

Carlos Ivorra Castillo

GEOMETRÍA

La geometría ilumina el intelecto y temple la mente. Todas sus pruebas son claras y ordenadas. Apenas caben errores en el razonamiento geométrico, pues está bien dispuesto y ordenado. Así, no es probable que la mente que se aplica a la geometría con regularidad cometa errores. De este modo, quien sabe geometría adquiere inteligencia.

IBN KHALDUN

Índice General

Introducción	ix
Capítulo I: La geometría absoluta	1
1.1 Axiomas de incidencia	2
1.2 Axiomas de ordenación	3
1.3 Ángulos y triángulos	6
1.4 Axiomas de congruencia	9
1.5 Suma de ángulos	13
1.6 Más propiedades de segmentos, ángulos y triángulos	16
1.7 Perpendiculares	18
1.8 El axioma de continuidad, círculos y circunferencias	21
Capítulo II: Medida de segmentos, ángulos y arcos	27
2.1 Longitud de segmentos. Números reales	27
2.2 Complementos sobre números reales	38
2.3 Amplitud de ángulos	43
2.4 Arcos y sectores circulares	46
Capítulo III: La geometría euclídea	49
3.1 El axioma de las paralelas	49
3.2 Semejanza de triángulos	53
3.3 Relaciones entre ángulos y arcos	59
3.4 Las razones trigonométricas	61
3.5 Propiedades de los triángulos	64
Capítulo IV: La geometría analítica	73
4.1 Vectores	73
4.2 Espacios afines	79
4.3 Coordenadas cartesianas y baricéntricas	84
4.4 Espacios euclídeos	88
4.5 Los giros y la medida de ángulos	97
4.6 Complementos sobre trigonometría	104
4.7 Circunferencias	105
4.8 Cónicas	106

Capítulo V: Números complejos	115
5.1 Definición y propiedades básicas	115
5.2 La clausura algebraica de \mathbb{C}	117
5.3 Construcciones con regla y compás	119
5.4 Polígonos regulares	124
5.5 Geometría discontinua	134
5.6 Apéndice: El teorema de Sylow	136
Capítulo VI: Biyecciones afines	139
6.1 El grupo afín y el grupo lineal	139
6.2 Homotecias	141
6.3 El teorema fundamental de la geometría afín	146
6.4 Isometrías y semejanzas	149
6.5 Clasificación de endomorfismos	153
6.6 Clasificación de isometrías	172
6.7 Aplicaciones	179
Capítulo VII: La geometría afín	183
7.1 Incidencia y paralelismo	183
7.2 Homotecias y traslaciones	186
7.3 Vectores y escalares	189
7.4 Los teoremas de Desargues y Pappus-Pascal	194
7.5 Axiomas de ordenación	200
Capítulo VIII: La geometría proyectiva	205
8.1 Espacios proyectivos	207
8.2 Homografías y coordenadas homogéneas	214
8.3 Perspectividades	219
8.4 Caracterización axiomática	224
8.5 Dualidad	229
8.6 Razones dobles y separación armónica	235
8.7 Espacios sobre cuerpos ordenados	245
Capítulo IX: Secciones cónicas	251
9.1 Clasificación de formas bilineales simétricas	251
9.2 Cónicas proyectivas y afines	256
9.3 La polaridad de una cónica	264
9.4 El teorema de Steiner	268
9.5 Propiedades de las cónicas proyectivas	272
9.6 Homografías entre cónicas	282
9.7 Cónicas sobre cuerpos ordenados	289
9.8 Complexificación	294

Capítulo X: La geometría parabólica	299
10.1 Espacios parabólicos	299
10.2 El plano euclídeo	309
10.3 El plano de Lorentz	317
10.4 Propiedades métricas de las cónicas	336
10.5 Espacios de dimensiones superiores	346
Capítulo XI: La geometría circular	349
11.1 La proyección estereográfica	349
11.2 Transformaciones circulares	352
11.3 Homografías en la esfera	355
11.4 Conservación de ángulos	357
11.5 El teorema de Feuerbach	360
Capítulo XII: La geometría hiperbólica	363
12.1 El plano hiperbólico	363
12.2 Medida de segmentos y ángulos	368
12.3 El modelo de Poincaré	372
12.4 Trigonometría hiperbólica	375
12.5 Las isometrías hiperbólicas	380
Capítulo XIII: La geometría elíptica	387
13.1 El plano elíptico	387
13.2 Biláteros y triángulos	390
13.3 Isometrías elípticas	394
13.4 Trigonometría elíptica	395
Bibliografía	397
Índice de Materias	398

Introducción

La geometría es, junto a la teoría de números, una de las ramas más antiguas de la matemática. Si por un momento restringimos el término para referirnos a lo que los antiguos griegos entendían como tal, podemos decir que su objeto de estudio está íntimamente arraigado en nuestra forma de concebir la realidad. Toda la información que recibimos del mundo que nos rodea, todo lo que vemos, oímos y tocamos, lo procesamos en primera instancia en términos geométricos. Sin embargo, no podemos considerar a las leyes formales que rigen el espacio tridimensional que percibimos como una parte de la física. Al contrario que las leyes físicas, las leyes de la geometría nos son dadas a priori, en cuanto que ninguna experiencia puede confirmar o refutar ninguna de ellas. Por ejemplo, podemos asegurar a priori que es imposible percibir una recta que posea dos paralelas por un mismo punto.¹ Nuestra intuición geométrica nos permite decidir inmediatamente la verdad o falsedad de un gran número de afirmaciones. A su vez, de todas ellas se sigue mediante razonamientos lógicos un cuerpo de teoremas no menos numeroso que, si nuestra intuición no alcanza a validar directamente, al menos los corrobora en instancias particulares.

Los antiguos griegos exploraron en profundidad este cuerpo de teoremas y llegaron a comprender en gran medida su estructura lógica. Tanto es así que en sus exposiciones más elaboradas, el modelo de las cuales son, sin duda, los Elementos de Euclides, no sólo se demuestran con un gran sentido del rigor todos los hechos no evidentes, sino que incluso los que cualquiera daría tranquilamente por obvios son demostrados a partir del mínimo número de principios a los que el autor pudo reducirlos.

Fermat y Descartes descubrieron que la geometría como teoría lógica es equivalente a una estructura algebraica, esencialmente al espacio vectorial \mathbb{R}^3 , en el sentido de que los puntos, rectas, planos, circunferencias, etc. pueden ser identificados con ciertos subconjuntos de \mathbb{R}^3 de modo que los teoremas geométricos sobre estos conceptos se corresponden con los teoremas algebraicos sobre sus conjuntos asociados. Así surgió la llamada geometría analítica y con ella la clave para una comprensión mucho más profunda de la geometría en general.

El álgebra es especialmente dada a encontrar principios profundos, poco evidentes por sí mismos pero enormemente iluminadores. El que una determinada afirmación se nos aparezca o no como evidente es una cuestión psicológica sin

¹Por supuesto, salvo que pervirtamos el significado de la palabra “recta” y lo confundamos, por ejemplo, con un concepto físico como pueda ser el de “trayectoria de un rayo de luz”.

ningún significado matemático, por lo que la geometría axiomática al estilo de Euclides se considera hoy, con razón, como algo superado. El tratamiento algebraico de la geometría, aparte de ser lógicamente más simple, nos abre las puertas de “otras geometrías”, es decir, de otras teorías algebraicas lo suficientemente cercanas a las de la geometría tradicional euclídea como para que sea justo englobarlas bajo el mismo nombre. El caso más elemental es la sustitución del exponente en \mathbb{R}^3 por cualquier otro número natural. No tenemos ninguna intuición que pueda aplicarse a \mathbb{R}^4 , pero el cambio de un 3 por un 4 apenas modifica la teoría algebraica, que de hecho se desarrolla sin dificultad y por el mismo precio en el espacio general \mathbb{R}^n . Otros casos menos triviales son las geometrías no euclídeas o las geometrías basadas en los números complejos.

La algebrización de la geometría no supone únicamente un cambio de lenguaje. En el siglo XIX la geometría, al igual que las demás ramas de la matemática, experimentó un desarrollo gigantesco en varias direcciones. Por un lado, Poncelet sentó las bases de la geometría proyectiva, que viene a demostrar que nuestra intuición nos proporciona una imagen sesgada de una estructura algebraica más regular de lo que los ojos nos muestran. Esta regularidad se pone de manifiesto postulando la existencia de puntos infinitos. Gracias a ellos, una hipérbola y una elipse pueden considerarse como una misma figura vista desde posiciones distintas (la primera con dos puntos en el infinito y la segunda con todos sus puntos finitos). Si postulamos la existencia de puntos imaginarios (en el sentido de los números complejos) la regularidad de la geometría se multiplica una vez más. Por otra parte, Gauss mostró las posibilidades del cálculo diferencial aplicado al estudio de las superficies. La geometría diferencial es hoy la aproximación más potente a la mayoría de las ramas de la geometría.

El objeto de este libro es presentar una panorámica de la geometría previa a la geometría diferencial. Más precisamente, de la geometría sin topología. Hay varias razones por las que consideramos útil conocer las técnicas no topológicas en geometría. Por una parte entre ellas se encuentran las técnicas genuinamente algebraicas, que son de gran valor en sí mismas y por sus posibilidades de aplicación. En muchos casos el álgebra suple con razonamientos conceptuales exquisitamente limpios lo que en un enfoque más directo se convertiría en una ristra de cálculos, concluyentes pero ciegos.

En segundo lugar, y a pesar de que más arriba la hayamos calificado de anticuada, es interesante conocer la geometría sintética, es decir, la geometría tradicional que parte de axiomas evidentes intuitivamente. Esta aproximación es la única que justifica que ciertas ramas del álgebra describen realmente las leyes de nuestra intuición. Formalmente es posible evitarla, pero con ello se incurre en una especie de “estafa legal”. Pensemos por ejemplo en un resultado tan clásico como el teorema de Pitágoras. Podemos definir la norma de un vector de \mathbb{R}^2 como $\|(x, y)\| = \sqrt{x^2 + y^2}$ y la perpendicularidad como $(x, y) \perp (x', y')$ si y sólo si $xx' + yy' = 0$, y a partir de aquí demostrar el teorema de Pitágoras, pero esa demostración, lógicamente irrefutable, no puede convencer a nadie de la validez del teorema si antes no se nos “demuestran” las definiciones, si antes no se nos convence de que si tomamos un papel cuadriculado y clavamos las dos puntas del compás en los ángulos opuestos de un rectángulo de lados

3 y 4, nos encontraremos con que podemos girarlo hasta situarlo sobre dos puntos que disten 5 unidades, o que si dibujamos los vectores $(2, 1)$ y $(-1, 2)$ obtendremos realmente segmentos perpendiculares. Al igual que este ejemplo podemos encontrar muchísimos más, en torno a la definición de la medida de ángulos, de las razones trigonométricas, de la ortogonalidad de vectores, del cálculo diferencial básico, etc.

Cubrir estas lagunas apenas cuesta cuatro de los trece capítulos que tiene este libro, por lo que creemos que merece la pena.

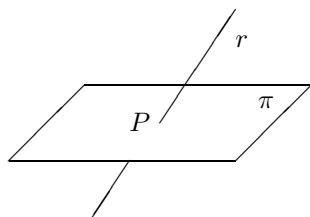
Capítulo I

La geometría absoluta

El objeto de estudio de la geometría es lo que se conoce como el *espacio*. No es posible definir este concepto, pero todos tenemos una idea intuitiva del mismo. Nuestra intención es obtener un modelo matemático del espacio, para lo cual iremos introduciendo axiomáticamente afirmaciones intuitivamente evidentes, hasta llegar a un punto en que podamos probar que sólo hay un objeto matemático que satisface dichas propiedades.

La primera aproximación a la caracterización matemática del espacio en el seno de la teoría de conjuntos será precisamente concebirlo como tal, es decir, considerar al espacio como un conjunto E a cuyos elementos llamaremos *puntos*. Un punto es una posición en el espacio, carente de toda extensión. Un punto ortográfico es una buena imagen de un punto geométrico, si bien hemos de tener presente que la pequeña mancha de tinta tiene una cierta extensión, de la que debemos hacer abstracción.

