2|d|v^n$, y la ecuación que deben cumplir las coordenadas de los puntos de la cuádrica es

$$\sum_{i=1}^{r} \lambda_i (v^i)^2 - 2|d|v^n = 0,$$

que es la expresión (ii).

1.3.Nota. *Si $F_j = O'' + W_j$, siendo W_j el espacio ortogonal al vector j-ésimo de la base $\{f_1, ..., f_n\}$ en que la hipercuádrica tiene la forma canónica (i) ó (ii), la simetría s_{F_j} lleva una hipercuádrica del tipo (i) en sí misma para $1 \leq j \leq n$ y lo mismo ocurre con una cuádrica del tipo (ii) para $1 \leq j \leq n-1$. Por ello las rectas $O'' + < \{f_j\} >$ se llaman direcciones principales de la hipercuádrica. Para una cuádrica del tipo (i), la simetría respecto del origen del sistema de coordenadas en que se escribe de esta forma lleva la hipercuádrica en si misma, y se dice que es un centro de simetría de la cuádrica.*

§2. Cónicas

Cuando $\dim(E) = 2$, las hipercuádricas se llaman cónicas. Para simplificar la notación, en este apartado, $\lambda_1 \equiv \lambda$, $\lambda_2 \equiv \mu$, $x^1 \equiv x$ y $x^2 \equiv y$.

Las posibles fromas de las cónicas, de acuerdo con el Teorema 1.2, son:

Caso (i):

i.1). Supongamos $r = 2$:

i.1.1) Si $\delta = 0$:

i.1.1.1) Si $\lambda\mu > 0$, entonces la ecuación de la cónica queda $\lambda x^2 + \mu y^2 = 0$ con λ y μ con el mismo signo, y como la única solución de esta ecuación es el $(0,0)$, la cónica se reduce a un punto.

i.1.1.2) Si $\lambda\mu < 0$, por ejemplo $\lambda = a^2 > 0$ y $\mu = -b^2 < 0$, entonces la ecuación queda $a^2x^2 - b^2y^2 = 0$, que tiene por solución el par de rectas $ax + by = 0$ y $ax = -by = 0$ que se cortan en el $(0,0)$, y la cónica se reduce a este par de rectas en este caso.

i.1.2) Si $\delta \neq 0$:

i.1.2.1) Si $\lambda\mu > 0$ (i.e. $\det(A) > 0$) y $\lambda\delta < 0$ y definimos $a^2 = -\frac{\delta}{\lambda}$ y $b^2 = -\frac{\delta}{\mu}$, la ecuación de la cónica queda en la forma

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

que es una *elipse* de semiejes a y b y ejes x e y. Si $\lambda = \mu$ (i.e. $a = b$) se trata de una circunferencia.

i.1.2.2) Si $\lambda\mu > 0$ y $\lambda\delta > 0$ y definimos $a^2 = \frac{\delta}{\lambda}$ y $b^2 = \frac{\delta}{\mu}$, la ecuación de la cónica queda en la forma $\frac{x^2}{a^2} + \frac{y^2}{b^2} = -1$, que no tiene solución. En este caso no existe la cónica (es el vacío).

i.1.2.3) Si $\lambda\mu < 0$ (i.e. $\det(A) < 0$), entonces $\frac{\delta}{\lambda}$ y $\frac{\delta}{\mu}$ tienen signos opuestos (por ejemplo, $0 < a^2 = -\frac{\delta}{\lambda}$ y $0 < b^2 = \frac{\delta}{\mu}$, y la ecuación de la cónica queda en la forma

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

que es la ecuación de una *hipérbola* de ejes x e y.

i.2) Si $r = 1$ (por ejemplo $\lambda \neq 0$ y $\mu = 0$) la ecuación de la cónica es $\lambda x^2 = -\delta$, que son las dos rectas $x = \pm\sqrt{-\delta/\lambda}$ si $\lambda\delta < 0$, la recta $x = 0$ si $\delta = 0$, o el vacío si $\lambda\delta > 0$.

Caso (ii):

Como vimos en la demostración del Teorema 1.2, ha de ser $r < 2$, por tanto $r = 1$, y la ecuación de la cónica queda, definiendo $a = \lambda/(2\beta)$, de la forma

$$y = ax^2,$$

que es la ecuación de una *parábola* de eje x contenida en el semiespacio $y \geq 0$ si $\lambda > 0$ o en el semiespacio $y \leq 0$ si $\lambda < 0$.

