Capítulo VIII

Espacios afines

El concepto del espacio vectorial real de tres dimensiones surge de un modo natural al describir las propiedades principales del espacio físico corriente, siempre que en este espacio se haya destacado un punto como el origen de coordenadas. Puesto que ningún punto del espacio físico tiene ventaja alguna ante los demás, conviene tener, además del concepto de espacio vectorial, un modelo matemático que corresponda al concepto del espacio en el que no se hayan fijado de antemano puntos algunos. Ejemplos de estos modelos son el espacio afín y el espacio euclídeo puntual. En este capítulo serán expuestos los elementos de la teoría de estos espacios.

§ 29. Espacios afines generales

29.1. Axiomática. Sea dado un espacio vectorial $\mathfrak{L}$ sobre un cuerpo conmutativo K. Se llama *espacio afín* $\mathfrak{A}$ *sobre el espacio vectorial* $\mathfrak{L}$ un conjunto arbitrario tal que a todo par de elementos X e Y del mismo le corresponde un vector de $\mathfrak{L}$, indicado en lo sucesivo por $\overline{XY}$, con la condición de que se cumplen las exigencias siguientes:

A_1: *para todo* $X \in \mathfrak{A}$ *y todo* $v \in \mathfrak{L}$ *existe un elemento* $Y \in \mathfrak{A}$, *y sólo uno, que satisface la relación* $\overline{XY} = v$;

A_2: *para cualesquiera* X, Y *y* Z *de* $\mathfrak{A}$ *es válida la igualdad* $\overline{XY} + \overline{YZ} = \overline{XZ}$.

Los elementos de un espacio afín se llaman *puntos* del mismo y los vectores de $\mathfrak{L}$ se llaman *vectores libres* del espacio $\mathfrak{A}$.

Para cualesquiera $X \in \mathfrak{A}$ y $v \in \mathfrak{L}$ indicaremos por $X \cdot v$ aquel punto $Y \in \mathfrak{A}$ para el cual $\overline{XY} = v$. Según el axioma A_1 un punto así de $\mathfrak{A}$ existe y es único. De aquí, en particular, se desprende que para cualesquiera X, $Y \in \mathfrak{A}$ y $u \in \mathfrak{L}$ se tiene

$$X \cdot \overline{XY} = Y \quad \text{y} \quad \overline{X \cdot (X \cdot u)} = u. \tag{1}$$

En virtud del axioma A_2 tenemos $\overline{XX}+\overline{XY}=\overline{XY}$ y, por ello,

$$\overline{XX}=o \quad \text{y} \quad X\cdot o=X \quad (X\in\mathfrak{A} \quad \text{y} \quad o\in\mathfrak{L}).$$

En vista del mismo axioma se tiene $\overline{XY}+\overline{YX}=\overline{XX}$ y, por ello,

$$\overline{XY}=-\overline{YX}. \tag{2}$$

Finalmente, para cualesquiera $X, O, B, C, D\in\mathfrak{A}$ se tiene

$$(X\cdot\overline{OB})\cdot\overline{CD}=X\cdot(\overline{OB}+\overline{CD}). \tag{3}$$

En efecto, sea $X\cdot\overline{OB}=Y$ e $Y\cdot\overline{CD}=Z$. Debido a A_2 y a (1), se tiene

$$\overline{OB}=\overline{XY}, \quad \overline{CD}=\overline{YZ} \quad \text{y} \quad \overline{OB}+\overline{CD}=\overline{XZ}$$

y por esto

$$(X\cdot\overline{OB})\cdot\overline{CD}=Z=X\cdot\overline{XZ}=X\cdot(\overline{OB}+\overline{CD}).$$

Pongamos en correspondencia a todo par (O, B) de puntos del espacio $\mathfrak{A}$ la aplicación de $\mathfrak{A}$ en $\mathfrak{A}$, indicada por $\overrightarrow{OB}$ y definida mediante la fórmula

$$\overrightarrow{OB}\colon X\to X\cdot\overline{OB} \quad (X\in\mathfrak{A}), \tag{4}$$

es decir, tomemos $X\cdot\overrightarrow{OB}=X\cdot\overline{OB}$. Las aplicaciones $\overrightarrow{OB}$ $(O, B\in\mathfrak{A})$ se llaman *traslaciones* (o también *desplazamientos* o *traslados*) del espacio $\mathfrak{A}$. La superposición de unas traslaciones $\overrightarrow{OB}$ y $\overrightarrow{CD}$ suele llamarse suma de estas traslaciones, es decir, se toma por definición que

$$X\cdot(\overrightarrow{OB}+\overrightarrow{CD})=(X\cdot\overrightarrow{OB})\cdot\overrightarrow{CD}=X\cdot(\overline{OB}+\overline{CD}). \tag{5}$$

El producto de un elemento $\lambda\in K$ por una traslación $\overrightarrow{OB}$ se define mediante la fórmula

$$X\cdot(\lambda\overrightarrow{OB})=X\cdot(\lambda\overline{OB}). \tag{6}$$

El conjunto de todas las traslaciones del espacio $\mathfrak{A}$, provisto de las operaciones de adición y de multiplicación por todos los $\lambda\in K$, se indicará por $L(\mathfrak{A})$. Las igualdades (5) y (6) muestran que la aplicación $\overline{OX}\to\overrightarrow{OX}$ es un isomorfismo de $\mathfrak{L}$ sobre $L(\mathfrak{A})$ y por esto $L(\mathfrak{A})$ es *un espacio vectorial sobre* K.

Fijemos ahora en $\mathfrak{A}$ un punto cualquiera O. Entonces la aplicación $X\to\overline{OX}$ $(X\in\mathfrak{A})$ establece una correspondencia biyectiva entre los puntos de $\mathfrak{A}$ y los vectores de $\mathfrak{L}$. El vector $\overline{OX}$ suele llamarse *radio vector* del punto X respecto al punto inicial O. El radio vector del punto desplazado $X\cdot\overrightarrow{BC}$ se expresa en términos del radio vector del punto dado inicialmente X mediante la fórmula

$$\overline{O(X\cdot\overrightarrow{BC})}=\overline{OX}+(\overline{OC}-\overline{OB}). \tag{7}$$

Efectivamente, tomando $X\cdot\overrightarrow{BC}=Y$, obtenemos $\overline{XY}=\overline{BC}$ y, por consiguiente,

$$\overline{O(X\cdot\overrightarrow{BC})}=\overline{OY}=\overline{OX}+\overline{XY}=\overline{OX}+\overline{BC}.$$

Supongamos ahora dado un espacio vectorial $\mathfrak{L}$ sobre un cuerpo conmutativo K. Llamemos puntos a los vectores de $\mathfrak{L}$ y pongamos en correspondencia a todo par (u, v) de ellos el vector $\overline{uv}=v-u$. Está claro que en este caso las exigencias A_1 y A_2 se cumplirán sin duda alguna y de esta forma el propio espacio vectorial $\mathfrak{L}$ se convertirá en un espacio afín sobre sí mismo. Indicaremos este espacio afín por $A(\mathfrak{L})$. Puesto que la igualdad $\overline{xy}=u$ equivale a la relación $y-x=u$, obtenemos para las traslaciones del espacio $A(\mathfrak{L})$ la fórmula

$$x\cdot\overrightarrow{bc}=x+(c-b). \tag{8}$$

Desde el punto de vista puramente algebraico, es conveniente a veces considerar, además del concepto de *un espacio afín sobre un espacio vectorial*, el concepto de un espacio afín *sobre un cuerpo conmutativo* que surge del modo siguiente.

Indiquemos por $\boldsymbol{A}$ el conjunto de todos los puntos de un espacio afín $\mathfrak{A}$ sobre un espacio vectorial $\mathfrak{L}$ sobre un cuerpo conmutativo principal K. Introducimos en $\boldsymbol{A}$ unas operaciones nuevas $\boldsymbol{S}(X, Y, Z)$ y $\boldsymbol{P}_\lambda(X, Y)$ tomando por definición

$$\boldsymbol{S}(X, Y, Z)=Z\cdot\overline{XY} \text{ y } \boldsymbol{P}_\lambda(X, Y)=X\cdot(\lambda\overline{XY}) \tag{9}$$

y convertiendo con ello $\mathfrak{A}$ en un álgebra nueva

$$\mathfrak{A}^*=(\boldsymbol{A};\ \boldsymbol{S},\ \{\boldsymbol{P}_\lambda\})\quad (\lambda\in K).$$

Es fácil comprobar que en el álgebra $\mathfrak{A}^*$ las operaciones $\boldsymbol{S}$ y $\boldsymbol{P}_\lambda$ poseen las propiedades siguientes:

A_1^*: *Para cualesquiera* $O, B, C, X\in\mathfrak{A}^*$ se tiene

$$\boldsymbol{S}(B, C, \boldsymbol{S}(O, B, X))=\boldsymbol{S}(O, C, X); \tag{10}$$

A_2^*: *Para cualesquiera* $O, B, X\in\mathfrak{A}^*$ y $\lambda\in K$ se tiene

$$\boldsymbol{S}(O, B, \boldsymbol{P}_\lambda(O, X))=\boldsymbol{P}_\lambda(B, \boldsymbol{S}(O, B, X)); \tag{11}$$

A_3^*: *Fijamos en* $\mathfrak{A}^*$ *un punto cualquiera* O *y mediante las operaciones* $\boldsymbol{S}$ *y* $\boldsymbol{P}_\lambda$ *introducimos para los elementos de* $\mathfrak{A}^*$ *unas operaciones nuevas*

$$X+_O Y=\boldsymbol{S}(O, X, Y) \text{ y } \lambda\cdot_O X=\boldsymbol{P}_\lambda(O, X), \tag{12}$$

llamadas adición local y multiplicación local de los elementos de $\mathfrak{A}^*$ *por los números* λ *(en el punto* O*). Respecto a estas operaciones locales* $+_O$ *y* $\lambda\cdot_O$ *los elementos de* $\mathfrak{A}^*$ *forman un espacio vectorial sobre* K *que se indica en lo sucesivo por* $\mathfrak{L}_O$.

En efecto, hemos visto que la aplicación $X \to \overline{OX}$ ($X \in \boldsymbol{A}$) es una correspondencia biyectiva entre los puntos de $\boldsymbol{A}$ y los vectores de $\mathfrak{L}$. Pasemos, de acuerdo con esta correspondencia, las operaciones vectoriales de $\mathfrak{L}$ a $\boldsymbol{A}$, es decir, tomemos por definición

$$X+Y=Z \Leftrightarrow \overline{OX}+\overline{OY}=\overline{OZ} \text{ y } \lambda X=Z \Leftrightarrow \lambda\overline{OX}=\overline{OZ}. \tag{13}$$

Entonces $\boldsymbol{A}$ se convierte en un espacio lineal sobre K y el punto O será el elemento nulo de este espacio. Puesto que

$$X\cdot\lambda\overline{AB}=Y \Leftrightarrow \overline{XY}=\lambda\overline{AB} \Leftrightarrow \overline{OY}-\overline{OX}=\lambda(\overline{OB}-\overline{OA}),$$

y puesto que de (13) se deduce que

$$\overline{OY}-\overline{OX}=\lambda(\overline{OB}-\overline{OA}) \Leftrightarrow Y-X=\lambda(B-A),$$

tenemos

$$X\cdot\lambda\overline{AB}=Y \Leftrightarrow Y=X+\lambda(B-A),$$

de donde

$$X\cdot\lambda\overline{AB}=X+\lambda(B-A) \tag{14}$$

y, por consiguiente,

$$\begin{gathered} \boldsymbol{S}(X,\ Y,\ Z)=Y+Z-X, \\ \boldsymbol{P}_\lambda(X,\ Y)=\lambda(Y-X)+X, \\ X+_O Y=X+Y \quad \text{y} \quad \lambda\cdot_O X=\lambda X. \end{gathered}$$

Después de esto las afirmaciones A_1^*, A_2^* y A_3^* se hacen evidentes.

Se llama espacio afín sobre un cuerpo conmutativo K *un conjunto de elementos* $\boldsymbol{A}$ *provisto de una operación ternaria* $\boldsymbol{S}(X, Y, Z)$ *y una serie de operaciones binarias* $\boldsymbol{P}_\lambda(X, Y)$ ($\lambda \in K$) *que cumplen sobre* $\boldsymbol{A}$ *las exigencias* A_1^*, A_2^* *y* A_3^*.

Un espacio afín sobre un cuerpo conmutativo K, al igual que un espacio vectorial sobre un cuerpo conmutativo K, es un *álgebra* (cuya signatura depende de K y es, en el caso general, infinita). Por esto para los espacios afines sobre un cuerpo conmutativo resultan definidos automáticamente una serie de conceptos como isomorfismo, endomorfismo, congruencia, etc.

Hemos visto como todo espacio afín $\mathfrak{A}$ sobre un espacio vectorial $\mathfrak{L}$ sobre un cuerpo conmutativo K puede ser convertido en un espacio afín $\mathfrak{A}^*$ sobre el cuerpo conmutativo K. Queremos demostrar ahora que de esta forma se puede obtener cualquier espacio afín sobre K.

Sea, pues, $\mathfrak{B}=(\boldsymbol{A},\ \boldsymbol{S},\ \boldsymbol{P}_\lambda)$ un espacio afín sobre un cuerpo conmutativo K. Representando la identidad (10) en la forma

$$C+_B(B+_O X)=C+_O X \tag{15}$$

y tomando $B=O$, obtenemos

$$C+_O(O+X)=C+_O X.$$

Esta es una ecuación en el espacio vectorial $\mathfrak{L}_O$ y, por ello,

$$O+_O X=X, \tag{16}$$

es decir, el punto O es el cero del espacio $\mathfrak{L}_O$.
Para cualesquiera puntos fijos O y B la aplicación

$$X\to \boldsymbol{S}(O,\ B,\ X)\quad (X\in\mathfrak{B})$$

del espacio $\mathfrak{B}$ en sí mismo se llamará traslación de $\mathfrak{B}$ y se indicará por $\overrightarrow{OB}$. Por consiguiente, tomamos por definición

$$X\cdot\overrightarrow{OB}=\boldsymbol{S}(O,\ B,\ X)=X+_O B. \tag{17}$$

Representando la relación (16) en la forma $X\cdot\overrightarrow{OO}=X$, vemos que la traslación $\overrightarrow{OO}$ es la aplicación idéntica de $\mathfrak{B}$ sobre sí mismo y que, en particular, $\overrightarrow{OO}=\overrightarrow{XX}$ para cualesquiera $O,\ X\in\mathfrak{B}$.

Al igual que antes, llamamos suma $\overrightarrow{OB}+\overrightarrow{CD}$ de traslaciones la superposición de las mismas, es decir, tomamos

$$X\cdot(\overrightarrow{OB}+\overrightarrow{CD})=(X\cdot\overrightarrow{OB})\cdot\overrightarrow{CD}=\boldsymbol{S}(C,\ D,\ \boldsymbol{S}(O,\ B,\ X)). \tag{18}$$

La identidad (10) significa que para las traslaciones es válida *la regla del triángulo*

$$\overrightarrow{OB}+\overrightarrow{BC}=\overrightarrow{OC}, \tag{19}$$

de la cual se desprende, en particular, que $\overrightarrow{OB}+\overrightarrow{BO}=\overrightarrow{OO}$, es decir, que las traslaciones $\overrightarrow{OB}$ y $\overrightarrow{BO}$ son aplicaciones *recíprocamente inversas*.

De (15) y (17) obtenemos

$$X\cdot\overrightarrow{CD}=X+_C D=X+_C(C+_B(D-_B C))=X+_B(D-_B C),$$

es decir,

$$\overrightarrow{CD}=\overrightarrow{B\,(D-_B C)}. \tag{20}$$

Combinando las fórmulas (20) y (19), llegamos a la relación

$$\overrightarrow{OB}+\overrightarrow{CD}=\overrightarrow{OB}+\overrightarrow{B\,(D-_B C)}=\overrightarrow{O\,(D-_B C)}$$

que muestra que *la suma de traslaciones es una traslación*.

Definamos, finalmente, la operación de multiplicación de los números $\lambda \in K$ por una traslación $\overrightarrow{OB}$, poniendo

$$\lambda \cdot \overrightarrow{OB} = \overrightarrow{O(\lambda \cdot_O B)} = \overrightarrow{OP_\lambda(O, B)},$$

es decir, aceptando que

$$X \cdot (\lambda \overrightarrow{OB}) = \boldsymbol{S}(O, \boldsymbol{P}_\lambda(O, B), X) = \boldsymbol{P}_\lambda(X, X \cdot \overrightarrow{OB}). \tag{21}$$

De la fórmula (21) se ve que siendo $\overrightarrow{OB} = \overrightarrow{CD}$, se tiene $\lambda \cdot \overrightarrow{OB} = \lambda \cdot \overrightarrow{CD}$, es decir, la operación de multiplicación de un número por una traslación está definida unívocamente.

TEOREMA 1. *El conjunto $L(\mathfrak{A}^*)$ de todas las traslaciones de un espacio afín $\mathfrak{A}^*$ sobre un cuerpo conmutativo K es un espacio vectorial sobre K respecto a las operaciones de adición de traslaciones y de multiplicación de un número por una traslación. El espacio $L(\mathfrak{A}^*)$ es isomorfo a cualquier de los subespacios $\mathfrak{L}_O$. Poniendo en correspondencia a todo par de puntos X, Y de $\mathfrak{A}^*$ la traslación $\overrightarrow{XY} \in L(\mathfrak{A}^*)$ convertimos $\mathfrak{A}^*$ en un espacio afín $\mathfrak{A}$ sobre el espacio vectorial $L(\mathfrak{A}^*)$. Si en el espacio $\mathfrak{A}$ se introducen mediante las fórmulas* (9) *las operaciones $\boldsymbol{S}$ y $\boldsymbol{P}_\lambda$, se obtiene de nuevo el espacio $\mathfrak{A}^*$.*

Las dos últimas afirmaciones son evidentes, ya que, por ejemplo, las fórmulas (9), aplicadas al espacio $\mathfrak{A}$, se convierten simplemente en las fórmulas (17) y (21). Demostremos las dos primeras afirmaciones.

Tomemos un punto cualquiera $O \in \mathfrak{A}^*$ y consideremos la aplicación

$$A: X \to \overrightarrow{OX} \quad (X \in \mathfrak{L}_O).$$

Puesto que cualquier traslación puede ser representada en la forma $\overrightarrow{OX}$, resulta que A es una aplicación de $\mathfrak{L}_O$ sobre $L(\mathfrak{A}^*)$. Veamos si esta aplicación A conserva las operaciones de adición y de multiplicación por números. Tenemos

$$Z \cdot \overrightarrow{O(X +_O Y)} = Z +_O (X +_O Y) = (Z +_O X) +_O Y = Z \cdot (\overrightarrow{OX} + \overrightarrow{OY}),$$

es decir, $\overrightarrow{O(X +_O Y)} = \overrightarrow{OX} + \overrightarrow{OY}$ y, por consiguiente,

$$(X +_O Y)A = XA + YA.$$

Análogamente, de (21) obtenemos $(\lambda \cdot_O X)A = \lambda \cdot XA$. Por consiguiente, la aplicación A es un isomorfismo de $\mathfrak{L}_O$ sobre $L(\mathfrak{A}^*)$ respecto a las operaciones de adición y de multiplicación por números y, por ello, $L(\mathfrak{A}^*)$, así como $\mathfrak{L}_0$, es un espacio vectorial sobre K.

Planteemos, finalmente, la pregunta: ¿bajo qué condiciones serán isomorfos dos espacios afines?

El concepto de isomorfismo suele introducirse para las álgebras arbitrarias con una misma signatura. Los espacios afines sobre un cuerpo conmutativo K han sido definidos como álgebras cuya signatura depende del cuerpo conmutativo K. Por ello para los espacios afines sobre *un mismo cuerpo conmutativo* está definido también el concepto de isomorfismo. En cuanto a los espacios afines sobre un espacio vectorial dado, éstos no han sido definidos como unas álgebras de signatura fija, sino como unos sistemas de una estructura más compleja para los cuales no hemos introducido por ahora el concepto de isomorfismo.

Introduzcamos la definición siguiente.

Se llama *homomorfismo* de un espacio afín $\mathfrak{A}$, definido sobre un espacio vectorial $\mathfrak{L}$, en un espacio afín $\mathfrak{B}$, definido sobre el mismo espacio vectorial $\mathfrak{L}$, toda aplicación puntual $\varphi\colon \mathfrak{A} \to \mathfrak{B}$ para la cual

$$\overline{XY} = \overline{X^\varphi Y^\varphi} \quad (X,\ Y \in \mathfrak{A}). \tag{22}$$

Está claro que dicha condición equivale a la siguiente:

$$(X \cdot v)^\varphi = X^\varphi \cdot v \quad (X \in \mathfrak{A} \quad \text{y} \quad v \in \mathfrak{L}).$$

Se ve de (22) que la aplicación φ transforma distintos puntos en distintos y que siendo φ un homomorfismo de $\mathfrak{A}$ sobre $\mathfrak{B}$, la aplicación inversa φ^{-1} es un homomorfismo de $\mathfrak{B}$ sobre $\mathfrak{A}$.

Un homomorfismo de $\mathfrak{A}$ sobre $\mathfrak{B}$ se llama isomorfismo si la aplicación inversa es un homomorfismo de $\mathfrak{B}$ sobre $\mathfrak{A}$. Luego, *para los espacios afines sobre un espacio vectorial fijo los conceptos de homomorfismo y de isomorfismo son equivalentes.*

Un isomorfismo de un espacio afín $\mathfrak{A}$, definido sobre un espacio vectorial $\mathfrak{L}$, sobre si mismo se llama *automorfismo* de $\mathfrak{A}$ sobre $\mathfrak{L}$.

TEOREMA 2. *Todos los espacios afines sobre un mismo espacio vectorial son isomorfos. Los automorfismos de un espacio afín sobre un espacio vectorial son las traslaciones del mismo.*

Sean $\mathfrak{A}$ y $\mathfrak{B}$ unos espacios afines sobre un espacio vectorial $\mathfrak{L}$. Tomemos unos puntos arbitrarios $A \in \mathfrak{A}$ y $B \in \mathfrak{B}$ y consideremos la aplicación $\varphi\colon Av \to Bv$ $(v \in \mathfrak{L})$. De los axiomas A_1 y A_2 deducimos directamente que φ es un isomorfismo de $\mathfrak{A}$ sobre $\mathfrak{B}$.

Además, si φ es un automorfismo del espacio afín $\mathfrak{A}$ sobre el espacio vectorial $\mathfrak{L}$, obtenemos de la fórmula (22) que $\overline{AX} = \overline{A^\varphi X^\varphi}$, de donde $X^\varphi = A^\varphi \cdot \overline{AX}$ y por ello

$$X^\varphi = X \cdot \overline{XA^\varphi} \cdot \overline{AX} = X \cdot (\overline{AX} + \overline{XA^\varphi}) = X \cdot \overline{AA^\varphi},$$

es decir, φ es la traslación de $\mathfrak{A}$ determinada por el vector $\overline{AA^\varphi}$. Recíprocamente, si φ es una traslación, se tiene que φ es de la forma $X^\varphi = X \cdot v$, donde v es un vector arbitrario fijo de $\mathfrak{L}$, y por lo tanto

$$\overline{X^\varphi Y^\varphi} = \overline{(X \cdot v)\,(Y \cdot v)} = \overline{XY}.$$

Un espacio afín sobre un cuerpo conmutativo K es un álgebra de signatura $\boldsymbol{S}$, $\boldsymbol{P}_\lambda$ $(\lambda \in K)$.

Un isomorfismo φ de un espacio afín $\mathfrak{A}$, definido sobre un cuerpo conmutativo K, sobre un espacio afín $\mathfrak{B}$, definido sobre el mismo cuerpo K, es una aplicación biyectiva de $\mathfrak{A}$ sobre $\mathfrak{B}$ que satisface las condiciones

$$\boldsymbol{S}(X, Y, Z)^{\varphi} = \boldsymbol{S}(X^{\varphi}, Y^{\varphi}, Z^{\varphi}),$$
$$\boldsymbol{P}_{\lambda}(X, Y)^{\varphi} = \boldsymbol{P}_{\lambda}(X^{\varphi}, Y^{\varphi})$$

o, que es equivalente, las condiciones

$$(Z \cdot \overline{XY})^{\varphi} = Z^{\varphi} \cdot \overline{X^{\varphi}Y^{\varphi}}, \tag{23}$$
$$(X \cdot \lambda\overline{XY})^{\varphi} = X^{\varphi} \cdot \lambda\overline{X^{\varphi}Y^{\varphi}}. \tag{24}$$

TEOREMA 3. *Para que una aplicación φ: $\mathfrak{A} \dashrightarrow \mathfrak{B}$ de un espacio afín $\mathfrak{A}$, definido sobre un cuerpo conmutativo K, sobre un espacio afín $\mathfrak{B}$, definido sobre el mismo cuerpo, sea un isomorfismo, es necesario y suficiente que la aplicación ψ: $\overline{XY} \to \overline{X^{\varphi}Y^{\varphi}}$ $(X, Y \in \mathfrak{A})$ sea un isomorfismo del espacio vectorial $L(\mathfrak{A})$ sobre el espacio vectorial $L(\mathfrak{B})$.*

NECESIDAD. Probemos, ante todo, que la aplicación ψ es unívoca. Sea $\overline{XY} = \overline{UV}$ $(X, Y, U, V \in \mathfrak{A})$ y, por consiguiente, $Z \cdot \overline{XY} = Z \cdot \overline{UV}$. De (23) obtenemos $Z^{\varphi} \cdot \overline{X^{\varphi}Y^{\varphi}} = Z^{\varphi} \cdot \overline{U^{\varphi}V^{\varphi}}$. Puesto que Z^{φ} recorre todo el espacio $\mathfrak{B}$, de la última igualdad tenemos $\overline{X^{\varphi}Y^{\varphi}} = \overline{U^{\varphi}V^{\varphi}}$ y por esto la aplicación ψ es unívoca. Representando ahora la relación (23) en la forma $(Zu)^{\varphi} = Z^{\varphi}u^{\psi}$, obtenemos para cualesquiera $v, w \in L(\mathfrak{A})$

$$Z^{\varphi}(v+w)^{\psi} = (Z(v+w))^{\varphi} = ((Zv)w)^{\varphi} = (Z^{\varphi}v^{\psi})w^{\psi} = Z^{\varphi}(v^{\psi}+w^{\psi}),$$

de donde

$$(v+w)^{\psi} = v^{\psi} + w^{\psi}. \tag{25}$$

Análogamente, representando (24) en la forma $(X \cdot \lambda u)^{\varphi} = X^{\varphi} \cdot \lambda u^{\psi}$, obtenemos

$$X^{\varphi}(\lambda u)^{\psi} = (X \cdot (\lambda u))^{\varphi} = X^{\varphi} \cdot \lambda u^{\psi},$$

de donde

$$(\lambda u)^{\psi} = \lambda u^{\psi}. \tag{26}$$

Las relaciones (25) y (26) significan que ψ es un isomorfismo de $L(\mathfrak{A})$ sobre $L(\mathfrak{B})$.

SUFICIENCIA Supongamos que dada la aplicación φ: $\mathfrak{A} \to \mathfrak{B}$ del conjunto $\mathfrak{A}$ sobre el conjunto $\mathfrak{B}$, la aplicación ψ: $\overline{XY} \to \overline{X^{\varphi}Y^{\varphi}}$ es un isomorfismo de $L(\mathfrak{A})$ sobre $L(\mathfrak{B})$. Puesto que en un isomorfismo de espacios vectoriales el vector nulo corresponde al vector nulo, de $X^{\varphi} = Y^{\varphi}$ resulta $\overline{X^{\varphi}Y^{\varphi}} = o$ y por esto $\overline{XY} = o$ y $X = Y$, es decir, la aplicación φ es biyectiva. Para demostrar ahora que para φ es válida la relación (24), tomemos en $\mathfrak{A}$ un punto U tal que $\overline{XY} = \overline{ZU}$.

Entonces se tiene

$$(Z\cdot\overline{XY})^{\varphi}=U^{\varphi}=Z^{\varphi}\cdot\overline{Z^{\varphi}U^{\varphi}}=Z^{\varphi}\cdot\overline{X^{\varphi}Y^{\varphi}} \tag{27}$$

y, en particular,

$$(Zu)^{\varphi}=Z^{\varphi}u^{\psi} \quad (u\in L(\mathfrak{A})).$$

Ahora resulta de (26) que

$$(X\cdot\lambda\overline{XY})^{\varphi}=X^{\varphi}\cdot(\lambda\overline{XY})^{\psi}=X^{\varphi}\lambda\overline{XY}^{\psi}=X^{\varphi}\lambda\overline{X^{\varphi}Y^{\varphi}}$$

que es lo que se quería demostrar.

COROLARIO 1. *Los espacios afines sobre un cuerpo conmutativo dado son isomorfos cuando, y sólo cuando, sus espacios vectoriales de traslaciones son de una misma dimensión.*

Efectivamente, según el teorema 3, un isomorfismo entre espacios afines sobre un cuerpo conmutativo equivale a un isomorfismo de los espacios vectoriales de traslaciones y un isomorfismo de espacios vectoriales equivale (véase el p. 4.3) a la coincidencia de sus dimensiones.

COROLARIO 2. *Los automorfismos de un espacio afín $\mathfrak{A}$ sobre un cuerpo conmutativo que dejan en su sitio un punto $O\in\mathfrak{A}$ son aplicaciones de tipo φ: $X\to O\cdot\overline{OX}^{\psi}$, donde ψ es un automorfismo fijo cualquiera del espacio vectorial $L(\mathfrak{A})$.*

En efecto, siendo φ un automorfismo del espacio $\mathfrak{A}$ que deja en su sitio el punto O, se tiene

$$X^{\varphi}=O\cdot\overline{OX^{\varphi}}=O\cdot\overline{O^{\varphi}X^{\varphi}}=O\cdot\overline{OX}^{\psi}.$$

Recíprocamente, si ψ es un automorfismo de $L(\mathfrak{A})$, se tiene

$$\overline{X^{\varphi}Y^{\varphi}}=\overline{(O\cdot\overline{OX}^{\psi})\,(O\cdot\overline{OY}^{\psi})}=\overline{OY}^{\psi}-\overline{OX}^{\psi}=(\overline{OY}-\overline{OX})^{\psi}=\overline{XY}^{\psi}$$

y por esto φ es, según el teorema 3, un automorfismo del espacio $\mathfrak{A}$.

TEOREMA 4. *Siendo O un punto arbitrario de un espacio afín $\mathfrak{A}$ sobre un cuerpo conmutativo K, todo automorfismo φ del espacio $\mathfrak{A}$ sobre K puede ser representado en la forma*

$$\varphi=\varphi_O\varphi_L, \tag{28}$$

donde φ_L es una traslación adecuada de $\mathfrak{A}$ y φ_O es un automorfismo de $\mathfrak{A}$ que deja en su sitio el punto O. Para todo automorfismo φ la descomposición de tipo (28) *es unívoca.*

Sea φ un automorfismo del espacio $\mathfrak{A}$. Entonces la aplicación $\varphi_O=\varphi\cdot\overrightarrow{O^{\varphi}O}$ es un automorfismo de $\mathfrak{A}$ y de las relaciones

$$O\cdot\varphi_O=O^{\varphi}\cdot\overrightarrow{O^{\varphi}O}=O$$

se ve que φ_O deja en su sitio el punto O. Poniendc $\varphi_L=\overrightarrow{OO^{\varphi}}$, obtenemos la igualdad (28). Sea ahora

$$\varphi=\varphi_O\varphi_L=\varphi_O^{(1)}\varphi_L^{(1)},$$

donde $\varphi_L^{(1)}$ es una traslación y $\varphi_O^{(1)}$ un automorfismo de $\mathfrak{A}$ que deja en su sitio a O. Tenemos entonces

$$O\cdot\varphi_L^{(1)}\varphi_L^{-1}=O\cdot\varphi_O^{(1)}\varphi_O^{-1}=O.$$

Puesto que la traslación $\varphi_L^{(1)}\varphi_L^{-1}$ deja en su sitio el punto O, se tiene $\varphi_L^{(1)}\varphi_L^{-1}=\overrightarrow{OO}$, de donde $\varphi_L^{(1)}=\varphi_L$ y $\varphi_O^{(1)}=\varphi_O$.

29.2. Variedades lineales. Sea $\mathfrak{A}$ un espacio afín sobre un cuerpo conmutativo K. El espacio vectorial $L(\mathfrak{A})$ formado por las traslaciones de $\mathfrak{A}$ lo indicaremos por $\mathfrak{L}$. A todo par de puntos $X, Y\in\mathfrak{A}$ corresponderá entonces un vector $\overline{XY}=\overrightarrow{XY}$ de $\mathfrak{L}$ y esta correspondencia satisfará los axiomas A_1 y A_2 del p. 29.1. Se llama *dimensión* del espacio afín $\mathfrak{A}$ la dimensión del correspondiente espacio vectorial $\mathfrak{L}$. En particular, un espacio afín de dimensión 0 está compuesto por un punto solo.

Unos puntos $X_0, X_1, \ldots, X_m$ del espacio $\mathfrak{A}$ se llaman *linealmente independientes*, si son linealmente independientes los vectores $\overline{X_0X_1}, \overline{X_0X_2}, \ldots, \overline{X_0X_m}$. Se dice que un punto $X\in\mathfrak{A}$ *depende linealmente* de la sucesión de puntos $X_0, X_1, \ldots, X_m$, si el vector $\overline{X_0X}$ puede ser expresado linealmente en términos de los vectores $\overline{X_0X_1}, \ldots, \overline{X_0X_m}$.