Hay dos conceptos más que se encuentran al mismo nivel elemental que el concepto de punto. Se trata de los conceptos de *recta* y *plano*. De nuevo es imposible definir la característica que diferencia a una línea recta de una línea curva o a una superficie plana de una superficie curva, pero intuitivamente todos sabemos distinguir las rectas y los planos de las restantes curvas y superficies.



Una recta r y un plano π que se cortan en un punto P .

Hemos de hacer una aclaración sobre la interpretación que vamos a dar a estas palabras: una hoja de papel proporciona una buena imagen de un plano, salvo por el hecho de que la hoja termina en unos bordes, mientras que para nosotros un plano será una superficie ilimitada. Si apoyamos la hoja en una

mesa, la superficie de la mesa representa una porción más amplia del mismo plano que representa la hoja. Similarmente, por una recta entenderemos una línea recta sin extremos, de modo que si trazamos una porción de recta con la ayuda de una regla, cualquier extensión de la misma por cualquiera de sus extremos será una porción mayor de la misma recta. Del mismo modo que hemos hecho con el espacio, podemos concebir las rectas y los planos como conjuntos de puntos, es decir, como *ciertos* subconjuntos de E que no podemos definir.

1.1 Axiomas de incidencia

Definición 1.1 Una *geometría* (tridimensional) está formada por un conjunto E al que llamaremos *espacio*—y a cuyos elementos llamaremos *puntos*— junto con dos familias no vacías de subconjuntos de E a cuyos elementos llamaremos respectivamente *rectas* y *planos*, de modo que se cumplan los siete axiomas indicados más abajo.

Diremos que una recta o plano X *pasa* por un punto P , o que X *incide* en el punto P , si $P \in X$.

Axioma A1 *Por cada par de puntos distintos P y Q pasa una única recta, que representaremos por PQ .*

Axioma A2 *Toda recta pasa al menos por dos puntos.*

Diremos que tres o más puntos son *colineales* si hay una recta que pasa por todos ellos.

Axioma A3 *Por cada tres puntos no colineales P, Q, R pasa un único plano, al que representaremos por PQR .*

Axioma A4 *Todo plano contiene tres puntos no colineales*

Axioma A5 *Si una recta r tiene dos puntos en común con un plano π , entonces r está contenida en π .*

Axioma A6 *Si dos planos tienen un punto en común, entonces tienen dos puntos en común.*¹

Diremos que cuatro o más puntos son *coplanares* si hay un plano que pasa por todos ellos.

Axioma A7 *Existen cuatro puntos distintos no coplanares.*

Veamos algunas consecuencias de estos axiomas básicos. Todas ellas son elementales, así que omitiremos los detalles de las pruebas.

Si una recta o un plano X no pasa por un punto P , diremos que P es un punto *exterior* a X .

¹Este axioma implica que la dimensión del espacio es ≤ 3 , mientras que el siguiente implica que es ≥ 3 .

Teorema 1.2 *Toda recta tiene un punto exterior contenido en un plano dado* (Por A4)

Teorema 1.3 *Todo plano tiene un punto exterior.* (Por A7)

Teorema 1.4 *Si P y Q son puntos de una recta r y R es exterior a r , entonces P , Q y R no son colineales.* (Por A1)

Teorema 1.5 *Si P , Q y R son puntos de un plano π no colineales y S es exterior a π , entonces P , Q , R y S no son coplanares.* (Por A3)

Teorema 1.6 *Dados dos planos distintos, entonces o bien no tienen puntos comunes o bien su intersección es una recta.* (Por A6 y A3)

Diremos que dos planos son *coincidentes* si son iguales, *secantes* (lat. ‘que se cortan’) si su intersección es una recta y *paralelos* (gr. ‘uno al lado del otro’) si no tienen puntos comunes.

Teorema 1.7 *Dos rectas distintas tienen como máximo un punto en común.* (Por A1)

Diremos que dos rectas son *coincidentes* si son iguales, *secantes* si su intersección es un punto, y si no tienen puntos comunes diremos que son *paralelas* si están contenidas en un mismo plano o que *se cruzan* en caso contrario.

Teorema 1.8 *Dados un plano y una recta, o bien no tienen puntos comunes, o bien tienen un único punto en común, o bien la recta está contenida en el plano.* (Por A5)

Diremos que una recta y un plano son *secantes* si se cortan en un punto y *paralelos* si no tienen puntos comunes.

Teorema 1.9 *Una recta y un punto exterior a ella están contenidos en un único plano.* (Por A2, A3 y A5)

Teorema 1.10 *Dos rectas secantes están contenidas en un único plano.*

1.2 Axiomas de ordenación

Si fijamos dos puntos A y B en una recta r y establecemos que A está a la izquierda de B , esto determina cuándo un punto cualquiera de r está a la izquierda o a la derecha de otro punto de r . Más aún, tiene sentido decir qué punto está más a la izquierda o más a la derecha de uno dado, es decir, los puntos de la recta quedan ordenados por el criterio de que un punto es menor cuanto más a la izquierda se encuentra. Es importante notar que la noción de izquierda y derecha es relativa, pues si giramos la recta, la izquierda se convierte en derecha y viceversa. Los axiomas siguientes recogen las propiedades necesarias para determinar estas nociones intuitivas.

Definición 1.11 Una geometría está *ordenada* si cada par ordenado de puntos distintos A y B tiene asociado una relación \leq_{AB} sobre los puntos de la recta AB de tal modo que se satisfacen los cinco axiomas siguientes:

Axioma B1 Para todo par de puntos distintos A y B , la relación \leq_{AB} es una relación de orden total sobre la recta AB , es decir, es reflexiva, antisimétrica, transitiva y todo par de puntos P y Q de AB cumple $P \leq_{AB} Q$ o bien $Q \leq_{AB} P$. Además, con este orden, la recta no tiene máximo ni mínimo, y para todo par de puntos $P <_{AB} Q$, existe un punto R tal que $P <_{AB} R <_{AB} Q$.

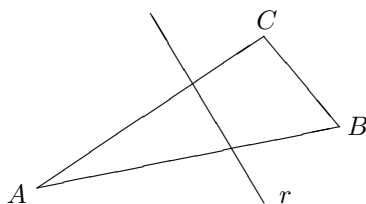
Axioma B2 Para todo par de puntos distintos A y B se cumple $A \leq_{AB} B$.

Axioma B3 Si A y B son dos puntos distintos y P, Q son dos puntos de la recta AB , entonces $P \leq_{AB} Q$ si y sólo si $Q \leq_{BA} P$.

Axioma B4 Si $A \neq B$ y $C \neq D$ son pares de puntos de una misma recta, entonces $\leq_{AB} = \leq_{CD}$ o bien $\leq_{AB} = \leq_{DC}$.

Los axiomas B3 y B4 afirman que en realidad sólo estamos considerando dos ordenaciones en cada recta, y una es la inversa de la otra. Diremos que un punto P está entre dos puntos A y B si los tres son colineales y $A \leq_{AB} P \leq_{AB} B$.

Axioma B5 Sean A, B, C tres puntos no colineales y r una recta contenida en el plano ABC pero que no pase por ninguno de ellos. Si r pasa por un punto situado entre A y B , entonces r pasa por un punto entre A y C o bien por un punto entre B y C , y sólo se da uno de los dos casos.



Teorema 1.12 Dados tres puntos A, B, C , se cumple que B está entre A y C si y sólo si está entre C y A . (Por B2)

Definición 1.13 Dados dos puntos distintos A y B , llamaremos *segmento* (lat. ‘corte’) de extremos A y B al conjunto de los puntos situados entre A y B . Lo representaremos por \overline{AB} .

Por el teorema anterior $\overline{AB} = \overline{BA}$. Observemos que un segmento s contiene a sus extremos, luego está contenido en una única recta, a la que llamaremos *prolongación* de s . Los extremos de s son el máximo y el mínimo para cualquiera de las ordenaciones de su prolongación, luego están determinados por s , es decir, dos segmentos son iguales si y sólo si tienen los mismos extremos.

Si a una recta le quitamos un punto, ésta queda dividida en dos partes. El teorema siguiente lo demuestra, al tiempo que caracteriza a cada una de estas partes. Para enunciarlo conviene adoptar el convenio de que $\overline{AA} = \{A\}$, si bien no consideraremos como segmentos a estos conjuntos de un solo punto.

Teorema 1.14 Sea r una recta y O un punto de r . Entonces la relación en $r \setminus \{O\}$ dada por $P \sim Q$ si y sólo si $O \notin \overline{PQ}$ es una relación de equivalencia con exactamente dos clases de equivalencia.

DEMOSTRACIÓN: Sea B un punto de r y sea A tal que $A <_{OB} O <_{OB} B$. Es fácil ver que $P \sim Q$ si y sólo si P y Q son ambos mayores que O o ambos menores que O respecto a la relación \leq_{OB} , de donde se sigue inmediatamente que la relación es de equivalencia. Además hay sólo dos clases, a saber, la clase de A y la clase de B . ■

Definición 1.15 Dado un punto O en una recta r llamaremos *semirrectas* en r con *origen* en O a cada una de las dos clases de equivalencia descritas en el teorema anterior junto con el punto O .

Claramente, dados dos puntos distintos O y A , existe una única semirrecta de origen O y que pasa por A . La representaremos por \overrightarrow{OA} . Es fácil ver que $\overrightarrow{OA} = \{P \in OA \mid O \leq_{OA} P\}$.

Una semirrecta está contenida en una única recta, a la que llamaremos su *prolongación*. Notemos que si s es una semirrecta, entonces el origen de s es el máximo o el mínimo de s respecto a cada una de las dos ordenaciones de su prolongación. Si es el mínimo entonces s no tiene máximo y viceversa. Por consiguiente una semirrecta determina a su origen, de modo que dos semirrectas iguales tienen el mismo origen.

Si s es una semirrecta, existe una única semirrecta distinta de s con la misma prolongación y el mismo origen. A ésta la llamaremos *semirrecta complementaria* de s . El único punto en común entre una semirrecta y su complementaria es el origen de ambas.

Del mismo modo que un punto divide a una recta en dos semirrectas, una recta divide a un plano en dos semiplanos. El teorema siguiente lo demuestra.

Teorema 1.16 Sea π un plano y r una recta contenida en π . Entonces la relación en $\pi \setminus r$ dada por $P \sim Q$ si y sólo si r no corta a \overline{PQ} es una relación de equivalencia con exactamente dos clases de equivalencia.

DEMOSTRACIÓN: La relación es obviamente reflexiva y simétrica, y es transitiva por el axioma B5. Tomemos un punto A de π exterior a r y un punto X en r . La recta AX está contenida en π y corta a r únicamente en X , luego si tomamos un punto B en AX tal que X esté entre A y B , ciertamente $B \in \pi \setminus r$. Más aún, $A \not\sim B$. De nuevo el axioma B5 implica que todo punto de $\pi \setminus r$ es equivalente a A o a B , luego sólo hay dos clases de equivalencia. ■

Definición 1.17 Sea π un plano y r una recta contenida en π . Llamaremos *semiplanos* en π de *frontera* r a las uniones de cada una de las dos clases de equivalencia descritas en el teorema anterior con la recta r .

Dada una recta r y un punto exterior A , existe un único semiplano de frontera r y que contiene a A . Lo representaremos por \overrightarrow{rA} , y es el conjunto de todos los puntos X del plano que contiene a r y a A tales que r no corta al segmento \overline{AX} .

Un semiplano s está contenido en un único plano, al que llamaremos su *prolongación*. La frontera de s está formada por los puntos X con la propiedad de que existe una semirrecta de origen X contenida en s cuya semirrecta complementaria no tiene más punto en s que el propio X . Por lo tanto s determina a su frontera, es decir, si dos semiplanos son iguales, sus fronteras son iguales.

Dado un semiplano s , existe un único semiplano distinto de s con la misma prolongación y con la misma frontera, al que llamaremos *semiplano complementario* de s . Los únicos puntos en común entre un semiplano y su complementario son los de la frontera.

Ejercicio: Probar que un plano divide al espacio en dos semiespacios. Definir los conceptos de prolongación de un semiespacio y semiespacio complementario.

Definición 1.18 Un conjunto de puntos F es *convexo* si cuando A , B son puntos de F entonces el segmento \overline{AB} está contenido en F .