Ejercicio. *: Reducir a la forma canónica la cónica:*

$$x^2 - y^2 - 2\sqrt{3}xy + 4x + 7 = 0,$$

y decir de qué cónica se trata.

§3 Cuádricas

Cuando $\dim(E) = 3$, las hipercuádricas se llaman cuádricas. Para simplificar la notación, en este apartado, $\lambda_1 \equiv \lambda$, $\lambda_2 \equiv \mu$, $\lambda_3 \equiv \nu$, $x^1 \equiv x$, $x^2 \equiv y$ y $x^3 \equiv z$.

Las posibles fromas de las ccuádricas, de acuerdo con el Teorema 1.2, son:

Caso (i):

i.1). Supongamos $r = 3$ (y, por tanto, $\lambda \neq 0$, $\mu \neq 0$ y $\nu \neq 0$):

i.1.1) Si $\delta = 0$, la ecuación queda $\lambda x^2 + \mu y^2 + \nu z^2 = 0$.

i.1.1.1) Si λ, μ y ν tienen el mismo signo, la cuádrica se reduce a un punto.

i.1.1.2) Si dos coeficientes tienen el mismo signo (por ejemplo $\lambda > 0$ y $\mu > 0$) y el otro opuesto ($\nu < 0$), entonces, tomando $\lambda = 1/(a^2)$, $\mu = 1/(b^2)$ y $\nu = -1/(c^2)$, la ecuación de la cuádrica se escribe

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0 \quad \text{i.e.} \quad \frac{z^2}{c^2} = \frac{x^2}{a^2} + \frac{y^2}{b^2},$$

que es la ecuación de un *cono* (para cada $(x, y, z) \in Q$, la recta $t(x, y, z)$ esta contenida en Q) con vértice el origen. Las intersecciones del cono Q con los planos $z = d \equiv$cte son elipses de semiejes $(ad)/c)$ y $(bd)/c$. Si $a = b$, se trata de un cono de revolución (es invariante por una rotación de eje z y todo él se obtiene por una "revolución alrededor del eje z" de la curva (recta) dada por las ecuaciones $z = x$ e $y = 0$; donde por "revolución alrededor del eje z" se entiende el conjunto de rotaciones de eje z y ángulo variando entre 0 y 2π).

i.1.2) Si $\delta \neq 0$, la ecuación queda $\lambda x^2 + \mu y^2 + \nu z^2 = -\delta$. Si dividimos por $-\delta$ y llamamos $\alpha = -\lambda/\delta$, $\beta = -\mu/\delta$ y $\gamma = -\nu/\delta$, la ecuación queda en la forma $\alpha x^2 + \beta y^2 + \gamma z^2 = 1$.

i.1.2.1) Si α, β y γ son positivos, tomando $\alpha = 1/(a^2)$, $\beta = 1/(b^2)$ y $\gamma = 1/(c^2)$, la ecuación de la cuádrica se escribe

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

que es la ecuación de un *elipsoide* (las intersecciones de Q con los planos $x =$cte, $y =$cte, $z =$cte son elipses) de semiejes a, b y c y ejes x, y, z. Si $a = b$, o $b = c$, o $a = c$ se trata de un elipsoide de revolución (si $a = b$ es invariante por una rotación de eje z y todo él se obtiene por una "revolución alrededor del eje z" de la curva (elipse) dada por las ecuaciones $(x^2/a^2) + (z^2/c^2) = 1$ e $y = 0$, si $b = c$ se obtiene por una revolución alrededor del eje x, ...). Si $a = b = c$, se trata de una esfera.

i.1.2.2) Si dos coeficientes (por ejemplo α y β) son positivos y el otro (γ) negativo, tomando $\alpha = 1/(a^2)$, $\beta = 1/(b^2)$ y $\gamma = -1/(c^2)$, la ecuación de la cuádrica se escribe

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1,$$

que es la ecuación de un *hiperboloide de una hoja* (las intersecciones de Q con los planos $x =$cte ó $y =$cte son hipérbolas, y las intersecciones con los planos $z =$cte son elipses) de ejes x, y, z. Si $a = b$, se trata de un hiperboloide de revolución de una hoja (si $a = b$ es invariante por una rotación de eje z y todo él se obtiene por una "revolución alrededor del eje z" de la curva (hipérbola) dada por las ecuaciones $(x^2/a^2) - (z^2/c^2) = 1$ e $y = 0$, y las curvas $z =$cte son circunferencias).