En estas definiciones el punto X_0 desempeña un papel especial. Pero de hecho, la dependencia lineal no está ligada al orden de los puntos en la sucesión $X_0, X_1, \ldots, X_m$. Sea O un punto arbitrario del espacio $\mathfrak{A}$. Entonces se tiene

$$\overline{X_0X_i}=\overline{OX_i}-\overline{OX_0}\quad(i=0,\ 1,\ \ldots,\ m)$$

y el hecho de que el sistema de puntos $X_0, \ldots, X_m$ sea linealmente independiente equivale a que para cualesquiera $\lambda_1, \ldots, \lambda_m\in K$ de la relación

$$\lambda_1(\overline{OX_1}-\overline{OX_0})+\ldots+\lambda_m(\overline{OX_m}-\overline{OX_0})=o$$

se desprenda $\lambda_1=\ldots=\lambda_m=0$. Tomando $\lambda_0=-\lambda_1-\ldots-\lambda_m$, vemos que *los puntos $X_0, X_1, \ldots, X_m$ son linealmente independientes cuando, y sólo cuando, para cualesquiera $\lambda_0, \lambda_1, \ldots, \lambda_m\in K$ de las relaciones*

$$\lambda_0+\lambda_1+\ldots+\lambda_m=0,$$
$$\lambda_0\overline{OX_0}+\lambda_1\overline{OX_1}+\ldots+\lambda_m\overline{OX_m}=o$$

se deduce que $\lambda_0=\lambda_1=\ldots=\lambda_m=0$.

En esta afirmación no destaca ya ninguno de los puntos $X_0, X_1, \ldots, X_m$ y por esto si la sucesión de puntos $X_0, X_1, \ldots, X_m$ es linealmente dependiente, también será linealmente dependiente la sucesión $X_{i_0}, X_{i_1}, \ldots, X_{i_m}$ cualquiera que sea la permutación $(i_0, i_1, \ldots, i_m)$ de los números $0, 1, \ldots, m$. Esto también se refiere al concepto de dependencia lineal de un punto X respecto a una sucesión $X_0, X_1, \ldots, X_m$.

En las definiciones dadas se supone que la sucesión $X_0, X_1, \ldots, X_m$ tiene no menos de dos términos ($m \geqslant 1$). Por definición, se acepta que *un sistema compuesto de un punto es linealmente dependiente y que el hecho de que un punto X dependa linealmente del sistema X_0 significa que $X = X_0$.*

De los teoremas del p. 4.2 sobre la dependencia lineal de los vectores obtenemos directamente que

a) *una sucesión de puntos $X_0, X_1, \ldots, X_m$ ($m \geqslant 1$) es linealmente dependiente cuando, y sólo cuando, al menos uno de sus términos depende linealmente de los demás;*

b) *si un punto X depende linealmente de los puntos $X_0, \ldots, X_m$ y todo punto X_i depende linealmente de los puntos $Y_0, \ldots, Y_k$, el punto X depende linealmente de los puntos $Y_0, \ldots, Y_k$;*

c) *si un punto X depende linealmente de los puntos $X_1, \ldots, X_m$, el punto X depende linealmente también de los puntos $X_0, X_1, \ldots, X_m$, donde X_0 es un punto cualquiera del espacio.*

De la propiedad c) se deduce que si un sistema de puntos $X_0, X_1, \ldots, X_m$ es linealmente independiente, cualquier subsistema del mismo que contiene más de un punto es también linealmente independiente. Esto permite introducir la definición siguiente: un conjunto arbitrario (finito o infinito) de puntos $\mathfrak{M}$ de un espacio afín $\mathfrak{A}$ se llama *linealmente independiente*, si todo su subconjunto finito que contiene más de un punto es linealmente independiente.

Análogamente se dice que un punto $X \in \mathfrak{A}$ *depende linealmente* de un conjunto de puntos $\mathfrak{M}$, si X depende linealmente de algún subconjunto finito de puntos de $\mathfrak{M}$.

Se llama *adherencia lineal* de un conjunto cualquiera no vacío $\mathfrak{M}$ de puntos de un espacio afín $\mathfrak{A}$ el conjunto formado por todos los puntos $X \in \mathfrak{A}$ que dependen linealmente del conjunto $\mathfrak{M}$. Un conjunto $\mathfrak{M}$ se llama *linealmente cerrado* si $\mathfrak{M}$ coincide con su adherencia lineal.

Los conjuntos linealmente cerrados de puntos de un espacio afín se llaman *planos* o *variedades lineales* del mismo. En otras palabras, un conjunto $\mathfrak{M}$ de puntos de un espacio afín $\mathfrak{A}$ se llama plano de $\mathfrak{A}$, si cualesquiera que sean $X_1, \ldots, X_m \in \mathfrak{M}$ este conjunto contiene también cualquier punto del espacio $\mathfrak{A}$ que dependa linealmente de los puntos $X_1, \ldots, X_m$.

De aquí se deduce directamente que todo punto y el propio espacio $\mathfrak{A}$ son indudablemente planos.

Dos planos $\mathfrak{M}$ y $\mathfrak{N}$ se llaman *incidentes* si uno de ellos está contenido en el otro. Si $\mathfrak{M} \subseteq \mathfrak{N}$, se dice también que $\mathfrak{M}$ *se encuentra sobre* $\mathfrak{N}$ (o que $\mathfrak{N}$ *pasa por* $\mathfrak{M}$).

Entre los planos de un espacio afín $\mathfrak{A}$ y los subespacios lineales del espacio vectorial $L(\mathfrak{A})$ existe una relación muy simple que también se emplea a veces para definir el concepto de plano.

TEOREMA 1. *Si para un conjunto $\mathfrak{M}$ de puntos de un espacio afín $\mathfrak{A}$ y un punto $P \in \mathfrak{M}$ el conjunto $L(\mathfrak{M})$ de los vectores de tipo $\overline{PX}$ ($X \in \mathfrak{M}$) es un subespacio lineal del espacio vectorial $L(\mathfrak{A})$, el conjunto $\mathfrak{M}$ es un plano de $\mathfrak{A}$. Para cualquier plano no vacío $\mathfrak{M}$ de un espacio afín $\mathfrak{A}$ y para cualquier punto $P \in \mathfrak{M}$ el conjunto $L(\mathfrak{M})$ de todos los vectores de tipo $\overline{PX}$ ($X \in \mathfrak{M}$) es un subespacio lineal de $L(\mathfrak{A})$ que no depende de la selección del punto P en $\mathfrak{M}$.*

Efectivamente, sea $P \in \mathfrak{M}$ y sea $L(\mathfrak{M})$ un subespacio de $L(\mathfrak{A})$. Si $X_1, \ldots, X_m \in \mathfrak{M}$ y un punto $X \in \mathfrak{A}$ depende linealmente de $X_1, \ldots, X_m$, se tiene entonces para unos valores convenientes $\lambda_1, \ldots, \lambda_m \in K$

$$\overline{PX} = \lambda_1 \overline{PX_1} + \ldots + \lambda_m \overline{PX_m} \in L(\mathfrak{M}),$$

es decir, para un punto $Y \in \mathfrak{M}$ convenientemente escogido se tiene $\overline{PX} = \overline{PY}$, de donde $X = Y \in \mathfrak{M}$.

Recíprocamente, sea $\mathfrak{M}$ un plano de $\mathfrak{A}$ y sean $P, X, Y \in \mathfrak{M}$ y $\lambda, \mu \in K$. Poniendo

$$Q = P \cdot (\lambda \overline{PX} + \mu \overline{PY}),$$

vemos que

$$\overline{PQ} = \lambda \overline{PX} + \mu \overline{PY},$$

es decir, el punto Q depende linealmente de P, X e Y y, por consiguiente, $Q \in \mathfrak{M}$ y $\overline{PQ} \in L(\mathfrak{M})$; es decir, $L(\mathfrak{M})$ es un subespacio lineal de $L(\mathfrak{A})$. Para cualesquiera $P, P_1, X \in \mathfrak{M}$ tenemos $\overline{P_1X} = \overline{PX} - \overline{PP_1}$ y, por consiguiente, el conjunto de los vectores de tipo $\overline{P_1X}$ ($X \in \mathfrak{M}$) coincide con el conjunto de los vectores de tipo $\overline{PX}$ ($X \in \mathfrak{M}$).

Señalemos dos importantes corolarios que se desprenden del teorema 1.

COROLARIO 1. *Sea O un punto fijo de un espacio afín $\mathfrak{A}$, sea $\mathfrak{M}$ un plano que pasa por O y sea P un punto arbitrario de $\mathfrak{A}$. Entonces el conjunto $\mathfrak{N} = \mathfrak{M} \cdot \overline{OP}$ formado por todos los puntos desplazados del plano $\mathfrak{M}$ es un plano de $\mathfrak{A}$ que pasa por el punto P. Recíprocamente, si $\mathfrak{N}$ es un plano que pasa por P, el conjunto $\mathfrak{M} = \mathfrak{N} \cdot \overline{OP}$ es un plano que pasa por O.*

En efecto, de $O \in \mathfrak{M}$ se deduce que $P = O \cdot \overline{OP} \in \mathfrak{N}$ y para todo punto $X \in \mathfrak{N}$ existe un punto $Y \in \mathfrak{M}$ tal que $X = Y \cdot \overline{OP}$. De aquí se tiene $\overline{YX} = \overline{OP}$ y $\overline{PX} = \overline{OY}$ y por lo tanto $L(\mathfrak{M}) = L(\mathfrak{M} \cdot \overline{OP})$.

COROLARIO 2. *Para todo punto fijo O de un espacio afín $\mathfrak{A}$ la aplicación $X \to \overline{OX}$ $(X \in \mathfrak{A})$ determina una correspondencia biyectiva entre los subespacios lineales del espacio vectorial $L(\mathfrak{A})$ y los planos del espacio $\mathfrak{A}$ que pasan por O.*

La demostración es evidente.

Hemos señalado en el p. 29.1 que si llamamos punto a todo vector v de un espacio vectorial $\mathfrak{L}$ y si tomamos $\overline{vw}=w-v$, se obtiene el espacio afín $A(\mathfrak{L})$ sobre $\mathfrak{L}$. ¿Qué conjuntos de vectores de $\mathfrak{L}$ serán los planos de $A(\mathfrak{L})$?

Tomando por el punto fijo O de $A(\mathfrak{L})$ el vector nulo, vemos que la aplicación $X \to \overline{OX}$, de la cual se trata en el corolario 2, se convierte en la aplicación idéntica $w \to w$. Luego, son planos del espacio $A(\mathfrak{L})$ que pasan por O todos los subespacios lineales del espacio $\mathfrak{L}$, y sólo ellos. Los planos que pasan por un punto arbitrario $p \in \mathfrak{L}$ son, según el corolario 1, los conjuntos de vectores de tipo $p+\mathfrak{M}$, donde $\mathfrak{M}$ es un subespacio lineal del espacio $\mathfrak{L}$.

Se llama *base* de un plano arbitrario $\mathfrak{M}$ de un espacio afín $\mathfrak{A}$ todo conjunto linealmente independiente de puntos, cuya adherencia lineal coincide con $\mathfrak{M}$. Está claro que un conjunto de puntos $A_i (i \in I)$ de un plano $\mathfrak{M}$ es una base del plano $\mathfrak{M}$ cuando, y sólo cuando, el conjunto de vectores $\overline{A_\alpha A_i}$ $(i \in I)$, donde A_α $(\alpha \in I)$ es un punto arbitrario fijo, es una base del espacio vectorial $L(\mathfrak{M})$. En virtud del teorema sobre las bases de los espacios vectoriales del p. 4.2 tenemos:

a) *todo plano de un espacio afín posee una base;*

b) *si un plano $\mathfrak{M}$ está contenido en un plano $\mathfrak{N}$, toda base del plano $\mathfrak{M}$ es una parte de una base convenientemente escogida de $\mathfrak{N}$;*

c) *todas las bases de cualquier plano fijo $\mathfrak{M}$ son de una misma potencia igual a* dim.$L(\mathfrak{M})+1$.

Se llama dimensión de un plano $\mathfrak{M}$ de un espacio afín $\mathfrak{A}$ la dimensión del espacio vectorial $L(\mathfrak{M})$ que le corresponde. De la propiedad c) se deduce que la dimensión de un plano es igual a la potencia de cualquier base suya disminuida en 1. El conjunto vacío de vectores de un espacio afín $\mathfrak{A}$ se llama a veces *plano vacío* de $\mathfrak{A}$ de base vacía. La dimensión del plano vacío se toma, por definición, igual a -1.

Los planos 0-dimensionales son los puntos del espacio $\mathfrak{A}$. Los planos unidimensionales se llaman *rectas* (o líneas rectas) del espacio $\mathfrak{A}$. Cualesquiera dos puntos distintos de una recta forman una base de la misma y por cualesquiera dos puntos distintos de un espacio $\mathfrak{A}$ pasa una recta, y sólo una. En general, cualesquiera $k+1$ puntos linealmente independientes de un plano de dimensión k forman una base de este plano y por esto por cualesquiera $k+1$ puntos linealmente independientes (k es un número finito) de un espacio afín $\mathfrak{A}$ pasa un plano k-dimensional de este espacio y sólo uno.

Siendo $\mathfrak{M}$ y $\mathfrak{N}$ dos planos cualesquiera de un espacio afín $\mathfrak{A}$, convendremos en indicar por $\mathfrak{M} \vee \mathfrak{N}$ la adherencia lineal del conjunto $\mathfrak{M} \cup \mathfrak{N}$. El plano $\mathfrak{M} \vee \mathfrak{N}$ se puede definir también como la intersección de todos los planos del espacio $\mathfrak{A}$ que contienen a los planos $\mathfrak{M}$ y $\mathfrak{N}$.

La operación $\vee$, considerada como una operación binaria definida sobre el conjunto de todos los planos de un espacio afín, es conmutativa, asociativa e idempotente (es decir, $\mathfrak{M} \vee \mathfrak{M} = \mathfrak{M}$); además, para cualesquiera planos $\mathfrak{M}$ y $\mathfrak{N}$ de un espacio $\mathfrak{A}$ se tiene

$$\mathfrak{M} \subseteq \mathfrak{N} \Leftrightarrow \mathfrak{M} \vee \mathfrak{N} = \mathfrak{N}.$$

Por ejemplo, siendo $X_1, \ldots, X_m$ un conjunto finito de puntos de un espacio $\mathfrak{A}$, se tiene que

$$\mathfrak{M} = X_1 \vee X_2 \vee \ldots \vee X_m$$

es la adherencia lineal de dicho conjunto, es decir, es el menor plano que pasa por los puntos $X_1, \ldots, X_m$. La dimensión de $\mathfrak{M}$ es igual al número máximo de puntos linealmente independientes en el conjunto $X_1, \ldots, X_m$ disminuido en 1. En particular,

$$\dim. (X_1 \vee X_2 \vee \ldots \vee X_m) \leqslant m-1.$$

Es fácil ver que tiene lugar el siguiente teorema general:

TEOREMA 2. *Cualesquiera que sean los planos $\mathfrak{M}$ y $\mathfrak{N}$ de un espacio afín $\mathfrak{A}$ se tiene*

$$\dim. (\mathfrak{M} \vee \mathfrak{N}) \leqslant \dim. \mathfrak{M} + \dim. \mathfrak{N} + 1. \tag{1}$$

Si los planos $\mathfrak{M}$ y $\mathfrak{N}$ poseen un punto común, se tiene

$$\dim. (\mathfrak{M} \vee \mathfrak{N}) + \dim. (\mathfrak{M} \cap \mathfrak{N}) = \dim. \mathfrak{M} + \dim. \mathfrak{N}. \tag{2}$$

Demostremos primero la igualdad (2). Sea $O \in \mathfrak{M} \cap \mathfrak{N}$. Entonces, la aplicación $X \longrightarrow \overline{OX}$ del espacio $\mathfrak{A}$ sobre el espacio vectorial $L(\mathfrak{A})$ transforma los planos de $\mathfrak{A}$ que pasan por el punto O en los subespacios lineales del espacio vectorial $L(\mathfrak{A})$, con la particularidad de que conserva las dimensiones y la relación de inclusión. Por esto la fórmula (2) es un corolario directo de la fórmula análoga que tiene lugar para los subespacios lineales de los espacios vectoriales (véase el teorema 1 del p. 6.1).

La estimación (1) se deduce de la fórmula (2). Efectivamente, si los planos $\mathfrak{M}$ y $\mathfrak{N}$ poseen un punto común, tenemos según (2)

$$\dim. (\mathfrak{M} \vee \mathfrak{N}) \leqslant \dim. \mathfrak{M} + \dim. \mathfrak{N}. \tag{3}$$

Sea $\mathfrak{M} \cap \mathfrak{N} = \varnothing$. Agregando el punto $O \in \mathfrak{M}$ a una base del plano $\mathfrak{N}$, obtenemos una base de $O \vee \mathfrak{N}$, es decir,

$$\dim. (O \vee \mathfrak{N}) = \dim. \mathfrak{N} + 1. \tag{4}$$

Puesto que los planos $\mathfrak{M}$ y $O \vee \mathfrak{N}$ poseen el punto común O, tenemos basándonos en (3)

$$\dim. (\mathfrak{M} \vee \mathfrak{N}) = \dim. (\mathfrak{M} \vee O \vee \mathfrak{N}) \leqslant \dim. \mathfrak{M} + \dim. (O \vee \mathfrak{N}),$$

y de aquí, debido a (4), obtenemos (1).

Por lo tanto, si los planos $\mathfrak{M}$ y $\mathfrak{N}$ no tienen puntos comunes, la dimensión de $\mathfrak{M} \vee \mathfrak{N}$ está comprendida en los límites siguientes:

$$\text{máx} (\dim. \mathfrak{M},\ \dim. \mathfrak{N}) < \dim. (\mathfrak{M} \vee \mathfrak{N}) \leqslant \dim. \mathfrak{M} + \dim. \mathfrak{N} + 1.$$

Es fácil comprobar, considerando ejemplos, que la dimensión de $\mathfrak{M} \vee \mathfrak{N}$ puede tomar efectivamente cualquier valor comprendido en estos límites.

A todo plano $\mathfrak{M}$ de un espacio afín arbitrario $\mathfrak{A}$ le corresponde el subespacio lineal $L(\mathfrak{M})$ del espacio vectorial $L(\mathfrak{A})$. Según el § 6 (véase el problema 3 de la pág. 113), en $L(\mathfrak{A})$ existe *un subespacio complementario* $\mathfrak{L}_1$ que satisface las condiciones

$$L(\mathfrak{M}) + \mathfrak{L}_1 = L(\mathfrak{A}) \quad \text{y} \quad L(\mathfrak{M}) \cap \mathfrak{L}_1 = o,$$

donde o es el vector nulo del espacio $L(\mathfrak{A})$. La dimensión de $\mathfrak{L}_1$, que es la misma para todos los subespacios complementarios, se llama *codimensión* de $L(\mathfrak{M})$ en $L(\mathfrak{A})$. Por definición, se llama *codimensión de un plano* $\mathfrak{M}$ *en un espacio* $\mathfrak{A}$ la codimensión del subespacio vectorial $L(\mathfrak{M})$ en el espacio vectorial $L(\mathfrak{A})$.

Esta definición puede darse de otra forma. Para todo plano $\mathfrak{M}$ diremos que un plano $\mathfrak{N}$ es *complementario de* $\mathfrak{M}$ *en el espacio* $\mathfrak{A}$, si $\mathfrak{M} \vee \mathfrak{N} = \mathfrak{A}$ y si la intersección $\mathfrak{M} \cap \mathfrak{N}$ consta sólo de un punto. Está claro que, en estas condiciones, el subespacio $L(\mathfrak{N})$ es complementario de $L(\mathfrak{M})$ en $L(\mathfrak{A})$ y por ello *la codimensión de un plano* $\mathfrak{M}$ *en el espacio* $\mathfrak{A}$ *es igual a la dimensión de cualquier plano complementario.*

Puesto que los planos recíprocamente complementarios poseen un punto común, obtenemos, aplicándoles la fórmula (2), la igualdad

$$\dim. \mathfrak{M} + \text{codim}. \mathfrak{M} = \dim. \mathfrak{A}. \tag{5}$$

Si el espacio $\mathfrak{A}$ es de dimensión finita, obtenemos de (5)

$$\text{codim}. \mathfrak{M} = \dim. \mathfrak{A} - \dim. \mathfrak{M}. \tag{6}$$

Para un espacio $\mathfrak{A}$ de dimensión finita obtenemos de las fórmulas (2) y (6) el resultado siguiente:

TEOREMA 3. *Para cualesquiera dos planos* $\mathfrak{M}$ *y* $\mathfrak{N}$ *que se intersecan de un espacio afín* $\mathfrak{A}$ *se tiene*

$$\text{codim}. (\mathfrak{M} \cap \mathfrak{N}) + \text{codim}. (\mathfrak{M} \vee \mathfrak{N}) = \text{codim}. \mathfrak{M} + \text{codim}. \mathfrak{N}. \tag{7}$$

No obstante, es fácil ver que la fórmula (7) es válida también para los espacios de dimensión infinita.

COROLARIO. *Si la intersección de los planos $\mathfrak{N}_1, \mathfrak{N}_2, \ldots, \mathfrak{N}_s$ de un espacio afín $\mathfrak{A}$ arbitrario es no vacía, se tiene*

$$\text{codim.}\,(\mathfrak{N}_1 \cap \ldots \cap \mathfrak{N}_s) \leqslant \text{codim.}\,\mathfrak{N}_1 + \ldots + \text{codim.}\,\mathfrak{N}_s. \tag{8}$$

Para $s=2$, la fórmula (8) se desprende directamente de (7) y en el caso general ella se obtiene aplicando sucesivamente la fórmula del caso particular indicado.

Si el plano $\mathfrak{M}$ se encuentra sobre el plano $\mathfrak{N}$ de un espacio $\mathfrak{A}$, entonces, considerando $\mathfrak{N}$ como el espacio afín sobre $L(\mathfrak{N})$ e indicando por $\text{codim.}_{\mathfrak{N}}\mathfrak{M}$ la codimensión del plano $\mathfrak{M}$ en el espacio $\mathfrak{N}$, obtenemos la fórmula

$$\text{codim.}\,_{\mathfrak{N}}\mathfrak{M} + \text{codim.}\,\mathfrak{N} = \text{codim.}\,\mathfrak{M} \tag{9}$$

que se demuestra fácilmente pasando a los subespacios del espacio vectorial $L(\mathfrak{A})$. De la fórmula (9) se deduce, en particular, que si la codimensión del plano $\mathfrak{M}$ en $\mathfrak{A}$ es finita y $\mathfrak{M} \subset \mathfrak{N}$, se tiene

$$\text{codim.}\,\mathfrak{N} \leqslant \text{codim.}\,\mathfrak{M} - 1. \tag{10}$$

Los planos de codimensión igual a 1 se llaman *hiperplanos*. En otras palabras, en un espacio afín de dimensión finita n se llaman hiperplanos los planos de dimensión $n-1$. Luego, por cada n puntos linealmente independientes de un espacio de este tipo pasa un hiperplano, y sólo uno.

Consideremos ahora un conjunto cualquiera finito de hiperplanos $\mathfrak{P}_1, \ldots, \mathfrak{P}_s$ de un espacio afín arbitrario $\mathfrak{A}$. Debido a la fórmula (8) tenemos

$$0 < \text{codim.}\,(\mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_s) \leqslant s. \tag{11}$$

Se dice que los hiperplanos $\mathfrak{P}_1, \ldots, \mathfrak{P}_s$ (para s no mayor que la dimensión de $\mathfrak{A}$) se encuentran *en posición general*, si tiene lugar la igualdad exacta

$$\text{codim.}\,(\mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_s) = s. \tag{12}$$

Probemos que para todo $t \leqslant s$ de (12) resulta

$$\text{codim.}\,(\mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_t) = t.$$

En efecto, si fuese

$$\text{codim.}\,(\mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_t) < t,$$

tendríamos, debido a (8) y a (11),

$$\text{codim.}\,((\mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_t) \cap (\mathfrak{P}_{t+1} \cap \ldots \cap \mathfrak{P}_s)) < t + (s-t) = s$$

lo que estaría en contradicción con (12).

TEOREMA 4. *Todo plano $\mathfrak{M}$ de una codimensión finita s es una intersección de s hiperplanos convenientemente escogidos.*

Tomemos una base de $\mathfrak{M}$ y complementémosla con unos puntos $A_1, \ldots, A_s$ hasta obtener una base de todo el espacio $\mathfrak{A}$. Sea $\mathfrak{P}_i$ la adherencia lineal del conjunto $\mathfrak{M} \cup \{A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_s\}$. Como $\mathfrak{P}_i \neq \mathfrak{A}$ y $\mathfrak{P}_i \vee A_i = \mathfrak{A}$, resulta que $\mathfrak{P}_i$ es un hiperplano $(i=1, \ldots, s)$ y que

$$\mathfrak{M} \subseteq \mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_s.$$

De $A_i \notin \mathfrak{P}_i$ tenemos

$$\mathfrak{A} \supset \mathfrak{P}_1 \supset \mathfrak{P}_1 \cap \mathfrak{P}_2 \supset \ldots \supset \mathfrak{P}_1 \cap \ldots \mathfrak{P}_s,$$

de donde, debido a (10),

$$\text{codim.}\,(\mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_s) \geqslant s.$$

Comparando esta desigualdad con (11), obtenemos

$$\text{codim.}\,(\mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_s) = s.$$

Puesto que las codimensiones de los planos $\mathfrak{M}$ y $\mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_s$ coinciden y el primero está contenido en el segundo, ambos planos, debido a (11), coinciden que es lo que se quería demostrar.

Para terminar, haremos unas observaciones relacionadas con la definición del concepto del plano aceptada más arriba. Esta definición puede ser enunciada de nuevo en la forma siguiente: un conjunto $\mathfrak{M}$ de puntos de un espacio afín $\mathfrak{A}$ se llama plano, si para todo número natural s y para toda sucesión $A_1, \ldots, A_s$ de puntos de $\mathfrak{M}$ cualquier punto del espacio $\mathfrak{A}$, que dependa linealmente de los puntos de la sucesión mencionada, pertenece a $\mathfrak{M}$. De hecho, es suficiente tomar en esta definición $s=3$ y para varios tipos de cuerpos conmutativos principales K es suficiente tomar incluso $s=2$. Tiene lugar concretamente el teorema siguiente:

TEOREMA 5. *Si un conjunto $\mathfrak{M}$ de puntos de un espacio afín $\mathfrak{A}$ contiene con tres cualesquiera puntos suyos todos los puntos de $\mathfrak{A}$ que dependen linealmente de éstos, el conjunto $\mathfrak{M}$ es un plano. Si $\mathfrak{A}$ es un espacio afín sobre un cuerpo conmutativo K, de característica diferente de 2, y si un conjunto $\mathfrak{M}$ con dos cualesquiera puntos suyos contiene todos los puntos de $\mathfrak{A}$ que dependen linealmente de éstos, el conjunto $\mathfrak{M}$ es un plano de $\mathfrak{A}$.*

Demostremos la segunda afirmación. Supongamos que un conjunto $\mathfrak{M}$ cumple las condiciones de esta afirmación y que $X_0, X_1, X_2 \ldots$ es una sucesión de puntos de $\mathfrak{M}$. Debemos demostrar que para cualesquiera $\lambda_1, \ldots, \lambda_s \in K$ la relación

$$\overline{X_0X} = \lambda_1 \overline{X_0X_1} + \ldots + \lambda_s \overline{X_0X_s}$$

implica $X \in \mathfrak{M}$. Para $s=1$ esto es válido por el enunciado mismo del teorema. Aplicando ahora la inducción, aceptamos que la implicación indicada es válida para un $s \geqslant 1$. Sea

$$\overline{X_0Y} = \lambda_1 \overline{X_0X_1} + \ldots + \lambda_{s+1} \overline{X_0X_{s+1}}.$$

Consideremos unos puntos X y Z que satisfacen las condiciones

$$\overline{X_0X}=\lambda_1\overline{X_0X_1}+\ldots+\lambda_s\overline{X_0X_s} \text{ y } \overline{X_0Z}=\lambda_{s+1}\overline{X_0X_{s+1}}.$$

Según la hipótesis de inducción, de estas relaciones resulta que $X\in\mathfrak{M}$ y $Z\in\mathfrak{M}$. Tomando ahora (fig. 5)

$$\overline{X_0B}=\frac{1}{2}(\overline{X_0X}+\overline{X_0Z}),$$

vemos que el punto B depende linealmente de X y de Z y que el punto Y depende linealmente de los puntos X_0 y B. Basándonos en las condiciones del teorema, obtenemos de aquí sucesivamente que $B\in\mathfrak{M}$ e $Y\in\mathfrak{M}$ que es lo que se quería demostrar. La primera afirmación del teorema 5 se demuestra análogamente, pero sin emplear el punto B, y por esto ella es válida independientemente de la característica del cuerpo conmutativo K. Es fácil ver que para cuerpos conmutativos de característica 2 la segunda afirmación no tiene lugar.

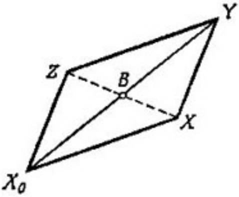

Fig. 5.

Hemos introducido en el p. 29.1 dos conceptos principales: el de un espacio afín sobre un espacio vectorial y el de un espacio afín sobre un cuerpo conmutativo. En este punto hemos considerado hasta el momento los espacios afines sobre los espacios vectoriales. Sea ahora $\mathfrak{A}$ un espacio afín sobre un cuerpo conmutativo K y sean $\boldsymbol{S}(X, Y, Z)$ y $\boldsymbol{P}_\lambda(X, Y)$ las operaciones principales de $\mathfrak{A}$. Hemos visto ya en el p. 29.1 que las traslaciones de $\mathfrak{A}$ son unos automorfismos del espacio $\mathfrak{A}$ sobre K de tipo especial. Demostremos ahora que *un punto X de un espacio $\mathfrak{A}$ depende linealmente de unos puntos $X_0, X_1, \ldots X_m$ de $\mathfrak{A}$ cuando, y sólo cuando, X se expresa en forma de un polinomio en $X_0, \ldots X_m$ mediante las operaciones $\boldsymbol{S}$ y $\boldsymbol{P}_\lambda$.*

En efecto, si

$$\overline{X_0X}=\lambda_1\overline{X_0X_1}+\ldots+\lambda_m\overline{X_0X_m},$$

donde $\lambda_1, \ldots, \lambda_m\in K$, se tiene según el p. 29.1

$$X=((X_0\cdot\lambda_1\overline{X_0X_1})\ldots)\cdot\lambda_m\overline{X_0X_m}. \tag{13}$$

Pero para cualesquiera $B, C, Z\in\mathfrak{A}$ y $\lambda\in K$ se tiene

$$Z\cdot\lambda\overline{BC}=\boldsymbol{S}(B, \boldsymbol{P}_\lambda(B, C), Z).$$

Transformando mediante esta fórmula el segundo miembro de la igualdad (13), obtenemos la expresión requerida de X en forma de un polinomio en $X_0, X_1, \ldots, X_m$. Recíprocamente, de

$$X=\boldsymbol{S}(Y_1, Y_2, Y_3) \text{ y } Z=\boldsymbol{P}_\lambda(Y_1, Y_2)$$

se deduce que

$$\overline{Y_3X} = \overline{Y_1Y_2} \quad \text{e} \quad \overline{Y_1Z} = \lambda\overline{Y_1Y_2},$$

es decir, que X y Z dependen linealmente de Y_1, Y_2, Y_3 y por esto, si X se expresa en forma de un polinomio en $X_0, X_1, \ldots, X_m$, el punto X depende linealmente de $X_0, X_1, \ldots, X_m$.

Según la definición, un conjunto $\mathfrak{M}$ de puntos de un espacio $\mathfrak{A}$ se llama plano, si este conjunto $\mathfrak{M}$ con cualesquiera puntos suyos $X_1, \ldots, X_m$ contiene también todos los puntos de $\mathfrak{A}$ que dependen linealmente de éstos, es decir, contiene los valores de todos los polinomios en $X_1, \ldots, X_m$. En otras palabras, los planos de un espacio $\mathfrak{A}$ son simplemente subálgebras del álgebra $\mathfrak{A}$ y las adherencias lineales son las subálgebras generadas por los elementos de los conjuntos correspondientes.

29.3. Planos paralelos. Hemos visto ya que para todo plano $\mathfrak{M}$ de un espacio afín $\mathfrak{A}$ el conjunto $L(\mathfrak{M})$ de todos los vectores de tipo $\overline{XY}$ $(X, Y \in \mathfrak{M})$ es un subespacio del espacio vectorial $L(\mathfrak{A})$. El subespacio $L(\mathfrak{M})$ suele llamarse frecuentemente *subespacio tangente* al plano $\mathfrak{M}$.

DEFINICIÓN. *Los planos $\mathfrak{M}$ y $\mathfrak{N}$ de un espacio $\mathfrak{A}$ se llaman paralelos (notación simbólica $\mathfrak{M} \parallel \mathfrak{N}$), si sus espacios vectoriales tangentes $L(\mathfrak{M})$ y $L(\mathfrak{N})$ son incidentes, es decir, si $L(\mathfrak{M}) \subseteq L(\mathfrak{N})$ o $L(\mathfrak{N}) \subseteq L(\mathfrak{M})$.*

De esta definición se desprenden inmediatamente varios corolarios importantes.