Es fácil probar que las rectas, los planos, los segmentos, las semirrectas, los semiplanos y los semiespacios son conjuntos convexos. Así mismo es claro que la intersección de conjuntos convexos es un conjunto convexo.

1.3 Ángulos y triángulos

Definición 1.19 Sean l_1 y l_2 dos semirrectas con origen común O y no contenidas en la misma recta. Sean r_1 y r_2 sus respectivas prolongaciones. Sea π el plano que las contiene. Es claro que l_1 está contenido en uno de los semiplanos en que r_2 divide a π y l_2 está contenido en uno de los semiplanos en que r_1 divide a π . Llamaremos *ángulo* (lat. ‘rincón’) de *vértice* (lat. ‘cumbre’) O y lados l_1 y l_2 a la intersección del semiplano de π respecto a r_2 que contiene a l_1 con el semiplano de π respecto a r_1 que contiene a l_2 . Lo representaremos $\widehat{l_1 l_2}$. Los puntos de l_1 y l_2 constituyen la *frontera* del ángulo.

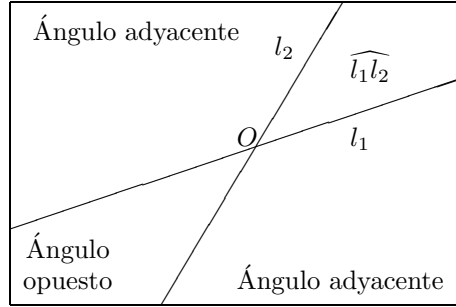
Observemos que $\widehat{l_1 l_2}$ contiene más puntos, aparte de los de sus lados. De hecho es un conjunto convexo, pues es la intersección de dos conjuntos convexos. Por lo tanto, si A y B son puntos en l_1 y l_2 respectivamente, entonces todos los puntos entre ellos están en el ángulo.

Si tres puntos A , O y B no son colineales, llamaremos \widehat{AOB} al ángulo de vértice O y lados \overrightarrow{OA} y \overrightarrow{OB} .

Un ángulo está contenido en un único plano, llamado su *sopORTE*. Es fácil probar que un ángulo determina su vértice y sus lados.

Dos rectas secantes dividen el plano que las contiene en cuatro ángulos con vértice común. Dos ángulos con el mismo vértice, un lado en común y los otros

lados formados por semirrectas complementarias se llaman ángulos *adyacentes*. Dos ángulos con el mismo vértice y cuyos lados son semirrectas complementarias se llaman ángulos *opuestos por el vértice*. Cada ángulo tiene exactamente dos ángulos adyacentes y un ángulo opuesto por el vértice.



Teorema 1.20 Sean A, O, B puntos no colineales en un plano π . Entonces una semirrecta de origen O y contenida en π está contenida en el ángulo \widehat{AOB} si y sólo si corta al segmento \overline{AB} .

DEMOSTRACIÓN: Sea s una semirrecta de origen O y contenida en π . Podemos suponer que no es uno de los lados de \widehat{AOB} . Si s corta a \overline{AB} en un punto X y P es cualquier otro punto de s (distinto de O), entonces la prolongación del segmento \overline{PX} es la prolongación de s , que corta a los lados del ángulo en el punto O , y éste no está en \overline{PX} . Por lo tanto P y X están en los mismos semiplanos respecto a los lados del ángulo, y como X está en el ángulo, P también. Esto prueba que s está contenida en \widehat{AOB} .

Recíprocamente, supongamos que s está contenida en \widehat{AOB} . Sea \bar{s} la prolongación de s y consideremos un punto C tal que O esté entre C y B . Los puntos A, B y C no son colineales, y \bar{s} no pasa por ninguno de ellos. Como \bar{s} pasa por O , que es un punto entre C y B , por el axioma B5 ha de pasar por un punto entre A y B o bien por un punto entre A y C . Ahora bien, s está contenida en \widehat{AOB} y puesto que al pasar por O cruza a la vez los dos lados de s , es claro que la semirrecta complementaria de s está contenida en el ángulo opuesto por el vértice a \widehat{AOB} . Sin embargo el segmento \overline{AC} está contenido en el ángulo \widehat{AOC} , que es uno de los adyacentes a \widehat{AOB} . Por lo tanto \bar{s} no puede cortar a \overline{AC} . Así pues, \bar{s} corta a \overline{AB} . Sin embargo, la semirrecta complementaria de s no puede cortar a este segmento, pues está fuera de \widehat{AOB} . Concluimos que es s quien corta al segmento. ■

Observemos que si un punto P está en un ángulo \widehat{AOB} entonces la semirrecta \overrightarrow{OP} está contenida en \widehat{AOB} .

Teorema 1.21 Dos ángulos con un lado en común y contenidos en el mismo semiplano respecto a él están contenidos el uno en el otro.

DEMOSTRACIÓN: Los ángulos serán de la forma $\widehat{AOB_1}$ y $\widehat{AOB_2}$. Supongamos que el segmento $\overline{AB_1}$ no corta a la recta OB_2 . Entonces todos los puntos

de $\overline{AB_1}$ están en el mismo semiplano que A respecto a OB_2 , y por hipótesis también están en el mismo semiplano que B_2 respecto a OA . Por lo tanto $\overline{AB_1}$ está contenido en $\widehat{AOB_2}$. Si P es cualquier punto de $\widehat{AOB_1}$ (que no esté en OA) la semirrecta \overrightarrow{OP} corta a $\overline{AB_1}$ en un punto X . Si $X = P$ ya tenemos que P está en $\widehat{AOB_2}$. En caso contrario, el segmento \overline{XP} no contiene a O , por lo que X y P están en los mismos semiplanos respecto a los lados de $\widehat{AOB_2}$, luego P está en $\widehat{AOB_2}$.

Supongamos ahora que $\overline{AB_1}$ corta a la recta OB_2 en un punto X . Entonces X está en $\widehat{AOB_1}$, luego está en el mismo semiplano que B_2 respecto a OA , luego en realidad X está en $\overrightarrow{OB_2}$. Esto implica que $\widehat{AOX} = \widehat{AOB_2}$, luego no perdemos generalidad si suponemos que $X = B_2$. Ahora, si P está en $\widehat{AOB_2}$ la semirrecta \overrightarrow{OP} corta a $\overline{AB_2}$, luego corta a $\overline{AB_1}$, luego está contenida en $\widehat{AOB_1}$, luego P está en $\widehat{AOB_1}$. ■

De la prueba del teorema anterior y del teorema 1.20 se deduce el hecho siguiente:

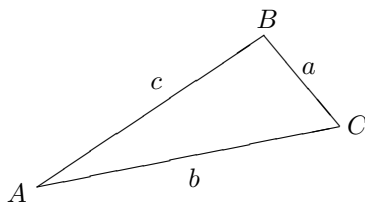
Teorema 1.22 Sean l_1, l_2 y l_3 semirrectas de origen O y tales que l_2 y l_3 estén contenidas en un mismo semiplano de frontera la prolongación de l_1 . Entonces $l_1 l_2$ está contenido en $l_1 l_3$ si y sólo si l_2 está contenida en $l_1 l_3$.

Definición 1.23 Sean A, B y C tres puntos no colineales. Llamaremos *triángulo de vértices A, B y C* a la intersección de los ángulos $\widehat{BAC}, \widehat{ABC}$ y \widehat{ACB} . Lo representaremos por \widehat{ABC} .

Los ángulos $\widehat{BAC}, \widehat{ABC}$ y \widehat{ACB} se llaman *ángulos* del triángulo \widehat{ABC} . Cuando no haya ambigüedad, nos referiremos a ellos como \hat{A}, \hat{B} y \hat{C} , respectivamente (es decir, los nombraremos por sus vértices). Los segmentos $\overline{AB}, \overline{AC}$ y \overline{BC} se llaman *lados* de \widehat{ABC} . Los tres lados de un triángulo forman su *frontera*.

Los lados \overline{AB} y \overline{AC} se llaman lados *contiguos* al ángulo \hat{A} , mientras que el lado \overline{BC} es el lado *opuesto* al ángulo \hat{A} (similarmente con los otros dos ángulos).

Normalmente llamaremos a, b y c a los lados de un triángulo \widehat{ABC} , de modo que a será el lado opuesto al ángulo \hat{A} , b será el lado opuesto a \hat{B} y c será el lado opuesto a \hat{C} .



1.4 Axiomas de congruencia

Continuamos introduciendo conceptos geométricos básicos ocupándonos de la congruencia de figuras. La idea subyacente es que dos figuras son congruentes si se diferencian a lo sumo en su posición en el espacio, es decir, si una puede convertirse en la otra mediante un movimiento. Aunque en principio el concepto de congruencia es aplicable a cualquier figura, de momento sólo necesitamos considerar congruencias de segmentos, ángulos y triángulos. Además, la congruencia de triángulos puede definirse en términos de las otras dos.

Definición 1.24 Una *geometría métrica* es una geometría ordenada junto con dos relaciones, que llamaremos de *congruencia* y las representaremos por \equiv , definidas respectivamente sobre los conjuntos de los segmentos y ángulos, y que cumplen los axiomas siguientes:

Axioma C1 Las dos congruencias son relaciones de equivalencia, es decir, son reflexivas, simétricas y transitivas.

Axioma C2 Dados tres puntos A , B y A' y una semirrecta s de origen A' , existe un único punto B' en s tal que $\overline{AB} \equiv \overline{A'B'}$.

Axioma C3 Sean A , B , C puntos colineales de modo que B esté entre A y C , sean A' , B' y C' otros tres puntos en las mismas condiciones. Entonces, si $\overline{AB} \equiv \overline{A'B'}$ y $\overline{BC} \equiv \overline{B'C'}$, también $\overline{AC} \equiv \overline{A'C'}$.

Axioma C4 Sea L un ángulo, s una semirrecta y π un semiplano cuya frontera sea la prolongación de s . Entonces existe un único ángulo L' contenido en π , con un lado igual a s y tal que $L \equiv L'$.

Diremos que dos triángulos T y T' son *congruentes* si existe una correspondencia entre sus vértices para la cual cada par de lados y ángulos correspondientes son congruentes. En lo sucesivo, cuando digamos que dos triángulos \widehat{ABC} y $\widehat{A'B'C'}$ son congruentes, se sobreentenderá que cumplen la definición para la correspondencia $A \rightarrow A'$, $B \rightarrow B'$, $C \rightarrow C'$, es decir, que $\overline{AB} \equiv \overline{A'B'}$, etc.

Axioma C5 Dado un triángulo \widehat{ABC} , un segmento $\overline{A'B'} \equiv \overline{AB}$ y un semiplano π de frontera la prolongación de $\overline{A'B'}$, existe un (único) triángulo $\widehat{A'B'C'}$ contenido en π y congruente con \widehat{ABC} .

Observemos que la unicidad del triángulo se sigue del axioma C4, pues las rectas $A'C'$ y $B'C'$ son únicas, y C' ha de ser su intersección. En el lenguaje tradicional de la geometría es costumbre hablar de ángulos, segmentos y triángulos ‘iguales’ en el sentido que aquí hemos dado a la palabra ‘congruentes’, mientras que para indicar que dos segmentos, ángulos o triángulos son iguales en el sentido conjuntista, es decir, que contienen los mismos puntos, se suele decir que son ‘coincidentes’. Nosotros usaremos estos términos excepto cuando pueda dar lugar a confusión.

Comencemos estudiando las propiedades de la congruencia de segmentos. El axioma C3 nos permite definir una suma:

Teorema 1.25 *Dados dos segmentos \overline{AB} y \overline{CD} existe un segmento \overline{PQ} con la propiedad de que existe un punto R entre P y Q de modo que $\overline{PR} \equiv \overline{AB}$ y $\overline{RQ} \equiv \overline{CD}$. La clase de congruencia de \overline{PQ} sólo depende de las clases de congruencia de \overline{AB} y \overline{CD} .*

DEMOSTRACIÓN: Tomamos por ejemplo $P = A$ y $R = B$. Ahora consideramos la semirrecta de AB con origen B y que no contiene a A . Por el axioma C2 existe en ella un punto Q tal que $\overline{BQ} \equiv \overline{CD}$. Es claro que P , Q y R cumplen lo pedido. El resto es consecuencia del axioma C3. ■

Definición 1.26 En las condiciones del teorema anterior, escribiremos

$$\overline{PQ} \equiv \overline{AB} + \overline{CD},$$

entendiendo la expresión como una igualdad entre clases de congruencia.