i.1.2.3) Si dos coeficientes (por ejemplo β y γ) son negativos y el otro (α) positivo, tomando $\alpha = 1/(a^2)$, $\beta = -1/(b^2)$ y $\gamma = -1/(c^2)$, la ecuación de la cuádrica se escribe

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

que es la ecuación de un *hiperboloide de dos hojas* (las intersecciones de Q con los planos y =cte ó z =cte son hipérbolas, y las intersecciones con los planos x =cte son elipses cuando no son vacias) de ejes x, y, z. Si $b = c$, se trata de un hiperboloide de revolución de dos hojas (si $b = c$ es invariante por una rotación de eje x y todo él se obtiene por una "revolución alrededor del eje x" de la curva (hipérbola) dada por las ecuaciones $(x^2/a^2) - (z^2/c^2) = 1$ e $y = 0$, y las curvas x =cte$> a$ son circunferencias).

i.1.2.4) Si los tres coeficientes son negativos, la cuádrica es vacía.

i.2) Si $r = 2$ (por ejemplo $\lambda \neq 0 \neq \mu$ y $\nu = 0$), queda la ecuación $\lambda x^2 + \mu y^2 + \delta = 0$, que es la ecuación reducida de una cónica en el plano xy. Los puntos de estas cuádrica están formados por las rectas perpendicuales al plano xy (rectas con las coordenadas x, y fijas y variando la z) cuyas intersecciones con ese plano son los puntos de la cónica $\lambda x^2 + \mu y^2 + \delta = 0$. Son los llamados cilindros de curva directriz una cónica (de la forma (i)) o cuya sección normal es esa cónica.

Caso (ii):

Como vimos en la demostración del Teorema 1.2, ha de ser $r < 3$, por tanto quedan las siguientes posibilidades:

ii.1) $r = 2$, por ejemplo $\lambda \neq 0 \neq \mu$ y $\nu = 0$. Queda la siguiente ecuación para la cuádrica; $\lambda x^2 + \mu y^2 = 2\beta z$ con $\beta > 0$.

ii.1.1) Si $\lambda\mu > 0$, tomando $\lambda/(2\beta) = \pm 1/(a^2)$ y $\mu/(2\beta) = \pm 1/(b^2)$ (con el signo $+$ si $\lambda > 0$ y con el signo $-$ si $\lambda < 0$, la ecuación de la cuádrica se escribe

$$z = \frac{x^2}{a^2} + \frac{y^2}{b^2} \quad \text{o} \quad z = -\frac{x^2}{a^2} - \frac{y^2}{b^2},$$

que es la ecuación de un *paraboloide elíptico* (las intersecciones de Q con los planos x =cte ó y =cte son parábolas, y las intersecciones con los planos z =cte son elipses) de ejes x, y. Si $a = b$, se trata de un paraboloide de revolución(si $a = b$ es invariante por una rotación de eje z y todo él se obtiene por una "revolución alrededor del eje z" de la curva (parábola) dada por las ecuaciones $z = (x^2)/(a^2)$ e $y = 0$, y las curvas z =cte son circunferencias).

ii.1.2) Si $\lambda\mu < 0$ (por ejemplo $\lambda > 0$ y $\mu < 0$), tomando $\lambda/(2\beta) = 1/(a^2)$ y $\mu/(2\beta) = -1/(b^2)$, la ecuación de la cuádrica se escribe

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

que es la ecuación de un *paraboloide hiperbólico* (las intersecciones de Q con los planos x =cte ó y =cte son parábolas, y las intersecciones con los planos z =cte son hipérbolas) de ejes x, y.

ii.2) $r = 1$, por ejmplo $\lambda \neq 0$ y $\mu = \nu = 0$. La ecuación de la cuádrica queda en la forma

$$\lambda x^2 = 2\beta z \quad \text{i.e.} \quad z = \frac{\lambda}{2\beta} x^2,$$

que es una parábola en el plano xz, i.e., es la cuádrica Q es un cilindro cuyas rectas generatrices son ortogonales al plano xz y cuyas curva directriz en ese plano es una parábola.