COROLARIO 1. *Dos planos $\mathfrak{M}_1$ y $\mathfrak{M}_2$ que se intersecan son paralelos cuando, y sólo cuando, uno de ellos está contenido en el otro.*

Efectivamente, si $\mathfrak{M}_1$ y $\mathfrak{M}_2$ tienen un punto común O, entonces $L(\mathfrak{M}_i)$ $(i = 1, 2)$ es el conjunto de vectores de tipo $\overline{OX}$ $(X \in \mathfrak{M}_i)$ y por ello la incidencia de los espacios tangentes $L(\mathfrak{M}_1)$ y $L(\mathfrak{M}_2)$ equivale a la incidencia de los planos $\mathfrak{M}_1$ y $\mathfrak{M}_2$.

COROLARIO 2. *Si los planos $\mathfrak{M}_1$, $\mathfrak{M}_2$ y $\mathfrak{M}_3$ son de una misma dimensión finita (o de una misma codimensión finita) y $\mathfrak{M}_1 \parallel \mathfrak{M}_3$ y $\mathfrak{M}_2 \parallel \mathfrak{M}_3$, se tiene $\mathfrak{M}_1 \parallel \mathfrak{M}_2$.*

Las dimensiones (las codimensiones) de los espacios tangentes $L(\mathfrak{M}_i)$ son finitas e iguales. Por ello la incidencia de $L(\mathfrak{M}_i)$ y $L(\mathfrak{M}_j)$ equivale a la coincidencia de los mismos y la afirmación del corolario se reduce a la observación trivial de que $L(\mathfrak{M}_1) = L(\mathfrak{M}_3)$ y $L(\mathfrak{M}_2) = L(\mathfrak{M}_3)$ implican $L(\mathfrak{M}_1) = L(\mathfrak{M}_2)$.

El corolario 2 significa que para cualquier k finito todos los planos k-dimensionales de un espacio afín arbitrario $\mathfrak{A}$ se descomponen en haces de planos paralelos entre sí. Estos haces se llaman a veces direcciones *k-dimensionales en $\mathfrak{A}$.*

COROLARIO 3. *En un espacio afín $\mathfrak{A}$ se tiene*

$$\mathfrak{M} \parallel \mathfrak{M} \cdot u$$

para cualquier plano $\mathfrak{M}$ y para cualquier traslación $u \in L(\mathfrak{A})$.

Efectivamente, el espacio tangente $L(\mathfrak{M}\cdot u)$ está formado por vectores de tipo $\overline{XuYu}$ $(X,\ Y\in\mathfrak{M})$. Pero

$$\overline{XuYu}=\overline{XuX}+\overline{XY}+\overline{YYu}=-u+\overline{XY}+u=\overline{XY}$$

y por ello $L(\mathfrak{M}\cdot u)=L(\mathfrak{M})$.

COROLARIO 4. *Por todo punto B de un espacio afín pasa un hiperplano $\mathfrak{P}$, y sólo uno, paralelo a cualquier hiperplano $\mathfrak{M}$ dado de antemano.*

Si $B\in\mathfrak{M}$, se ve del corolario 1 que el único hiperplano que pasa por B y es paralelo a $\mathfrak{M}$, es el propio hiperplano $\mathfrak{M}$. Podemos aceptar, pues, que $B\notin\mathfrak{M}$. En virtud del corolario 3. el hiperplano $\mathfrak{M}\cdot\overline{OB}$ $(O\in\mathfrak{M})$ pasa por B y es paralelo a $\mathfrak{M}$. Si existe otro hiperplano $\mathfrak{N}$ que pasa por B y tal que $\mathfrak{N}\parallel\mathfrak{M}$, tenemos, según el corolario 2, $\mathfrak{N}\parallel\mathfrak{M}\cdot\overline{OB}$ y entonces, en virtud del corolario 1, $\mathfrak{N}$ y $\mathfrak{M}\cdot\overline{OB}$ son incidentes. Pero los hiperplanos incidentes coinciden que es lo que se quería demostrar.

TEOREMA 1 *Si un hiperplano $\mathfrak{P}$ y un plano $\mathfrak{N}$ de un espacio afín $\mathfrak{A}$ no se intersecan, resulta que $\mathfrak{P}$ es paralelo a $\mathfrak{N}$.*

Queremos demostrar que $L(\mathfrak{N})\subseteq L(\mathfrak{P})$. Supongamos, al contrario, que para unos puntos $A,\ B\in\mathfrak{N}$ el vector $\overline{AB}$ no pertenece a $L(\mathfrak{P})$. Tomemos un punto cualquiera $O\in\mathfrak{P}$ (fig. 6) y consideremos el vector $\overline{OA}$. Puesto que $\mathfrak{P}$ es un hiperplano, se tiene

$$L(\mathfrak{A})=L(\mathfrak{P})+L(A\vee B)$$

y por ello para $X\in\mathfrak{P}$ y $U\in A\vee B$ convenientemente escogidos tenemos $\overline{OA}=\overline{OX}-\overline{AU}$, de donde $\overline{OU}=\overline{OX}$ y, por consiguiente, $U=X$, lo que contradice a la condición de que $\mathfrak{P}\cap\mathfrak{N}=\varnothing$.

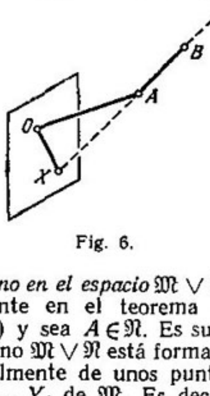

Fig. 6.

TEOREMA 2. *Para que dos planos $\mathfrak{M}$ y $\mathfrak{N}$ que no se intersecan sean paralelos, es necesario y suficiente que uno de ellos sea un hiperplano en el espacio $\mathfrak{M}\vee\mathfrak{N}$.*

La suficiencia está contenida directamente en el teorema 1. Demostremos la necesidad. Sea $L(\mathfrak{N})\subseteq L(\mathfrak{M})$ y sea $A\in\mathfrak{N}$. Es suficiente demostrar que $\mathfrak{M}\vee\mathfrak{N}\subseteq\mathfrak{M}\vee A$. El plano $\mathfrak{M}\vee\mathfrak{N}$ está formado por aquellos puntos X que dependen linealmente de unos puntos $A,\ X_1,\ \ldots,\ X_s$ de $\mathfrak{N}$ y unos puntos $Y_1,\ \ldots,\ Y_t$ de $\mathfrak{M}$. Es decir,

$$\overline{AX}=\lambda_1\overline{AY_1}+\ldots+\lambda_t\overline{AY_t}+\mu_1\overline{AX_1}+\ldots+\mu_s\overline{AX_s}=$$
$$=\overline{AY}+(\mu_1+\ldots+\mu_s)\overline{AO}+\overline{OZ}\quad (Y\in\mathfrak{N}\text{ y } O,\ Z\in\mathfrak{M}).$$

Tenemos, por hipótesis, $\overline{AY}\in L(\mathfrak{N})\subseteq L(\mathfrak{M})$ y, por consiguiente, $\overline{AY}=\overline{OU}$ $(U\in\mathfrak{M})$. De aquí se desprende que $\overline{AX}\in L(\mathfrak{M}\vee A)$ y por lo tanto $X\in\mathfrak{M}\vee A$.

COROLARIO 1. *Si los planos $\mathfrak{M}$ y $\mathfrak{N}$ son de dimensión finita y no se intersecan, estos planos son paralelos cuando, y sólo cuando,*

$$\dim. (\mathfrak{M} \vee \mathfrak{N}) = \text{máx} (\dim. \mathfrak{M}, \ \dim. \mathfrak{N}) + 1. \tag{1}$$

La demostración es obvia. En particular, de este corolario sacamos la conclusión de que dos rectas que no se intersecan son paralelas cuando, y sólo cuando, se encuentran sobre un mismo plano bidimensional. Análogamente, una recta y un plano bidimensional que no se intersecan son paralelos cuando, y sólo cuando, se encuentran en un mismo plano tridimensional, etc.

COROLARIO 2. *Cualquiera que sea el número finito $k > 0$, por todo punto B de un espacio afín $\mathfrak{A}$ pasa un plano $\mathfrak{P}$ k-dimensional, y sólo uno, paralelo a un plano k-dimensional $\mathfrak{M}$ dado de antemano.*

Podemos limitarnos a considerar el caso en que $B \notin \mathfrak{M}$. Debido a la fórmula (1), el plano $\mathfrak{P}$ debe ser un hiperplano del espacio $\mathfrak{M} \vee B$ y por ello el asunto se reduce al corolario 4 de la definición de los planos paralelos.

29.4. Funcionales lineales. Una función $f : \mathfrak{A} \to K$, definida sobre el conjunto de todos los puntos de un espacio afín $\mathfrak{A}$ y con valores en el cuerpo conmutativo principal K, se llama *funcional* sobre $\mathfrak{A}$.

El concepto de una funcional sobre $\mathfrak{A}$ es un caso particular del concepto de una función que está definida sobre un conjunto arbitrario y cuyos valores pertenecen a un cuerpo conmutativo dado K. Para estas funciones hemos definido en el p. 4.1 los conceptos de una suma y de un producto de un número por una función, resultando que el conjunto de todas las funciones forma, respecto a estas operaciones, un espacio lineal, cuya dimensión coincide con la potencia del conjunto que sirve como el dominio de definición.

Una funcional f, definida sobre un espacio afín $\mathfrak{A}$, se llama *lineal* (o *afín*) sobre $\mathfrak{A}$, si para cualesquiera $O, P, Q, R \in \mathfrak{A}$ y $\lambda, \mu \in K$ que cumplen la relación

$$\overline{OR} = \lambda \overline{OP} + \mu \overline{OQ} \tag{1}$$

tiene lugar la igualdad

$$f(R) = \lambda (f(P) - f(O)) + \mu (f(Q) - f(O)) + f(O). \tag{2}$$

Demostremos que para toda funcional lineal f es válida la implicación

$$\overline{X_0 X} = \sum_{i=1}^{s} \lambda_i \overline{X_0 X_i} \Rightarrow f(X) = \sum_{i=1}^{s} \lambda_i (f(X_i) - f(X_0)) + f(X_0), \tag{3}$$

cualesquiera que sean $X_0, X_1, \ldots, X_s, X \in \mathfrak{A}$ y $\lambda_1, \ldots, \lambda_s \in K$.

Para $s = 2$ la implicación (3) coincide la implicación (1) $\Rightarrow$ (2) que define la linealidad de f. Aceptemos ahora por inducción que la implicación (3) es válida para un valor fijo $s \geqslant 2$. Sea $\overline{X_0 Y} = \sum_{i=1}^{s+1} \lambda_i \overline{X_0 X_i}$. Considerando un punto X que cumple la condición $\overline{X_0X} = \sum_{i=1}^{s} \lambda_i \overline{X_0X_i}$, obtenemos (3). Después, de la relación $\overline{X_0Y} = \overline{X_0X} + \lambda_{s+1} \overline{X_0X_{s+1}}$ y de la implicación $(1) \Rightarrow (2)$ obtenemos

$$f(Y) = (f(X) - f(X_0)) + \lambda_{s+1}(f(X_{s+1}) - f(X_0)) + f(X)_0 =$$
$$= \sum_{i=1}^{s+1} \lambda_i (f(X_i) - f(X_0)) + f(X_0)$$

que es lo que se quería demostrar.

Está claro que todas las funcionales constantes son lineales. Es también evidente que la suma de unas funcionales lineales y el producto de un número por una funcional lineal son de nuevo funcionales lineales. Por esto el conjunto de todas las funcionales lineales sobre un espacio afín $\mathfrak{A}$ es un espacio lineal. Este espacio se indica por $\mathfrak{A}_f$ y se llama *conjugado* de $\mathfrak{A}$. El conjunto de todas las funcionales constantes es un subespacio lineal de dimensión uno del espacio $\mathfrak{A}_f$.

LEMA. *Si f es una funcional lineal sobre un espacio afín $\mathfrak{A}$ y para unos puntos A, $B \in \mathfrak{A}$ se tiene $f(A) \neq f(B)$, entonces en la recta $A \vee B$ existe un punto X tal que $f(X) = \alpha$ cualquiera que sea $\alpha \in K$.*

Tenemos para un punto arbitrario $X \in A \vee B$

$$\overline{AX} = \lambda \cdot \overline{AB} \quad (\lambda \in K), \tag{4}$$

de donde, en virtud de (2),

$$f(X) = \lambda (f(B) - f(A)) + f(A).$$

Por ello, tomando para λ el valor

$$\lambda = (\alpha - f(A))(f(B) - (f(A))^{-1},$$

obtenemos de (4) el punto requerido $X \in A \vee B$ tal que $f(X) = \alpha$.

TEOREMA 1 *Para toda funcional lineal no constante f sobre un espacio afín $\mathfrak{A}$, el conjunto de los puntos $X \in \mathfrak{A}$ en los que f toma un valor fijo $\alpha \in K$ es un hiperplano de $\mathfrak{A}$. Recíprocamente, para todo hiperplano $\mathfrak{M}$ y cualquier $\alpha \in K$ existe una funcional lineal f, cuyos valores son iguales a α sobre $\mathfrak{M}$ y diferentes de α fuera de $\mathfrak{M}$.*

Demostremos la primera afirmación. Sea $\mathfrak{M}$ el conjunto de los puntos $X \in \mathfrak{A}$ en los que $f(X) = \alpha$. Puesto que la funcional f no es constante sobre $\mathfrak{A}$, tendremos $f(A) \neq f(B)$ para unos puntos A, $B \in \mathfrak{A}$ convenientemente escogidos. Según el lema, de aquí se deduce que en la recta $A \vee B$ existe un punto X perteneciente a $\mathfrak{M}$, de modo que el conjunto $\mathfrak{M}$ no es vacío. Si $X_0, X_1, \ldots, X_s \in \mathfrak{M}$ y X es un punto que depende linealmente de éstos, se tiene

$$f(X_0) = f(X_1) = \ldots = f(X_s) = \alpha \tag{5}$$

y para unos valores adecuados $\lambda_1, \ldots, \lambda_s \in K$ se tiene

$$\overline{X_0X} = \lambda_1 \cdot \overline{X_0X_1} + \ldots + \lambda_s \cdot \overline{X_0X_s}. \tag{6}$$

De las fórmulas (6), (3) y (5) encontramos $f(X) = \alpha$, es decir, $X \in \mathfrak{M}$. En otras palabras, el conjunto $\mathfrak{M}$ es *un plano* de $\mathfrak{A}$. Resta demostrar que codim. $\mathfrak{M} = 1$. Como la funcional f no es constante, existe un punto A tal que $f(A) \neq \alpha$ y por ello $A \notin \mathfrak{M}$. Tomemos un punto cualquiera $X \in \mathfrak{A}$ y demostremos que $X \in A \vee \mathfrak{M}$. Si $f(X) = \alpha$, tenemos $X \in \mathfrak{M}$ y no hay nada que demostrar. Sea, por esto, $f(X) \neq \alpha$ y sea $M \in \mathfrak{M}$. En la recta $A \vee M$ (fig. 7) existe, según el lema, un punto B en el cual $f(B) \neq f(X)$ y, por ello, en la recta $B \vee X$ existe un punto $N \in \mathfrak{M}$. Vemos que X depende linealmente de B y de N y que B depende linealmente de A y M. Luego, X depende linealmente de A, M y N y $X \in \mathfrak{M} \vee A$.

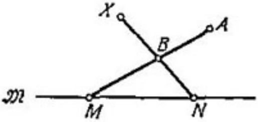

Fig. 7.

Demostremos ahora la segunda afirmación del teorema. Sea $\mathfrak{A} = \mathfrak{M} \vee A$ y sea $M \in \mathfrak{M}$. Para todo punto $X \in \mathfrak{A}$ existe entonces una descomposición de tipo

$$\overline{MX} = \lambda \cdot \overline{MA} + \overline{MU} \quad (\lambda \in K \text{ y } U \in \mathfrak{M})$$

y sólo una. Introduciendo la funcional $f_0(X) = \lambda$, vemos que ella es igual a 0 sobre $\mathfrak{M}$ y es diferente de 0 fuera de $\mathfrak{M}$ y que la funcional $f(X) = f_0(X) + \alpha$ es igual a α sobre $\mathfrak{M}$ y es diferente de α fuera de $\mathfrak{M}$. Por lo tanto, resta sólo demostrar que la funcional f_0 es lineal. Sea

$$\begin{gathered} \overline{X_0X_{s+1}} = \lambda_1 \cdot \overline{X_0X_1} + \ldots + \lambda_s \cdot \overline{X_0X_s}, \\ \overline{MX_i} = \alpha_i \cdot \overline{MA} + \overline{MU_i} \quad (U_i \in \mathfrak{M};\ i = 0,\ 1,\ \ldots,\ s) \end{gathered} \tag{7}$$

y, por consiguiente, $f_0(X_i) = \alpha_i$ $(i = 0, 1, \ldots, s)$. Como $\overline{X_0X_j} = \overline{MX_j} - \overline{MX_0}$, de las fórmulas (7) obtenemos

$$\overline{MX_{s+1}} = MX_0 + \sum \lambda_i (\overline{MX_i} - \overline{MX_0}) = (\sum (\lambda_i\alpha_i - \lambda_i\alpha_0) + \alpha_0) \cdot \overline{MA} + \overline{MU},$$

de donde

$$f_0(X_{s+1}) = \sum \lambda_i (f(X_i) - f(X_0)) + f(X_0)$$

que es lo que se quería demostrar.

El conjunto de aquellos puntos $X \in \mathfrak{A}$ en los que la funcional dada f o el sistema de funcionales f_i dado se anula se llama *variedad radical* de la funcional (o del sistema de funcionales). Por esto el teorema 1 equivale a la afirmación de que la variedad radical de una funcional lineal no constante es un hiperplano y de que todo hiperplano es la variedad radical de una funcional lineal adecuada. Si la funcional es constante, su variedad radical es, obviamente, o bien vacía o bien coincidente con $\mathfrak{A}$ (funcional nula). La variedad radical de un sistema de funcionales es la intersección de las variedades radicales correspondientes a las funcionales del sistema dado. Por lo tanto, la variedad radical de un sistema arbitrario de funcionales lineales es un plano (posiblemente vacío o coincidente con todo el espacio $\mathfrak{A}$).

Hemos demostrado en el p. 29.2 que todo plano $\mathfrak{M}$ de codimensión finita $s \geqslant 1$ puede ser representado en la forma $\mathfrak{M} = \mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_s$, donde $\mathfrak{P}_1, \ldots, \mathfrak{P}_s$ son unos hiperplanos adecuados. Indicando por f_i la funcional lineal, cuya variedad radical es $\mathfrak{P}_i$, vemos que $\mathfrak{M}$ es el conjunto de aquellos puntos $X \in \mathfrak{A}$, para los cuales

$$f_1(X) = 0, \ \ldots, \ f_s(X) = 0. \tag{8}$$

Considerando el problema recíproco, preguntémonos ¿para qué funcionales lineales $f_1, \ldots, f_s$ la correspondiente variedad radical (8) resultará ser un plano de codimensión s?

TEOREMA 2. *Para que la variedad radical de un sistema de funcionales lineales $f_1, \ldots, f_s$ sea un plano de codimensión s, es necesario y suficiente que las funcionales $f_1, \ldots, f_s$ sean linealmente independientes (en $\mathfrak{A}_f$) y que posean al menos un punto radical común.*

NECESIDAD. Indiquemos por $\mathfrak{P}_i$ la variedad radical de la funcional f_i. Si

$$f_s(X) = \alpha_1 f_1(X) + \ldots + \alpha_{s-1} f_{s-1}(X) \quad (X \in \mathfrak{A}),$$

se tiene $\mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_{s-1} \subseteq \mathfrak{P}_s$. Luego, resulta que codim. $(\mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_s) =$ $=$ codim. $(\mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_{s-1}) \neq s$ y, por consiguiente, las condiciones del teorema son necesarias.

SUFICIENCIA. Tomemos

$$\mathfrak{M} = \mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_{s-1} \cap \mathfrak{P}_s.$$

Tenemos, según el p. 29.2, codim. $\mathfrak{M} \leqslant s$. Supongamos que codim. $\mathfrak{M} < s$. Existen entonces en $\mathfrak{A}$ unos puntos $A_1, \ldots, A_{s-1}$ tales que

$$\mathfrak{A} = \mathfrak{M} \vee A_1 \vee \ldots \vee A_{s-1}. \tag{9}$$

Consideremos todas las funciones definidas sobre el conjunto $\{A_1, \ldots, A_{s-1}\}$ con valores en un cuerpo conmutativo K. Estas funciones forman un espacio lineal, cuya dimensión es igual a $s-1$, si los puntos $A_1, \ldots, A_{s-1}$ son diferentes, y es menor que $s-1$, si algunos de estos puntos coinciden. Es decir, el espacio de las funciones definidas sobre $\{A_1, \ldots, A_{s-1}\}$ es de dimensión no mayor que $s-1$ y por ello al menos una de s cualesquiera funciones definidas sobre el conjunto señalado puede ser expresada linealmente en términos de las demás. En particular, si consideramos las funcionales $f_1, \ldots, f_s$ sólo sobre el conjunto $\{A_1, \ldots, A_{s-1}\}$, una de ellas puede ser sin duda alguna expresada linealmente en términos de las restantes. Sea, por ejemplo,

$$f_s(X) = \alpha_1 f_1(X) + \ldots + \alpha_{s-1} f_{s-1}(X). \qquad (10)$$

Esta igualdad es válida para todos los puntos $X = A_1, \ldots, A_{s-1}$. Basándonos en que las funcionales consideradas son lineales y se anulan sobre el plano $\mathfrak{M}$ y empleando la relación (9), veremos fácilmente que la igualdad (10) es válida para todos los puntos $X \in \mathfrak{A}$, es decir, que las funcionales $f_1, \ldots, f_s$ son linealmente dependientes y con ello queda demostrada la suficiencia de las condiciones. En efecto, de (9) se deduce que para todo punto $X \in \mathfrak{A}$ existen unos puntos $O, Y \in \mathfrak{M}$ y unos números $\lambda_1, \ldots, \lambda_{s-1}$ tales que

$$\overline{OX} = \lambda_1 \overline{OA_1} + \ldots + \lambda_{s-1} \overline{OA_{s-1}} + \overline{OY};$$

luego, $f_i(O) = f_i(Y) = 0$ y en virtud e (3) tenemos

$$f_i(X) = \lambda_1 f_i(A_1) + \ldots + \lambda_{s-1} f_i(A_{s-1}) \quad (i = 1, \ldots, s). \qquad (11)$$

Teniendo en cuenta que la igualdad (10) es válida para $X = A_j$ $(j = 1, \ldots, s-1)$, obtenemos de (11)

$$f_s(X) = \sum_{j=1}^{s-1} \lambda_j f_s(A_j) = \sum_{j=1}^{s-1} \lambda_j \sum_{k=1}^{s-1} \alpha_k f_k(A_j) =$$

$$= \sum_{k=1}^{s-1} \alpha_k \sum_{j=1}^{s-1} \lambda_j f_k(A_j) = \alpha_1 f_1(X) + \ldots + \alpha_{s-1} f_{s-1}(X)$$

que es lo que se quería demostrar.

COROLARIO 1 *Supongamos que un plano $\mathfrak{M}$ de un espacio afín $\mathfrak{A}$ es de codimensión finita s. Entonces el conjunto $\mathfrak{M}^\perp$ de todas las funcionales lineales sobre $\mathfrak{A}$ iguales a 0 sobre $\mathfrak{M}$, es un espacio vectorial de dimensión s.*

Efectivamente, si $\mathfrak{M}^\perp$ contiene unas funcionales $f_1, \ldots, f_{s+1}$ linealmente independientes y si $\mathfrak{P}_i$ es el hiperplano radical de la funcional f_i, tenemos

$$\mathfrak{M} \subseteq \mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_{s+1}$$

y, por ello,

$$\text{codim.}\, \mathfrak{M} \geqslant s + 1.$$

En cambio, si todas las funcionales lineales de $\mathfrak{M}^\perp$ se expresan linealmente en términos de $f_1, \ldots, f_{s-1}$, se tiene

$$\mathfrak{M} = \mathfrak{P}_1 \cap \ldots \cap \mathfrak{P}_{s-1}$$

y, por consiguiente, codim. $\mathfrak{M} \leqslant s - 1$.

COROLARIO 2. *Para todo espacio afín $\mathfrak{A}$ la dimensión del espacio conjugado $\mathfrak{A}_f$ es mayor en 1 que la dimensión de $\mathfrak{A}$.*

Sea $\mathfrak{B}=\{A_i \mid i\in I\}$ una base de $\mathfrak{A}$ y sea F una función cualquiera que está definida sobre $\mathfrak{B}$ y toma valores en K. Fijemos un punto cualquiera $O\in\mathfrak{B}$. Para todo punto $X\in\mathfrak{A}$ existe un sistema $\lambda_{i_1}, \ldots, \lambda_{i_s}\in K$ definido unívocamente que cumple la condición

$$\overline{OX}=\lambda_{i_1}\overline{OA_{i_1}}+\ldots+\lambda_{i_s}\overline{OA_{i_s}}.$$

Definamos mediante la función F la funcional

$$f(X)=\sum\lambda_{i_k}(F(A_{i_k})-F(O))+F(O).$$

Se comprueba fácilmente que la funcional f es lineal sobre $\mathfrak{A}$ y que $F\longrightarrow f$ es un isomorfismo del espacio vectorial de las funciones, definidas sobre $\mathfrak{B}$, sobre el espacio vectorial $\mathfrak{A}_f$. Por lo tanto, la dimensión de $\mathfrak{A}_f$ es igual a la potencia de la base $\mathfrak{B}$, es decir, es mayor en 1 que la dimensión de $\mathfrak{A}$.

Complementos y ejemplos

1. Sea $\mathfrak{A}$ un espacio afín sobre un espacio vectorial $\mathfrak{L}=L(\mathfrak{A})$. Para todo plano $\mathfrak{M}\subseteq\mathfrak{A}$, indiquemos por $\mathfrak{M}^{\perp}$ el conjunto de las funcionales lineales sobre $\mathfrak{A}$ que son iguales a 0 sobre $\mathfrak{M}$. Recíprocamente, para todo conjunto $\mathfrak{U}$ de funcionales lineales, indiquemos por $\mathfrak{U}^{\top}$ la variedad radical de $\mathfrak{U}$. Entonces, para un espacio $\mathfrak{A}$ de dimensión arbitraria es válida la siguiente *ley de dualidad:* para todo plano no vacío $\mathfrak{M}\subseteq\mathfrak{A}$ se tiene $(\mathfrak{M}^{\perp})^{\top}=\mathfrak{M}$ y la dimensión de $\mathfrak{M}$ es igual a la codimensión de $\mathfrak{M}^{\perp}$ en el espacio lineal $\mathfrak{A}_f$ de todas las funcionales lineales sobre $\mathfrak{A}$. Recíprocamente, si un subespacio lineal $\mathfrak{U}\subset\mathfrak{A}_f$ no contiene funcionales constantes diferentes de cero, se tiene $\mathfrak{U}^{\top}\neq\varnothing$, $(\mathfrak{U}^{\top})^{\perp}=\mathfrak{U}$ y la dimensión de $\mathfrak{U}$ es igual a la codimensión de $\mathfrak{U}^{\top}$ en $\mathfrak{A}$.

2. HOMOMORFISMOS DE ESPACIOS AFINES. Los espacios afines sobre un cuerpo conmutativo K son unas álgebras de signatura fija y, por ello, se puede hablar de *homomorfismos, congruencias* y *espacios cocientes* de los espacios afines sobre K. Es fácil ver que siendo θ una congruencia sobre un espacio afín $\mathfrak{A}$, las clases de equivalencia según θ son los planos que se obtienen uno de otro mediante traslaciones y que son, por ende, paralelos. Recíprocamente, si $\mathfrak{M}$ es un plano, todos los planos paralelos a éste, que se obtienen de $\mathfrak{M}$ mediante traslaciones, forman el sistema de clases de equivalencia según una congruencia $\theta_{\mathfrak{M}}$ sobre $\mathfrak{A}$. La dimensión del espacio cociente $\mathfrak{A}/\theta_{\mathfrak{M}}$ es igual a la codimensión de $\mathfrak{M}$.

3. De los axiomas A_1^*, A_2^* y A_3^* se ve que la clase de todos los espacios afines sobre un cuerpo conmutativo fijo K se define mediante identidades, es decir, que esta clase es *una variedad de álgebras*. Por ello se puede hablar de espacios afines *libres* sobre un cuerpo conmutativo K provistos de un sistema dado de generadores libres. Puesto que una base de un espacio de dimensión finita es un sistema de generadores del mismo, el número de sus elementos disminuido en 1 coincide con la dimensión de $\mathfrak{A}$ y como todos los espacios de una dimensión dada son isomorfos, todo espacio afín $\mathfrak{A}$ sobre K es un espacio libre y los elementos de cualquier base de $\mathfrak{A}$ son los generadores libres de $\mathfrak{A}$.

4. Indiquemos por $\mathfrak{A}\times\mathfrak{B}$ el producto cartesiano de los espacios afines $\mathfrak{A}$ y $\mathfrak{B}$ sobre un cuerpo conmutativo fijo K (comprendido como el producto cartesiano de las *álgebras* $\mathfrak{A}$ y $\mathfrak{B}$). Demuéstrese que el espacio vectorial de las traslaciones $L(\mathfrak{A}\times\mathfrak{L})$ es isomorfo a la suma directa de los espacios $L(\mathfrak{A})$ y $L(\mathfrak{B})$.

En particular, un espacio afín de dimensión finita n es isomorfo al producto cartesiano de n espacios afines de una dimensión.

5. Indiquemos por G el grupo de los automorfismos de un espacio afín $\mathfrak{A}$. Para todo punto $O \in \mathfrak{A}$ indiquemos por G_O el conjunto de todos los automorfismos de $\mathfrak{A}$ que dejan inmóvil el punto O. Sea, finalmente, D el conjunto de todas las traslaciones de $\mathfrak{A}$. Es fácil comprobar que G_O es un subgrupo de G y que D es un subgrupo abeliano invariante de G. De acuerdo con el teorema 4 del p. 29.1, tenemos la descomposición semidirecta

$$G = G_O \cdot D \quad (G_O \cap D = 1),$$

de la cual se deduce, en particular, que el grupo cociente G/D es isomorfo a G_O. Si el espacio $\mathfrak{A}$ es de una dimensión, el grupo G_O es isomorfo al grupo multiplicativo del cuerpo conmutativo K y es por lo tanto abeliano, mientras que el grupo G es metaabeliano.

6. Compruébese, en las notaciones del complemento anterior, la igualdad

$$\overrightarrow{AO} \cdot G_O \cdot \overrightarrow{OA} = \overrightarrow{OA}^{-1} \cdot G_O \cdot \overrightarrow{OA} = G_A.$$

7. Como hemos visto en el teorema 5 del p. 29.2, si la característica del cuerpo conmutativo K es diferente de 2, todo conjunto $\mathfrak{M}$ de puntos de un espacio afín sobre K que con dos cualesquiera puntos distintos suyos contiene también toda la recta que pasa por éstos, es un plano. Para cuerpos de característica 2 esto, en general, no es válido. Sea K el cuerpo conmutativo formado por dos elementos 0 y 1, sea $\mathfrak{L}$ el espacio vectorial de pares (x, y) $(x, y \in K)$ y sea $\mathfrak{A}$ el espacio afín sobre $\mathfrak{L}$ (fig. 8). Se comprueba fácilmente que $\mathfrak{A}$ tiene en total 4 puntos y 6 rectas. El conjunto $\mathfrak{M} = \{(0, 0), (0, 1), (1, 0)\}$ no es un plano, aunque con cualquier par de sus puntos contiene la recta que pasa por los mismos (que coincide con este par).

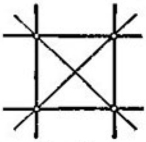

Fig. 8

§ 30. Coordenadas afines

30.1. Coordenadas de un punto. Sea $\mathfrak{A}$ un espacio afín de dimensión finita n sobre un cuerpo conmutativo K. Según el p. 29.2, toda sucesión $(A_0, A_1, \ldots, A_n)$ formada por $n+1$ puntos linealmente independientes del espacio $\mathfrak{A}$ se llama *base* de $\mathfrak{A}$. Una sucesión $(A_0, v_1, \ldots, v_n)$, formada por un punto cualquiera $A_0 \in \mathfrak{A}$ y por unos vectores linealmente independientes $v_1, \ldots, v_n$ del espacio vectorial $L(\mathfrak{A})$, se llama *base de referencia* (hablando con más precisión, base de referencia n-dimensional) de $\mathfrak{A}$. El punto A_0 se llama origen de la base de referencia. Está claro que a toda base $(A_0, A_1, \ldots, A_n)$ le corresponde la base de referencia $(A_0, \overline{A_0A_1}, \ldots, \ldots, \overline{A_0A_n})$ y que a toda base de referencia $(A_0, v_1, \ldots, v_n)$ le corresponde la base $(A_0, A_0v_1, \ldots, A_0v_n)$. Por esta razón no se hace frecuentemente diferencia entre los conceptos de base y de base de referencia.