De este modo tenemos definida la suma de dos (clases de) segmentos cualesquiera. Es obvio que esta suma es asociativa y conmutativa. El hecho siguiente no es exactamente una consecuencia inmediata de la definición de suma:

Teorema 1.27 *Si $\overline{PQ} \equiv u + v$ entonces existe un punto R entre P y Q tal que $\overline{PR} \equiv u$ y $\overline{RQ} \equiv v$.*

DEMOSTRACIÓN: Por el axioma C2 existe un punto R en la semirrecta \overrightarrow{PQ} tal que $\overline{PR} \equiv u$. Del mismo modo, existe un punto R' en la semirrecta de PQ de origen R y que no contiene a P de modo que $\overline{RR'} \equiv v$. Por definición de suma tenemos que $\overline{PR'} \equiv u + v \equiv \overline{PQ}$, luego la unicidad del axioma C2 implica que $R' = Q$. ■

De aquí se sigue que la suma de segmentos es simplificable:

Teorema 1.28 *Dados tres segmentos u , v y w , si $u + v \equiv u + w$ entonces $v \equiv w$.*

DEMOSTRACIÓN: Sea $u + v \equiv u + w \equiv \overline{PQ}$. Entonces existe un punto R entre P y Q tal que $\overline{PR} \equiv u$ y $\overline{RQ} \equiv v$. También tiene que existir un punto R' tal que $\overline{PR'} \equiv u$ y $\overline{R'Q} \equiv w$, y por la unicidad de C2 ha de ser $R = R'$, luego $v \equiv \overline{RQ} = \overline{R'Q} \equiv w$. ■

Definición 1.29 Diremos que un segmento u es menor que un segmento v (y lo representaremos por $u < v$) si existe un segmento w tal que $v \equiv u + w$. En tal caso el teorema anterior afirma que w es único (salvo congruencia) y lo llamaremos *resta* o *diferencia* de u y v , y lo representaremos $w \equiv v - u$.

De las propiedades de la suma se sigue inmediatamente que la desigualdad de segmentos depende sólo de las clases de congruencia y es una relación de orden estricto. También es fácil probar lo siguiente:

Teorema 1.30 Si $\overline{AB} < \overline{AC}$ y ambos segmentos están situados sobre una semirrecta de origen A , entonces B está entre A y C .

Nos ocupamos ahora de la congruencia de ángulos y triángulos. Comenzamos con dos criterios de igualdad de triángulos.

Teorema 1.31 (Criterio lado-ángulo-lado) Si dos triángulos $T = \widehat{ABC}$ y $T' = \widehat{A'B'C'}$ cumplen $\overline{AB} \equiv \overline{A'B'}$, $\overline{AC} \equiv \overline{A'C'}$ y $\hat{A} \equiv \hat{A'}$ entonces $T \equiv T'$.

DEMOSTRACIÓN: Por el axioma C5 el triángulo T es igual a un triángulo $T'' = \widehat{A'B'C''}$ contenido en mismo semiplano que T' respecto a la recta $A'B'$ y con todos sus lados y ángulos congruentes con los de T . En particular, usando el axioma C1 resulta que $\overline{A'C'} \equiv \overline{A'C''}$ y $\widehat{B'A'C'} \equiv \widehat{B'A'C''}$.

Estos últimos ángulos tienen un lado en común y están contenidos en el mismo semiplano respecto a dicho lado, luego por el axioma C4 son coincidentes. En particular la semirrecta $\overrightarrow{A'C'}$ coincide con $\overrightarrow{A'C''}$. Ahora el axioma C2 implica que $C' = C''$, por lo que T' coincide con T'' , luego T y T' son iguales. ■

Teorema 1.32 (Criterio ángulo-lado-ángulo) Si dos triángulos $T = \widehat{ABC}$ y $T' = \widehat{A'B'C'}$ cumplen $\overline{AB} \equiv \overline{A'B'}$, $\hat{A} \equiv \hat{A'}$ y $\hat{B} \equiv \hat{B'}$ entonces $T \equiv T'$.

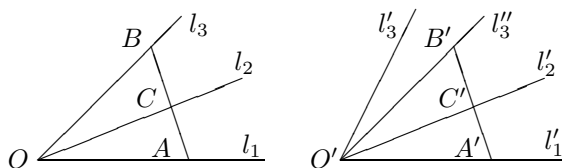
DEMOSTRACIÓN: Razonamos igual que en el teorema anterior, con lo que obtenemos un triángulo $T'' = \widehat{A'B'C''}$ igual a T y que comparte con T' el lado $\overline{A'B'}$ y los ángulos \hat{A} y $\hat{B'}$. Esto implica que $\overline{A'C'} \equiv \overline{A'C''}$ y $\overline{B'C'} \equiv \overline{B'C''}$, pero C' y C'' son los respectivos puntos donde se cortan estas semirrectas, luego $C' = C''$ y concluimos igualmente. ■

Más adelante probaremos que si dos triángulos tienen dos ángulos iguales entonces tienen los tres ángulos iguales, con lo que el criterio anterior cubrirá cualquier caso en que dos triángulos tengan iguales dos ángulos y un lado. Más delicado es probar el criterio lado-lado-lado. Antes necesitamos un par de resultados. En primer lugar demostramos para ángulos lo que el axioma C3 afirma para segmentos.

Teorema 1.33 Sean l_1, l_2 y l_3 semirrectas de origen O tales que las dos últimas están contenidas en un mismo semiplano con frontera la prolongación de la primera. Sean l'_1, l'_2 y l'_3 semirrectas de origen O' en las mismas condiciones. Supongamos que el ángulo $\widehat{l_1 l_2}$ está contenido en $\widehat{l_1 l_3}$. Si $\widehat{l_1 l_2} \equiv \widehat{l'_1 l'_2}$ y $\widehat{l_2 l_3} \equiv \widehat{l'_2 l'_3}$ entonces $\widehat{l'_1 l'_2}$ está contenido en $\widehat{l'_1 l'_3}$ y $\widehat{l_1 l_3} \equiv \widehat{l'_1 l'_3}$.

DEMOSTRACIÓN: Por el axioma C4 existe una semirrecta l''_3 de origen O' y contenida en el mismo semiplano respecto a l'_1 que las demás y de modo que $\widehat{l_1 l_3} \equiv \widehat{l'_1 l''_3}$. Hemos de probar que $l''_3 = l'_3$.

Tomemos puntos $A \in l_1$ y $B \in l_3$. Por el teorema 1.20 sabemos que la semirrecta l_2 corta a \overline{AB} en un punto C . Consideremos puntos $A' \in l'_1$ y $B' \in l'_3$ tales que $\overline{O'A'} \equiv \overline{OA}$ y $\overline{O'B'} \equiv \overline{OB}$.



Por el criterio LAL de igualdad de triángulos tenemos que $\widehat{AOB} \equiv \widehat{A'O'B'}$, luego $\overline{AB} \equiv \overline{A'B'}$. Sea C' el punto de $\overline{A'B'}$ que cumple $\overline{AC} \equiv \overline{A'C'}$. Como $\overline{AC} < \overline{AB}$ resulta que C' está entre A' y B' . Veamos que está en l'_2 . Esto se debe a que los triángulos \widehat{AOC} y $\widehat{A'O'C'}$ son iguales, porque tienen iguales el ángulo $\hat{A} \equiv \hat{A'}$ y los lados $\overline{OA} \equiv \overline{O'A'}$ y $\overline{AC} \equiv \overline{A'C'}$. Por lo tanto $\widehat{AOC} \equiv \widehat{A'O'C'}$, y por la unicidad del axioma C4 ha de ser $l'_2 = \overline{O'C'}$. Esto implica que $l'_1 l'_2$ está contenido en $l'_1 l'_3$.

Finalmente observamos que los triángulos \widehat{BOC} y $\widehat{B'O'C'}$ tienen iguales el ángulo $\hat{B} \equiv \hat{B'}$ y sus lados adyacentes, luego $l'_2 l'_3 \equiv l_2 l_3 \equiv l'_2 l'_3$, y por consiguiente $l'_3 = l_3$. ■

Definición 1.34 Un triángulo es *equilátero* (lat. ‘de lados iguales’) si sus tres lados son iguales. Un triángulo es *isósceles* (gr. ‘de piernas iguales’) si tiene al menos dos lados iguales. Un triángulo es *escaleno* (gr. ‘oblicuo’) si sus lados son desiguales dos a dos.

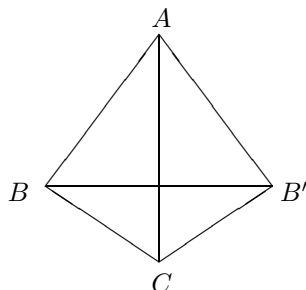
Probamos ahora un resultado con una prueba mucho más elemental de lo que podría preverse:

Teorema 1.35 (Criterio del triángulo isósceles) Si en un triángulo \widehat{ABC} se cumple $\overline{CA} \equiv \overline{CB}$ entonces $\hat{A} \equiv \hat{B}$.

DEMOSTRACIÓN: Basta aplicar el criterio LAL a los triángulos (coincidentes) \widehat{ACB} y \widehat{BCA} (es decir, tomando $A' = B$, $B' = A$ y $C' = C$). ■

Observemos que el teorema anterior implica que los triángulos equiláteros tienen también sus tres ángulos iguales. Finalmente podemos probar:

Teorema 1.36 (Criterio lado-lado-lado) Si $T = \widehat{ABC}$ y $T' = \widehat{A'B'C'}$ cumplen $\overline{AB} \equiv \overline{A'B'}$, $\overline{AC} \equiv \overline{A'C'}$ y $\overline{BC} \equiv \overline{B'C'}$, entonces $T \equiv T'$.



DEMOSTRACIÓN: Trasladando uno de los triángulos podemos suponer que $\overline{AC} = \overline{A'C'}$ y que los vértices B y B' se encuentran en semiplanos distintos respecto a este lado común.

En estos términos $\overline{AB} \equiv \overline{A'B'}$, $\overline{BC} \equiv \overline{B'C'}$ y hemos de probar que $\widehat{AB'C} \equiv \widehat{ABC}$. Para ello basta probar que $\widehat{AB'C} \equiv \widehat{ABC}$, pues entonces el criterio LAL nos da el resultado. Si C está en el segmento $\overline{BB'}$ concluimos por el teorema anterior aplicado a

$\widehat{ABB'}$. En caso contrario aplicamos el teorema anterior a los triángulos $\widehat{ABB'}$ y $\widehat{CBB'}$, con lo que obtenemos $\widehat{ABB'} \equiv \widehat{AB'B}$ y $\widehat{CBB'} \equiv \widehat{CB'B}$. El teorema 1.33 implica que $\widehat{AB'C} \equiv \widehat{ABC}$. ■

1.5 Suma de ángulos

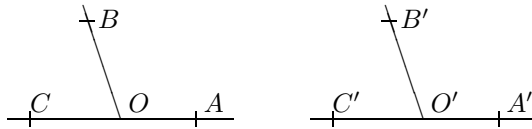
El teorema 1.33 nos permite definir una suma de ángulos de forma similar a como hemos definido la suma de segmentos. Conviene definir primero la ordenación de los ángulos.

Definición 1.37 Diremos que un ángulo A es menor que un ángulo B (y lo representaremos por $A < B$) si existen ángulos A' y B' congruentes con A y B respectivamente, con un lado en común, situados en un mismo semiplano respecto a dicho lado y de modo que A' está (estrictamente) contenido en B' .

Si A y B son dos ángulos no congruentes, por el axioma C4 existen ángulos A' y B' en las condiciones de la definición, por el teorema 1.21 uno de ellos estará contenido en el otro y por el teorema 1.33 el resultado de la comparación depende sólo de las clases de congruencia de A y B . Ahora es fácil probar que la relación que acabamos de definir es ciertamente una relación de orden total estricto sobre (las clases de congruencia de) todos los ángulos, así como que si dos ángulos comparten un lado y están contenidos en un mismo semiplano respecto a dicho lado, entonces el menor estará contenido en el mayor.