Fijemos en el espacio $\mathfrak{A}$ una base de referencia $R = (O, \overline{OA_1}, \ldots, \overline{OA_n})$ que será llamada *base de referencia coordenada*. Para todo $X \in \mathfrak{A}$, su radio vector $\overline{OX}$ puede ser representado entonces unívocamente en la forma

$$\overline{OX}=\xi^1\cdot\overline{OA_1}+\ldots+\xi^n\cdot\overline{OA_n}\quad(\xi^i\in K).\tag{1}$$

Los números $\xi^1, \ldots, \xi^n$ se llaman *coordenadas del punto* X en la base de referencia $\boldsymbol{R}$ (o en la base $(O, A_1, \ldots, A_n)$) y la fila $\xi=(\xi^1, \ldots, \xi^n)$ se llama *fila coordenada* del punto X en la base de referencia señalada. Estas coordenadas se llaman a veces coordenadas afines. Al cambiar la base de referencia coordenada (incluso, por ejemplo, al cambiar el orden en el que siguen los vectores), cambia también la fila coordenada del punto. La ley de variación será considerada más tarde, en el p. 30.4.

Sean X_0, X_1 y X_2 unos puntos arbitrarios del espacio. Se dice que el punto X_1 divide al par de puntos (X_0, X_2) según la razón λ $(\lambda\in K)$, si $\overline{X_0X_1}=\lambda\overline{X_1X_2}$. De aquí se deduce, en particular, que si el punto X_1 divide al par de puntos (X_0, X_2) según una razón determinada, los tres puntos se hallan sobre una misma recta. Recíprocamente, si los puntos X_0, X_1, X_2 se hallan sobre una misma recta y $X_1\neq X_2$, el punto X_1 divide indudablemente al par (X_0, X_2) según la razón $\lambda=\overline{X_0X_1}:\overline{X_1X_2}$, ya que los vectores $\overline{X_0X_1}$ y $\overline{X_1X_2}$ deben ser linealmente dependientes.

¿Cómo determinar — dados el número $\lambda\in K$ y las coordenadas $(\xi_0^1, \ldots, \xi_0^n)$ y $(\xi_2^1, \ldots, \xi_2^n)$ de los puntos X_0 y X_2 — las coordenadas $(\xi_1^1, \ldots, \xi_1^n)$ del punto X_1 que divide al par (X_0, X_2) según la razón λ? Tenemos, por hipótesis,

$$\overline{OX_i}=\xi_i^1\cdot\overline{OA_1}+\ldots+\xi_i^n\cdot\overline{OA_n}\quad(i=0,\ 1,\ 2),$$

de donde

$$\overline{X_iX_{i+1}}=\overline{OX}_{i+1}-\overline{OX}_i=(\xi_{i+1}^1-\xi_i^1)\cdot\overline{OA}_1+\ldots+(\xi_{i+1}^n-\xi_i^n)\cdot\overline{OA}_n.$$

De la relación $\overline{X_0X_1}=\lambda\overline{X_1X_2}$ resulta

$$\xi_1^j-\xi_0^j=\lambda(\xi_2^j-\xi_1^j)$$

y por ello

$$\xi_1^j=\frac{\xi_0^j+\lambda\xi_2^j}{1+\lambda}\quad(j=1,\ \ldots,\ n).$$

Está claro que el número dado λ debe ser diferente de -1, ya que en el caso contrario de $\overline{X_0X_1}=-\overline{X_1X_2}$ se tiene $X_0=X_2$. Tomando $\lambda=1$, obtenemos las fórmulas

$$\xi_1^j=\frac{\xi_0^j+\xi_2^j}{2}\quad(j=1,\ \ldots,\ n)$$

para las coordenadas del *punto medio* X_1 del par (X_0, X_2).

Veamos ahora cómo se refleja la independencia lineal de unos puntos arbitrarios $X_1, \ldots, X_s$ del espacio $\mathfrak{A}$ en las filas coordenadas de los mismos. Sea

$$\overline{OX_i}=\xi_i^1\cdot\overline{OA_1}+\ldots+\xi_i^n\cdot\overline{OA_n}\quad(i=1,\ \ldots,\ s),\tag{2}$$

de modo que $(\xi_i^1, \ldots, \xi_i^n)$ es la fila coordenada del punto X_i en la base de referencia escogida. La independencia lineal de estos puntos equivale, por definición, a la independencia lineal de los vectores $\overline{X_1X_2}, \ldots, \overline{X_1X_s}$. Tenemos de la igualdad (2)

$$\overline{X_1X_i}=\overline{OX_i}-\overline{OX_1}=(\xi_i^1-\xi_1^1)\cdot\overline{OA_1}+\ldots+(\xi_i^n-\xi_1^n)\cdot\overline{OA_n},$$

es decir, la fila $(\xi_i^1-\xi_1^1, \ldots, \xi_i^n-\xi_1^n)$ es la fila coordenada del vector $\overline{X_1X_i}$ en la base $\overline{OA_1}, \ldots, \overline{OA_n}$. Por esto el número máximo de puntos linealmente independientes del sistema $X_1, \ldots, X_s$, que se obtiene agregando la unidad al número máximo de vectores linealmente independientes del sistema $\overline{X_1X_2}, \ldots, \overline{X_1X_s}$, es igual a

$$1+\text{rango}\begin{bmatrix}\xi_2^1-\xi_1^1 & \ldots & \xi_2^n-\xi_1^n\\ \cdot & \cdot\ \cdot\ \cdot & \cdot\\ \xi_s^1-\xi_1^1 & \ldots & \xi_s^n-\xi_1^n\end{bmatrix}=1+\text{rango}\,\|\xi_i^j-\xi_1^j\|.$$

Es fácil ver que

$$1+\text{rango}\begin{bmatrix}\xi_2^1-\xi_1^1 & \ldots & \xi_2^n-\xi_1^n\\ \cdot & \cdot\ \cdot\ \cdot & \cdot\\ \xi_s^1-\xi_1^1 & \ldots & \xi_s^n-\xi_1^n\end{bmatrix}=\text{rango}\begin{bmatrix}1 & \xi_1^1 & \ldots & \xi_1^n\\ 1 & \xi_2^1 & \ldots & \xi_2^n\\ \cdot & \cdot & \cdot & \cdot\\ 1 & \xi_s^1 & \ldots & \xi_s^n\end{bmatrix}.$$

Efectivamente, restando de la segunda fila, ..., de la n-ésima fila de la matriz del segundo miembro su primera fila, no alteramos su rango y, por ello,

$$\text{rango}\begin{bmatrix}1 & \xi_1^1 & \ldots & \xi_1^n\\ 1 & \xi_2^1 & \ldots & \xi_2^n\\ \cdot & \cdot & \cdot & \cdot\\ 1 & \xi_s^1 & \ldots & \xi_s^n\end{bmatrix}=\text{rango}\begin{bmatrix}1 & \xi_1^1 & \ldots & \xi_1^n\\ 0 & \xi_2^1-\xi_1^1 & \ldots & \xi_2^n-\xi_1^n\\ \cdot & \cdot & \cdot & \cdot\\ 0 & \xi_s^1-\xi_1^1 & \ldots & \xi_s^n-\xi_1^n\end{bmatrix}=$$
$$=1+\text{rango}\,\|\xi_i^j-\xi_1^j\|$$

que es lo que se quería demostrar.

Sea $(\xi^1, \ldots, \xi^n)$ la fila coordenada de un punto X. Con frecuencia la fila $(1, \xi^1, \ldots, \xi^n)$ suele llamarse entonces fila coordenada *ampliada* del punto X. El resultado que hemos obtenido puede ser expresado ahora en los términos siguientes: *el número máximo de puntos linealmente independientes en el sistema $X_1, \ldots, X_s$ es igual al rango de la matriz formada por las filas coordenadas ampliadas de los puntos de este sistema.*

En particular, *para que los puntos $X_1, \ldots, X_n, X_{n+1}$ de un espacio afín de n dimensiones sean linealmente independientes es necesario y suficiente que*

$$\begin{vmatrix}1 & \xi_1^1 & \ldots & \xi_1^n\\ 1 & \xi_2^1 & \ldots & \xi_2^n\\ \cdot & \cdot & \cdot & \cdot\\ 1 & \xi_{n+1}^1 & \ldots & \xi_{n+1}^n\end{vmatrix}\neq 0, \tag{3}$$

donde $\xi_i^1, \ldots, \xi_i^n$ *son las coordenadas del punto* X_i $(i=1, \ldots, n+1)$.

El determinante que figura en (3) a la izquierda del signo $\neq$ se llama a veces *determinante volumen* del sistema de puntos $X_1, \ldots, X_{n+1}$. Su anulación significa que los puntos $X_1, \ldots, X_{n+1}$ se hallan sobre un mismo hiperplano del espacio $\mathfrak{A}$.

Hemos introducido en el p. 29.4 el concepto de una funcional lineal $f(X)$ en un espacio afín $\mathfrak{A}$. Veamos cómo pueden expresarse los valores de esta funcional en términos de las coordenadas del punto X. Sea $(\xi^1, \ldots, \xi^n)$ la fila coordenada de un punto X en una base de referencia $\boldsymbol{R}=(O, \overline{OA}_1, \ldots, \overline{OA}_n)$. De (1) y de la relación que caracteriza la linealidad de $f(X)$ obtenemos entonces

$$f(X)=\alpha_1\xi^1+\ldots+\alpha_n\xi^n+\alpha_0, \tag{4}$$

donde $\alpha_i=f(A_i)-f(O)$ y $\alpha_0=f(O)$ son unos números fijos que dependen solamente de la funcional f y de la base de referencia $\boldsymbol{R}$. Viceversa, tomando unos números $\alpha_0, \alpha_1, \ldots, \alpha_n \in K$ totalmente arbitrarios y definiendo la funcional f mediante la fórmula (4), podemos comprobar fácilmente que f es lineal sobre $\mathfrak{A}$.

Se dice que *el polinomio* $\alpha_1\xi^1+\ldots+\alpha_n\xi^n+\alpha_0$ *en las variables* $\xi^1, \ldots, \xi^n$ *representa la funcional* f *en la base de referencia* $\boldsymbol{R}$. De lo expuesto se ve que la correspondencia entre los polinomios lineales en n variables $\xi^1, \ldots, \xi^n$ y las funcionales lineales sobre $\mathfrak{A}$ es biyectiva; además, si la funcional f está representada por el polinomio $\alpha_1\xi^1+\ldots+\alpha_n\xi^n+\alpha_0$ y la funcional g por el polinomio $\beta_1\xi^1+\ldots+\beta_n\xi^n+\beta_0$, se tiene

$$\lambda f(X)+\mu g(X)=(\lambda\alpha_1+\mu\beta_1)\xi^1+\ldots+(\lambda\alpha_n+\mu\beta_n)\xi^n+\lambda\alpha_0+\mu\beta_0. \tag{5}$$

En otras palabras, poniendo en correspondencia a toda funcional lineal sobre $\mathfrak{A}$ el polinomio en $\xi^1, \ldots, \xi^n$ que la representa, obtenemos *una aplicación isomorfa* del espacio vectorial $\mathfrak{A}_f$ de todas las funcionales lineales, definidas sobre $\mathfrak{A}$, sobre el espacio vectorial de todos los polinomios lineales (no homogéneos) en $\xi^1, \ldots \xi^n$ con coeficientes del cuerpo conmutativo principal K.

De (5) se ve que las operaciones de adición de las funcionales lineales y de multiplicación de las mismas por un número se reducen a las operaciones correspondientes con las filas de coeficientes de los polinomios lineales. De aquí deducimos, en particular, que el número máximo de funcionales lineales linealmente independientes, contenidas en el sistema $f_1(X), \ldots, f_s(X)$, es igual al rango de la matriz formada por los coeficientes de las formas lineales que representan a las funcionales señaladas.

30.2. Ecuaciones de planos. Aceptaremos que en un espacio $\mathfrak{A}$ de n dimensiones sobre un cuerpo conmutativo K se ha fijado una base de referencia $\boldsymbol{R}=(O, \overline{OA}_1, \ldots, \overline{OA}_n)$ y en lo sucesivo entenderemos por coordenadas de los puntos de $\mathfrak{A}$ las coordenadas de éstos en la base de referencia $\boldsymbol{R}$.

Sea $P(\xi^1, \ldots, \xi^n)$ *una condición* que relaciona unos números arbitrarios $\xi^1, \ldots, \xi^n$ de K. Luego, para todo sistema concreto de valores de las variables $\xi^1, \ldots, \xi^n \in K$ la condición P puede ser verídica o falsa. Se dice que *la condición P define en el espacio $\mathfrak{A}$ un conjunto de puntos $\mathfrak{M}$, si cualquier punto $X \in \mathfrak{A}$ pertenece a $\mathfrak{M}$ cuando, y sólo cuando, las coordenadas $(\xi^1, \ldots, \xi^n)$ del punto X satisfacen la condición P.*

De aquí resulta, en particular, que si las condiciones $P(\xi^1, \ldots$ $\ldots, \xi^n)$ y $Q(\xi^1, \ldots, \xi^n)$ definen en el espacio $\mathfrak{A}$ los conjuntos respectivos $\mathfrak{M}$ y $\mathfrak{N}$, la conjunción de estas condiciones $P(\xi^1, \ldots$ $\ldots, \xi^n)$ *y* $Q(\xi^1, \ldots, \xi^n)$ define la intersección de los conjuntos señalados, la disjunción P *o* Q define la unión $\mathfrak{M} \cup \mathfrak{N}$ de los mismos y la negación *no* P define el complemento $C\mathfrak{M} = \mathfrak{A} \setminus \mathfrak{M}$.

Sea $f(\xi^1, \ldots, \xi^n)$ una función que está definida sobre K y toma valores en K. La condición de tipo

$$f(\xi^1, \ldots, \xi^n) = 0 \tag{1}$$

se llama *ecuación* en las variables $\xi^1, \ldots, \xi^n$ y la conjunción de las ecuaciones

$$f_i(\xi^1, \ldots, \xi^n) = 0 \quad (i = 1, 2, \ldots, s) \tag{2}$$

se llama *sistema de ecuaciones* en las variables $\xi^1, \ldots, \xi^n$. Si el sistema de ecuaciones (2) define en el espacio $\mathfrak{A}$ un conjunto $\mathfrak{M}$, se dice también que (2) es el sistema de ecuaciones para el conjunto $\mathfrak{M}$. El conjunto $\mathfrak{M}$ de puntos definido por el sistema de ecuaciones (2) es, según la observación hecha anteriormente, la intersección de los conjuntos definidos por cada una de las ecuaciones del sistema (2) por separado.

En un cuerpo conmutativo arbitrario la condición $\alpha\beta = 0$ equivale a la disjunción $\alpha = 0$ *o* $\beta = 0$. Por esto, si las ecuaciones

$$f(\xi^1, \ldots, \xi^n) = 0 \quad \text{y} \quad g(\xi^1, \ldots, \xi^n) = 0$$

definen (por separado) en el espacio $\mathfrak{A}$ los conjuntos respectivos $\mathfrak{M}$ y $\mathfrak{N}$, la ecuación

$$f(\xi^1, \ldots, \xi^n) \cdot g(\xi^1, \ldots, \xi^n) = 0 \tag{3}$$

define en $\mathfrak{A}$ la unión de los conjuntos $\mathfrak{M}$ y $\mathfrak{N}$.

Un conjunto $\mathfrak{M}$ de puntos de un espacio afín $\mathfrak{A}$ se llama *hipersuperficie algebraica*, si existe un polinomio $f(\xi^1, \ldots, \xi^n)$ en las variables $\xi^1, \ldots, \xi^n$ con coeficientes del cuerpo conmutativo K tal que $\mathfrak{M}$ se define por la ecuación (1).

El grado del polinomio f respecto al conjunto de las variables $\xi^1, \ldots, \xi^n$ se llama *grado* de la hipersuperficie $\mathfrak{M}$. Puesto que un mismo conjunto $\mathfrak{M}$ puede tener varias ecuaciones diferentes, una misma hipersuperficie algebraica puede tener diferentes grados. Es fácil darse cuenta, en particular, de que si una hipersuperficie tiene el grado n, cualquier número mayor también es grado de ella. El menor de los grados de una hipersuperficie algebraica dada se llama *orden* de la misma.

Como el grado de un producto de polinomios es igual a la suma de los grados de los factores, de la fórmula (3) se deduce que la unión de unas hipersuperficies algebraicas de ordenes s y t es de nuevo una superficie algebraica de orden no mayor que $s+t$.

Se dice que una hipersuperficie algebraica $\mathfrak{M}$ se descompone, si es la unión de dos hipersuperficies algebraicas no vacías diferentes de la dada.

La intersección de un número finito de hipersuperficies algebraicas se llama *variedad algebraica*.

Los conceptos de hipersuperficie algebraica y de su orden han sido definidos mediante el tipo de las ecuaciones a las que satisfacen las coordenadas de un punto arbitrario de la hipersuperficie. Al cambiar la base de referencia coordenada, también cambiarán las coordenadas del punto y con ellas las ecuaciones del conjunto considerado. En lugar de hipersuperficie algebraica de orden s, sería más correcto hablar por esto *de una hipersuperficie algebraica de orden s en una base de referencia coordenada dada* y, análogamente, *de la descomposición en la base de referencia coordenada dada*, etc. Sin embargo, demostraremos en adelante que la propiedad de un conjunto de ser una hipersuperficie algebraica de un orden dado no depende, de hecho, de la selección de la base de referencia coordenada.

Las propiedades de las variedades algebraicas arbitrarias constituyen el tema de estudio de una asignatura especial, la Geometría algebraica. En este libro serán estudiadas solamente las hipersuperficies de primer orden (los hiperplanos) y sus intersecciones (los planos).

Es poco cómodo que las definiciones de *algebraicidad* y de *orden* dependan, como hemos señalado, de la selección de la base de referencia coordenada. Para deshacerse de esta dependencia se emplea el siguiente procedimiento. Sea $f(X)$ una funcional dada en un espacio afín $\mathfrak{A}$ de dimensión n. Escojamos en $\mathfrak{A}$ una base de referencia coordenada cualquiera $\boldsymbol{R}=(O, \overline{OA}_1, \ldots, \overline{OA}_n)$ y pongamos en correspondencia a toda sucesión $\xi^1, \ldots, \xi^n$ de números de K el número $f(\xi^1, \ldots, \xi^n)$ igual a $f(X)$, donde X es el punto de coordenadas $\xi^1, \ldots, \xi^n$. Resulta así que a toda funcional se le pone en correspondencia una función en n variables $\xi^1, \ldots, \xi^n$ que *representa* a la funcional f en la base de referencia coordenada dada $\boldsymbol{R}$. Recíprocamente, definiendo para la función $f(\xi^1, \ldots, \xi^n)$ la funcional $f(X)$ mediante la igualdad

$$f(X)=f(\xi^1, \ldots, \xi^n),$$

vemos que cualquier función en n variables representa a una funcional en la base de referencia señalada. Está claro que *la variedad radical de la funcional* $f(X)$, compuesta de aquellos puntos $X \in \mathfrak{A}$ para los cuales $f(X)=0$, coincide con el conjunto que la ecuación $f(\xi^1, \ldots, \xi^n)=0$ define en la base de referencia coordenada dada. Sin embargo, en la definición del concepto de una fincional y de su variedad radical no figuran las bases de referencia coordenadas y, por ello, es preferible representar las ecuaciones de un conjunto de puntos en la forma $f(X)=0$, donde f es una funcional. ¿Cómo entonces caracterizar aquellas funcionales cuyos valores se representan por polinomios? La respuesta es la siguiente. Hemos visto ya en el p. 30.1 que los polinomios *lineales* representan a las funcionales lineales. Consideremos ahora las funcionales $F(X_1, \ldots, X_s)$ en s puntos variables. Una funcional $F(X_1, \ldots, X_s)$ se llama *lineal respecto al i-ésimo argumento* X_i, si se convierte en una funcional lineal en X_i para cualesquiera valores fijos de las restantes variables. Una funcional $F(X_1 \ldots, X_s)$ se llama *polilineal* si es lineal respecto a cada uno de sus argumentos. Hemos llegado ahora al momento central: una funcional $f(X)$ en una variable X se llama *funcional* (o *forma*) *de orden s* (o *de grado s*), si existe una funcional polilineal $F(X_1, \ldots, X_s)$ en s variables tal que

$$f(X)=F(X, \ldots, X).$$

Se comprueba fácilmente que habiendo sido fijada una base de referencia coordenada cualquiera, la funcional $f(X)$ es de orden s cuando, y sólo cuando, sus valores se representan por un polinomio adecuado de orden s en las coordenadas del punto X. La demostración la omitimos aquí debido a su evidencia.

Después de estas consideraciones generales pasamos ahora al problema principal de este parágrafo que es el estudio de las ecuaciones de los planos.

Supongamos, pues, que en un espacio afín dado de dimensión finita n se ha fijado una base de referencia coordenada arbitraria $\boldsymbol{R}=(O, \overline{OA}_1, \ldots, \overline{OA}_n)$. Una ecuación de primer grado en las variables $\xi^1, \ldots, \xi^n$ es una ecuación de tipo

$$\alpha_1\xi^1+\ldots+\alpha_n\xi^n=\beta, \tag{4}$$

donde al menos uno de los coeficientes $\alpha_1, \ldots, \alpha_n$ es diferente de cero. Introduciendo la funcional

$$f(X)=\alpha_1\xi^1+\ldots+\alpha_n\xi^n-\beta,$$

vemos que el conjunto de puntos $\mathfrak{M}$ representado por la ecuación (4) es la variedad radical de la funcional $f(X)$. La funcional $f(X)$ es, según el p. 30.1, lineal y no es constante, de modo que $\mathfrak{M}$ es un hiperplano en $\mathfrak{A}$. Luego, *una ecuación lineal arbitraria en unas variables* $\xi^1, \ldots, \xi^n$ *representa un hiperplano en un espacio afín.*

Consideremos un sistema cualquiera de ecuaciones lineales

$$\left.\begin{array}{l}\alpha_1^1\xi^1+\alpha_2^1\xi^2+\ldots+\alpha_n^1\xi^n=\beta^1,\\ \ldots\ldots\ldots\ldots\ldots\ldots\\ \alpha_1^s\xi^1+\alpha_2^s\xi^2+\ldots+\alpha_n^s\xi^n=\beta^s.\end{array}\right\} \tag{5}$$

Indicando por $\mathfrak{M}_i$ el conjunto de los puntos cuyas coordenadas satisfacen la i-ésima ecuación de (5), vemos que el sistema de ecuaciones (5) define la intersección de hiperplanos $\mathfrak{M}=\mathfrak{M}_1\cap\ldots\cap\mathfrak{M}_s$. Según el p. 29.2, esta intersección o bien es vacía o bien es un plano cuya codimensión no pasa de s, de modo que la dimensión de $\mathfrak{M}$ es no menor que $n-s$. ¿Cuál es el valor exacto de la dimensión de $\mathfrak{M}$?

Introduzcamos las funcionales lineales

$$f_i(X)=\alpha_1^i\xi^1+\ldots+\alpha_n^i\xi^n-\beta^i \quad (i=1,\ \ldots,\ s).$$

El plano $\mathfrak{M}$ es la variedad radical del sistema de estas funcionales. Si $\mathfrak{M}\neq\varnothing$, la dimensión de $\mathfrak{M}$ es igual, por lo visto en el teorema 2 del p. 29.4, al número máximo de funcionales linealmente independientes en el sistema $f_1,\ \ldots,\ f_s$, es decir (véase el p. 29.4), la dimensión de $\mathfrak{M}$ es igual al rango de la matriz $\|\alpha_j^i\|$ formada por los coeficientes de las variables del sistema de ecuaciones (5).

Resta aclarar en qué caso $\mathfrak{M}=\varnothing$ y en qué caso $\mathfrak{M}\neq\varnothing$, es decir, bajo qué condiciones el sistema (5) es compatible. Pero la respuesta a esta última pregunta viene dada por el teorema de Kronecker—Capelli (véase el p. 5.3): para que el sistema de ecuaciones (5) sea compatible es necesario y suficiente que el rango de la matriz principal de este sistema coincida con el rango de su matriz ampliada. Hemos obtenido de esta forma el teorema siguiente:

TEOREMA. *En un espacio afín $\mathfrak{A}$ de n dimensiones todo plano $(n-r)$-dimensional $(0\leqslant r\leqslant n)$ $\mathfrak{M}$ puede ser representado, en cualquier base de referencia coordenada, por un sistema de ecuaciones de tipo*

$$\alpha_1^i\xi^1+\alpha_2^i\xi^2+\ldots+\alpha_n^i\xi^n=\beta^i \quad (i=1,\ \ldots,\ r) \tag{6}$$

tal que el rango de la matriz principal $\|\alpha_j^i\|$ es igual a r. Un sistema arbitrario de ecuaciones lineales de tipo (4) *representa en $\mathfrak{A}$ un plano $(n-r)$-dimensional siempre que los rangos de las matrices principal y ampliada del sistema* (4) *coincidan y sean iguales a r. En cambio, si dichos rangos son diferentes, el sistema* (4) *es incompatible y, por consiguiente, representa el plano vacío.*

Consideremos el problema siguiente. Dadas las filas coordenadas $(\beta_i^1,\ \ldots,\ \beta_i^n)$ de unos puntos B_i $(i=0,\ 1,\ \ldots,\ s)$ de un espacio $\mathfrak{A}$, hallar la ecuación del plano mínimo $\mathfrak{M}=B_0\vee B_1\vee\ldots\vee B_s$ que pasa por los puntos dados.

Resolvámoslo primero en la forma vectorial. El plano $\mathfrak{M}$ se compone, según el p. 29.2, de los puntos $X \in \mathfrak{A}$ tales que el vector $\overline{B_0X}$ puede ser representado en la forma

$$\overline{B_0X} = \lambda_1 \overline{B_0B_1} + \ldots + \lambda_s \overline{B_0B_s} \quad (\lambda_1, \ \ldots, \ \lambda_s \in K),$$

es decir,

$$\overline{OX} = \lambda_1 (\overline{OB_1} - \overline{OB_0}) + \ldots + \lambda_s (\overline{OB_s} - \overline{OB_0}) + \overline{OB_0}. \tag{7}$$

La ecuación (7) se llama a veces ecuación de un plano *en la forma vectorial paramétrica*. Introduciendo aquí en lugar de los radios vectores $\overline{OX}$ y $\overline{OB_j}$ sus filas coordenadas correspondientes, obtenemos las relaciones

$$\xi^i = (\beta_1^i - \beta_0^i)\lambda_1 + \ldots + (\beta_s^i - \beta_0^i)\lambda_s + \beta_0^i \quad (i = 1, \ \ldots, \ n). \tag{8}$$

Poniendo $\beta_j^i - \beta_0^i = \gamma_j^i$, podemos representar estas relaciones en la forma

$$\xi^i = \gamma_1^i \lambda_1 + \ldots + \gamma_s^i \lambda_s + \beta_0^i \quad (i = 1, \ \ldots, \ n), \tag{9}$$

que expresa las coordenadas de un punto arbitrario $X \in \mathfrak{M}$ en términos de los parámetros independientes auxiliares $\lambda_1, \ \ldots, \ \lambda_s$. Las ecuaciones (9) se llaman *ecuaciones coordenadas paramétricas* del plano $\mathfrak{M}$. Puesto que $\mathfrak{M} = B_0 \vee B_1 \vee \ldots \vee B_s$, la dimensión de $\mathfrak{M}$ es igual al número máximo de puntos linealmente independientes que figuran en el sistema $B_0, \ B_1, \ \ldots, \ B_s$ disminuido en 1, es decir, es igual al rango de la matriz $C = \| \gamma_j^i \|$ formada por los coeficientes de las variables $\lambda_1, \ \ldots, \ \lambda_s$ de las ecuaciones paramétricas.

Para obtener de las ecuaciones paramétricas (9) de un plano $\mathfrak{M}$ sus ecuaciones generales de tipo (5), es suficiente eliminar del sistema (9) los parámetros $\lambda_1, \ \ldots, \ \lambda_s$. Esto se puede hacer, por ejemplo, del modo siguiente. Buscamos el rango de la matriz $\| \gamma_j^i \|$. Supongamos que éste es igual a r; luego, entre las ecuaciones (9) existen unas ecuaciones i_1-ésima, i_2-ésima, $\ldots$, i_r-ésima y entre los parámetros $\lambda_1, \ \ldots \ \lambda_s$ existen unos parámetros $\lambda_{j_1}, \ \ldots, \ \lambda_{j_r}$ $(i_1 < \ldots < i_r; \ j_1 < \ldots < j_r)$ tales que $\det \left\| \gamma_{j_\beta}^{i_\alpha} \right\| \neq 0$. Resolviendo estas ecuaciones respecto a las incógnitas $\lambda_{j_1}, \ \ldots, \ \lambda_{j_r}$, obtendremos para éstas unas expresiones lineales en términos de $\xi_1, \ \ldots, \ \xi_n$ y de los restantes parámetros λ_j $(j \notin \{j_1, \ \ldots, \ j_r\})$. Introduzcamos ahora los valores obtenidos de los parámetros en cada una de las ecuaciones restantes del sistema (9). Obtendremos así unas relaciones lineales entre las variables $\xi_1, \ \ldots, \ \xi_n$ y los parámetros λ_j, $j \notin \{j_1, \ \ldots, \ j_r\}$. Los parámetros λ_j aparecerán, de hecho, en estas relaciones con coeficientes nulos (debido a que r es el rango de la matriz $\| \gamma_j^i \|$) y, por ello, las relaciones obtenidas serán unas relaciones de tipo (5) que representarán el plano $\mathfrak{M}$.

En los razonamientos realizados hemos supuesto que $0 \leqslant r < n$. Si $r = n$, el plano $\mathfrak{M}$ pasará por $n+1$ puntos linealmente independientes y coincidirá con todo el espacio $\mathfrak{A}$; si se quiere, su «ecuación general» puede ser representada en la forma

$$0 \cdot \xi^1 + \ldots + 0 \cdot \xi^n = 0. \tag{10}$$

Consideremos el problema recíproco: ¿cómo hallar las ecuaciones paramétricas de un plano $\mathfrak{M}$ dado por unas ecuaciones generales de tipo (5)? Supongamos que la dimensión de $\mathfrak{M}$ es $r < n$. El sistema (5) contiene entonces sólo r ecuaciones independientes y los coeficientes que tienen en estas ecuaciones unas r variables forman una matriz de determinante diferente de 0. Supongamos, para concretar, que son independientes las r primeras ecuaciones y que es diferente de cero el determinante formado por los coeficientes de las variables $\xi^1, \ldots, \xi^r$. Resolviendo las r primeras ecuaciones respecto a las variables $\xi^1, \ldots, \xi^r$, obtendremos entonces un sistema de ecuaciones de tipo

$$\xi^i = \gamma^i_{r+1}\xi^{r+1} + \ldots + \gamma^i_n\xi^n + \beta^i \quad (i = 1, \ldots, r) \tag{11}$$

equivalente al sistema (9). Está claro que el sistema (11) es equivalente a las ecuaciones paramétricas

$$\begin{aligned} \xi^i &= \gamma^i_{r+1}\lambda_1 + \ldots + \gamma^i_n\lambda_{n-r} + \beta^i & (i &= 1, \ldots, r), \\ \xi^j &= \lambda_{j-r} & (j &= r+1, \ldots, n). \end{aligned}$$

Hemos considerado detalladamente el problema de la determinación de las ecuaciones del plano mínimo que pasa por los puntos dados. Está claro que dicho problema representa un caso particular de un problema más general sobre la determinación del plano mínimo que pasa por los planos dados $\mathfrak{M}$ y $\mathfrak{N}$. Sin embargo, este problema más general puede ser reducido de un modo formal al primer problema. En efecto, los planos $\mathfrak{M}$ y $\mathfrak{N}$ pueden ser representados en la forma

$$\begin{aligned} \mathfrak{M} &= B_0 \vee B_1 \vee \ldots \vee B_s \quad (\dim.\mathfrak{M} = s), \\ \mathfrak{N} &= C_0 \vee C_1 \vee \ldots \vee C_t \quad (\dim.\mathfrak{N} = t), \end{aligned}$$

donde $B_0, B_1, \ldots, B_s$ son unos puntos linealmente independientes de $\mathfrak{M}$ y $C_0, C_1, \ldots, C_t$ son unos puntos linealmente independientes de $\mathfrak{N}$. Entonces

$$\mathfrak{M} \vee \mathfrak{N} = B_0 \vee \ldots \vee B_s \vee C_0 \vee \ldots \vee C_t$$

y el problema ha quedado reducido a la determinación del plano mínimo que pasa por los puntos $B_0, \ldots, B_s, C_0, \ldots, C_t$. ¿Cómo determinar los puntos $B_0, \ldots, B_s$ si el plano $\mathfrak{M}$ viene dado por sus ecuaciones generales de tipo (5)? Uno de los métodos (que no es el más breve) es el siguiente. Reducimos el sistema (5) a la forma (11). Indudablemente podemos tomar entonces para los puntos B_i los puntos de las filas coordenadas siguientes

$$\begin{aligned}
(B_0) &= (\beta^1, \quad \beta^2, \quad \ldots, \beta^r, \quad 0, 0, \ldots, 0),\\
(B_1) &= (\beta^1+\gamma^1_{r+1}, \beta^2+\gamma^2_{r+1}, \ldots, \beta^r+\gamma^r_{r+1}, 1, 0, \ldots, 0),\\
(B_2) &= (\beta^1+\gamma^1_{r+2}, \beta^2+\gamma^2_{r+2}, \ldots, \beta^r+\gamma^r_{r+2}, 0, 1, \ldots, 0),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
(B_s) &= (\beta^1+\gamma^1_n, \beta^2+\gamma^2_n, \ldots, \beta^r+\gamma^r_n, 0, 0, \ldots, 1)\\
&(s+r=n).
\end{aligned}$$

Consideremos otro problema particular. Supongamos que en un espacio $\mathfrak{A}$ se ha escogido una base de referencia coordenada $\boldsymbol{R}=(O, \overline{OA_1}, \ldots, \overline{OA_n})$. Los planos de tipo $O \vee A_{i_1} \vee \ldots \vee A_{i_r}$ $(1 \leqslant i_1 < \ldots < i_r \leqslant n)$ se llaman entonces *planos coordenados r-dimensionales*. Poniendo

$$\mathfrak{P}_i = O \vee A_1 \vee \ldots \vee A_{i-1} \vee A_{i+1} \vee \ldots \vee A_n, \tag{12}$$

vemos que $\mathfrak{P}_1, \mathfrak{P}_2, \ldots, \mathfrak{P}_n$ son todos *los hiperplanos coordenados* posibles. De (12) resulta que un punto X de coordenadas $(\xi^1, \ldots, \xi^n)$ pertenece a $\mathfrak{P}_i$ cuando, y sólo cuando,

$$\xi^i = 0. \tag{13}$$

En otras palabras, *la ecuación* (13) *es la ecuación del hiperplano* $\mathfrak{P}_i$. El plano coordenado $O \vee A_{i_1} \vee \ldots \vee A_{i_r} = \mathfrak{M}$ es la intersección de todos los hiperplanos $\mathfrak{P}_j$ que tienen el índice diferente de $i_1, \ldots, i_r$. Por esto las ecuaciones del plano $\mathfrak{M}$ pueden ser representadas en la forma

$$\xi^\alpha = \xi^\beta = \ldots = \xi^\gamma = 0,$$

donde $\alpha, \beta, \ldots, \gamma$ es la sucesión de aquellos números naturales del conjunto $1, 2, \ldots, n$ que no figuran en el conjunto $\{i_1, \ldots, i_r\}$. Por ejemplo, las ecuaciones del i-ésimo eje coordenado $O \vee A_i$ pueden ser representadas en la forma

$$\xi^1 = \ldots = \xi^{i-1} = \xi^{i+1} = \ldots = \xi^n = 0.$$

A título de ilustración consideremos el siguiente ejemplo numérico. En el espacio afín R^3 corriente real de tres dimensiones está dada mediante dos ecuaciones

$$\left.\begin{aligned} \xi^1 + 2\xi^2 - \xi^3 &= 1,\\ \xi^1 + 2\xi^2 - 2\xi^3 &= 2 \end{aligned}\right\} \tag{14}$$

una variedad lineal $\mathfrak{M}$.