Teorema 1.38 Si L y L' son ángulos iguales, S es un ángulo adyacente a L y S' es un ángulo adyacente a L' , entonces S y S' también son iguales.

DEMOSTRACIÓN: Digamos que $L = \widehat{AOB}$ y que $S = \widehat{COB}$, donde los puntos C, O, A están alineados. Sea O' el vértice de L' y S' , tomemos B' en el lado común entre ambos y de modo que $\overline{OB} \equiv \overline{O'B'}$. Sea A' en el otro lado de L' de modo que $\overline{OA} \equiv \overline{O'A'}$ y sea C' en el otro lado de S' de modo que $\overline{OC} \equiv \overline{O'C'}$.



Es claro entonces que $\overline{CA} \equiv \overline{C'A'}$, $\widehat{AOB} \equiv \widehat{A'O'B'}$ (por el criterio LAL), de donde $\widehat{AB} \equiv \widehat{A'B'}$, de donde $\widehat{CAB} \equiv \widehat{C'A'B'}$ (por el mismo criterio), de donde $\widehat{BC} \equiv \widehat{B'C'}$, de donde $\widehat{COB} \equiv \widehat{C'O'B'}$ (por el criterio LLL). Esto implica que $\widehat{COB} \equiv \widehat{C'O'B'}$, es decir, $S \equiv S'$. ■

Como consecuencia inmediata tenemos:

Teorema 1.39 Los dos ángulos adyacentes a un ángulo dado son iguales entre sí. Dos ángulos opuestos por el vértice son iguales entre sí.

(Notemos que dos ángulos opuestos por el vértice son adyacentes a un mismo ángulo.)

Definición 1.40 Dos ángulos son *suplementarios* si uno es congruente con un ángulo adyacente al otro.

Es obvio que la relación de ser suplementarios depende sólo de las clases de congruencia de los ángulos, así como que es simétrica. Si dos ángulos suplementarios tienen un lado en común y están en semiplanos opuestos respecto a éste, entonces son adyacentes.

Teorema 1.41 Si un ángulo A es menor que un ángulo B , entonces el suplementario de B es menor que el suplementario de A .

DEMOSTRACIÓN: Podemos suponer que A y B tienen un lado en común y están contenidos en un mismo semiplano respecto a éste. Digamos $A = \widehat{l_1 l_2}$ y $B = \widehat{l_1 l_3}$. Sea l'_1 la semirrecta complementaria de l_1 . Entonces los suplementarios de A y B son $\widehat{l'_1 l_2}$ y $\widehat{l'_1 l_3}$. Basta probar que l_3 está contenida en $\widehat{l'_1 l_2}$. Por el teorema 1.20 esto equivale a que l'_1 y l_2 estén en semiplanos distintos respecto a l_3 , lo que a su vez equivale a que l_1 y l_2 estén en el mismo semiplano respecto a l_3 , y por definición de ángulo esto equivale a que l_2 esté contenida en $\widehat{l_1 l_3}$, lo cual es cierto por hipótesis. ■

Definición 1.42 Un *semihaz de semirrectas* (lat. fascem = ‘manejo’) es el conjunto de todas las semirrectas con un origen común O contenidas en un semiplano S que tenga a O en su frontera.

Notemos que un semihaz determina su origen y su semiplano. El origen O divide a la frontera de S en dos semirrectas s y t a las que llamaremos *extremos* del semihaz. Cada semirrecta l en un semihaz que sea distinta de sus extremos determina dos ángulos suplementarios \widehat{sl} y \widehat{tl} . Esto nos permite definir dos ordenaciones totales en el semihaz: una dada por

$$l \leq_{st} l' \quad \text{si y sólo si} \quad \widehat{sl} \leq \widehat{sl'}$$

(con el convenio adicional de que $s \leq_{st} l \leq_{st} t$, para cualquier semirrecta l distinta de s y t), y otra \leq_{ts} definida análogamente. El teorema anterior prueba que ambas ordenaciones son mutuamente inversas.

En particular, si l_1 , l_2 y l_3 son tres semirrectas de un semihaz, diremos que l_2 está *entre* l_1 y l_3 si $l_1 \leq_{st} l_2 \leq_{st} l_3$, donde el orden de los extremos s y t es irrelevante. Es fácil ver que esto sucede si y sólo si l_1 y l_3 están en semiplanos distintos respecto a l_2 .

Ahora conviene adoptar el convenio siguiente:

Definición 1.43 Llamaremos *ángulos llanos* a los semiplanos. Extendemos la congruencia de ángulos a los ángulos llanos estipulando que todos ellos son congruentes entre sí y no son congruentes con ningún ángulo en sentido estricto. Extendemos la relación de orden estipulando que un ángulo llano es mayor que cualquier ángulo en sentido estricto.

Notemos que un ángulo llano no tiene definidos un vértice y unos lados. Podemos considerar como tales a un punto cualquiera de su frontera y las semirrectas que éste determina, pero hay infinitos ángulos llanos con el mismo vértice y los mismos lados (todos los semiplanos con una misma recta como frontera). Los ángulos llanos no tienen suplementario. Si s y t son los extremos de un semihaz de semirrectas contenidas en el semiplano S , entonces convendremos en que \widehat{st} representa precisamente a S (pero esto sólo tiene sentido con respecto a un semihaz prefijado). Con este convenio se cumple el teorema siguiente (la demostración es muy simple):

Teorema 1.44 Sean l_1 y l_2 dos semirrectas distintas en un semihaz prefijado. Entonces $\widehat{l_1 l_2}$ es la unión de todas las semirrectas del semihaz que están entre l_1 y l_2 .

Todas estas consideraciones nos permiten estudiar cómodamente la suma de ángulos.

Definición 1.45 Diremos que un ángulo $A = \widehat{l_1 l_2}$ es la *suma* de dos ángulos B y C si existe una semirrecta l_3 de origen el vértice de A y contenida en A tal que $B \equiv \widehat{l_1 l_3}$ y $C \equiv \widehat{l_3 l_2}$. Lo representaremos por $A \equiv B + C$.

Como en el caso de los segmentos, es claro que la relación $A \equiv B + C$ puede verse como una igualdad entre clases de congruencia de los ángulos. Por ejemplo, notemos que si A , B y C cumplen la definición anterior y $A' \equiv A$, entonces $B < A$, luego existe un ángulo B' con un lado en común con A y contenido en A . Por el teorema 1.33, los otros lados de B' y A' forman un ángulo C' congruente con C , luego A' también es una suma de B y C . Igualmente se prueba que todas las sumas de B y C son congruentes.

Convenimos en que un ángulo llano A es la suma de dos ángulos B y C si y sólo si éstos son suplementarios. Observemos que la definición general de suma es aplicable a este caso tomando como vértice de A cualquier punto de su frontera.

La suma de ángulos presenta una diferencia importante con la de segmentos, y es que no todo par de ángulos tiene una suma. Concretamente:

Teorema 1.46 Dos ángulos B y C admiten una suma si y sólo si C es menor o igual que el suplementario de B .

DEMOSTRACIÓN: Sea A una suma de B y C . Es claro que existe un semihaz de semirrectas contenidas en el semiplano de A y con un extremo l_1 igual a un lado de A (incluso si A es un semiplano). Sea l_3 el otro lado de A . Por definición de suma, existe una semirrecta l_2 entre l_1 y l_3 de modo que $\widehat{l_1 l_2} \equiv B$ y $\widehat{l_2 l_3} \equiv C$.

Si llamamos l_4 a la semirrecta complementaria de l_1 tenemos

$$l_1 \leq_{l_1 l_4} l_2 \leq_{l_1 l_4} l_3 \leq_{l_1 l_4} l_4,$$

luego $C \equiv \widehat{l_2 l_3} \leq \widehat{l_2 l_4}$, y este último ángulo es el suplementario de B .

Recíprocamente, si C es menor o igual que el suplementario de B , entonces podemos tomar $B = \widehat{l_1 l_2}$, donde l_1 es un extremo de un semihaz de semirrectas al cual pertenece l_2 . Por hipótesis $C \leq \widehat{l_4 l_2}$, donde l_4 es la semirrecta complementaria de l_1 , luego existe una semirrecta l_3 entre l_2 y l_4 tal que $C \equiv \widehat{l_2 l_3}$. Es claro entonces que $\widehat{l_1 l_3}$ es una suma de B y C . ■

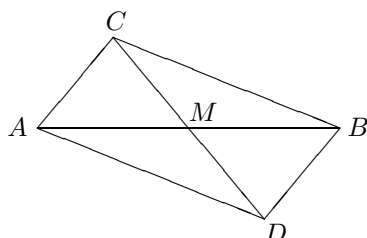
Se comprueba sin dificultad que la suma de ángulos es asociativa, es decir, si $A + B$ es sumable con C , entonces A también es sumable con $B + C$ y $(A + B) + C \equiv A + (B + C)$. Así mismo es conmutativa y simplificable. Si $B < C$ entonces existe un único ángulo D (salvo congruencia) tal que $C \equiv B + D$. Lo representaremos por $D \equiv C - B$.

1.6 Más propiedades de segmentos, ángulos y triángulos

Con los resultados de que disponemos hasta el momento ya podemos probar con cierta agilidad muchas propiedades intuitivamente evidentes acerca de segmentos, ángulos y triángulos. Recogemos aquí las que nos harán falta más adelante para estudiar la perpendicularidad, el paralelismo y las circunferencias, entre otras nociones.

Teorema 1.47 *Todo segmento \overline{AB} contiene un único punto M que cumple $\overline{AM} \equiv \overline{MB}$. Se le llama punto medio del segmento.*

DEMOSTRACIÓN: Sea C un punto cualquiera fuera de la recta AB . El triángulo \widehat{ACB} es congruente con un único triángulo \widehat{BDA} contenido en el semiplano complementario del que contiene al primero.



Los puntos C , A y D no pueden ser colineales, pues entonces los ángulos \widehat{CAB} y \widehat{BAD} serían suplementarios, luego también lo serían los ángulos (iguales a éstos) \widehat{ABD} y \widehat{CBA} , luego C , B y D serían también colineales, y así los cuatro puntos estarían alineados, en contra de la elección de C .

Tenemos, pues, dos triángulos \widehat{ACD} y \widehat{BDC} , que tienen sus lados iguales, luego también sus ángulos. Además \overline{CD} corta a la recta AB en un punto M (porque C y D están en semiplanos distintos). Los triángulos \widehat{ACM} y \widehat{BDM} son iguales por el criterio LAL , luego $\overline{AM} \equiv \overline{MB}$. El punto M ha de estar entre A y B , pues en caso contrario habría dos segmentos iguales con extremo M y el otro extremo al mismo lado de M .

La unicidad es fácil de probar: Si hubiera dos puntos medios M_1 y M_2 , podemos suponer que M_1 está entre A y M_2 . Entonces

$$\overline{AM_2} = \overline{AM_1} + \overline{M_1 M_2},$$

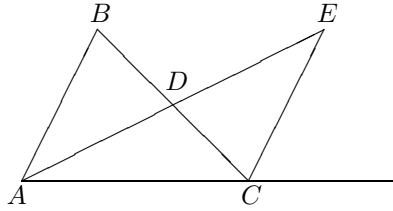
luego $\overline{AB} = 2\overline{AM_2} = 2\overline{AM_1} + 2\overline{M_1M_2} = \overline{AB} + 2\overline{M_1M_2}$, lo cual es absurdo. ■

Teorema 1.48 *Dado un ángulo $\widehat{l_1l_2}$, existe una única semirrecta l contenida en él, tal que $\widehat{l_1l} \equiv \widehat{ll_2}$. Se la llama bisectriz del ángulo.*

DEMOSTRACIÓN: Sea O el vértice del ángulo. Tomemos un punto A en l_1 y sea B en l_2 tal que $OA \equiv OB$. Sea M el punto medio de AB y l la semirrecta de origen O y que pasa por M . Claramente l está contenida en el ángulo. Los triángulos $\triangle OAM$ y $\triangle OBM$ tienen los tres lados iguales, luego también los ángulos. En particular $\widehat{l_1l} \equiv \widehat{ll_2}$. La unicidad se prueba como en el caso de los segmentos o, alternatively, se prueba que una bisectriz ha de pasar por el punto medio de AB . ■

Teorema 1.49 *Todo ángulo de un triángulo es menor que el suplementario de cualquier otro de los ángulos.*

DEMOSTRACIÓN: Sea el triángulo $\triangle ABC$. Vamos a probar que el suplementario de \hat{C} es mayor que \hat{B} . Sea D el punto medio del lado \overline{BC} . Consideremos la semirrecta \overrightarrow{AD} y, sobre ella, sea E el punto que cumple $\overline{AD} \equiv \overline{DE}$. Los triángulos $\triangle ABD$ y $\triangle ECD$ son iguales por el criterio LAL , luego $\widehat{DCE} \equiv \hat{B}$ y está contenido en el ángulo adyacente a \hat{C} . ■



El teorema siguiente generaliza al criterio del triángulo isósceles.