¿Cuál es la dimensión de $\mathfrak{M}$? Componiendo la matriz ampliada de este sistema

$$\begin{bmatrix} 1 & 2 & -1 & 1\\ 1 & 2 & -2 & 2 \end{bmatrix},$$

vemos que contiene un determinante de segundo orden diferente de cero, cuyos elementos están indicados con las cifras gruesas. Es más, este determinante pertenece a la matriz principal del sistema. Por esto el rango de la matriz ampliada, que es igual a dos, coincide con el rango de la matriz principal. De aquí sacamos la conclusión de que la codimensión de $\mathfrak{M}$ es igual a 2 y de que la dimensión de $\mathfrak{M}$ es igual a 1, es decir, $\mathfrak{M}$ es una recta.

Resolviendo el sistema (14) respecto a ξ^1 y ξ^3, encontramos

$$\begin{aligned} \xi^3 &= -1, \\ \xi^1 &= -2\xi^2. \end{aligned} \tag{15}$$

Podemos aceptar aquí que ξ^2 es un parámetro, de modo que (15) puede ser considerado como las ecuaciones paramétricas de la recta $\mathfrak{M}$. Dando al parámetro ξ^2 los valores 0 y 1, obtenemos de (15) que los puntos de coordenadas $(0, 0, -1)$ y $(-2, 1, -1)$ se hallan sobre la recta $\mathfrak{M}$.

Consideremos otro ejemplo. ¿Qué conjuntos de puntos del espacio R^3 son representados por las ecuaciones

a) $\xi^1\xi^2\xi^3 = 0$,

b) $(\xi^1)^2 - \xi^1\xi^2 + \xi^1\xi^3 + \xi^1 - \xi^2 + \xi^3 = 0$?

Está claro que en el caso a) el conjunto incógnito es la colección de los tres hiperplanos coordenados de ecuaciones 1) $\xi^1 = 0$, 2) $\xi^2 = 0$ y 3) $\xi^3 = 0$. En el caso b), el primer miembro se descompone en los factores $\xi^1 + 1$ y $\xi^1 - \xi^2 + \xi^3$ y, por ello, el conjunto incógnito es la colección de los hiperplanos de ecuaciones $\xi^1 + 1 = 0$ y $\xi^1 - \xi^2 + \xi^3 = 0$.

30.3. Ecuaciones de hiperplanos y de rectas. Queremos examinar aquí más detalladamente las ecuaciones de los hiperplanos y de las rectas, así como las condiciones de paralelismo de los mismos, expresadas en la forma coordenada.

TEOREMA 1. *Todo hiperplano de un espacio afín de n dimensiones puede ser representado mediante una ecuación de primer grado*

$$\alpha_1\xi^1 + \ldots + \alpha_n\xi^n + \alpha_0 = 0. \tag{1}$$

Para que un hiperplano representado por la ecuación (1) *coincida con un hiperplano representado por la ecuación*

$$\beta_1\xi^1 + \ldots + \beta_n\xi^n + \beta_0 = 0, \tag{2}$$

es necesario y suficiente que sean proporcionales los coeficientes correspondientes:

$$\frac{\alpha_1}{\beta_1} = \ldots = \frac{\alpha_n}{\beta_n} = \frac{\alpha_0}{\beta_0}. \tag{3}$$

Los hiperplanos representados por las ecuaciones (1) *y* (2) *son paralelos cuando, y sólo cuando, sus respectivos coeficientes* *principales son proporcionales:*

$$\frac{\alpha_1}{\beta_1} = \ldots = \frac{\alpha_n}{\beta_n}. \tag{4}$$

El teorema contiene tres afirmaciones. La primera ha sido ya demostrada en el p. 30.2. Para que los hiperplanos (1) y (2) coincidan es necesario y suficiente que el conjunto de las ecuaciones (1) y (2) defina un hiperplano. Esto equivale, según el teorema del p. 30.2, a que

$$\text{rango} \begin{bmatrix} \alpha_1 & \ldots & \alpha_n & \alpha_0 \\ \beta_1 & \ldots & \beta_n & \beta_0 \end{bmatrix} = 1$$

y esto significa precisamente que tiene lugar (3).

Los hiperplanos (1) y (2) son paralelos, si coinciden, y entonces tenemos (3), o si no se intersecan. En el último caso las ecuaciones (1) y (2) deben ser incompatibles y por esto, según el teorema de Kronecker—Capelli, se tiene

$$\text{rango} \begin{bmatrix} \alpha_1 & \ldots & \alpha_n \\ \beta_1 & \ldots & \beta_n \end{bmatrix} = 1,$$

es decir, se tiene (4). Luego, la condición (4) es verídica en ambos casos.

Consideremos el problema: *dada la ecuación* (1) *de un hiperplano* $\mathfrak{P}$ *y las coordenadas* $\xi_0^1, \ldots, \xi_0^n$ *de un punto* A, *hallar la ecuación del hiperplano que es paralelo a* $\mathfrak{P}$ *y que pasa por el punto* A.

Sea (2) la ecuación del hiperplano $\mathfrak{P}'$ que buscamos. Puesto que este hiperplano es paralelo al hiperplano (1), deben cumplirse las condiciones (4). Por esto multiplicando la ecuación (2) por un coeficiente de proporcionalidad adecuado, podemos representar la ecuación del hiperplano que buscamos $\mathfrak{P}'$ en la forma

$$\alpha_1\xi^1 + \ldots + \alpha_n\xi^n + \beta_0' = 0. \tag{5}$$

El punto A se halle, por hipotesis, sobre $\mathfrak{P}'$. Por ello

$$\alpha_1\xi_0^1 + \ldots + \alpha_n\xi_0^n + \beta_0' = 0. \tag{6}$$

Restando de (5) término por término (6), obtenemos la ecuación deseada

$$\alpha_1(\xi^1 - \xi_0^1) + \ldots + \alpha_n(\xi^n - \xi_0^n) = 0.$$

Está claro que el hiperplano definido por la ecuación (1) pasa por el origen $(0, \ldots, 0)$ de la base de referencia coordenada cuando, y sólo cuando, $\alpha_0 = 0$. Consideremos el caso en que $\alpha_0 \neq 0$. Dividiendo por $-\alpha_0$ todos los coeficientes de la ecuación (1), podemos reducirla a la forma

$$\frac{\xi^1}{\gamma_1} + \ldots + \frac{\xi^n}{\gamma_n} = 1. \tag{7}$$

Los números $\gamma_1, \ldots, \gamma_n$ admiten una interpretación geométrica simple. Efectivamente, busquemos el punto de intersección del hiperplano (7) y del eje coordenado $O \vee A_i$. Las ecuaciones de este eje son

$$\xi^1 = \xi \ldots = \xi^{i-1} = \xi^{i+1} = \ldots = \xi^n = 0. \tag{8}$$

Resolviendo conjuntamente (7) y (8), obtenemos $\xi^i = \gamma_i$. Por consiguiente, el hiperplano (7) intercepta en el eje $O \vee A_i$ el vector $\gamma_i \overline{OA}_i$. Por esta razón la ecuación (7) suele llamarse generalmente *ecuación segmentaria de un hiperplano.*

Examinemos también el siguiente problema: *¿bajo qué condiciones la ecuación* (1) *representa un hiperplano que pasa por el plano coordenado* $O \vee A_1 \vee \ldots \vee A_r$ $(1 \leqslant r < n)$? Este plano esta formado por los puntos de coordenadas $(\xi^1, \ldots, \xi^r, 0, \ldots, 0)$. Introduciendo estas coordenadas en la ecuación (1), obtenemos $\alpha_1\xi^1 + \ldots + \alpha_r\xi^r + \alpha_0 = 0$. Esta relación debe satisfacerse para cualesquiera valores de los números $\xi^1, \ldots, \xi^r$ del cuerpo conmutativo K. Por consiguiente, $\alpha_1 = \ldots = \alpha_r = \alpha_0 = 0$. Análogamente se comprueba que el hiperplano (1) pasa por el plano coordenado r-dimensional $O \vee A_{i_1} \vee \ldots \vee A_{i_r}$, cuando, y sólo cuando, $\alpha_{i_1} = \ldots = \alpha_{i_r} = \alpha_0 = 0$.

Manteniendo invariables en la ecuación (1) los coeficientes principales $\alpha_1, \ldots, \alpha_n$ y dando diferentes valores al término independiente α_0, obtenemos ecuaciones de hiperplanos paralelos. Por esto, *si en la ecuación* (1) *se tiene* $\alpha_{i_1} = \ldots = \alpha_{i_r} = 0$, *mientras que el término independiente* α_0 *es arbitrario, el hiperplano* (1) *es paralelo al plano coordenado* $O \vee A_{i_1} \vee \ldots \vee A_{i_r}$. En particular, el hiperplano (1) es paralelo al eje coordenado $O \vee A_i$ cuando, y sólo cuando, $\alpha_i = 0$, es decir, cuando la ecuación (1) no contiene explícitamente la i-ésima coordenada.

Como sabemos, por cualesquiera n puntos linealmente independientes $B_1, \ldots, B_n$ de un espacio afín de n dimensiones pasa un hiperplano $B_1 \vee \ldots \vee B_n$ y sólo uno. ¿Cómo escribir la ecuación de este hiperplano, si se conocen las filas coordenadas de los puntos $B_1, \ldots, B_n$? Supongamos que $(\beta_1^1, \ldots, \beta_1^n), \ldots, (\beta_n^1, \ldots, \beta_n^n)$ son las filas coordenadas de los puntos $B_1, \ldots, B_n$. Para que un punto arbitrario X de coordenadas $\xi^1, \ldots, \xi^n$ se halle sobre el hiperplano $B_1 \vee \ldots \vee B_n$ es necesario y suficiente que el sistema de puntos $X, B_1, \ldots, B_n$ sea linealmente dependiente. Según el p. 30.1, el sistema de puntos señalado será linealmente dependiente cuando, y sólo cuando, las coordenadas de estos puntos satisfacen la condición

$$\begin{vmatrix} 1 & \xi^1 & \ldots & \xi^n \\ 1 & \beta_1^1 & \ldots & \beta_1^n \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \beta_n^1 & \ldots & \beta_n^n \end{vmatrix} = 0. \tag{9}$$

Desarrollando el determinante según los elementos de la primera fila, vemos que la ecuación (9) es lineal respecto a las variables $\xi^1, \ldots, \xi^n$ y, por ello, representa el plano deseado.

El término independiente de la ecuación (9) es igual a det $\|\beta_i^j\|$. Por consiguiente, el hiperplano que pasa por los puntos linealmente independientes $B_1, \ldots, B_n$ pasará por el origen de coordenadas cuando, y sólo cuando, det $\|\beta_i^j\| = 0$.

Consideremos ahora más detalladamente las ecuaciones de las rectas. Supongamos dado en un espacio afín $\mathfrak{A}$ de n dimensiones un par de puntos diferentes $B_0(\beta_0^1, \ldots, \beta_0^n)$ y $B_1(\beta_1^1, \ldots, \beta_1^n)$, donde se indican entre paréntesis las coordenadas de estos puntos en una base de referencia coordenada fija $\boldsymbol{R} = (O, \overline{OA_1}, \ldots, \overline{OA_n})$. El punto X pertenece a la recta $B_0 \vee B_1$ cuando, y sólo cuando, existe un número λ tal que $\overline{B_0X} = \lambda\overline{B_0B_1}$, es decir, cuando

$$\xi^i - \beta_0^i = \lambda(\beta_1^i - \beta_0^i) \quad (i = 1, \ldots, n). \tag{10}$$

Determinando λ de cada una de las ecuaciones (10) e igualando los resultados, obtenemos

$$\frac{\xi^1 - \beta_0^1}{\beta_1^1 - \beta_0^1} = \frac{\xi^2 - \beta_0^2}{\beta_1^2 - \beta_0} = \ldots = \frac{\xi^n - \beta_0^n}{\beta_1^n - \beta_0^n}. \tag{11}$$

Recíprocamente, siendo $\xi^1, \ldots, \xi^n$ las coordenadas de un punto X que satisfacen las igualdades (11) e indicando por λ el valor común de las razones (11), obtenemos (10), de modo que $X \in B_0 \vee B_1$. Por consiguiente, *las ecuaciones* (11) *son las ecuaciones de la recta que pasa por dos puntos dados.*

Notemos que en las razones (11) algunos de los denominadores pueden anularse. Puesto que queremos que las relaciones (11) sean equivalentes a las condiciones (10), debemos tomar $\xi^i - \beta_0^i = 0$ siempre que $\beta_1^i - \beta_0^i = 0$.

Poniendo $\beta_1^i - \beta_0^i = \mu^i$ y $\beta_0^i = \beta^i$, podemos representar las ecuaciones (11) en la forma

$$\frac{\xi^1 - \beta^1}{\mu^1} = \frac{\xi^2 - \beta^2}{\mu^2} = \ldots = \frac{\xi^n - \beta^n}{\mu^n}. \tag{12}$$

Los números $\mu^1, \ldots, \mu^n$ se llaman *coeficientes directores* de la recta y las ecuaciones (12) se llaman *ecuaciones de la recta en coeficientes directores.*

TEOREMA 2. *Para que la recta* (12) *sea paralela a la recta*

$$\frac{\xi^1 - \gamma^1}{\nu^1} = \frac{\xi^2 - \gamma^2}{\nu^2} = \ldots = \frac{\xi^n - \gamma^n}{\nu^n} \tag{13}$$

es necesario y suficiente que sus coeficientes directores sean proporcionales:

$$\frac{\mu^1}{\nu^1} = \frac{\mu^2}{\nu^2} = \ldots = \frac{\mu^n}{\nu^n}. \tag{14}$$

Las rectas (12) *y* (13) *coinciden cuando, y sólo cuando, se cumplen las condiciones de proporcionalidad* (14) *y las condiciones*

$$\frac{\beta^1-\gamma^1}{\nu^1}=\frac{\beta^2-\gamma^2}{\nu^2}=\ldots=\frac{\beta^n-\gamma^n}{\nu^n},$$

que significan que el punto de coordenadas $\beta^1, \ldots, \beta^n$ *que pertenece a la recta* (12) *se halle sobre la recta* (13).

De las ecuaciones (12) se ve que los puntos $B_0(\beta^1, \ldots, \beta^n)$ y $B_1(\beta^1+\mu^1, \ldots, \beta^n+\mu^n)$ se hallan sobre la recta que representan estas ecuaciones. Análogamente, sobre la recta (13) se hallan los puntos $C_0(\gamma^1, \ldots, \gamma^n)$ y $C_1(\gamma^1+\nu^1, \ldots, \gamma^n+\nu^n)$. Según el p. 29.3, las rectas $B_0\vee B_1$ y $C_0\vee C_1$ son paralelas cuando, y sólo cuando, se tenga $\overline{C_0C_1}=\lambda\overline{B_0B_1}$ para un $\lambda\in K$, es decir, cuando las coordenadas de los vectores $\overline{B_0B_1}$ y $\overline{C_0C_1}$ sean proporcionales. Puesto que las coordenadas de estos vectores son iguales, respectivamente, a $\mu^1, \ldots, \mu^n$ y $\nu^1, \ldots, \nu^n$, obtenemos de aquí (14). Como para la coincidencia de las rectas es suficiente que tengan un punto común y que sean paralelas, la segunda afirmación del teorema 2 se desprende directamente de la primera afirmación.

Además de las ecuaciones de tipo (13), una recta en un espacio afín de n dimensiones puede ser definida también mediante un sistema general de ecuaciones lineales de tipo

$$\alpha_1^i\xi^1+\ldots+\alpha_n^i\xi^n+\alpha_0^i=0 \quad (i=1, \ldots, s), \tag{15}$$

donde los rangos de la matriz principal y de la ampliada son iguales a $n-1$. ¿Cómo conociendo las ecuaciones de una recta en la forma general (15) hallar sus ecuaciones en la forma canónica (13)? Para ello es suficiente hallar las coordenadas de dos cualesquiera puntos diferentes de la recta (15) y recurrir después a la fórmula (11); pero se puede proceder también del modo siguiente. El rango del sistema (15) es, por hipótesis, igual a $n-1$. Luego, $n-1$ coordenadas variables, digamos $\xi^1, \ldots, \xi^{n-1}$, pueden ser expresadas mediante la restante coordenada variable ξ^n. Así obtenemos unas expresiones de tipo

$$\xi^i=\mu^i\xi^n+\beta^i \quad (i=1, \ldots, n-1),$$

de donde

$$\frac{\xi^1-\beta^1}{\mu^1}=\ldots=\frac{\xi^{n-1}-\beta^{n-1}}{\mu^{n-1}}=\frac{\xi^n}{1}.$$

Estas serán precisamente las ecuaciones canónicas de la recta (15).

Para concluir determinemos las condiciones de paralelismo de una recta y de un hiperplano en la forma coordenada.

TEOREMA 3. *Un hiperplano dado mediante la ecuación*

$$\alpha_1\xi^1+\ldots+\alpha_n\xi^n+\alpha_0=0 \tag{16}$$

es paralelo a una recta dada mediante las ecuaciones

$$\frac{\xi^1-\beta^1}{\mu^1}=\ldots=\frac{\xi^n-\beta^n}{\mu^n} \tag{17}$$

cuando, y sólo cuando,

$$\alpha_1\mu^1+\ldots+\alpha_n\mu^n=0. \tag{18}$$

Indicando por λ el valor común de las razones de (17), obtenemos

$$\xi^i=\mu^i\lambda+\beta^i \quad (i=1,\ \ldots,\ n). \tag{19}$$

Introduciendo en (16) estos valores en lugar de las coordenadas variables ξ^i, llegamos a la ecuación

$$(\alpha_1\mu^1+\ldots+\alpha_n\mu^n)\,\lambda=-(\alpha_1\beta^1+\ldots+\alpha_n\beta^n+\alpha_0). \tag{20}$$

Si resulta que $\alpha_1\mu^1+\ldots+\alpha_n\mu^n\neq 0$, encontraremos primero de (20) un valor único para λ y, después, de (19) unos valores, también únicos, para las coordenadas ξ^i del punto de intersección de la recta (17) y del hiperplano (16). Por consiguiente, en este caso la recta y el hiperplano son, sin duda alguna, no paralelos.

Sea $\alpha_1\mu^1+\ldots+\alpha_n\mu^n=0$. Si el segundo miembro de (20) es diferente de cero, la ecuación (20) no tiene soluciones para λ, la recta (17) y el hiperplano (16) no se cortan y, por ello, son paralelos. Por otra parte, si el segundo miembro de (20) es igual a cero, la ecuación (20) se cumple para cualesquiera valores de λ, es decir, todos los puntos de la recta se hallan sobre el hiperplano y, por consiguiente, la recta y el hiperplano son de nuevo paralelos.

30.4. Transformación de coordenadas afines. ¿Cómo varían las coordenadas $\xi^1, \ldots, \xi^n$ de un punto arbitrario X de un espacio afín $\mathfrak{A}$ de n dimensiones, si pasamos de la base de referencia coordenada fija $\boldsymbol{R}=(O,\ \overline{OA}_1,\ \ldots,\ \overline{OA}_n)$ a otra base de referencia $\boldsymbol{R}'=(O',\ \overline{O'A'_1},\ \ldots,\ \overline{O'A'_n})$? Consideremos primero el caso en el que la base de referencia $\boldsymbol{R}'$ se obtiene de la base de referencia $\boldsymbol{R}$ mediante una traslación determinada por el vector $\overline{OO'}$. En este caso se tiene

$$\overline{O'A'_i}=\overline{OA_i} \quad (i=1,\ \ldots,\ n). \tag{1}$$

Tenemos, por hipótesis,

$$\overline{OX}=\xi^1\cdot\overline{OA_1}+\ldots+\xi^n\cdot\overline{OA}_n. \tag{2}$$

Indicando por $\alpha^1, \ldots, \alpha^n$ las coordenadas del origen nuevo O' (en la base de referencia "antigua" $\boldsymbol{R}$) y por $\xi^{1'}, +\ldots, \xi^{n'}$ las coordenadas nuevas del punto X (en la base de referencia $\boldsymbol{R}'$), obtenemos

$$\overline{OO'}=\alpha^1\cdot\overline{OA_1}+\ldots+\alpha^n\cdot\overline{OA}_n, \tag{3}$$

$$\overline{O'X}=\xi^{1'}\cdot\overline{O'A'_1}+\ldots+\xi^{n}\cdot\overline{O'A'_n}. \tag{4}$$

Sumando término a término las igualadades (3) y (4), empleando las relaciones (1) y comparando el resultado obtenido con las fórmulas (2), llegamos a las fórmulas que buscamos

$$\xi^i = \xi^{i'} + \alpha^i \quad (i = 1, \ldots, n) \tag{5}$$

que representan *la ley de transformación de las coordenadas en una traslación del origen.*

Consideremos ahora el caso en el que el origen de la base de referencia coordenada nueva coincide con el origen de la base de referencia antigua. Los sistemas de los vectores coordenados $\overline{OA_1}$, $\ldots$, $\overline{OA_n}$ y $\overline{O'A'_1}$, $\ldots$, $\overline{O'A'_n}$ representan dos bases del espacio vectorial $L(\mathfrak{A})$, de modo que los vectores coordenados nuevos $\overline{O'A'_i}$ pueden ser expresados mediante los antiguos por fórmulas de tipo

$$\begin{gathered}\overline{OA'_i} = \tau_i^1 \cdot \overline{OA_1} + \ldots + \tau_i^n \cdot \overline{OA_n} \\ (i = 1, \ldots, n),\end{gathered} \tag{6}$$

donde el determinante de la matriz $T = \| \tau_j^i \|$ es diferente de cero (p. 8.3).

Observando que $O = O'$, introducimos en la fórmula (4) las expresiones (6) en lugar de los vectores $\overline{OA'_i}$. Comparando el resultado de la sustitución con la igualdad (2), llegamos a las relaciones

$$\xi^i = \xi^{1'} \tau_1^i + \ldots + \xi^{n'} \tau_n^i \quad (i = 1, \ldots, n)$$

que pueden ser representadas brevemente en la forma matricial conocida

$$(\xi^1, \ldots, \xi^n) = (\xi^{1'}, \ldots, \xi^{n'}) \cdot T. \tag{7}$$

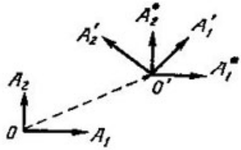

Fig. 9.

Finalmente el paso de una base de referencia $\boldsymbol{R} = (O, \overline{OA_1}, \ldots, \overline{OA_n})$ a otra base cualquiera de referencia $\boldsymbol{R}' = (O', \overline{O'A'_1}, \ldots, \overline{O'A'_n})$ puede ser realizado en dos pasos (véase la fig. 9 para el caso en el que $n = 2$): 1) trasladamos la base de referencia inicial en el vector $\overline{OO'}$, obteniendo así una base de referencia $\boldsymbol{R}^*$ de origen en el punto O' y 2) pasamos de la base de referencia $\boldsymbol{R}^*$ a la base de referencia $\boldsymbol{R}'$ sin cambiar el origen de coordenadas. Empleando sucesivamente las fórmulas (5) y (7), obtenemos las relaciones definitivas

$$(\xi^1, \ldots, \xi^n) = (\xi^{1'}, \ldots, \xi^{n'}) \cdot T + (\alpha^1, \ldots, \alpha^n), \tag{8}$$

donde $T = \| \tau_j^i \|$ es la matriz del cambio de la base $\overline{OA_1}$, $\ldots$, $\overline{OA_n}$ por la base $\overline{O'A'_1}$, $\ldots$, $\overline{O'A'_n}$ (que está formada por los coeficientes de las expresiones lineales de los vectores $\overline{O'A'_i}$ en términos de los vectores $\overline{OA_1}$, $\ldots$, $\overline{OA_n}$), $(\alpha^1, \ldots, \alpha^n)$ son las coordenadas del origen nuevo y $(\xi_1, \ldots, \xi^n)$ y $(\xi^{1'}, \ldots, \xi^{n'})$ son, respectivamente, las coordenadas antiguas y nuevas de un punto arbitrario X.

Por razones puramente formales la fila $(1, \xi^1, \ldots, \xi^n)$ ha sido llamada en el p. 30.1 fila coordenada ampliada del punto X siendo $(\xi^1, \ldots, \xi^n)$ la fila corriente de sus coordenadas. De la fórmula (8) se desprende directamente que

$$(1, \xi^1, \ldots, \xi^n) = (1, \xi^{1'}, \ldots, \xi^{n'}) \cdot T_{R'R}, \tag{9}$$

donde

$$T_{R'R} = \begin{vmatrix} 1 & \alpha^1 & \ldots & \alpha^n \\ 0 & \tau_1^1 & \ldots & \tau_n^1 \\ . & . & . & . \\ 0 & \tau_1^n & \ldots & \tau_n \end{vmatrix}. \tag{10}$$

La matriz $T_{R'R}$ se llama *matriz del cambio de la base de referencia* $\boldsymbol{R}$ *por la base de referencia* $\boldsymbol{R}'$. De (10) resulta que el determinante de la matriz $T_{R'R}$ coincide con el determinante de la matriz $\|\tau_j^{i}\|$ que es la matriz del cambio de la base coordenada antigua del espacio vectorial $L(\mathfrak{A})$ por la base nueva. Recíprocamente, cualquiera que sea la matriz T_0 de tipo (10) con el determinante diferente de cero y cualquiera que sea la base de referencia $\boldsymbol{R}$ dada de antemano, las fórmulas (9) y los números $\alpha^1, \ldots, \alpha^n$ permiten determinar unívocamente una base de referencia $\boldsymbol{R}'$ tal que la matriz T_0 sea la matriz del cambio de $\boldsymbol{R}$ por $\boldsymbol{R}'$.

De la fórmula (9) se desprende el siguiente corolario importante.

COROLARIO. *Sean* $\boldsymbol{R}$, $\boldsymbol{R}'$ *y* $\boldsymbol{R}''$ *tres bases arbitrarias de referencia, sea* $T_{R'R}$ *la matriz del cambio de* $\boldsymbol{R}$ *por* $\boldsymbol{R}'$ *y sea* $T_{R''R'}$ *la matriz del cambio de* $\boldsymbol{R}'$ *por* $\boldsymbol{R}''$. *Entonces la matriz del cambio de* $\boldsymbol{R}$ *por* $\boldsymbol{R}''$ *será igual al producto* $T_{R''R'}T_{R'R}$. *En particular, si* $T_{R'R}$ *es la matriz del cambio de la base de referencia* $\boldsymbol{R}$ *por la base de referencia* $\boldsymbol{R}'$, *la matriz del cambio de* $\boldsymbol{R}'$ *por* $\boldsymbol{R}$ *es la matriz inversa* $T^{-1}_{R'R}$.

En efecto, indicando por $(\xi^1, \ldots, \xi^n)$, $(\xi^{1'}, \ldots, \xi^{n'})$ y $(\xi^{1''}, \ldots, \xi^{n''})$ las coordenadas de un punto arbitrario X en las bases de referencia $\boldsymbol{R}$, $\boldsymbol{R}'$ y $\boldsymbol{R}''$, respectivamente, tenemos, según (9),

$$(1, \xi^1, \ldots, \xi^n) = (1, \xi^{1''}, \ldots, \xi^{n''}) \cdot T_{R''R}$$

y, al mismo tiempo, se tiene

$$(1, \xi^1, \ldots, \xi^n) = (1, \xi^{1'}, \ldots, \xi^{n'}) \cdot T_{R'R} = \\ = (1, \xi^{1''}, \ldots, \xi^{n''}) \cdot T_{R''R'}T_{R'R}.$$

Por consiguiente, para cualesquiera $\xi^{1''}, \ldots, \xi^{n''} \in K$ resulta

$$(1, \xi^{1''}, \ldots, \xi^{n''}) \cdot T_{R''R} = (1, \xi^{1''}, \ldots, \xi^{n''}) \cdot T_{R''R'}T_{R'R}.$$

Tomando aquí para $\xi^{1''}, \ldots, \xi^{n}$ las coordenadas en la base de referencia $\boldsymbol{R}''$ del vértice de esta base de referencia, obtenemos

$$T_{R''R} = T_{R''R'}T_{R'R}. \tag{11}$$

Si la base de referencia $\boldsymbol{R}''$ coincide con la base de referencia $\boldsymbol{R}'$, la matriz $T_{R''R}$ es, obviamente, la matriz unidad y la relación (11) se convierte en

$$E = T_{RR'}T_{R'R},$$

de donde resulta

$$T_{R'R} = T_{RR'}^{-1}. \tag{12}$$

Hasta el momento nos hemos interesado sólo por la ley de variación de las coordenadas de los puntos. Preguntémonos ¿cómo varían las ecuaciones de los conjuntos de puntos cuando se pasa de una base de referencia coordenada a otra?

Sea dada una ecuación

$$f(\xi^1, \ldots, \xi^n) = 0 \tag{13}$$

de un conjunto $\mathfrak{M}$ en una base de referencia $\boldsymbol{R}$. Esto significa, por definición, que un punto arbitrario X pertenece a $\mathfrak{M}$ cuando, y sólo cuando, sus coordenadas $\xi^1, \ldots, \xi^n$ satisfacen la relación (13). Al pasar a una base de referencia coordenada nueva $\boldsymbol{R}'$, tenemos según (8)

$$\xi^i = \sum_{j=1}^{n} \xi^{j'}\tau_j^i + \alpha^i \quad (i = 1, \ldots, n).$$

Introduciendo en (13) estas expresiones para $\xi^1, \ldots, \xi^n$, obtenemos

$$f(\sum \xi^{j'}\tau_j^1 + \alpha^1, \ldots, \sum \xi^{j'}\tau_j^n + \alpha^n) = 0. \tag{14}$$

Puesto que las coordenadas nuevas $\xi^{1'}, \ldots, \xi^{n'}$ de un punto arbitrario X satisfacen la relación (14) cuando, y sólo cuando, $X \in \mathfrak{M}$, resulta que (14) es la ecuación que buscamos del conjunto $\mathfrak{M}$ en la base de referencia coordenada nueva. El primer miembro de la igualdad (14) es una función $g(\xi^{1'}, \ldots, \xi^{n'})$ de las variables $\xi^{1'}, \ldots, \xi^{n'}$ que se distingue, en general, por su forma de la función inicial f. Sin embargo, si f es un polinomio de un grado s en las variables $\xi^1, \ldots, \xi^n$, se ve de (14) que g también será un polinomio de grado s en las variables $\xi^{1'}, \ldots, \xi^{n'}$.

Igual que en el caso de los espacios vectoriales (véase el p. 5.1), el problema sobre la transformación de las coordenadas afines está ligado estrechamente al problema de la determinación de todos los automorfismos de un espacio afín $\mathfrak{A}$ sobre un cuerpo conmutativo fijo K. Sea n la dimensión del espacio $\mathfrak{A}$ y sea $\mathcal{A}$ un automorfismo de $\mathfrak{A}$ sobre K (véase el p. 29.1). Tomemos en $\mathfrak{A}$ una base de referencia coordenada $\boldsymbol{R}=(O, \overline{OA_1}, \ldots, \overline{OA_n})$. Sus vértices $O, A_1, \ldots, A_n$ forman en $\mathfrak{A}$ un sistema linealmente independiente de puntos. Com la relación de independencia lineal se conserva en los automorfismos, los puntos $O\mathcal{A}, A_1\mathcal{A}, \ldots, A_n\mathcal{A}$ pueden ser considerados como los vértices de una base de referencia coordenada nueva $\boldsymbol{R}\mathcal{A}=(O\mathcal{A}, \overline{O\mathcal{A}A_1\mathcal{A}}, \ldots, \overline{O\mathcal{A}A_n\mathcal{A}})$. Indiquemos por $(X)_R^1, \ldots, (X)_R^n$ las coordenadas de un punto arbitrario X calculadas en la base de referencia $\boldsymbol{R}$ y sea

$$[X]_R=(1, (X)_R^1, \ldots, (X)_R^n)$$

la fila ampliada de coordenadas del punto X en la base de referencia $\boldsymbol{R}$. La fórmula (2) toma entonces la forma siguiente:

$$\overline{OX}=(X)_R^1\cdot\overline{OA_1}+\ldots+(X)_R^n\cdot\overline{OA_n}. \qquad (15)$$

Como los isomorfismos conservan las relaciones de dependencia lineal entre vectores, obtenemos de (15)

$$\overline{O\mathcal{A}X\mathcal{A}}=(X)_R^1\cdot\overline{O\mathcal{A}A_1\mathcal{A}}+\ldots+(X)_R^n\cdot\overline{O\mathcal{A}A_n\mathcal{A}}. \qquad (16)$$

En otras palabras, las coordenadas de $X\mathcal{A}$ en la base de referencia $\boldsymbol{R}\mathcal{A}$ coinciden con las coordenadas de X en la base de referencia $\boldsymbol{R}$, es decir, en notación abreviada

$$[X\mathcal{A}]_{R\mathcal{A}}=[X]_R.$$

Aplicando la fórmula (9), obtenemos

$$[X]_R=[X\mathcal{A}]_{R\mathcal{A}}=[X\mathcal{A}]\cdot T_{R\mathcal{A}R}$$

y por esto, debido a (12), resulta

$$[X\mathcal{A}]_R=[X]\,T_{RR\mathcal{A}}. \qquad (17)$$

La matriz $T_{RR\mathcal{A}}$ se llama *matriz del automorfismo* $\mathcal{A}$ en la base de referencia $\boldsymbol{R}$ y se indica por $[\mathcal{A}]_R$ o simplemente por $[\mathcal{A}]$, si la base de referencia $\boldsymbol{R}$ se conoce de antemano. Por esta razón la relación (17) puede ser representada en la forma definitiva

$$[X\mathcal{A}]=[X]\,[\mathcal{A}]. \qquad (18)$$

Sean $\mathcal{A}$ y $\mathcal{B}$ dos automorfismos arbitrarios dados del espacio $\mathfrak{A}$. Aplicando la fórmula (18), obtenemos

$$[X(\mathcal{A}\mathcal{B})]=[X]\,[\mathcal{A}\mathcal{B}],$$
$$[(X\mathcal{A})\mathcal{B})]=[X\mathcal{A}][\mathcal{B}]=[X]\,[\mathcal{A}]\,[\mathcal{B}],$$

de donde resulta

$$[X]\,[\mathcal{A}\mathcal{B}]=[X]([\mathcal{A}]\,[\mathcal{B}]).$$

Puesto que esta igualdad debe cumplirse para cualquier punto X, tenemos

$$[\mathcal{A}\mathcal{B}] = [\mathcal{A}]\,[\mathcal{B}]$$

y, por consiguiente,

$$[\mathcal{A}^{-1}] = [\mathcal{A}]^{-1}.$$

Consideremos un ejemplo. Sea $\mathcal{A}$ una traslación de $\mathfrak{A}$ de vector $\overline{OO'}$ de coordenadas $\alpha^1, \ldots, \alpha^n$ en la base de referencia $R = (O, \overline{OA_1}, \ldots, \overline{OA_n})$. De las fórmulas (9) y (5) obtenemos entonces

$$[\mathcal{A}] = T_{RR\mathcal{A}} = \begin{bmatrix} 1 & -\alpha^1 & -\alpha^2 & \ldots & -\alpha^n \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}.$$

En cambio, si el automorfismo $\mathcal{A}$ deja inmóvil el origen de la base de referencia coordenada $\boldsymbol{R}$, la matriz $[\mathcal{A}]$ se descompone en 1 y en la matriz $\|\tau_j^i\|$ del cambio en el espacio vectorial $L(\mathfrak{A})$.

Ejemplos y problemas

1. Hállense las condiciones necesarias y suficientes de intersección de dos rectas dadas por ecuaciones en coeficientes directores.
2. Hállense las ecuaciones paramétricas del plano dado por el sistema de ecuaciones

$$\left.\begin{aligned} x_1 + x_2 - 2x_3 + 3x_4 &= 1, \\ x_1 + 2x_2 - x_3 + 2x_4 &= 3, \\ x_1 - x_2 - 4x_3 - 5x_4 &= -3. \end{aligned}\right\}$$

3. Demuéstrese que todo plano $\mathfrak{M}$ de un espacio afín es el mismo un espacio afín cuya dimensión es igual a la dimensión de $\mathfrak{M}$.
4. Demuéstrese que un plano $\mathfrak{M}$ de un espacio afín diferente de un punto es paralelo a cualquier plano que no corte el primero cuando, y sólo cuando, $\mathfrak{M}$ es un hiperplano.

§ 31. Cuerpos convexos

Hasta el momento no hemos considerado conceptos geométricos tan esenciales como son los conceptos de la relación "hallarse entre", de semiplano y de convexidad. La razón de esto estriba en que en los espacios afines sobre un cuerpo conmutativo arbitrario estos conceptos no pueden ser introducidos de un modo natural. Para tratarlos es necesario limitar la clase de cuerpos conmutativos y pasar al estudio de los espacios afines sobre cuerpos conmutativos ordenados, por ejemplo, sobre el cuerpo de los números reales. Las propiedades principales de estos espacios constituyen precisamente el tema que se estudia en este parágrafo.

31.1. Rayos. Recordemos que un cuerpo conmutativo K se llama *ordenado*, si para los elementos de K, además de las operaciones de adición y de multiplicación, se introduce también una relación de orden $\leqslant$ sometida a dos condiciones:

$$\alpha \leqslant \beta \Rightarrow \alpha+\gamma \leqslant \beta+\gamma, \tag{1}$$
$$0 \leqslant \alpha \ \text{ y } \ 0 \leqslant \beta \Rightarrow 0 \leqslant \alpha\beta \tag{2}$$
$$(\alpha,\ \beta,\ \gamma \in K).$$

El ejemplo más importante de un cuerpo conmutativo ordenado es el cuerpo de los números reales que es el que debe tenerse en primer orden en cuenta en todos los razonamientos ulteriores de este parágrafo y del siguiente.

No está de más observar que de (1) y (2) se desprenden directamente las siguientes propiedades:

Para cualquier elemento α de un cuerpo conmutativo ordenado se tiene

$$0 \leqslant \alpha^2 \ \text{ y } \ -\alpha^2 \leqslant 0 \tag{3}$$

y, en particular, $0 \leqslant 1$ y $-1 \leqslant 0$.

Para unos elementos arbitrarios α, β y γ de un cuerpo conmutativo ordenado se tiene

$$\alpha \leqslant \beta \Rightarrow -\beta \leqslant -\alpha, \tag{4}$$
$$0 \leqslant \alpha \ \text{ y } \ \beta \leqslant \gamma \Rightarrow \alpha\beta \leqslant \alpha\gamma. \tag{5}$$

Se toma por definición que la relación $\alpha < \beta$ equivale a la conjunción $\alpha \leqslant \beta$ y $\alpha \neq \beta$ y que las relaciones $\alpha \geqslant \beta$ y $\alpha > \beta$ significan lo mismo que las relaciones $\beta \leqslant \alpha$ y $\beta < \alpha$. El elemento no negativo de los elementos α y $-\alpha$ se llama *valor absoluto de α* y se indica por $|\alpha|$. De las relaciones de (1) a (5) se deduce fácilmente que

$$|-\alpha| = |\alpha|, \quad |\alpha+\beta| \leqslant |\alpha|+|\beta| \ \text{ y } \ |\alpha\beta| = |\alpha|\cdot|\beta|.$$

Consideremos ahora un espacio afín $\mathfrak{A}$ sobre un cuerpo conmutativo *ordenado* K. Se dice que en el espacio afín $\mathfrak{A}$ el punto X *se halla entre* los puntos A y B, si

$$\overline{AX} = \lambda\cdot\overline{AB} \ \text{ y } \ 0 \leqslant \lambda \leqslant 1 \quad (\lambda \in K). \tag{6}$$

Está claro que si X se halla entre A y B, los puntos A, X y B se hallan sobre una misma recta. Además, la relación «hallarse entre» es simétrica: *si X se halla entre A y B, X se halla entre B y A*. Efectivamente, de (6) obtenemos

$$\overline{AB} = \overline{AX} + \overline{XB} = \lambda\cdot\overline{AB} + \overline{XB},$$

y por esto

$$\overline{BX} = (1-\lambda)\cdot\overline{BA}, \quad 0 \leqslant 1-\lambda \leqslant 1.$$

Sean A, B y C unos puntos, diferentes dos a dos, que se hallan sobre una misma recta. Para un valor adecuado $\lambda \in K$ tenemos entonces $\overline{AC} = \lambda \cdot \overline{AB}$. El número λ satisface a una de las condiciones, y sólo a una de ellas: a) $\lambda < 0$; b) $0 \leqslant \lambda \leqslant 1$ y c) $\lambda > 1$. Es fácil comprobar que en el caso a) el punto A se halla entre B y C, que en el caso c) el punto B se halla entre A y C y que en el caso b) el punto C se halla, por definición, entre A y B.

Por consiguiente, *de cualesquiera tres puntos que se encuentran sobre una misma recta uno, y sólo uno, se halla entre los otros dos.*

Análogamente se comprueba también la segunda propiedad principal de la relación «hallarse entre»: *si el punto X se halla entre los puntos A y B y el punto Y se halla entre A y X, el punto Y se halla entre A y B.*

Empleando el concepto «entre», podemos introducir en toda recta dos relaciones naturales de orden llamadas también direcciones de la recta. Tomemos sobre la recta dada $\mathfrak{M}$ dos puntos distintos cualesquiera A y B e introduzcamos para los puntos de $\mathfrak{M}$ una relación binaria $\leqslant_{AB}$ que depende de A y de B tomando, por definición, que $X \leqslant_{AB} Y$ es verídica si

$$\overline{AX} = \lambda \cdot \overline{AB}, \ \overline{AY} = \mu \cdot \overline{AB} \quad \text{y} \quad \lambda \leqslant \mu \ (\lambda, \mu \in K), \tag{7}$$

es decir, ordenando los puntos de la recta $\mathfrak{M}$ según el orden en el que se encuentran sus coordenadas en la base de referencia $(A, \overline{AB})$.

Está claro que los órdenes $\leqslant_{AB}$ y $\leqslant_{BA}$ *son duales* es decir, que

$$X \leqslant_{AB} Y \Longleftrightarrow Y \leqslant_{BA} X.$$

Por otra parte, el orden $\leqslant_{PQ}$ definido sobre $\mathfrak{M}$ mediante otro par cualquiera de puntos P, Q coincide con el orden $\leqslant_{AB}$, si $P \leqslant_{AB} Q$, y coincide con el orden $\leqslant_{BA}$, si $Q \leqslant_{AB} P$. Por consiguiente, *entre los órdenes de tipo $\leqslant_{PQ}$ sobre una recta $\mathfrak{M}$ existen solamente dos órdenes diferentes*. Estos órdenes se llaman *órdenes* o *direcciones naturales* de la recta. Los automorfismos del espacio $\mathfrak{A}$ sobre K conservan la relación «hallarse entre» y también conservan los órdenes naturales. Sin embargo, gracias a la existencia sobre una recta de dos órdenes naturales, los automorfismos del espacio $\mathfrak{A}$ pueden transformar un orden natural sobre la recta $\mathfrak{M}$ en el otro orden natural sobre la misma recta. Por esto, es invariante el par de relaciones duales de órdenes y no la relación de orden.

TEOREMA. *Si una funcional lineal $f(X)$ no es constante sobre una recta $\mathfrak{M}$, la aplicación $f(X) \to X$ $(X \in \mathfrak{M})$ es una aplicación biyectiva de K sobre $\mathfrak{M}$ en la que el orden $\leqslant$, definido sobre el cuerpo conmutativo K, se transforma en uno de los dos órdenes naturales definidos sobre $\mathfrak{M}$.*

Fijemos sobre la recta unos puntos O y A tales que $f(O) \neq f(A)$ y sea

$$\overline{OX} = \lambda \cdot \overline{OA} \quad (X \in \mathfrak{M} \quad \text{y} \quad \lambda \in K).$$

La aplicación $\lambda \to X$, es, debido a (7), una aplicación biyectiva de K sobre $\mathfrak{M}$ que transforma el orden dado sobre K en un orden natural sobre $\mathfrak{M}$. De la linealidad de f se deduce que

$$f(X)=\lambda(f(A)-f(O))+f(O)=\alpha\lambda+\beta.$$

Puesto que $\alpha\neq 0$, la aplicación $f(X)\to\lambda$ es una aplicación biyectiva de K sobre K que o bien conserva el orden o bien lo invierte. La aplicación $f(X)\to X$, por ser la composición de las aplicaciones $f(X)\to\lambda$ y $\lambda\to X$, también es biyectiva y transforma el orden sobre K en un orden natural sobre $\mathfrak{M}$.

Cualesquiera que sean dos puntos A y B de un espacio $\mathfrak{A}$ se llama *segmento* $[A, B]$ el conjunto de todos los puntos que se hallan entre A y B. Los puntos A y B se llaman *extremos del segmento* $[A, B]$. De la definición de la relación «hallarse entre» se deduce que los extremos del segmento pertenecen al segmento y de la simetría de esta relación se deduce que $[A, B]=[B, A]$. Está claro que el segmento $[A, A]$ está formado sólo por el punto A.

Se dice que un número λ se halla entre los números α y β ($\alpha, \beta, \lambda\in K$), si $\alpha\leqslant\lambda\leqslant\beta$ o $\beta\leqslant\lambda\leqslant\alpha$. El conjunto de los números que se hallan entre α y β se llama segmento numérico y se indica por $[\alpha, \beta]$. Del teorema anterior obtenemos el siguiente corolario:

COROLARIO. *El conjunto de los valores que toma una funcional lineal $f(X)$ sobre un segmento $[A, B]$ es un segmento numérico $[f(A), f(B)]$. El conjunto de los valores que toma f sobre la recta $\mathfrak{M}$ o bien coincide con todo el cuerpo conmutativo K (si f no es constante sobre $\mathfrak{M}$) o bien está formado por un solo número (si f es constante sobre $\mathfrak{M}$).*

Introduzcamos ahora el concepto de semirrecta o de rayo. Se dice que el punto X se encuentra al mismo lado del punto O que el punto A, si X se halla entre O y A o si A se halla entre O y X.

Un conjunto de puntos $\mathfrak{M}_0$ del espacio $\mathfrak{A}$ se llama *rayo*, si en $\mathfrak{M}_0$ existen unos puntos distintos O y A tales que $\mathfrak{M}_0$ está formado por todos los puntos X que se encuentran al mismo lado del punto O que el punto A. El punto O se llama vértice del rayo que será indicado ahora por $\mathfrak{M}_{OA}$.

De esta definición se deduce que todos los puntos del rayo $\mathfrak{M}_{OA}$ pertenecen a la recta $O\vee A$. Si tomamos como base de referencia coordenada sobre la recta $O\vee A$ la base de referencia $(O, \overline{OA})$, a todo $\lambda\in K$ le corresponderá un punto X de coordenada lineal λ tal que $\overline{OX}=\lambda\overline{OA}$. Está claro que el rayo $\mathfrak{M}_{OA}$ consta de todos los puntos de la recta $O\vee A$ que tienen coordenada no negativa. En particular, de aquí se ve que todo rayo tiene solamente un vértice O y que si sobre el rayo $\mathfrak{M}_{OA}$ se toma un punto cualquiera B diferente de O, los rayos $\mathfrak{M}_{OA}$ y $\mathfrak{M}_{OB}$ coinciden. Es fácil comprobar también que al tomar en una recta un punto arbitrario O, la recta se descompone exactamente en dos rayos diferentes de vértice O.

Supongamos que una funcional lineal $f(X)$ no es constante sobre una recta $\mathfrak{M}$. En virtud del teorema 1, sobre la recta $\mathfrak{M}$ existe un punto O, sólo uno, en el que la funcional f se anula. Puesto que la aplicación $f(X) \to X$ transforma el orden de K en un orden natural de $\mathfrak{M}$, los conjuntos $\mathfrak{M}_0$ y $\mathfrak{M}_1$ de puntos de $\mathfrak{M}$ con $\lambda \leqslant 0$ y $\lambda \geqslant 0$, respectivamente, son precisamente aquellos rayos en los que el punto O divide a la recta $\mathfrak{M}$.

31.2. Semiespacios. El análogo del concepto de rayo en el caso de planos multidimensionales de un espacio $\mathfrak{A}$ es el concepto de semiplano y, en particular, el concepto de semiespacio que pasamos ahora a definir.

Sea de nuevo $\mathfrak{A}$ un espacio afín sobre un cuerpo conmutativo ordenado K y sea $\mathfrak{P}$ un hiperplano arbitrario de $\mathfrak{A}$. Si $\mathfrak{P}$ contiene unos puntos X e Y, $\mathfrak{P}$ contiene también todos los puntos de la recta $X \vee Y$. Por esto, cualquier segmento $[A, B]$ de $\mathfrak{A}$ o bien está contenido íntegramente en $\mathfrak{P}$ o bien contiene no más de un punto de $\mathfrak{P}$. Se dice que los puntos A y B *se hallan a distintos lados* del hiperplano $\mathfrak{P}$, si ambos no pertenecen a $\mathfrak{P}$, pero el segmento $[A, B]$ contiene un punto de $\mathfrak{P}$. En todos los demás casos se dice que los puntos A y B *se hallan a un mismo lado* de $\mathfrak{P}$ lo que se expresa simbólicamente así: $A \equiv B(\mathfrak{P})$. En particular, $A \equiv A(\mathfrak{P})$ y de $A \equiv B(\mathfrak{P})$ se deduce que $B \equiv A(\mathfrak{P})$ para cualesquiera A y B. Además, si $A \in \mathfrak{P}$, se tiene $A \equiv B(\mathfrak{P})$ para cualquier punto B.

TEOREMA 1. *Sea $f(X)$ una funcional lineal no constante sobre el espacio $\mathfrak{A}$ y sea $\mathfrak{P}$ un hiperplano formado por los puntos X tales que $f(X) = 0$. Entonces se tiene para cualesquiera A y B*

$$A \equiv B(\mathfrak{P}) \Longleftrightarrow f(A) \cdot f(B) \geqslant 0. \tag{1}$$

Supongamos que A y B se hallan a distintos lados de $\mathfrak{P}$. El conjunto de los valores que toma $f(X)$ en los puntos del segmento $[A, B]$ es el segmento númerico $[f(A), f(B)]$. Tenemos, por hipótesis, $f(A) \neq 0$, $f(B) \neq 0$ y $0 \in [f(A), f(B)]$. Por consiguiente $f(A) \cdot f(B) < 0$. Recíprocamente, supongamos que para unos puntos A y B se tiene $f(A) \cdot f(B) < 0$. Entonces $0 \in [f(A), f(B)]$, es decir, para un punto $X \in [A, B]$ se tiene $f(X) = 0$, de donde $X \in \mathfrak{P}$ y por ello los puntos A y B se hallan a distintos lados de $\mathfrak{P}$.

COROLARIO. *Si el punto A no pertenece al hiperplano $\mathfrak{P}$ y $A \equiv B(\mathfrak{P})$ y $A \equiv C(\mathfrak{P})$, se tiene $B \equiv C(\mathfrak{P})$.*

Efectivamente, si los puntos A, B y C satisfacen las exigencias indicadas, tenemos según el teorema 1

$$f(A) \neq 0, \quad f(A) f(B) \geqslant 0 \quad \text{y} \quad f(A) f(C) \geqslant 0,$$

de donde $(f(A))^2 f(B) f(C) \geqslant 0$ y, por consiguiente, $f(B) f(C) \geqslant 0$.

Introduzcamos ahora el concepto principal: cualesquiera que sean un hiperplano $\mathfrak{P}$ y un punto arbitrario $A \notin \mathfrak{P}$, el conjunto de los puntos que se hallan a un mismo lado de $\mathfrak{P}$ que el punto A

se llama *semiespacio definido por el hiperplano* $\mathfrak{P}$ *y por el punto* A. Convengamos en indicar provisionalmente este semiespacio por $\mathfrak{A}_{\mathfrak{P}A}$.

Se ve de la definición que $\mathfrak{A}_{\mathfrak{P}A}$ contiene indudablemente el punto A y el hiperplano $\mathfrak{P}$. El conjunto que se obtiene de $\mathfrak{A}_{\mathfrak{P}A}$ omitiendo el hiperplano $\mathfrak{P}$ se llama *semiespacio abierto* definido por el hiperplano $\mathfrak{P}$ y por el punto A.

TEOREMA 2 *Si los puntos* A *y* B *no pertenecen al hiperplano* $\mathfrak{P}$ *y el segmento* $[A, B]$ *no contiene puntos de* $\mathfrak{P}$, *los semiespacios* $\mathfrak{A}_{\mathfrak{P}A}$ *y* $\mathfrak{A}_{\mathfrak{P}B}$ *coinciden. Si los puntos* A *y* C *se hallan a distintos lados de* $\mathfrak{P}$, *se tiene* $\mathfrak{A}_{\mathfrak{P}A} \neq \mathfrak{A}_{\mathfrak{P}C}$ *y cualquier semiespacio* $\mathfrak{A}_{\mathfrak{P}D}$ *coincide con* $\mathfrak{A}_{\mathfrak{P}A}$ *o con* $\mathfrak{A}_{\mathfrak{P}C}$.

Consideremos una funcional lineal no constante $f(X)$ que se anula sobre $\mathfrak{P}$. Puesto que el segmento $[A, B]$ no contiene puntos de $\mathfrak{P}$, los números $f(A)$ y $f(B)$ son de un mismo signo debido a la fórmula (1). Análogamente se comprueba que los números $f(A)$ y $f(C)$ tienen signos diferentes. Del teorema 1 resulta que para cualquier punto $D \notin \mathfrak{P}$ siendo $f(D) > 0$ el conjunto $\mathfrak{A}_{\mathfrak{P}D}$ coincide con la colección $\mathfrak{A}^+$ de los puntos X tales que

$$f(X) \geqslant 0 \tag{2}$$

y siendo $f(D) < 0$ el conjunto $\mathfrak{A}_{\mathfrak{P}D}$ coincide con la colección $\mathfrak{A}^-$ de los puntos X tales que

$$f(X) \leqslant 0. \tag{3}$$

Por consiguiente, un semiespacio arbitrario $\mathfrak{A}_{\mathfrak{P}D}$ coincide con $\mathfrak{A}^+$ o con $\mathfrak{A}^-$. Si $f(A) > 0$, de las relaciones mencionadas obtenemos

$$\mathfrak{A}_{\mathfrak{P}A} = \mathfrak{A}_{\mathfrak{P}B} = \mathfrak{A}^+ \quad \text{y} \quad \mathfrak{A}_{\mathfrak{P}C} = \mathfrak{A}^-;$$

en cambio, si $f(A) < 0$, se tiene

$$\mathfrak{A}_{\mathfrak{P}A} = \mathfrak{A}_{\mathfrak{P}B} = \mathfrak{A}^- \quad \text{y} \quad \mathfrak{A}_{\mathfrak{P}C} = \mathfrak{A}^+.$$

COROLARIO 1. *Para toda funcional lineal no constante* $f(X)$ *el conjunto de los puntos* X *que satisfacen la desigualdad*

$$f(X) \geqslant 0$$

y el conjunto de los puntos X *que satisfacen la desigualdad*

$$f(X) \leqslant 0$$

representan los dos semiespacios definidos por el hiperplano

$$f(X) = 0. \tag{4}$$

Los conjuntos de los puntos X *definidos por la desigualdad estricta*

$$f(X) > 0 \tag{5}$$

y, respectivamente, por la desigualdad estricta

$$f(X) < 0 \tag{6}$$

representan los semiespacios abiertos en los que el hiperplano (4) *divide al espacio afín dado* $\mathfrak{A}$.

Efectivamente, la ecuación (4) determina, según el p. 29.4, un hiperplano $\mathfrak{P}$ que define los semiespacios (2) y (3). Los conjuntos de los puntos definidos por las desigualdades estrictas (5) y (6) se obtienen de los semiespacios (2) y (3) omitiendo las soluciones de la ecuación (4) y son, por consiguiente, semiespacios abiertos.

COROLARIO 2. *Supongamos que en un espacio* $\mathfrak{A}$ *de dimensión finita* n *se ha escogido una base de referencia coordenada cualquiera. El conjunto de los puntos* X *cuyas coordenadas* $\xi^1, \ldots, \xi^n$ *satisfacen una desigualdad determinada*

$$\alpha_1\xi^1 + \ldots + \alpha_n\xi^n \geqslant \alpha_0$$

en la que al menos uno de los coeficientes principales $\alpha_1, \ldots, \alpha_n$ *es diferente de cero, representa entonces un semiespacio definido por el hiperplano*

$$\alpha_1\xi^1 + \ldots + \alpha_n\xi^n = \alpha_0. \tag{7}$$

El conjunto de los puntos X *cuyas coordenadas satisfacen la desigualdad estricta*

$$\alpha_1\xi^1 + \ldots + \alpha_n\xi^n > \alpha_0$$

representa un semiespacio abierto definido por el hiperplano (7).

La funcional definida mediante la fórmula

$$f(X) = \alpha_1\xi^1 + \ldots + \alpha_n\xi^n - \alpha_0$$

es, por lo visto en el p. 30.1, lineal y no constante sobre $\mathfrak{A}$. Aplicando a f el corolario 1, obtenemos las afirmaciones requeridas.

Hemos considerado hasta el momento los casos, en cierto sentido, extremos: rayos o semirrectas y semiespacios. No obstante, es fácil definir también el concepto de semiplano para un plano cualquiera $\mathfrak{M}$ contenido en el espacio $\mathfrak{A}$. En efecto, sabemos que $\mathfrak{M}$ puede ser considerado como un subespacio sobre el mismo cuerpo conmutativo K sobre el cual está definido el espacio $\mathfrak{A}$. Aplicando lo expuesto anteriormente al espacio afín $\mathfrak{M}$ sobre K, obtenemos el concepto de semiespacio del espacio $\mathfrak{M}$. Estos semiespacios se llaman *semiplanos* del plano $\mathfrak{M}$.

TEOREMA 3. *Supongamos que el plano* $\mathfrak{M}$ *y el hiperplano* $\mathfrak{P}$ *no son paralelos y sea* A *un punto de* $\mathfrak{M}$ *que no pertenece a* $\mathfrak{P}$. *La intersección del semiespacio* $\mathfrak{A}_{\mathfrak{P}A}$ *y del plano* $\mathfrak{M}$ *será entonces un semiplano en* $\mathfrak{M}$ *definido por el punto* A *y el hiperplano* $\mathfrak{M} \cap \mathfrak{P}$ *del espacio* $\mathfrak{M}$.

Todo semiplano de $\mathfrak{M}$ *es la intersección de* $\mathfrak{M}$ *con un semiespacio adecuado* $\mathfrak{A}_{\mathfrak{P}A}$ *del espacio* $\mathfrak{A}$.

Sea $f(X)$ una funcional lineal no constante que se anula sobre $\mathfrak{P}$ y sea $f(A) > 0$. El semiespacio $\mathfrak{A}_{\mathfrak{P}A}$ es entonces el conjunto de las soluciones de la desigualdad $f(X) \geqslant 0$ y la intersección $\mathfrak{M} \cap \mathfrak{A}_{\mathfrak{P}A}$

es el conjunto de los puntos $Y \in \mathfrak{M}$ que satisfacen la desigualdad $f(Y) \geqslant 0$. Puesto que la contracción de f sobre $\mathfrak{M}$ es una funcional lineal sobre $\mathfrak{M}$, resulta, según el corolario 1, que el conjunto $\mathfrak{M} \cap \mathfrak{A}_{\mathfrak{P}A}$ es un semiplano en $\mathfrak{M}$. Análogamente se demuestra la segunda afirmación del teorema 3.

Volviendo de nuevo a los semiespacios del espacio $\mathfrak{A}$ plantéemonos el problema: *¿de qué modo pueden situarse con respecto uno al otro dos semiespacios $\mathfrak{A}_{\mathfrak{P}A}$ y $\mathfrak{A}_{\mathfrak{Q}B}$?*

Consideremos por separado cada uno de los tres casos que pueden darse aquí:

1) Los hiperplanos $\mathfrak{P}$ y $\mathfrak{Q}$ coinciden. Entonces los semiespacios $\mathfrak{A}_{\mathfrak{P}A}$ y $\mathfrak{A}_{\mathfrak{Q}B}$ o bien coinciden o bien la unión de ellos es todo el espacio $\mathfrak{A}$ y la intersección es el hiperplano $\mathfrak{P}$.

2) Los hiperplanos $\mathfrak{P}$ y $\mathfrak{Q}$ no se cortan y, por consiguiente, son paralelos. Aquí pueden darse los siguientes subcasos:

a) $\mathfrak{P} \notin \mathfrak{A}_{\mathfrak{Q}B}$ y $\mathfrak{Q} \notin \mathfrak{A}_{\mathfrak{P}A}$; entonces $\mathfrak{A}_{\mathfrak{P}A} \cap \mathfrak{A}_{\mathfrak{Q}B} = \varnothing$.

b) $\mathfrak{P} \in \mathfrak{A}_{\mathfrak{Q}B}$ y $\mathfrak{Q} \notin \mathfrak{A}_{\mathfrak{P}A}$; entonces $\mathfrak{A}_{\mathfrak{Q}B} \supset \mathfrak{A}_{\mathfrak{P}A}$.

c) $\mathfrak{P} \notin \mathfrak{A}_{\mathfrak{Q}B}$ y $\mathfrak{Q} \in \mathfrak{A}_{\mathfrak{P}A}$; entonces $\mathfrak{A}_{\mathfrak{P}A} \supset \mathfrak{A}_{\mathfrak{Q}B}$.

d) $\mathfrak{P} \in \mathfrak{A}_{\mathfrak{Q}B}$ y $\mathfrak{Q} \in \mathfrak{A}_{\mathfrak{P}A}$; entonces $\mathfrak{A}_{\mathfrak{P}A} \cap \mathfrak{A}_{\mathfrak{Q}B} \neq \varnothing$.

Los conjuntos del último tipo se llaman a veces *hipercapas*. La intersección de cualquier recta con una hipercapa es, como puede verse fácilmente, vacía o coincide con toda la recta o es un segmento de la recta.

3) Los hiperplanos $\mathfrak{P}$ y $\mathfrak{Q}$ se cortan. El conjunto $\mathfrak{P} \cap \mathfrak{Q}$ es, según el p. 29.2, un plano de codimensión 2. En este caso la intersección de los semiespacios $\mathfrak{A}_{\mathfrak{P}A} \cap \mathfrak{A}_{\mathfrak{Q}B}$ se llama *hiperángulo* de arista $\mathfrak{P} \cap \mathfrak{Q}$.

31.3. Conjuntos convexos. Un conjunto $\mathfrak{S}$ de puntos de un espacio afín $\mathfrak{A}$ sobre un cuerpo conmutativo ordenado K se llama *convexo*, si

$$X \in \mathfrak{S} \text{ e } Y \in \mathfrak{S} \Rightarrow [X, Y] \in \mathfrak{S}. \tag{1}$$

De aquí se deduce que los planos del espacio $\mathfrak{A}$, así como los segmentos, rayos, semiplanos y semiplanos abiertos de $\mathfrak{A}$ son conjuntos convexos.

De la definición (1) se ve directamente que *la intersección de cualquier familia de conjuntos convexos es un conjunto convexo.* En particular, la intersección de cualquier familia de semiespacios es un conjunto convexo.

La intersección de todos los conjuntos convexos que contienen un conjunto fijo de puntos $\mathfrak{M}$ se llama *adherencia convexa del conjunto* $\mathfrak{M}$ y se indica por $\operatorname{Conv}\mathfrak{M}$. De la observación hecha anteriormente se desprende que la adherencia convexa de un conjunto cualquiera $\mathfrak{M}$ es el menor conjunto convexo que contiene a $\mathfrak{M}$. Está claro que $\mathfrak{M}$ es convexo cuando, y sólo cuando, $\operatorname{Conv}\mathfrak{M} = \mathfrak{M}$.

También es evidente que de $\mathfrak{M} \subseteq \mathfrak{N}$ resulta $\operatorname{Conv}\mathfrak{M} \subseteq \operatorname{Conv}\mathfrak{N}$. Todo punto $X \in \mathfrak{A}$ constituye por sí mismo un conjunto convexo y, por consiguiente, $\operatorname{Conv} X = X$. Sin embargo, ya para dos puntos X e Y tenemos, como puede verse fácilmente, la fórmula

$$\operatorname{Conv}\{X,\ Y\} = [X,\ Y]. \tag{2}$$

El teorema que sigue ofrece la expresión para la adherencia convexa de un conjunto cualquiera.