Teorema 1.50 *Los ángulos de un triángulo satisfacen las mismas desigualdades que sus respectivos lados opuestos.*

DEMOSTRACIÓN: Sea $\triangle ABC$ un triángulo y supongamos, por ejemplo, que $\overline{BC} < \overline{AB}$. Entonces existe un punto D en \overline{AB} tal que $\overline{BD} \equiv \overline{BC}$.

Entonces $\hat{A} = \widehat{CAD} < \widehat{CDB}$, por el teorema anterior, pues el segundo es el suplementario de un ángulo de $\triangle ADC$; a su vez $\widehat{CDB} = \widehat{DCB}$, porque el triángulo $\triangle DCB$ es isósceles, y por último $\widehat{DCB} < \widehat{ACB} = \hat{C}$.

El recíproco es obvio: Si $\hat{A} < \hat{C}$ no puede ser $\overline{AB} < \overline{BC}$ por la parte ya probada, y tampoco puede darse la igualdad por el criterio del ángulo isósceles. ■

En particular tenemos que un triángulo es equilátero si y sólo si tiene sus tres ángulos iguales, es isósceles si y sólo si tiene dos ángulos iguales y es escaleno si y sólo si tiene sus tres ángulos desiguales.

1.7 Perpendiculares

Definición 1.51 Un ángulo es *recto* si es su propio suplementario.

Teorema 1.52 *Existen ángulos rectos.*

DEMOSTRACIÓN: Sea $\widehat{l_1 l_2}$ un ángulo cualquiera. Sea O su vértice, sea A un punto en l_1 y sea B el punto de l_2 que cumple $\overline{OA} \equiv \overline{OB}$. Sea M el punto medio de \overline{AB} . Entonces los ángulos \widehat{OMA} y \widehat{OMB} son adyacentes y por otra parte son iguales, pues los triángulos correspondientes tienen los lados iguales. Por consiguiente ambos son ángulos rectos. ■

Es obvio que todos los ángulos rectos son congruentes entre sí y que todo ángulo congruente con un ángulo recto es recto. La existencia de ángulos rectos generaliza el teorema 1.48 al caso de ángulos llanos.

Definición 1.53 Dos rectas son *perpendiculares* (lat. *perpendicularum* = ‘plomada’) si son secantes y uno de los ángulos que forman—y por consiguiente los cuatro— es recto. Dos semirrectas o segmentos son perpendiculares si lo son sus prolongaciones.

Un ángulo es *agudo* (lat. ‘con punta’) si es menor que un ángulo recto. Un ángulo es *obtuso* (lat. ‘sin punta’) si es mayor que un ángulo recto.

Es claro que el suplementario de un ángulo agudo es un ángulo obtuso y viceversa. El teorema 1.49 implica que todo triángulo tiene al menos dos ángulos agudos, pues si tiene uno obtuso su suplementario es agudo, y los otros dos son menores que éste. Esto nos permite clasificar los triángulos en *obtusángulos*, *rectángulos* y *acutángulos* según si tienen, respectivamente, un ángulo obtuso, un ángulo recto o si todos sus ángulos son agudos. En un triángulo rectángulo, los lados perpendiculares se llaman *catetos* (gr. ‘perpendiculares’) y el lado situado bajo el ángulo recto se llama *hipotenusa* (gr. ‘tendido por debajo’).

Teorema 1.54 *Dada una recta r y un punto P contenidos en un plano π , existe una única recta perpendicular a r que pasa por P y está contenida en π .*

DEMOSTRACIÓN: Si el punto P está en r es evidente, pues basta transportar un ángulo recto sobre una de las semirrectas que P determina en r . La unicidad también es clara. Supongamos ahora que P no está en r .

Sea A un punto de r , sea P' el único punto del semiplano de π opuesto al que contiene a P y que cumple que el ángulo que r forma con $\overrightarrow{AP'}$ es igual al que forma con \overrightarrow{AP} así como que $\overline{AP} \equiv \overline{AP'}$.

Si P , A y P' están alineados entonces la recta PP' forma con r dos ángulos adyacentes iguales, luego es perpendicular a r y pasa por P . Si no están alineados entonces la recta PP' corta a r en un punto B distinto de A . Los triángulos $\triangle ABP$ y $\triangle ABP'$ son iguales, luego también lo son los ángulos \widehat{ABP} y $\widehat{ABP'}$, que además son adyacentes. Por lo tanto la recta PP' es perpendicular a r y pasa por P . Si hubiera dos perpendiculares a r que pasaran por P , formarían un triángulo con dos ángulos rectos, lo cual es imposible. ■

Observemos que si P es un punto exterior a una recta r y Q es el punto donde la perpendicular a r por P corta a r , entonces \overline{PQ} es menor que \overline{PR} para cualquier otro punto R de r . En efecto, el triángulo \overline{PQR} es rectángulo, y por el teorema 1.50 el lado mayor es la hipotenusa \overline{PR} .

Teorema 1.55 *Todo lado de un triángulo es menor que la suma de los otros dos y mayor que su diferencia.*

DEMOSTRACIÓN: Consideremos un triángulo \widehat{ABC} y tracemos la perpendicular por A a la recta BC . Distinguiamos tres casos.

1) Si la perpendicular corta a BC en B o en C (por ejemplo en B), entonces el triángulo es rectángulo y \overline{BC} es menor que la hipotenusa \overline{AC} , y en particular es menor que $\overline{AC} + \overline{AB}$.

2) Si la perpendicular corta a BC fuera del segmento \overline{BC} , digamos en un punto P de modo que B está entre P y C , entonces $\overline{BC} < \overline{PC}$, que es el cateto de un triángulo rectángulo de hipotenusa \overline{AC} , luego $\overline{BC} < \overline{AC} < \overline{AC} + \overline{AB}$.

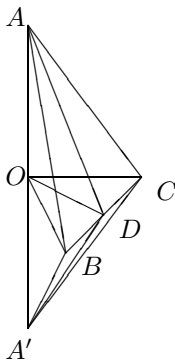
3) Si la perpendicular corta a BC en un punto X entre B y C , entonces tenemos dos triángulos rectángulos \widehat{AXB} y \widehat{AXC} , de hipotenusas \overline{AB} y \overline{AC} , luego

$$\overline{BC} \equiv \overline{BX} + \overline{XC} < \overline{AB} + \overline{AC}.$$

La segunda propiedad es consecuencia de la primera. Si llamamos a , b y c a los tres lados y, por ejemplo, $b \leq c$, entonces $a + b > c$ implica $a > c - b$. ■

Ahora pasamos a ocuparnos de la perpendicularidad entre rectas y planos. El resultado básico es el siguiente:

Teorema 1.56 *Si una recta es perpendicular a dos rectas contenidas en un plano, entonces es perpendicular a todas las rectas contenidas en dicho plano y que pasan por el punto de corte.*



DEMOSTRACIÓN: Sea O el punto de corte entre la recta y el plano. Sea A otro punto de la recta y A' el simétrico de A respecto a O (es decir, el que cumple $\overline{AO} \equiv \overline{OA'}$). Consideremos una recta cualquiera contenida en el plano que pase por O . Fijemos en particular una de sus semirrectas de origen O . Ésta estará contenida en uno de los cuatro ángulos en que las rectas de la hipótesis dividen al plano. Digamos que este ángulo es \widehat{BOC} . Entonces la semirrecta corta al segmento \overline{BC} en un punto D . En estos términos tenemos que los ángulos \widehat{AOB} y \widehat{AOC} son rectos, y queremos probar que \widehat{AOD} también lo es.

Los triángulos \widehat{AOC} y $\widehat{A'OC}$ son iguales por el criterio LAL y lo mismo sucede con \widehat{AOB} y $\widehat{A'OB}$. Esto implica que los triángulos \widehat{ABC} y $\widehat{A'BC}$ también son iguales, por el criterio LLL. De aquí pasamos a que los triángulos \widehat{ADB} y

$\widehat{A'DB}$ también son iguales, esta vez por el criterio *LAL*, lo que nos da finalmente la igualdad de los triángulos \widehat{AOB} y $\widehat{A'OB}$, pues sus lados son iguales. Esto implica que los ángulos \widehat{AOB} y $\widehat{A'OB}$ son iguales, a la vez que adyacentes, luego son rectos. ■

Como consecuencia inmediata obtenemos:

Teorema 1.57 *La unión de todas las rectas perpendiculares a una recta r que pasan por uno de sus puntos es un plano.*

DEMOSTRACIÓN: Sea P un punto de r . Consideremos dos planos que pasen por r , tracemos en ellos sendas perpendiculares a r por P y sea π el plano que las contiene. Entonces r es perpendicular a dos rectas de π , luego por el teorema anterior es perpendicular a todas las rectas que pasan por P y están contenidas en π . Falta probar que toda recta perpendicular a r por P está contenida en π . Sea s una de estas rectas. Entonces el plano que contiene a r y a s corta a π en una recta que, según sabemos, es perpendicular a r . Como en un mismo plano sólo puede haber una perpendicular, dicha intersección es s , luego s está contenida en π . ■

Definición 1.58 Diremos que una recta es *perpendicular* a un plano π si lo corta en un punto P y es perpendicular a todas las rectas contenidas en π que pasan por P .

El teorema anterior prueba que por cada punto de una recta pasa un único plano perpendicular a la misma. El recíproco es fácil de probar:

Teorema 1.59 *Dado un plano π y un punto A , existe una única recta perpendicular a π que pasa por A .*

DEMOSTRACIÓN: Supongamos primero que A está en π . Entonces consideremos dos rectas perpendiculares contenidas en π que se corten en A . Sus respectivos planos perpendiculares se cortarán en una recta que será perpendicular a las dos elegidas, luego a todas las de π .

Supongamos ahora que A no está en π . Tomemos una recta cualquiera r contenida en π , sea D el punto donde la perpendicular a r por A corta a r . Sean B y C puntos en r situados en semirrectas opuestas respecto a D . Sea s la perpendicular a r por D contenida en π . Sea A' el punto situado en el semiplano de frontera s complementario del que contiene a A y que hace $\overrightarrow{DA'}$ forme el mismo ángulo con s que \overrightarrow{DA} y además $\overline{DA'} \equiv \overline{DA}$. Sea O el punto en que $\overline{AA'}$ corta a s (la figura de la prueba del teorema 1.56 ilustra también la situación actual).

Por construcción $\widehat{ODA} \equiv \widehat{ODA'}$, de donde $\widehat{ADB} \equiv \widehat{A'DB}$ y $\widehat{ADC} \equiv \widehat{AD'C}$ (recordar que ambos son rectángulos). Por lo tanto tenemos que $\widehat{ABO} \equiv \widehat{A'BO}$ y $\widehat{ACO} \equiv \widehat{A'CO}$, luego ambos son rectángulos, y así la recta AA' es perpendicular a OB y OC , luego a π . ■

1.8 El axioma de continuidad, círculos y circunferencias

Para demostrar los hechos básicos sobre círculos y circunferencias necesitamos un axioma adicional que nos garantice la existencia de ciertos puntos de intersección. Se trata del hecho siguiente:

Axioma D *Supongamos que una recta r está dividida en dos partes disjuntas no vacías s_1 y s_2 con la propiedad de que si P y Q son puntos en s_i , entonces $\overline{PQ} \subset s_i$. Entonces s_1 y s_2 son dos semirrectas complementarias (salvo que a una de ellas le falta el origen, que sólo pertenece a la otra).*

Definición 1.60 Dado un plano π , un punto O en π y un segmento r , llamaremos *círculo* de centro O y radio r al conjunto de todos los puntos P de π tales que $\overline{OP} \leq r$. Llamaremos *circunferencia* (lat. circumferre = ‘llevar alrededor’) de centro O y radio r en π al conjunto de todos los puntos P de π tales que $\overline{OP} \equiv r$.