TEOREMA 1. *Escojamos en el espacio $\mathfrak{A}$ un punto O. La adherencia convexa* $\operatorname{Conv}\mathfrak{M}$ *de un conjunto arbitrario de puntos $\mathfrak{M}$ es el conjunto de todos los puntos $X \in \mathfrak{A}$ tales que para cada uno de ellos existe en $\mathfrak{M}$ un sistema finito de puntos $M_1, \ldots, M_s$ ligados a X por la relación*

$$\overline{OX} = \lambda_1 \cdot \overline{OM_1} + \ldots + \lambda_s \cdot \overline{OM_s}, \tag{3}$$

donde $\lambda_1, \ldots, \lambda_s$ son unos números adecuados de K sometidos a las condiciones

$$\lambda_1 \geqslant 0,\ \ldots,\ \lambda_s \geqslant 0, \tag{4}$$

$$\lambda_1 + \ldots + \lambda_s = 1. \tag{5}$$

Indiquemos por $\mathfrak{M}_c$ el conjunto de los puntos X que cumplen las exigencias (3), (4) y (5) y demostremos que $\mathfrak{M}_c$ es convexo. Sean $A, B \in \mathfrak{M}_c$, de modo que para convenientes $M_i \in \mathfrak{M}$ y $\alpha_j, \beta_j \in K$ se tenga

$$\overline{OA} = \alpha_1 \cdot \overline{OM_1} + \ldots + \alpha_s \cdot \overline{OM_s},$$
$$\overline{OB} = \beta_1 \cdot \overline{OM_1} + \ldots + \beta_s \cdot \overline{OM_s},$$

donde

$$\alpha_i \geqslant 0,\ \beta_i \geqslant 0 \quad (i = 1,\ \ldots,\ s),$$
$$\alpha_1 + \ldots + \alpha_s = \beta_1 + \ldots + \beta_s = 1.$$

Para un punto cualquiera $X \in [A,\ B]$ tenemos

$$\overline{OX} = \lambda \cdot \overline{OA} + \mu \cdot \overline{OB} \quad (\lambda \geqslant 0,\ \mu \geqslant 0 \text{ y } \lambda + \mu = 1)$$

y, por consiguiente,

$$\overline{OX} = (\lambda\alpha_1 + \mu\beta_1) \cdot \overline{OM_1} + \ldots + (\lambda\alpha_s + \mu\beta_s) \cdot \overline{OM_s}.$$

Puesto que aquí

$$\lambda\alpha_i + \mu\beta_i \geqslant 0 \quad (i = 1,\ \ldots,\ s)$$

y además

$$\sum (\lambda\alpha_i + \mu\beta_i) = \lambda \sum \alpha_i + \mu \sum \beta_i = 1,$$

resulta que $X \in \mathfrak{M}_c$ y el conjunto $\mathfrak{M}_c$ es convexo.

Resta demostrar que todo conjunto convexo $\mathfrak{N}$ que contiene al conjunto $\mathfrak{M}$ contiene también todos los puntos X que satisfacen las condiciones (3), (4) y (5). En estas condiciones figura el número natural s y para $s=1$ la afirmacion que acabamos de enunciar es trivial. Apliquemos ahora la inducción según s. Supongamos que para un valor de s cualquier punto X que satisface las condiciones (3), (4) y (5) pertenece a $\mathfrak{N}$. Consideremos el punto Y que satisfaga las condiciones

$$\overline{OY}=\alpha_1\cdot\overline{OM_1}+\ldots+\alpha_s\cdot\overline{OM_s}+\alpha_{s+1}\cdot\overline{OM_{s+1}}\quad (M_i\in\mathfrak{M}),$$
$$\alpha_1+\ldots+\alpha_s+\alpha_{s+1}=1,\ \alpha_i\geqslant 0\quad (i=1,\ \ldots,\ s+1).$$

Siendo aquí $\alpha_s=0$ o $\alpha_{s+1}=0$, tenemos, por la hipótesis de inducción, $Y\in\mathfrak{N}$. Supongamos por esto que $\alpha_s\cdot\alpha_{s+1}\neq 0$, de modo que tomando $\mu=\alpha_1+\ldots+\alpha_s$ tendremos $\mu>0$. El punto X definido por la relación

$$\overline{OX}=\frac{\alpha_1}{\mu}\overline{OM_1}+\ldots+\frac{\alpha_s}{\mu}\overline{OM_s}$$

satisface las condiciones (3), (4) y (5) con s fijo y, por consiguiente, $X\in\mathfrak{N}$. Tenemos ahora

$$\overline{OY}=\mu\cdot\overline{OX}+\alpha_{s+1}\cdot\overline{OM_{s+1}}\quad (\mu+\alpha_{s+1}=1,\ \mu\geqslant 0\ \text{y}\ \alpha_{s+1}\geqslant 0),$$

de modo que $Y\in[X,\ M_{s+1}]$. Como $\mathfrak{N}$ es convexo y $X,\ M_{s+1}\in\mathfrak{N}$, tenemos $Y\in\mathfrak{N}$ que es lo que se quería demostrar.

En cada una de las condiciones (3) figura solamente un número finito de puntos del conjunto. Por ello, del teorema 1 resulta que *la adherencia convexa de un conjunto infinito de puntos es la unión de las adherencias convexas de todos los subconjuntos finitos del conjunto dado.*

Se dice que *la dimensión de un conjunto convexo $\mathfrak{S}$ es igual a r*, si $\mathfrak{S}$ está contenido en un plano r-dimensional y no está contenido en ningún plano de menor dimensión. Análogamente se define también la codimensión de un conjunto convexo.

COROLARIO. *Si el número máximo de puntos linealmente independientes del conjunto $\mathfrak{M}$ es igual a $r+1$, la dimensión de* $\operatorname{Conv}\mathfrak{M}$ *es igual a r.*

Supongamos, por ejemplo, que todos los puntos del conjunto $\mathfrak{M}$ dependen linealmente de los puntos linealmente independientes $A_0,\ A_1,\ \ldots,\ A_r$ de este conjunto. Entonces $\mathfrak{M}$ está contenido en el plano $\mathfrak{N}=A_0\vee A_1\vee\ldots\vee A_r$ de dimensión r y por ello $\operatorname{Conv}\mathfrak{M}\subseteq\operatorname{Conv}\mathfrak{N}$. Pero un plano es un conjunto convexo y, por consiguiente, $\operatorname{Conv}\mathfrak{M}\subseteq\mathfrak{N}$.

Un conjunto convexo $\mathfrak{S}$ se llama *propio* en el espacio $\mathfrak{A}$, si la codimensión de $\mathfrak{S}$ es igual a cero. Los demás conjuntos convexos se llaman *impropios* en $\mathfrak{A}$. Como ejemplos de conjuntos convexos propios podemos indicar el mismo espacio $\mathfrak{A}$, así como los semiespacios y los semiespacios abiertos de éste. Los planos de codimensión no nula y sus semiplanos son conjuntos convexos impropios en $\mathfrak{A}$.

Está claro que todo conjunto convexo $\mathfrak{S}$ es un conjunto convexo propio en el plano que sirve de adherencia lineal de $\mathfrak{S}$.

Se dice que un punto S se halla *estrictamente dentro* del segmento $[A, B]$, si S es diferente de A y de B y $S \in [A, B]$. Un punto S se llama interior de un conjunto convexo $\mathfrak{S}$, si en toda recta que pasa por S existe un segmento que contiene a S estrictamente dentro de sí y pertenece íntegramente a $\mathfrak{S}$.

Se dice que un punto S *es tangente* a un conjunto convexo $\mathfrak{S}$, si existe una recta que pasa por S tal que cualquier segmento de ella que contenga al punto S estrictamente dentro de sí contenga al menos un punto de $\mathfrak{S}$. En particular, todos los puntos del conjunto $\mathfrak{S}$ son tangentes a $\mathfrak{S}$. Pero pueden también existir puntos tangentes a $\mathfrak{S}$ que no pertenezcan a $\mathfrak{S}$. Si todos los puntos tangentes a $\mathfrak{S}$ pertenecen a $\mathfrak{S}$, se dice que $\mathfrak{S}$ es un conjunto *cerrado*. Un conjunto convexo $\mathfrak{S}$ se llama *abierto*, si todo punto S suyo es un punto interior de $\mathfrak{S}$.

TEOREMA 2. *La intersección de una familia arbitraria de conjuntos convexos cerrados es un conjunto convexo cerrado. La intersección de una familia finita cualquiera de conjuntos convexos abiertos es un conjunto convexo abierto.*

La primera afirmación se deduce directamente de la definición de los conjuntos cerrados. Demostremos la segunda afirmación. Sean $\mathfrak{M}$ y $\mathfrak{N}$ unos conjuntos convexos abiertos. El conjunto $\mathfrak{M} \cap \mathfrak{N}$ es entonces convexo. Si es vacío, no hay nada que demostrar, ya que los conjuntos vacíos son, por definición, abiertos. Sea $S \in \mathfrak{M} \cap \mathfrak{N}$ y sea $\mathfrak{B}$ una recta cualquiera del espacio $\mathfrak{A}$ que pasa por el punto S. Sobre esta recta existen, por hipótesis, unos segmentos que contienen a S estrictamente dentro de sí y que pertenecen a los respectivos conjuntos $\mathfrak{M}$ y $\mathfrak{N}$. Pero la intersección de estos segmentos será entonces el segmento deseado que pertenece a $\mathfrak{M} \cap \mathfrak{N}$ y que contiene al punto S estrictamente dentro de sí.

Introduzcamos otro concepto importante. Un punto S se llama *punto frontera* de un conjunto convexo $\mathfrak{S}$, si S es tangente a $\mathfrak{S}$, pero no es interior de $\mathfrak{S}$. En otras palabras, se dice que S es un punto frontera de $\mathfrak{S}$, si existe una recta que pasa por S tal que todo segmento de la misma que contenga al punto S estrictamente dentro de sí contenga al menos un punto de $\mathfrak{S}$ y un punto que no sea de $\mathfrak{S}$. El conjunto de todos los puntos frontera de un conjunto convexo $\mathfrak{S}$ se llama *frontera* del conjunto $\mathfrak{S}$.

TEOREMA 3. *Un semiespacio arbitrario $\mathfrak{A}_{\mathfrak{P}A}$, definido en el espacio $\mathfrak{A}$ por un hiperplano $\mathfrak{P}$ y por un punto A que no se halla sobre $\mathfrak{P}$, así como el semiespacio abierto correspondiente $\mathfrak{A}^0_{\mathfrak{P}A} = \mathfrak{A}_{\mathfrak{P}A} \setminus \mathfrak{P}$, son unos conjuntos convexos cerrado y abierto, respectivamente, en $\mathfrak{A}$. La frontera de ambos conjuntos es el hiperplano $\mathfrak{P}$. Todo conjunto convexo $\mathfrak{S}$ impropio en $\mathfrak{A}$ no contiene puntos interiores.*

Sea $P \in \mathfrak{P}$. Consideremos la recta $A \vee P$ y en ella un segmento arbitrario $[B, C]$ que contiene al punto P estrictamente dentro de sí. Puesto que la recta $A \vee P$ no se halla sobre $\mathfrak{P}$, los extremos del segmento $[B, C]$ tampoco pertenecen a $\mathfrak{P}$ y, como $P \in [B, C]$, los puntos B y C se hallan a distintos lados del hiperplano $\mathfrak{P}$. Por consiguiente, uno de los extremos B o C pertenece al semiespacio abierto $\mathfrak{A}^0_{\mathfrak{P}A}$ y el otro extremo no pertenece a $\mathfrak{A}_{\mathfrak{P}A}$. Luego, el punto P es un punto frontera tanto de $\mathfrak{A}_{\mathfrak{P}A}$ como de $\mathfrak{A}^0_{\mathfrak{P}A}$.

Demostremos que cualquier punto $D \notin \mathfrak{A}_{\mathfrak{P}A}$ no es tangente a $\mathfrak{A}_{\mathfrak{P}A}$ y, por consiguiente, tampoco es tangente a $\mathfrak{A}^0_{\mathfrak{P}A}$. Sea $\mathfrak{M}$ una recta cualquiera que pasa por el punto D. Si esta recta no se corta con $\mathfrak{P}$, pertenece íntegramente al semiespacio $\mathfrak{A}^0_{\mathfrak{P}D} = \mathfrak{A}_{\mathfrak{P}D} / \mathfrak{P}$ y, por ello, todo segmento de la misma no contiene puntos de $\mathfrak{A}_{\mathfrak{P}A}$. Si $\mathfrak{M}$ se corta con $\mathfrak{P}$ en un punto P, existen en $\mathfrak{M}$ unos puntos B y C diferentes de D y P tales que $D \in [B, P]$ y $C \in [D, P]$. Por consiguiente, el segmento $[B, C]$ se halla sobre $\mathfrak{M}$, contiene a D estrictamente dentro de sí y no contiene puntos del semiespacio $\mathfrak{A}_{\mathfrak{P}A}$. Por esto, el punto D no es tangente a $\mathfrak{A}_{\mathfrak{P}A}$. Al mismo tiempo hemos demostrado que todos los puntos del semiespacio abierto $\mathfrak{A}^0_{\mathfrak{P}A}$ son interiores y, por consiguiente, cualquier semiespacio abierto es un conjunto convexo abierto.

Para demostrar la última afirmación del teorema es suficiente indicar por $\mathfrak{P}$ el hiperplano que contiene, por hipótesis, al conjunto $\mathfrak{S}$ y observar que todo segmento de la recta $A \vee P$ ($P \in \mathfrak{S}$ y $A \notin \mathfrak{P}$) que contiene estrictamente dentro de sí al punto P contiene también necesariamente puntos que no pertenecen al hiperplano $\mathfrak{P}$.

De los teoremas 2 y 3 obtenemos un corolario importante.

COROLARIO 1. *La intersección de cualquier familia de semiespacios de un espacio $\mathfrak{A}$ es un conjunto convexo cerrado de $\mathfrak{A}$. La intersección de cualquier familia finita de semiespacios abiertos del espacio $\mathfrak{A}$ es un conjunto convexo abierto de $\mathfrak{A}$.*

Recordando que cualquiera que sea la funcional lineal no constante $f(X)$ el conjunto de soluciones de la desigualdad $f(X) \geqslant 0$ es un semiespacio y el conjunto de soluciones de la desigualdad estricta $f(X) > 0$ es un semiespacio abierto, obtenemos otro corolario.

COROLARIO 2. *Sean $f_1(X), \ldots, f_s(X)$ unas funcionales lineales no constantes sobre un espacio afín $\mathfrak{A}$. Entonces el conjunto de las soluciones X del sistema de desigualdades*

$$f_i(X) \geqslant 0 \quad (i = 1, \ldots, s)$$

es un conjunto convexo cerrado de $\mathfrak{A}$ y el conjunto de las soluciones del sistema de desigualdades estrictas

$$f_i(X) > 0 \quad (i = 1, \ldots, s)$$

es un conjunto convexo abierto de $\mathfrak{A}$.

Un hiperplano $\mathfrak{P}$ se llama hiperplano *soporte* de un conjunto convexo $\mathfrak{S}$, si $\mathfrak{P} \cap \mathfrak{S} \neq \varnothing$ y todos los puntos de $\mathfrak{S}$ se hallan a un mismo lado de $\mathfrak{P}$. Las intersecciones de los hiperplanos soportes con el conjunto $\mathfrak{S}$ se llaman *facetas* de $\mathfrak{S}$. De estas definiciones se desprende directamente que las facetas de los conjuntos convexos pertenecen a las fronteras de estos conjuntos y son ellas mismas conjuntos convexos. Las facetas de los conjuntos convexos cerrados son conjuntos cerrados.

Complementos y ejemplos

Sea $\mathfrak{A}$ un espacio afín arbitrario de dimensión finita $n \geqslant 1$ sobre un cuerpo conmutativo ordenado K.

1. La adherencia convexa de unos puntos $A_1, \ldots, A_{r+1}$ linealmente independientes se llama *símplice r-dimensional* de $\mathfrak{A}$. Si $r < n$, este símplice se halla en el plano r-dimensional $A_1 \vee \ldots \vee A_{r+1}$ y por lo tanto es un conjunto convexo *impropio* en $\mathfrak{A}$. Los símplices n-dimensionales se llaman simplemente *símplices* de $\mathfrak{A}$. Los puntos $A_1, \ldots, A_{n+1}$ se llaman vértices del símplice $\operatorname{Conv}\{A_1, \ldots, A_{n+1}\}$.

2. Descríbanse los símplices de los espacios unidimensionales, bidimensionales y tridimensionales.

3. Para todo r, $0 \leqslant r \leqslant n-1$, todas las facetas r-dimensionales de un símplice $A = \operatorname{Conv}\{A_1, \ldots, A_{n+1}\}$ son unos símplices $\operatorname{Conv}\{A_{i_1}, \ldots, A_{i_{r+1}}\}$ $(1 \leqslant i_1 < \ldots < i_{r+1} \leqslant n+1)$.

4. Todo conjunto convexo $\mathfrak{S}$ compuesto sólo de puntos frontera de un símplice A pertenece íntegramente a un símplice $(n-1)$-dimensional $A^i = \operatorname{Conv}\{A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_{n+1}\}$.

5. Si los puntos $A_1, \ldots, A_{r+1}$ son linealmente independientes y para unos puntos $B_1, \ldots, B_{r+1}$ se tiene

$$\operatorname{Conv}\{A_1, \ldots, A_{r+1}\} = \operatorname{Conv}\{B_1, \ldots, B_{r+1}\},$$

los conjuntos $\{A_1, \ldots, A_{r+1}\}$ y $\{B_1, \ldots, B_{r+1}\}$ coinciden.

6. La adherencia convexa $C = \operatorname{Conv}\{C_1, \ldots, C_s\}$ de un sistema finito arbitrario de puntos $C_1, \ldots, C_s$ de $\mathfrak{A}$ se llama *poliedro convexo* de $\mathfrak{A}$.

Un sistema de puntos se llama *convexamente irreducible*, si ningún punto del sistema no está contenido en la adherencia convexa de los demás puntos.

Demuéstrese que siendo $C = \operatorname{Conv}\{C_1, \ldots, C_s\}$ un poliedro convexo, existen unos puntos $C_{i_1}, \ldots, C_{i_t}$, $1 \leqslant i_1 < \ldots < i_t \leqslant s$ tales que el sistema de los puntos $C_{i_1}, \ldots, C_{i_t}$ es convexamente irreducible y

$$\operatorname{Conv}\{C_1, \ldots, C_s\} = \operatorname{Conv}\{C_{i_1}, \ldots, C_{i_t}\}$$

7. Un símplice S se llama símplice *diagonal* de un poliedro convexo s-dimensional C generado por un sistema convexamente irreducible de puntos $C_1, \ldots, C_{s+1}$, si los vértices del símplice pertenecen al conjunto $\{C_1, \ldots, C_{s+1}\}$.

Se dice que un símplice r-dimensional A y un símplice s-dimensional B están en posición regular (uno con respecto al otro), si o no se cortan o se cortan por un conjunto que es una faceta tanto de un símplice como del otro.

Tiene lugar la siguiente proposición:

Existe un conjunto de símplices diagonales de un poliedro $\mathfrak{S}$ tal que todo punto del poliedro pertenece a la unión de estos símplices y cualesquiera dos símplices de este conjunto están en posición regular.

§ 32. Espacios euclídeos puntuales

En los parágrafos anteriores hemos estudiado las propiedades de los espacios afines definidos sobre un espacio vectorial dado $\mathfrak{L}$. Supongamos ahora que el espacio vectorial $\mathfrak{L}$ es unitario o euclídeo. Un espacio afín sobre un espacio vectorial unitario (respectivamente, euclídeo) $\mathfrak{L}$ se llama espacio puntual unitario (respectivamente, euclídeo) sobre $\mathfrak{L}$. Se llama dimensión de un espacio puntual unitario sobre $\mathfrak{L}$ la dimensión del correspondiente espacio vectorial unitario $\mathfrak{L}$. En lo que sigue dedicaremos la atención principal al estudio de los espacios puntuales euclídeos sobre el cuerpo de todos los números reales. Puesto que el cuerpo de los números reales es ordenado, en los espacios puntuales euclídeos está definido el concepto de convexidad (p. 31.3).

32.1. Longitud de una quebrada. Sea $\mathfrak{U}_n$ un espacio puntual unitario de n dimensiones sobre un espacio vectorial unitario $\mathfrak{L}$ definido sobre el cuerpo conmutativo K. A todo par de puntos A y B de $\mathfrak{U}_n$ corresponde un vector unívoco $\overline{AB}$ de $\mathfrak{L}$. Puesto que el espacio vectorial $\mathfrak{L}$ es unitario, en $\mathfrak{L}$ está definido el concepto de longitud del vector $\overline{AB}$. Indicaremos por $\rho(A, B)$ esta longitud llamándola *distancia* del punto A al punto B. Por consiguiente, tenemos por definición

$$\rho(A, B) = \|\overline{AB}\| = \sqrt{(\overline{AB}, \overline{AB})}. \qquad (1)$$

De las propiedades de las longitudes de los vectores (p. 17.2) se desprende que para cualesquiera puntos A y B se tiene

1° $\qquad \rho(A, B) = \rho(B, A).$

2° $\qquad \rho(A, B) = 0 \Leftrightarrow A = B.$

Como para cualesquiera puntos A, B y C de $\mathfrak{U}_n$ se tiene $\overline{AC} = \overline{AB} + \overline{BC}$, de la estimación para la longitud de una suma de vectores resulta

3° $\qquad \rho(A, B) + \rho(B, C) \geqslant \rho(A, C).$

Un conjunto arbitrario M tal que a todo par A, B de sus elementos corresponde un número real no negativo $\rho(A, B)$ que satisface las exigencias 1°, 2° y 3° se llama *espacio métrico* de métrica ρ. Por esto la definición (1) convierte un espacio puntual unitario en un espacio métrico.

Aplicando sucesivamente la desigualadad 3°, llegamos fácilmente a la desigualdad más general

$$\rho(A_1, A_2) + \rho(A_2, A_3) + \ldots + \rho(A_{s-1}, A_s) \geqslant \rho(A_1, A_s), \qquad (2)$$

válida para cualquier sucesión finita de puntos $A_1, \ldots, A_s$ de un espacio puntual unitario.

Si un espacio puntual $\mathfrak{U}_n$ es euclídeo, el campo principal K — que es el cuerpo de los números reales — es ordenado y, por ello, todo par de puntos A y B de $\mathfrak{U}_n$ determina un segmento $[A, B]$ (p. 31.1). El número $\rho(A, B)$ se llama *longitud* del segmento $[A, B]$. Dada una sucesión cualquiera de puntos $A_1, A_2, \ldots, A_s$, la sucesión de los segmentos

$$[A_1, A_2], [A_2, A_3], \ldots, [A_{s-1}, A_s]$$

se llama *quebrada* que une A_1 y A_s y el segmento $[A_1, A_s]$ se llama segmento *resultante* de la quebrada. La suma de las longitudes de los segmentos de una quebrada se llama *longitud de la quebrada.* La desigualdad (2) significa que en un espacio euclídeo cualquiera la longitud de una quebrada es no menor que la longitud de su segmento resultante.

¿Bajo qué condiciones en la desigualdad (2) tiene lugar el signo de igualdad? Según el p. 17.2 (para $s=2$), en un espacio vectorial unitario de las relaciones $\overline{A_1A_2} \neq o$ y

$$\|\overline{A_1A_2}\| + \|\overline{A_2A_3}\| + \ldots + \|\overline{A_{s-1}A_s}\| = \|\overline{A_1A_2} + \ldots + \overline{A_{s-1}A_s}\| \qquad (3)$$

se deduce que los vectores A_iA_{i+1} dependen linealmente de $\overline{A_1A_2}$ y por lo tanto

$$\overline{A_iA_{i+1}} = \lambda_i \cdot \overline{A_1A_2} \qquad (i = 1, \ldots, s-1).$$

Introduciendo estos valores en (3) y dividiendo por $\|\overline{A_1A_2}\|$, llegamos a la igualdad

$$|\lambda_1| + |\lambda_2| + \ldots + |\lambda_{s-1}| = |\lambda_1 + \lambda_2 + \ldots + \lambda_{s-1}|. \qquad (4)$$

Si el espacio es euclídeo, los números $\lambda_1, \ldots, \lambda_{s-1}$ son reales y la igualdad (4) es verídica cuando, y sólo cuando, estos números son de un mismo signo. Está claro que en este caso los puntos $A_1, A_2, \ldots, A_s$ se hallan sobre una misma recta y que los segmentos sucesivos $[A_1, A_2], [A_2, A_3], \ldots, [A_{s-1}, A_s]$ de la quebrada se intersecan sólo en los correspondientes puntos extremos. Es decir, *en un espacio puntual euclídeo la longitud de una quebrada es igual a la longitud del segmento resultante cuando, y sólo cuando, la quebrada es una partición del segmento resultante.*

Veamos ahora como determinar la distancia entre los puntos A y B de un espacio puntual unitario $\mathfrak{U}_n$ si se conocen las coordenadas de estos puntos. Sea $(O, e_1, \ldots, e_n)$ una base de referencia coordenada de $\mathfrak{U}_n$ cuyos vectores forman un sistema ortonormal en el espacio vectorial $\mathfrak{L}$. Indicando por $\alpha_1, \ldots, \alpha_n$ y $\beta_1, \ldots, \beta_n$ las coordenadas de los puntos A y B en la base de referencia señalada, tendremos

$$\overline{AB} = \overline{OB} - \overline{OA} = (\beta_1 - \alpha_1) e_1 + \ldots + (\beta_n - \alpha_n) e_n$$

de modo que

$$\rho(A, B)=\sqrt{|\beta_1-\alpha_1|^2+\ldots+|\beta_n-\alpha_n|^2}. \tag{5}$$

Si el espacio $\mathfrak{U}_n$ es euclídeo, las coordenadas de los puntos serán números reales y, por consiguiente, la fórmula (5) puede ser representada en la forma

$$\rho(A, B)=\sqrt{(\beta_1-\alpha_1)^2+\ldots+(\beta_n-\alpha_n)^2}. \tag{6}$$

Al deducir estas fórmulas hemos aceptado que los vectores de la base de referencia coordenada forman un sistema ortonormal. Estas bases de referencia se llamarán *ortonormales*. Si en lugar de una base de referencia coordenada ortonormal se toma una cualquiera, la fórmula para la distancia entre dos puntos adquiere una forma más compleja que aquí no la daremos.

32.2. Angulo entre rectas. Consideremos en un espacio puntual unitario $\mathfrak{U}_n$ de n dimensiones dos rectas cualesquiera $\mathfrak{X}$ y $\mathfrak{Y}$. Tomemos en cada una de estas rectas un par de diferentes puntos A_1, A_2 y A_3, A_4, respectivamente. Los vectores $\overline{A_1A_2}$ y $\overline{A_3A_4}$ que corresponden a estos pares satisfacen la desigualdad de Cauchy—Buniakovski

$$\frac{|(\overline{A_1A_2}, \overline{A_3A_4})|}{\|\overline{A_1A_2}\|\cdot\|\overline{A_3A_4}\|}\leqslant 1.$$

Por lo tanto, existe un número real φ que satisface la exigencia

$$\cos\varphi=\frac{|(\overline{A_1A_2}, \overline{A_3A_4})|}{\|\overline{A_1A_2}\|\cdot\|\overline{A_3A_4}\|} \quad \left(0\leqslant\varphi\leqslant\frac{\pi}{2}\right). \tag{1}$$

Este número φ no depende de cómo se escogan los pares indicados de puntos en las rectas $\mathfrak{X}$ y $\mathfrak{Y}$. Efectivamente, si B_1, B_2 y B_3, B_4 son otros pares análogos, se tiene

$$\overline{B_1B_2}=\lambda\cdot\overline{A_1A_2} \quad \text{y} \quad \overline{B_3B_4}=\mu\cdot\overline{A_3A_4}$$

para unos valores adecuados $\lambda, \mu\in K$ y por ello

$$\frac{|(\overline{B_1B_2}, \overline{B_3B_4})|}{\|\overline{B_1B_2}\|\cdot\|\overline{B_3B_4}\|}=\frac{|\lambda\mu|}{|\lambda|\cdot|\mu|}\,\frac{|(\overline{A_1A_2}, \overline{A_3A_4})|}{\|\overline{A_1A_2}\|\cdot\|\overline{A_3A_4}\|}=\cos\varphi.$$

Luego, el número real φ que cumple las exigencias (1) depende sólo de las rectas $\mathfrak{X}$ y $\mathfrak{Y}$ y se llama *magnitud del ángulo* entre las rectas $\mathfrak{X}$ y $\mathfrak{Y}$ o simplemente *ángulo* entre las rectas $\mathfrak{X}$ y $\mathfrak{Y}$.

Las rectas $\mathfrak{X}$ y $\mathfrak{Y}$ se llaman *perpendiculares*, si el ángulo entre ellas es igual a $\pi/2$. De las relaciones (1) se deduce que *las rectas $\mathfrak{X}$ y $\mathfrak{Y}$ son perpendiculares cuando, y sólo cuando, cualquier vector que se halla sobre una de estas rectas es ortogonal a cualquier vector que se halla sobre la otra recta.*

Notemos también que *el ángulo entre las rectas* $\mathfrak{X}$ *y* $\mathfrak{Y}$ *es igual a cero cuando, y sólo cuando, las rectas son paralelas.*

En efecto, si las rectas son paralelas, los vectores que se hallan sobre estas rectas son linealmente dependientes y por ello en la fórmula (1) tenemos $\overline{A_3A_4}=\lambda\cdot\overline{A_1A_2}$ y mediante simplificaciones directas obtenemos $\cos\varphi=1$ y $\varphi=0$. Recíprocamente, sea $\varphi=0$ y, por consiguiente,

$$|(\overline{A_1A_2},\ \overline{A_3A_4})|=\|\overline{A_1A_2}\|\cdot\|\overline{A_3A_4}\|.$$

Este es el caso en el que en la desigualdad de Cauchy — Buniakovski tiene lugar el signo de igualdad. Por consiguiente, $\overline{A_3A_4}=\lambda\cdot\overline{A_1A_2}$, de modo que las rectas $\mathfrak{X}$ y $\mathfrak{Y}$ son paralelas.

Fijemos en el espacio unitario $\mathfrak{U}_n$ una base de referencia coordenada ortonormal $(O,\ e_1,\ \ldots,\ e_n)$. Como sabemos las ecuaciones de las rectas $\mathfrak{X}$ y $\mathfrak{Y}$ pueden ser representadas en la forma

$$\frac{\xi_1-\xi_1^0}{l_1}=\ldots=\frac{\xi_n-\xi_n^0}{l_n}=t, \tag{2}$$

$$\frac{\xi_1-\eta_1^0}{m_1}=\ldots=\frac{\xi_n-\eta_n^0}{m_n}=t, \tag{3}$$

donde t es un parámetro y $l_1,\ \ldots,\ l_n$ y $m_1,\ \ldots,\ m_n$ son los coeficientes directores. Supongamos que el punto A_i se obtiene para el valor $t=t_i$ $(i=1,\ 2,\ 3,\ 4)$. Indicando por $\alpha_{i1},\ \ldots,\ \alpha_{in}$ las coordenadas del punto $A_i\,(i=1,\ 2,\ 3,\ 4)$, tendremos

$$\alpha_{is}=l_st_i+\xi_s^0\quad (i=1,\ 2;\ s=1,\ \ldots,\ n),$$
$$\alpha_{is}=m_st_i+\eta_s^0\quad (i=3,\ 4;\ \ s=1,\ \ldots,\ n)$$

y, por consiguiente,

$$\overline{A_1A_2}=(t_2-t_1)(l_1e_1+\ldots+l_ne_n),$$
$$\overline{A_3A_4}=(t_4-t_3)(m_1e_1+\ldots+m_ne_n).$$

Introduciendo en la fórmula (1) estos valores, obtenemos

$$\cos\varphi=\frac{|l_1m_1+\ldots+l_nm_n|}{\sqrt{|l_1|^2+\ldots+|l_n|^2}\ \sqrt{|m_1|^2+\ldots+|m_n|^2}}. \tag{4}$$

Esta es la fórmula estandard para el coseno del ángulo entre unas rectas dadas por las ecuaciones canónicas (2) y (3) en un sistema de coordenadas ortonormal. Tomando la recta (3) como la recta coordenada s-ésima, definida por las ecuaciones

$$\frac{\xi_1}{0}=\ldots=\frac{\xi_s}{1}=\ldots=\frac{\xi_n}{0}$$

e indicando por φ_s el ángulo entre la recta (1) y esta recta coordenada, obtenemos

$$\cos\varphi_s=\frac{|l_s|}{\sqrt{|l_1|^2+\ldots+|l_n|^2}}$$

y, en particular,

$$\cos^2\varphi_1+\ldots+\cos^2\varphi_n=1. \tag{5}$$

Hemos visto ya en el p. 30.3 que la condición de paralelismo de las rectas (2) y (3) puede ser representada en la forma

$$\frac{l_1}{m_1}=\ldots=\frac{l_n}{m_n}.$$

Tomando en la fórmula (4) $\varphi=\frac{\pi}{2}$, llegamos a la relación

$$l_1m_1+\ldots+l_nm_n=0, \tag{6}$$

que representa *la condición de perpendicularidad de las rectas dadas por las ecuaciones canónicas en una base de referencia coordenada ortonormal.*

En los espacios puntuales euclídeos se introduce también, además del concepto de ángulo entre rectas, el concepto de ángulo entre rayos. Las ecuaciones canónicas de los rayos en un espacio euclídeo n-dimensional puntual también se representan en la forma (2) y (3), donde $\xi_1^0, \ldots, \xi_n^0$ y $\eta_1^0, \ldots, \eta_n^0$ son las coordenadas de los vértices de estos rayos, y las coordenadas $\xi_1, \ldots, \xi_n$ de un punto arbitrario de estos rayos se obtienen de las ecuaciones indicadas dando al parámetro t unos valores no negativos arbitrarios. Puesto que el campo principal es en este caso el cuerpo de los números reales, existe un número real único φ que cumple las exigencias

$$\cos\varphi=\frac{l_1m_1+\ldots+l_nm_n}{\sqrt{l_1^2+\ldots+l_n^2}\sqrt{m_1^2+\ldots+m_n^2}} \qquad (0\leqslant\varphi\leqslant\pi), \tag{7}$$

que se denomina *ángulo entre los rayos* señalados. Según estas definiciones el ángulo entre cualesquiera rectas está comprendido siempre en los límites 0 y $\frac{\pi}{2}$, mientras que el ángulo entre unos rayos puede ser también obtuso. Supongamos, por ejemplo, que un rayo está dado por sus ecuaciones canónicas

$$\frac{\xi_1-\xi_1^0}{l_1}=\ldots=\frac{\xi_n-\xi_n^0}{l_n}=t \qquad (t\geqslant 0).$$

La ecuación del rayo opuesto que con el rayo dado forma una recta integra es entonces de la forma

$$\frac{\xi_1-\xi_1^0}{l_1}=\ldots=\frac{\xi_n-\xi_n^0}{l_n}=t \qquad (t\leqslant 0)$$

o en forma canónica

$$\frac{\xi_1-\xi_1^0}{-l_1}=\ldots=\frac{\xi_n-\xi_n^0}{-l_n}=t \qquad (t\geqslant 0)$$

Calculando el ángulo entre estos rayos opuestos según la fórmula (7), obtenemos $\cos\varphi = -1$ y por lo tanto $\varphi = \pi$.