Cada círculo tiene asociada una circunferencia (la del mismo centro y radio), a la que se llama también su *frontera*. Los puntos del círculo que no pertenecen a la circunferencia se llaman *interiores*, mientras que los puntos que no pertenecen al círculo se llaman puntos *exteriores* a él.

También se llama *radio* de un círculo o circunferencia a cualquier segmento que una su centro con uno de los puntos de la circunferencia. Es claro que todos los radios son congruentes entre sí. Un segmento que une dos puntos de una circunferencia se llama *cuerda* de la misma. Una cuerda que pase por el centro se llama *diámetro* (gr. ‘medida transversal’). Es fácil ver que un diámetro es igual a dos veces el radio.

Es fácil ver que cada círculo o circunferencia contiene al menos tres puntos no colineales, con lo que determina el plano que lo contiene, al que llamaremos su *soporte*. Veamos que también determinan su centro y su radio (éste último salvo congruencia):

Dados dos puntos A y B en un plano π , la perpendicular (en π) al segmento \overline{AB} por su punto medio se llama *mediatriz* de \overline{AB} . Es inmediato comprobar que la mediatriz de un segmento \overline{AB} contiene exactamente a los puntos que equidistan de sus extremos, es decir, que cumplen $\overline{AX} \equiv \overline{BX}$. Por lo tanto, si unimos dos puntos de una circunferencia y trazamos la mediatriz del segmento obtenido, ésta pasa por su centro. Si tomamos tres puntos (no colineales) en la circunferencia y hacemos lo mismo con dos pares de ellos, las rectas que obtendremos se cortarán precisamente en el centro, luego éste está unívocamente determinado por la circunferencia. Así mismo, el radio es congruente con cualquier segmento que una el centro con un punto de la circunferencia, luego también está determinado.

Por otra parte, un círculo determina su circunferencia (un punto P de un círculo está en su circunferencia si y sólo si hay un segmento \overline{AB} que lo contiene

de modo que los puntos de \overline{AP} son interiores al círculo y los restantes son exteriores). Concluimos que dos círculos o circunferencias (en un mismo plano) son iguales si y sólo si tienen el mismo centro y sus radios son congruentes.

Estudiamos ahora las intersecciones entre rectas y circunferencias.

Teorema 1.61 *Sea ω un círculo y s una recta contenida en el plano soporte de ω y que pase por un punto interior de ω . Entonces s corta a ω en un segmento \overline{PQ} , y a la circunferencia de ω sólo en los puntos P y Q .*

DEMOSTRACIÓN: Sea O el centro de ω y r su radio. Sea A el punto donde la perpendicular a s por O corta a s . Es claro que $\overline{OA} < \overline{OP}$ para todo punto P de s distinto de A , luego por hipótesis $\overline{OA} < r$.

Sea t una de las semirrectas en que A divide a s . Si s pasa por O el resultado es inmediato. Supongamos lo contrario. Dividamos s en dos partes X e Y . El conjunto X está formado por la semirrecta complementaria de t más aquellos puntos P en t tales que existe un punto Q posterior a P (desde A) de modo que $\overline{OQ} < r$. El conjunto Y está formado por los puntos de s que no están en X . No es vacío, pues cualquier punto B sobre t tal que $\overline{AB} > 2r$ cumple $\overline{OB} > 2r - \overline{OA} > r$ (por el teorema 1.55), luego está en Y .

Por la propia definición de X e Y es inmediato que cumplen las hipótesis del axioma D . Por lo tanto existe un punto C tal que X es la semirrecta \overrightarrow{CA} (salvo quizá el origen C). Es claro que C ha de estar en s .

Veamos que no puede ser $\overline{OC} < r$ ni $\overline{OC} > r$, con lo que tendremos que C está en la circunferencia. Si $\overline{OC} < r$ podemos tomar un punto D en s posterior a C y de modo que $\overline{CD} < r - \overline{OC}$. El teorema 1.55 implica entonces que $\overline{OD} < r$, con lo que todos los puntos entre C y D están en $X = \overrightarrow{CA}$, lo cual es imposible.

Similarmente, si $\overline{OC} > r$ entonces tomamos D en s anterior a C y de modo que $\overline{DC} < \overline{OC} - r$. Entonces llegamos a que $\overline{OD} > r$, luego cualquier punto posterior a D cumple lo mismo, luego D está en Y y esto también es imposible.

Aplicando lo anterior a las dos semirrectas en que A divide a s tenemos que s corta a la circunferencia en dos puntos P y Q .

Notemos ahora que Si M y N son dos puntos de s al mismo lado de A , digamos $A < M < N$, entonces el triángulo \widehat{OMA} es rectángulo, luego \widehat{OMA} es agudo, luego su adyacente \widehat{OMN} es obtuso, luego el teorema 1.50 implica que $\overline{OM} < \overline{ON}$. Esto prueba que sólo P y Q están en la circunferencia y que \overline{PQ} es la intersección de s con ω . ■

Dada una recta s y un círculo ω de centro O y radio r (ambos en el mismo plano), consideramos el punto A donde la perpendicular a s por O corta a s . Como ya hemos notado antes, es claro que $\overline{OA} < \overline{OP}$ para todo punto P de s distinto de A . Por lo tanto, si $\overline{OA} < r$ tenemos que la intersección de s con ω es un segmento \overline{PQ} , de modo que P y Q son los únicos puntos en común entre s y la circunferencia de ω ; si $\overline{OA} = r$ entonces A está en la circunferencia de ω y todos los demás puntos de s son exteriores; y si $\overline{OA} > r$, todos los puntos de s son exteriores.

Definición 1.62 Diremos que una recta r es *secante* (lat. ‘que corta’) a una circunferencia ω si tienen dos puntos en común. Diremos que r es *tangente* (lat. ‘que toca’) si tienen un punto en común y diremos que r es *exterior* a ω si no tienen puntos en común.

Hemos probado que estas definiciones cubren todas las posibilidades. También es claro lo siguiente:

Teorema 1.63 *Por un punto de una circunferencia pasa una única recta tangente. Ésta es concretamente la perpendicular al radio con extremo dicho punto.*

Ejercicio: Probar que los círculos son convexos.

Ejercicio: Probar que una circunferencia no contiene tres puntos colineales.

Ejercicio: Probar que por tres puntos no alineados pasa una única circunferencia.

Ejercicio: Probar que si una recta r es tangente a una circunferencia ω , entonces ω está contenida en un semiplano respecto a r .

Veamos ahora la intersección entre dos circunferencias:

Teorema 1.64 *Sean ω_1 y ω_2 circunferencias contenidas en un mismo plano, de centros respectivos O_1 y O_2 y radios r_1 y r_2 (con $r_1 \leq r_2$). Entonces:*

1. Si $\overline{O_1O_2} > r_1 + r_2$ entonces ω_1 y ω_2 no tienen puntos en común ni sus círculos tampoco.
2. Si $\overline{O_1O_2} = r_1 + r_2$ entonces ω_1 y ω_2 tienen un único punto en común, al igual que sus círculos.
3. Si $r_2 - r_1 < \overline{O_1O_2} < r_1 + r_2$ entonces ω_1 y ω_2 tienen dos puntos en común.
4. Si $\overline{O_1O_2} = r_2 - r_1$ entonces ω_1 y ω_2 tienen un único punto en común, y el círculo de ω_1 está contenido en el de ω_2 .
5. Si $\overline{O_1O_2} < r_2 - r_1$ entonces ω_1 y ω_2 no tienen puntos en común y el círculo de ω_1 está contenido en el de ω_2 .

DEMOSTRACIÓN: Todos los casos son sencillos excepto el tercero. Sea A el punto de la semirrecta $\overrightarrow{O_1O_2}$ tal que $\overline{O_1A} \equiv r_1$ y sea B el punto de la semirrecta $\overrightarrow{O_2O_1}$ tal que $\overline{O_2B} \equiv r_2$. Las hipótesis garantizan que A está en el interior de ω_2 y que B está en el interior de ω_1 . Como los círculos son convexos, el segmento \overline{AB} está en la intersección de ambos. Fijamos un semiplano respecto a la recta O_1O_2 . Para cada punto P de \overline{AB} la semirrecta de origen P , contenida en el semiplano elegido y perpendicular a O_1O_2 parte de un punto en ambos círculos, luego corta a las circunferencias ω_1 y ω_2 en dos puntos C_1 y C_2 , respectivamente. Llamaremos X al conjunto de todos los puntos P de \overline{AB} para los que existe un punto Q tal que $A \leq P < Q < B$ y $\overline{QC_1} \leq \overline{QC_2}$, más los puntos de la semirrecta

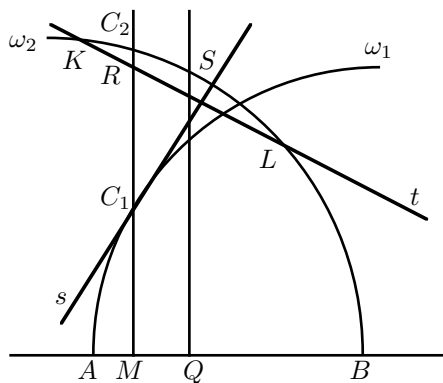
de origen A que no contiene a B . Llamaremos Y al conjunto de puntos de la recta AB que no están en X . El conjunto Y no es vacío, pues contiene todos los puntos posteriores a B desde A .

Es inmediato que se cumplen las hipótesis del axioma D, luego X e Y son semirrectas de origen un punto M , que claramente tiene que estar en \overline{AB} . Vamos a probar que para este M se cumple $C_1 = C_2$. Supongamos que $\overline{MC_1} < \overline{MC_2}$.

Sea s la tangente a ω_1 por C_1 , sea R un punto entre C_1 y C_2 . Sea t una recta que pase por R y corte a la recta AB en cualquier punto distinto de M . Entonces t corta a ω_2 en dos puntos K y L , de modo que los puntos del segmento \overline{KL} distintos de sus extremos son interiores a ω_2 . Sea S un punto de este segmento que esté en el mismo semiplano que B respecto a MR y que sea anterior a un eventual corte entre KL y la tangente s . Tracemos desde S la perpendicular a AB , que cortará a esta recta en un punto Q . Como las rectas MR y SQ tienen una perpendicular común (AB) no pueden cortarse (o formarían un triángulo con dos ángulos rectos). Esto implica que Q está entre M y B . Ahora notamos que el segmento \overline{RS} no corta a s , luego S está en el mismo semiplano que S respecto a s , es decir, en el semiplano opuesto a M , o sea, en el que no está ω_2 . Por lo tanto \overrightarrow{QS} corta a ω_1 antes de llegar a S , pero como S es interior a ω_2 , resulta que \overrightarrow{QS} corta a ω_2 después de llegar a S , o sea, $\overline{MC_{1Q}} < \overline{MC_{2Q}}$. Por consiguiente cualquier punto entre M y Q está en X , lo cual es imposible.

El mismo argumento prueba que si $\overline{MC_2} < \overline{MC_1}$ existe un punto Q entre A y M de modo que todos los puntos entre Q y M cumplen la misma desigualdad, con lo que están en Y y esto también es contradictorio.

Concluimos que existe un punto de intersección entre las dos circunferencias en el semiplano seleccionado en un principio. Cambiando de semiplano obtenemos otro más. La unicidad es inmediata, pues dos puntos de corte en el mismo semiplano darían lugar a dos triángulos distintos con lados iguales, uno de ellos común y ambos en el mismo semiplano respecto a él, lo cual es imposible. ■



Una consecuencia inmediata del teorema anterior es el resultado siguiente sobre existencia de triángulos:

Teorema 1.65 Sean r_1 , r_2 y r_3 tres segmentos que cumplan las desigualdades $r_2 \leq r_3$ y $r_3 - r_2 < r_1 < r_2 + r_3$ (o simplemente $r_1 < r_2 + r_3$ si $r_2 = r_3$). Entonces existe un triángulo de lados r_1 , r_2 y r_3 .