32.3. Proyecciones ortogonales. Empleando el concepto de perpendicularidad de rectas es fácil definir la relación de perpendicularidad de unos planos k-dimensional y l-dimensional en un espacio unitario $\mathfrak{U}_n$. Se dice que la recta $\mathfrak{X}$ del espacio $\mathfrak{U}_n$ es perpendicular a un plano $\mathfrak{M}$ k-dimensional de $\mathfrak{U}_n$, lo que se indica simbólicamente por $\mathfrak{X} \perp \mathfrak{M}$, si $\mathfrak{X}$ es perpendicular a cualquier recta que pertenezca a $\mathfrak{M}$.

Puesto que la perpendicularidad de las rectas equivale a la ortogonalidad de los vectores no nulos que se hallan sobre ellas, se tiene

$$\mathfrak{X} \perp \mathfrak{M} \Leftrightarrow \overline{\mathfrak{X}} \perp \overline{\mathfrak{M}}, \tag{1}$$

donde $\overline{\mathfrak{M}}$ significa el subespacio tangente al plano $\mathfrak{M}$ (en las notaciones del p. 29.3 se tiene $\overline{\mathfrak{M}} = L(\mathfrak{M})$).

Una recta $\mathfrak{X}$ se llama perpendicular trazada desde el punto A hacia el plano $\mathfrak{M}$, si $\mathfrak{X}$ pasa por A, es perpendicular a $\mathfrak{M}$ y se corta con $\mathfrak{M}$ en un punto P. El punto P se llama base de la perpendicular trazada desde A hacia $\mathfrak{M}$ o proyección de A sobre $\mathfrak{M}$.

TEOREMA 1. *En un espacio unitario $\mathfrak{U}_n$ de n dimensiones desde todo punto A que no se halle sobre un plano arbitrario k-dimensional $\mathfrak{M}$ $(1 \leqslant k \leqslant n-1)$ se puede trazar una perpendicular hacia $\mathfrak{M}$, y sólo una.*

UNICIDAD. Supongamos que existen dos rectas diferentes $A \vee B$ y $A \vee C$ $(B \neq C;\ B,\ C \in \mathfrak{M})$ perpendiculares a $\mathfrak{M}$. Estas rectas deben ser entonces perpendiculares a la recta $B \vee C$ que se halla sobre el plano $\mathfrak{M}$.

Supongamos que las rectas $A \vee B$, $A \vee C$ y $B \vee C$ — que se hallan obviamente sobre el plano bidimensional $A \vee B \vee C$ — tienen, en una base de referencia ortonormal del plano $A \vee B \vee C$, los coeficientes directores (l_1, l_2), (m_1, m_2) y (n_1, n_2), respectivamente. La condición de ortogonalidad de las rectas $A \vee B$ y $B \vee C$ implica la igualdad $l_1 n_1 + l_2 n_2 = 0$ (véase (6)); análogamente la ortogonalidad de $A \vee C$ y de $B \vee C$ implica la igualdad $m_1 n_1 + m_2 n_2 = 0$. Vemos que el sistema

$$\left.\begin{aligned} l_1 x_1 + l_2 x_2 &= 0, \\ m_1 x_1 + m_2 x_2 &= 0 \end{aligned}\right\}$$

tiene una solución no trivial (n_1, n_2). Por consiguiente,

$$\begin{vmatrix} l_1 & l_2 \\ m_1 & m_2 \end{vmatrix} = 0,$$

es decir, $\frac{l_1}{m_1} = \frac{l_2}{m_2}$. Luego, las rectas $A \vee B$ y $A \vee C$ son paralelas (véase el p. 30.3) y tienen un punto común A. De aquí se desprende que ellas coinciden. La contradicción obtenida concluye la demostración de la unicidad.

EXISTENCIA. Supongamos primero que $\mathfrak{M}$ es un hiperplano, es decir, que $k=n-1$. En $\mathfrak{U}_n$ existe, según el p. 17.5, un subespacio vectorial de una dimensión $\overline{\mathfrak{M}}^{\perp}$ ortogonal a $\overline{\mathfrak{M}}$. El conjunto $A\cdot\overline{\mathfrak{M}}^{\perp}$ es una recta que pasa por A y es perpendicular a $\mathfrak{M}$. Resta probar solamente que $A\cdot\overline{\mathfrak{M}}^{\perp}$ se interseca con $\mathfrak{M}$. Pero, esto es evidente, ya que en el caso contrario la recta $A\cdot\overline{\mathfrak{M}}^{\perp}$ sería paralela a $\mathfrak{M}$ y por lo tanto en $\mathfrak{M}$ existiría una recta paralela (y no perpendicular) a $A\cdot\overline{\mathfrak{M}}^{\perp}$.

Si ahora $\mathfrak{M}$ no es un hiperplano de $\mathfrak{U}_n$, consideremos el subespacio $\mathfrak{N}=A\vee\mathfrak{M}$. El conjunto $\mathfrak{M}$ es un hiperplano del espacio unitario $\mathfrak{N}$. Luego, según lo anterior, desde A se puede trazar dentro de $\mathfrak{N}$ una perpendicular hacia $\mathfrak{M}$. Es obvio que esta perpendicular será también una perpendicular dentro del espacio principal $\mathfrak{U}_n$.

En el teorema (1) se resuelve el problema sobre la posibilidad de trazar desde el punto A una perpendicular de modo que *corte* el plano $\mathfrak{M}$. Veamos cuál es la situación, si omitimos esta exigencia. Según el p. 17.5, el conjunto $\overline{\mathfrak{M}}^{\perp}$ de todos los vectores del espacio $\mathfrak{U}_n$ ortogonales a $\overline{\mathfrak{M}}$ es un subespacio vectorial de dimensión $n-k$. Por esto el conjunto $A\cdot\overline{\mathfrak{M}}^{\perp}$ será el conjunto de los puntos que se hallan sobre todas las rectas que pasan por A y que son perpendiculares a $\mathfrak{M}$. Puesto que $\overline{\mathfrak{M}}^{\perp}$ es un subespacio vectorial de dimensión $n-k$, el conjunto $A\cdot\overline{\mathfrak{M}}^{\perp}$ es un plano $(n-k)$-dimensional. La intersección de $A\cdot\overline{\mathfrak{M}}^{\perp}$ y de $\mathfrak{M}$ no puede contener ninguna recta $\mathfrak{X}$, ya que en el caso contrario cualquier vector que se halla sobre $\mathfrak{X}$ pertenecería simultáneamente a los espacios $\overline{\mathfrak{M}}$ y $\overline{\mathfrak{M}}^{\perp}$, lo cual es imposible. Por consiguiente, la intersección de $\mathfrak{M}$ y $A\cdot\overline{\mathfrak{M}}^{\perp}$ o es vacía o consta de un solo punto. Si $A\in\mathfrak{M}$, la intersección de $\mathfrak{M}$ y $A\cdot\overline{\mathfrak{M}}^{\perp}$ consta, obviamente, del punto A. En cambio, si $A\notin\mathfrak{M}$, la base de la perpendicular trazada desde A hacia $\mathfrak{M}$ pertenece simultáneamente a $A\cdot\overline{\mathfrak{M}}^{\perp}$ y a $\mathfrak{M}$ y, por ello, el conjunto $\mathfrak{M}\cap A\cdot\overline{\mathfrak{M}}^{\perp}$ está formado solamente por la base de la perpendicular mencionada. El plano $A\cdot\overline{\mathfrak{M}}^{\perp}$ suele llamarse a veces *complemento ortogonal a* $\mathfrak{M}$ *trazado por el punto* A. En particular, si $\mathfrak{M}$ es un hiperplano y $A\notin\mathfrak{M}$, el complemento ortogonal $A\cdot\overline{\mathfrak{M}}^{\perp}$ es la recta que es perpendicular al hiperplano $\mathfrak{M}$ y que lo intercepta en el punto A. Esta recta se llama perpendicular al hiperplano $\mathfrak{M}$ levantada desde el punto A.

Se llama *proyección* (con más precisión, *proyección ortogonal*) de un punto A sobre un plano cualquiera $\mathfrak{M}$ el punto de intersección de los planos $\mathfrak{M}$ y $A\cdot\overline{\mathfrak{M}}^{\perp}$. En otras palabras, si $A\notin\mathfrak{M}$, se llama proyección de A sobre $\mathfrak{M}$ la base de la perpendicular trazada desde A hacia $\mathfrak{M}$. Si $A\in\mathfrak{M}$, se llama proyección de A sobre $\mathfrak{M}$ el propio punto A. La proyección de A sobre $\mathfrak{M}$ se indica a veces por $\mathrm{Pr}_{\mathfrak{M}}A$ o por $A_{\mathfrak{M}}$. La aplicación $\mathrm{Pr}_{\mathfrak{M}}:\mathfrak{U}_n\rightarrow\mathfrak{M}$ se llama proyección (ortogonal) del espacio $\mathfrak{U}_n$ sobre el plano $\mathfrak{M}$.

Veamos cómo se expresan las coordenadas de la proyección del punto A sobre el plano $\mathfrak{M}$ en términos de las coordenadas de A. Tomemos en el espacio $\mathfrak{M}$ una base ortonormal $e_1, \ldots, e_k$ y complementémosla con los vectores $e_{k+1}, \ldots, e_n$ hasta obtener una base otronormal $e_1, \ldots, e_k, \ldots, e_n$ de todo el espacio $\mathfrak{U}_n$. Fijemos en $\mathfrak{M}$ un punto cualquiera O y tomemos como base de referencia coordenada de $\mathfrak{U}_n$ la base de referencia $(O, e_1, \ldots, e_n)$. Sean $\alpha_1, \ldots, \alpha_n$ las coordenadas del punto A en esta base de referencia. Demostremos que la proyección de A sobre $\mathfrak{M}$ será el punto P con las coordenadas $\alpha_1, \ldots, \alpha_k, 0, \ldots, 0$ (véase la fig. 10 para $k=2$).

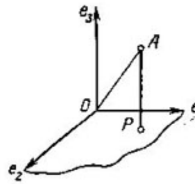

Fig. 10.

En efecto, el plano $\mathfrak{M}$ consta de los puntos cuyas filas coordenadas son de la forma $(\xi_1, \ldots, \xi_k, 0, \ldots, 0)$. Por esto todo vector x perteneciente a $\overline{\mathfrak{M}}$ puede ser representado en la forma $\xi_1e_1+\ldots+\xi_ke_k$. Pero $\overline{PA}=\alpha_{k+1}e_{k+1}+\ldots+\alpha_ne_n$ y por lo tanto se tiene $(x, \overline{PA})=0$, es decir, la recta $P\vee A$ es perpendicular al plano $\mathfrak{M}$ que es lo que se quería demostrar. Luego, si

$$\overline{OA}=\alpha_1e_1+\ldots+\alpha_ke_k+\ldots+\alpha_ne_n,$$

resulta

$$\overline{OA}_{\mathfrak{M}}=\alpha_1e_1+\ldots+\alpha_ke_k\,, \tag{2}$$

donde

$$\mathfrak{M}=O\vee Oe_1\vee\ldots\vee Oe_k.$$

De la fórmula (2) se desprende directamente el siguiente teorema importante.

TEOREMA 2. *La proyección ortogonal* $\mathrm{Pr}_{\mathfrak{M}}:\mathfrak{U}_n\rightarrow\mathfrak{M}$ *es una aplicación lineal, es decir, para cualesquiera puntos* $A^{(1)}, A^{(2)}, A^{(3)}, A^{(4)}, A^{(5)}$ *y* $A^{(6)}$ *y para cualesquiera números* $\lambda, \mu\in K$ *la relación*

$$\overline{A^{(1)}A^{(2)}}=\lambda\cdot\overline{A^{(3)}A^{(4)}}+\mu\cdot\overline{A^{(5)}A^{(6)}} \tag{3}$$

implica

$$\overline{A^{(1)}_{\mathfrak{M}}A^{(2)}_{\mathfrak{M}}}=\lambda\cdot\overline{A^{(3)}_{\mathfrak{M}}A^{(4)}_{\mathfrak{M}}}+\mu\cdot\overline{A^{(5)}_{\mathfrak{M}}A^{(6)}_{\mathfrak{M}}}. \tag{4}$$

Efectivamente, indicando por $\alpha_1^{(i)}, \ldots, \alpha_n^{(i)}$ las coordenadas del punto $A^{(i)}$ $(i=1, \ldots, 6)$ y tomando en consideración la condición

(3), obtenemos las igualdades

$$\alpha_s^{(2)}-\alpha_s^{(1)}=\lambda(\alpha_s^{(4)}-\alpha_s^{(3)})+\mu(\alpha_s^{(6)}-\alpha_s^{(5)}) \qquad (s=1, \ldots, n). \tag{5}$$

Por otra parte, debido a (2), tenemos

$$\overline{A_{\mathfrak{M}}^{(i)}A_{\mathfrak{M}}^{(i+1)}}=(\alpha_1^{(i+1)}-\alpha_1^{(i)})e_1+\ldots+(\alpha_k^{(i+1)}-\alpha_k^{(i)})e_k \quad (i=2, 4, 6). \tag{6}$$

Comparando (6) y (5) llegamos a la relación (4).

Se llama proyección $\mathfrak{S}_{\mathfrak{M}}$ de un conjunto arbitrario de puntos $\mathfrak{S}$ sobre un plano $\mathfrak{M}$ el conjunto de las proyecciones sobre $\mathfrak{M}$ de todos los puntos de $\mathfrak{S}$.

TEOREMA 3. *La proyección de una recta $\mathfrak{X}$ sobre un plano cualquiera $\mathfrak{M}$ es una recta o un punto. Si $\mathfrak{X}^{(1)}, \ldots, \mathfrak{X}^{(s)}$ son unas rectas de un espacio unitario $\mathfrak{U}_n$ y $\mathfrak{M}$ es un plano cualquiera de $\mathfrak{U}_n$, se tiene*

$$(\mathfrak{X}^{(1)}\vee\ldots\vee\mathfrak{X}^{(s)})_{\mathfrak{M}}=\mathfrak{X}_{\mathfrak{M}}^{(1)}\vee\ldots\vee\mathfrak{X}_{\mathfrak{M}}^{(s)}. \tag{7}$$

Tomemos en cada una de las rectas $\mathfrak{X}^{(i)}$ un par de puntos diferentes $A^{(i)}$, $B^{(i)}$ $(i=1, \ldots, s)$. El plano $\mathfrak{X}^{(1)}\vee\ldots\vee\mathfrak{X}^{(s)}$ está formado por aquellos puntos C para los cuales

$$\overline{A^{(1)}C}=\lambda_1\cdot\overline{A^{(1)}B^{(1)}}+\ldots+\lambda_s\cdot\overline{A^{(1)}B^{(s)}}+$$
$$+\mu_2\cdot\overline{A^{(1)}A^{(2)}}+\ldots+\mu_s\cdot\overline{A^{(1)}A^{(s)}} \tag{8}$$

para unos valores adecuados $\lambda_1, \ldots, \lambda_s, \mu_2, \ldots, \mu_s\in K$ (véase el p. 29.2). Basándonos en (4), obtenemos de aquí

$$\overline{A_{\mathfrak{M}}^{(1)}C_{\mathfrak{M}}}=\lambda_1\cdot\overline{A_{\mathfrak{M}}^{(1)}B_{\mathfrak{M}}^{(1)}}+\ldots+\lambda_s\cdot\overline{A_{\mathfrak{M}}^{(1)}B_{\mathfrak{M}}^{(s)}}+$$
$$+\mu_2\cdot\overline{A_{\mathfrak{M}}^{(1)}A_{\mathfrak{M}}^{(2)}}+\ldots+\mu_s\cdot\overline{A_{\mathfrak{M}}^{(1)}A_{\mathfrak{M}}^{(s)}}, \tag{9}$$

de modo que

$$C_{\mathfrak{M}}=(A_{\mathfrak{M}}^{(1)}\vee B_{\mathfrak{M}}^{(1)})\vee\ldots\vee(A_{\mathfrak{M}}^{(s)}\vee B_{\mathfrak{M}}^{(s)}).$$

Recíprocamente, si

$$D\in(A_{\mathfrak{M}}^{(1)}\vee B_{\mathfrak{M}}^{(1)})\vee\ldots\vee(A_{\mathfrak{M}}^{(s)}\vee B_{\mathfrak{M}}^{(s)}),$$

para valores adecuados $\lambda_1, \ldots, \lambda_s, \mu_2, \ldots, \mu_s\in K$ se tiene

$$\overline{A_{\mathfrak{M}}^{(1)}D}=\lambda_1\cdot\overline{A_{\mathfrak{M}}^{(1)}B_{\mathfrak{M}}^{(1)}}+\ldots+\lambda_s\cdot\overline{A_{\mathfrak{M}}^{(1)}B_{\mathfrak{M}}^{(s)}}+$$
$$+\mu_2\cdot\overline{A_{\mathfrak{M}}^{(1)}A_{\mathfrak{M}}^{(2)}}+\ldots+\mu_s\cdot\overline{A_{\mathfrak{M}}^{(1)}A_{\mathfrak{M}}^{(s)}}. \tag{10}$$

Tomemos en el espacio $\mathfrak{X}^{(1)}\vee\ldots\vee\mathfrak{X}^{(s)}$ un punto C tal que se cumpla la igualdad (8). Comparando (9) y (10) vemos que $C_{\mathfrak{M}}=D$ y por ello $D\in(\mathfrak{X}^{(1)}\vee\ldots\vee\mathfrak{X}^{(s)})_{\mathfrak{M}}$. Vemos, por consiguiente, que

$$(\mathfrak{X}^{(1)}\vee\ldots\vee\mathfrak{X}^{(s)})_{\mathfrak{M}}=(A_{\mathfrak{M}}^{(1)}\vee B_{\mathfrak{M}}^{(1)})\vee\ldots\vee(A_{\mathfrak{M}}^{(s)}\vee B_{\mathfrak{M}}^{(s)}).$$

Para $s=1$ obtenemos $\mathfrak{X}_{\mathfrak{M}}^{(1)}=A_{\mathfrak{M}}^{(1)}\vee B_{\mathfrak{M}}^{(1)}$. Por consiguiente $\mathfrak{X}_{\mathfrak{M}}^{(i)}=$ $=A_{\mathfrak{M}}^{(i)}\vee B_{\mathfrak{M}}^{(i)}$ y quedan demostradas ambas afirmaciones del teorema 3.

COROLARIO. *En un espacio unitario la proyección de un plano $\mathfrak{N}$ sobre otro plano es de nuevo un plano cuya dimensión no sobrepasa la dimensión de $\mathfrak{N}$.*

Para la demostración es suficiente representar el plano $\mathfrak{N}$ en forma de la adherencia lineal de las rectas independientes que pasan por un punto y aplicar la fórmula (7).

En los espacios puntuales euclídeos están definidos los conceptos de segmento y de cuerpo convexo. Mediante la fórmula (4) se comprueba fácilmente el teorema siguiente.

TEOREMA 4 *En un espacio puntual euclídeo la proyección de un segmento cualquiera $[A, B]$ sobre un plano $\mathfrak{M}$ es un segmento $[A_{\mathfrak{M}}, B_{\mathfrak{M}}]$ y por ello la proyección de cualquier conjunto convexo es un conjunto convexo.*

En efecto, si $C \in [A, B]$, se tiene $\overline{AC} = \lambda \cdot \overline{AB}$, $0 \leqslant \lambda \leqslant 1$. De aquí se deduce, debido a (4), que $\overline{A_{\mathfrak{M}}C_{\mathfrak{M}}} = \lambda \cdot \overline{A_{\mathfrak{M}}B_{\mathfrak{M}}}$ y por lo tanto $C_{\mathfrak{M}} \in [A_{\mathfrak{M}}, B_{\mathfrak{M}}]$ y $[A, B]_{\mathfrak{M}} \subseteq [A_{\mathfrak{M}}, B_{\mathfrak{M}}]$. Recíprocamente, si $D \in [A_{\mathfrak{M}}, B_{\mathfrak{M}}]$, para un valor conveniente de λ $(0 \leqslant \lambda \leqslant 1)$ tenemos $\overline{A_{\mathfrak{M}}D} = \lambda \cdot \overline{A_{\mathfrak{M}}B_{\mathfrak{M}}}$. Tomando en el segmento $[A, B]$ un punto C tal que $\overline{AC} = \lambda \cdot \overline{AB}$ y comparando (4) con la igualdad $\overline{A_{\mathfrak{M}}D} = \lambda \cdot \overline{A_{\mathfrak{M}}B_{\mathfrak{M}}}$, obtenemos $D = C_{\mathfrak{M}}$ y, por consiguiente, $[A_{\mathfrak{M}}, B_{\mathfrak{M}}] \subseteq [A, B]_{\mathfrak{M}}$ que es lo que se quería demostrar.

Veamos cómo pueden calcularse las coordenadas de la proyección de un punto dado sobre un hiperplano dado, si se conocen las coordenadas del punto y la ecuación del hiperplano. Sea $(O, e_1, \ldots, e_n)$ una base de referencia coordenada ortonormal de un espacio unitario, sea B un punto de coordenadas $\beta_1, \ldots, \beta_n$ y sea

$$\alpha_1\xi_1 + \ldots + \alpha_n\xi_n = \alpha \tag{11}$$

la ecuación del hiperplano dado. Consideremos dos puntos distintos cualesquiera C y D de coordenadas respectivas $\gamma_1, \ldots, \gamma_n$ y $\delta_1, \ldots, \delta_n$ que se hallan sobre el hiperplano (11). Tenemos entonces

$$\overline{CD} = (\delta_1 - \gamma_1) e_1 + \ldots + (\delta_n - \gamma_n) e_n. \tag{12}$$

Las coordenadas de C y de D deben satisfacer la ecuación (11). Introduciéndolas en la ecuación (11) y restando término por término las dos igualdades obtenidas, encontramos

$$\alpha_1 (\delta_1 - \gamma_1) + \ldots + \alpha_n (\delta_n - \gamma_n) = 0. \tag{13}$$

De (12) y (13) vemos que la recta

$$\frac{\xi_1 - \beta_1}{\alpha_1} = \ldots = \frac{\xi_n - \beta_n}{\alpha_n} \tag{14}$$

es perpendicular a cualquier vector $\overline{CD}$ que se halle sobre el plano

(11) y, por ende, es perpendicular a este plano. Puesto que la recta (14) pasa por el punto B, *las igualdades* (14) *son las ecuaciones canónicas de la recta que pasa por el punto dado B de coordenadas $\beta_1, \ldots, \beta_n$ y que es perpendicular al plano dado* (11).

Resolviendo conjuntamente las ecuaciones (11) y (14), encontramos las coordenadas

$$\xi_i = \frac{\alpha-(\alpha_1\beta_1+\ldots+\alpha_n\beta_n)}{\alpha_1^2+\ldots+\alpha_n^2}\alpha_i+\beta_i \; (i=1, \ldots, n)$$

de la proyección de B sobre el hiperplano (11).

32.4 Angulo entre un plano y una recta. Consideremos un plano arbitrario $\mathfrak{M}$ k-dimensional perteneciente a un espacio unitario $\mathfrak{U}$ de n dimensiones y una recta cualquiera $\mathfrak{X}$ de este espacio que intercepta el plano $\mathfrak{M}$ en el punto A. La proyección de la recta $\mathfrak{X}$ sobre el plano $\mathfrak{M}$ es una recta $\mathfrak{X}_{\mathfrak{M}}$ que pasa por el punto A. El ángulo entre la recta $\mathfrak{X}$ y su proyección $\mathfrak{X}_{\mathfrak{M}}$ se llama *ángulo entre la recta $\mathfrak{X}$ y el plano* $\mathfrak{M}$ (fig. 11). Si la recta $\mathfrak{X}$ se halla sobre el plano $\mathfrak{M}$, se tiene $\mathfrak{X}=\mathfrak{X}_{\mathfrak{M}}$ y el ángulo entre $\mathfrak{X}$ y $\mathfrak{M}$ es igual a cero. Si $\mathfrak{X}$ no se halla sobre el plano $\mathfrak{M}$, existe el único plano $(k+1)$-dimensional $\mathfrak{N}$ que pasa por $\mathfrak{M}$ y $\mathfrak{X}$. El plano $\mathfrak{M}$ es un hiperplano en el espacio $\mathfrak{N}$ y por ello se puede trazar en $\mathfrak{N}$ sólo una perpendicular $\mathfrak{Y}$ a $\mathfrak{M}$ por el punto A. Sea B un punto cualquiera de $\mathfrak{X}$ diferente de A y sea C la proyección de B sobre $\mathfrak{M}$. La recta $B \vee C$ se halla en el espacio $\mathfrak{N}$ y es perpendicular a $\mathfrak{M}$. Todas las perpendiculares a un mismo hiperplano son paralelas (véase el p. 32.3) y por esto las rectas $\mathfrak{Y}$, $\mathfrak{X}$ y $A \vee C$ (véase la fig. 11) se hallan sobre un mismo plano bidimensional. De aquí se deduce que *el ángulo φ entre $\mathfrak{X}$ y $\mathfrak{M}$ y el ángulo ψ entre $\mathfrak{X}$ y $\mathfrak{Y}$ dan en suma $\pi/2$.*

Fig. 11.

Resta considerar el caso en que el plano k-dimensional $\mathfrak{M}$ y la recta $\mathfrak{X}$ no se intersecan. Tomemos sobre el plano $\mathfrak{M}$ un punto cualquiera A. En el espacio $\mathfrak{U}$ existe, según el p. 30.3, una recta única $\mathfrak{X}'$ que pasa por el punto A y que es paralela a la recta $\mathfrak{X}$. El ángulo entre la recta $\mathfrak{X}'$ y el plano $\mathfrak{M}$ se llama en este caso ángulo entre la recta $\mathfrak{X}$ y el plano $\mathfrak{M}$. Es fácil comprobar que el ángulo entre la recta $\mathfrak{X}$ y el plano $\mathfrak{M}$ definido de esta forma no depende de cómo se escoja el punto auxiliar A.

Supongamos que en un espacio $\mathfrak{U}$ se ha fijado una base de referencia ortonormal $(O, e_1, \ldots, e_n)$ y que en esta base de referencia están dados un hiperplano $\mathfrak{M}$ por medio de la ecuación

$$\alpha_1\xi_1+\ldots+\alpha_n\xi_n=\alpha \tag{1}$$

y una recta $\mathfrak{X}$ mediante sus esuaciones canónicas

$$\frac{\xi_1-\xi_1^0}{l}=\ldots=\frac{\xi_n-\xi_n^0}{l_n}. \tag{2}$$

Tomemos en el hiperplano $\mathfrak{M}$ un punto A y sean $\sigma_1, \ldots, \sigma_n$ sus coordenadas. Entonces las ecuaciones de la recta $\mathfrak{X}'$ que es paralela a $\mathfrak{X}$ y que pasa por A son

$$\frac{\xi_1-\sigma_1}{l_1}=\ldots=\frac{\xi_n-\sigma_n}{l_n}, \tag{3}$$

y las ecuaciones de la perpendicular $\mathfrak{Y}$ levantada en el punto A hacia el hiperplano $\mathfrak{M}$ serán

$$\frac{\xi_1-\sigma_1}{\alpha_1}=\ldots=\frac{\xi_n-\sigma_n}{\alpha_n}. \tag{4}$$

Por lo visto en el p. 32.2, para el ángulo ψ entre $\mathfrak{X}'$ y $\mathfrak{Y}$ se tiene

$$\cos\psi=\frac{|\alpha_1 l_1+\ldots+\alpha_n l_n|}{\sqrt{|\alpha_1|^2+\ldots+|\alpha_n|^2}\ \sqrt{|l_1|^2+\ldots+|l_n|^2}}$$

y, por consiguiente, el ángulo φ entre $\mathfrak{M}$ y $\mathfrak{X}$ satisface las relaciones

$$\operatorname{sen}\varphi=\frac{|\alpha_1 l_1+\ldots+\alpha_n l_n|}{\sqrt{|\alpha_1|^2+\ldots+|\alpha_n|^2}\ \sqrt{|l_1|^2+\ldots+|l_n|^2}} \quad \left(0\leqslant\varphi\leqslant\frac{\pi}{2}\right). \tag{5}$$

Igual que en el espacio euclídeo corriente, el ángulo entre una recta y un plano en un espacio unitario posee la propiedad de extremalidad.

TEOREMA. *Supongamos que en un espacio unitario $\mathfrak{U}_n$ de n dimensiones una recta $\mathfrak{X}$ se corta con un plano k-dimensional $\mathfrak{M}$ en un punto A formando un ángulo φ diferente de 0 y de $\pi/2$. Entonces el ángulo entre $\mathfrak{X}$ y cualquier recta $\mathfrak{Z}$, que pasa en el plano $\mathfrak{M}$ por A y es diferente de la proyección de $\mathfrak{X}$ sobre $\mathfrak{M}$, es mayor que φ.*

Tomemos sobre $\mathfrak{X}$ un punto cualquiera B diferente de A y tracemos desde B hacia $\mathfrak{M}$ una perpendicular indicando su base por C. Indiquemos por D la base de la perpendicular trazada desde el punto B hacia la recta $\mathfrak{Z}$ (véase la fig. 11). Puesto que la recta $B \vee C$ es perpendicular al plano $\mathfrak{M}$, los ángulos BCD y BCA son rectos. El ángulo BDA es recto por construcción. De los triángulos rectángulos BCA, BDA y BCD obtenemos, respectivamente

$$\operatorname{sen}\varphi=\frac{BC}{BA},\ \operatorname{sen}(\mathfrak{X},\ \mathfrak{Z})=\frac{BD}{BA}\ \text{y}\ BD=\sqrt{BC^2+DC^2}>BC.$$

Luego, $\operatorname{sen}\varphi<\operatorname{sen}(\mathfrak{X},\ \mathfrak{Z})$ y $\varphi<\angle(\mathfrak{X},\ \mathfrak{Z})$ que es lo que se quería demostrar

Ejemplos y problemas

En los problemas que se ofrecen a continuación se supone que en cada uno de los casos se ha escogido y se ha fijado una base de referencia ortonormal del espacio puntual euclídeo y que los puntos de este espacio se dan por las coordenadas calculadas en esta base

1. Hállese la proyección ortogonal del segmento $[(0, 0, 0, 0), (4, -1, -3, 4)]$ sobre el plano $\mathfrak{M} = A_0 \vee A_1 \vee A_2 \vee A_3$, donde las filas coordenadas de los puntos A_0, A_1, A_2 y A_3 son, respectivamente, $(0, 0, 0, 0,)$, $(1, 1, 1, 1)$, $(1, 2, 2, -1)$ y $(1, 0, 0, 3)$.

2. Se llama distancia de un punto A a un plano $\mathfrak{M}$ el mínimo de las distancias entre el punto dado y los puntos de $\mathfrak{M}$. Demuéstrese que esta distancia es igual a la longitud del vector $\overline{A\,\mathrm{Pr}_{\mathfrak{M}} A}$.

3. Hallénse las longitudes de los lados y los ángulos interiores del triángulo cuyos vértices están dados por las filas coordenadas $(2, 4, 2, 4, 2)$, $(6, 4, 4, 4, 6)$ y $(5, 7, 5, 7, 2)$.

4. Hállese el ángulo entre la recta $A_0 \vee B$ y el plano $A_0 \vee A_1 \vee A_2$, donde las filas coordenadas de los puntos A_0, A_1, A_2, y B son iguales, respectivamente, a $(0, 0, 0, 0)$, $(3, 4, -4, -1)$, $(0, 1, -1, 2)$ y $(2, 2, 1, 1)$.

INDICE ALFABÉTICO