DEMOSTRACIÓN: Tomamos un segmento \overline{AB} igual a r_1 . Por el teorema anterior, la circunferencia de centro A y radio r_1 corta a la circunferencia de centro B y radio r_2 en un punto C que nos da el triángulo buscado. ■

Ejercicio: Usar el axioma D para probar que por un punto exterior a una circunferencia pasan dos tangentes a la misma.

Capítulo II

Medida de segmentos, ángulos y arcos

Dedicamos este capítulo a introducir y desarrollar unos conceptos muy intuitivos, pero también muy técnicos. De hecho se trata de un punto en el que la intuición ‘engaña’ un poco, pues una cuestión aparentemente muy simple encierra una sutileza. Se trata del concepto de medida de un segmento. En la concepción de la geometría que tenían los griegos representaba un papel muy importante la noción de proporción o razón entre dos segmentos. La idea es que la razón de dos segmentos u y v es, por ejemplo, $4 : 7$ si al dividir el segundo en siete partes iguales y sumar cuatro de estas partes obtenemos un segmento igual al primero. En términos de medidas esto significa que si tomamos a v como unidad de medida, entonces u mide $4/7$ unidades. Los griegos suponían que todo segmento puede medirse de este modo con respecto a una unidad prefijada o, dicho de otro modo, que dados dos segmentos siempre guardan una determinada razón entre sí. Aquí veremos que esto no es así, cosa que los griegos nunca llegaron a asimilar.

2.1 Longitud de segmentos. Números reales

El primer paso para detallar las ideas que acabamos de exponer es, naturalmente, probar que es posible dividir segmentos en partes iguales.

Teorema 2.1 *Para todo número natural $n \geq 2$, todo segmento se puede dividir en n partes iguales.*

DEMOSTRACIÓN: Sea \overline{AB} un segmento contenido en una recta r . Llamamos X al conjunto formado por la semirrecta complementaria de \overline{AB} más los puntos P de \overline{AB} tales que $n\overline{AP} < \overline{AB}$. Sea Y el conjunto de los puntos de r que no están en X . Es obvio que se cumplen las hipótesis del axioma D, luego existe un punto C en r tal que X e Y son semirrectas de origen C . Veamos que $n\overline{AC} \equiv \overline{AB}$.

Supongamos que $n\overline{AC} < \overline{AB}$. Sea entonces D un punto entre A y B tal que $n\overline{AC} \equiv \overline{AD}$. Sea m un número natural tal que $2^m > n$. Sea u el segmento que resulta de dividir m veces por la mitad el segmento \overline{DB} . Entonces tenemos $nu < 2^m u = \overline{DB}$. En particular $u < \overline{CB}$, luego podemos tomar un punto E en \overline{CB} tal que $\overline{CE} \equiv u$. Así,

$$n\overline{AE} \equiv n\overline{AC} + n\overline{CE} < \overline{AD} + \overline{DB} \equiv \overline{AB}.$$

Esto significa que E está en X , pero es posterior a C , lo cual es imposible. De modo similar se prueba que $n\overline{AC} > \overline{AB}$ lleva a contradicción. Por lo tanto $n\overline{AC} = \overline{AB}$. ■

Aunque la medida de ángulos la abordaremos cuando hayamos acabado con la de segmentos, la similitud de las pruebas aconseja incluir aquí este teorema:

Teorema 2.2 *Para todo número natural $n \geq 2$, todo ángulo se puede dividir en n partes iguales.*

DEMOSTRACIÓN: El teorema es cierto incluso para ángulos llanos, pero podemos reducir este caso al de ángulos menores del modo siguiente: Para dividir un ángulo llano en n partes dividimos un ángulo recto en n partes y tomamos el doble del resultado. Así pues, partamos de un ángulo en sentido estricto $L = \widehat{AOB}$. Cada punto P del segmento \overline{AB} distinto de A determina un ángulo $L_P = \widehat{AOP}$ contenido en L y, recíprocamente, todo ángulo contenido en L con un lado igual a \overrightarrow{OA} es de la forma L_P . Además, si $A < P < Q \leq B$, entonces el ángulo L_Q contiene a \overrightarrow{OP} , luego es mayor que L_P . Esto implica, más en general, que $P <_{AB} Q$ si y sólo si $L_P < L_Q$.

A partir de aquí la prueba sigue el mismo argumento que la del teorema anterior. Tomamos como X el conjunto de los puntos de la semirrecta complementaria a \overline{AB} más los puntos P de \overline{AB} tales que $nL_P < L$. El conjunto Y está formado por los puntos de la recta AB que no están en X , bien porque $nL_P \geq L$, bien porque L_P no se puede sumar n veces consigo mismo.

Es fácil ver que se cumplen las hipótesis del axioma D , con lo que obtenemos un punto C de manera que X es una semirrecta de origen C . Veamos que existe nL_C .

Dividiendo por la mitad el suplementario de L un número suficiente de veces (los detalles son los mismos que en el teorema anterior) encontramos un ángulo M tal que nM sea menor que dicho suplementario (y podemos exigir que sea menor que L_C). Existe un punto $P <_{AB} C$ tal que $L_C - M \equiv L_P$, luego P está en X y por lo tanto existe $n(L_C - M) < L$, y como nM es menor que el suplementario de L , también existe $n(L_C - M) + nM \equiv nL_C$. A partir de aquí, la prueba de que $nL_C \equiv L$ es formalmente idéntica a la del teorema anterior. ■

La definición de la medida de un segmento descansa fuertemente en la propiedad siguiente, que nos permitirá eludir la existencia de una razón respecto a una unidad.

Teorema 2.3 (Propiedad de Arquímedes) *Para todo par de segmentos u y v existe un número natural n tal que $nu > v$.*

DEMOSTRACIÓN: Podemos suponer que $u = \overline{AB}$ y $v = \overline{AC}$, así como que ambos están sobre una misma semirrecta s de origen A . Supongamos que el resultado es falso, es decir, que $nu < v$ para todo número natural n (observemos que si se diera la igualdad para un n , entonces $n+1$ cumpliría el teorema. Para cada n , sea A_n el punto de s que cumple que $\overline{AA_n} \equiv nu$. Llamemos X al conjunto de todos los puntos de la semirrecta complementaria a s más los P puntos de s que cumplen $P <_{AB} A_n$ para algún n . Sea Y el conjunto de los puntos de s que no están en X (por ejemplo C). Es evidente que se cumplen las hipótesis del axioma D, luego existe un punto D en s de modo que X e Y son las semirrectas de origen D . Sea E un punto de s tal que $\overline{ED} \equiv u$. Entonces $E < ABD$, luego E ha de estar en X . Por consiguiente existe un número n tal que $E < ABA_n$, lo que significa que $\overline{AE} < \overline{AA_n} \equiv nu$, de donde $\overline{AD} \equiv \overline{AE} + \overline{ED} < nu + u = (n+1)u \equiv \overline{AA_{n+1}}$, pero entonces A_{n+1} es un punto de X posterior a D , lo cual es contradictorio. ■

La propiedad de Arquímedes puede leerse como que todo segmento puede hacerse arbitrariamente grande sumándolo consigo mismo un número suficiente de veces, pero también implica claramente que todo segmento se puede hacer arbitrariamente pequeño dividiéndolo en un número suficiente de partes iguales. A su vez de aquí se sigue que si dividimos un segmento en partes suficientemente pequeñas y sumamos un número adecuado de éstas, podemos formar un segmento igual a cualquier otro prefijado con un error menor que la longitud de las partes que empleamos, es decir, con un error tan pequeño como se desee. Nuestra definición de medida de segmentos se basará en este hecho.

Definición 2.4 Dado un segmento u y un número racional positivo p/q , llamaremos $(p/q)u$ al segmento que resulta de dividir u en q partes iguales y sumar p de ellas.

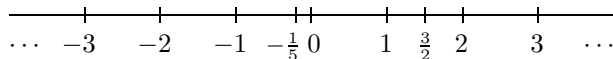
Claramente $(p/q)u$ está definido salvo congruencia. La simplificabilidad de la suma de segmentos permite probar por argumentos puramente algebraicos que la definición no depende de la fracción escogida como representante del número racional, así como que se cumplen las propiedades siguientes:

1. $(rs)u \equiv r(su)$, $r(u+v) \equiv ru+rv$, $(r+s)u \equiv ru+su$,
2. Si $ru \equiv rv$ entonces $u \equiv v$,
3. Si $ru \equiv su$, entonces $r = s$,
4. Si $r < s$ entonces $ru < su$,
5. Si $u < v$ entonces $ru < rv$,

para todos los números racionales positivos r , s y todos los segmentos u y v .

En estos términos, que la razón entre dos segmentos u y v sea un número racional positivo r significa que $u = rv$. Con esto podemos hacer una primera

aproximación al problema de la medida. Tomemos una recta cualquiera y en ella fijemos dos puntos arbitrarios P_0 y P_1 . Entonces a cada número racional positivo r le podemos asignar unívocamente un punto P_r de la semirrecta $\overrightarrow{P_0P_1}$, a saber, el único que cumple $\overline{P_0P_r} \equiv r \overline{P_0P_1}$. Es útil convenir en asignar a los números racionales negativos puntos en la semirrecta complementaria, de modo que si $r < 0$ entonces P_r está determinado por la relación $\overline{P_0P_r} \equiv -r \overline{P_0P_1}$.



Es importante notar que esta asignación de números racionales a algunos puntos de una recta depende de la elección arbitraria de los puntos P_0 y P_1 o, en otros términos, de la elección de P_0 , de la unidad de medida $u = \overline{P_0P_1}$ y de la orientación de la recta (es decir, de la semirrecta que tomamos como positiva). Cuando en una recta hemos hecho estas elecciones, diremos que tenemos una *recta graduada*. Si en una recta graduada consideramos el orden para el cual $P_0 < P_1$, entonces el orden en \mathbb{Q} se corresponde claramente con el orden de la recta.

El problema es que no todo punto de la recta tiene asignado un número racional. Sin embargo tenemos lo siguiente:

Teorema 2.5 *Si P y Q son dos puntos distintos de una recta graduada, existe un número racional r tal que $P < P_r < Q$.*

DEMOSTRACIÓN: Si P y Q están en semirrectas distintas respecto a P_0 es obvio. Podemos suponer que $P_0 < P < Q$. El caso contrario es análogo. Sea $u = \overline{P_0P_1}$ y $v = \overline{P_0Q}$. Por la propiedad de Arquímedes existe un número natural n tal que $u < nv$ o, equivalentemente, $(1/n)u < v$. De nuevo por la propiedad de Arquímedes existe un número m tal que $(m/n)u > \overline{P_0P}$. Podemos tomar el mínimo que cumpla esto. Entonces

$$\frac{m-1}{n} u \leq \overline{OP}, \quad \text{luego} \quad \frac{m}{n} u < \overline{P_0P} + \frac{1}{n} u < \overline{P_0P} + \overline{PQ} = \overline{P_0Q}.$$

Por consiguiente si llamamos $r = m/n$ tenemos que $\overline{P_0P} < P_r < \overline{P_0Q}$, es decir, $P < P_r < Q$. ■

La idea clave para definir la medida de un segmento es la siguiente: medir un segmento s con respecto a una unidad u significa cuantificar cómo es de grande s supuesto que sabemos cómo es de grande u . En el mejor de los casos esta información puede codificarse con un número racional, pero si no es así, para conocer el tamaño relativo de un segmento s con respecto a v es suficiente saber qué números racionales r hacen que $ru < s$ y qué números racionales hacen $ru \geq s$. Equivalentemente, si situamos a s en una recta graduada con la unidad u y con un extremo en P_0 , el problema es saber entre qué puntos P_r se encuentra el otro extremo de s .

En términos puramente conjuntistas, para conocer exactamente la posición en una recta graduada de un punto P es suficiente conocer el conjunto

$$\alpha_P = \{r \in \mathbb{Q} \mid P_r < P\}.$$

Ahora veremos que podemos tratar a cada uno de estos conjuntos de números racionales como a un solo número. La definición siguiente recoge las propiedades esenciales de los conjuntos α_P .

Definición 2.6 Una *sección inicial abierta* de \mathbb{Q} es un subconjunto $\alpha \subset \mathbb{Q}$ que cumpla:

1. Si $r \in \alpha$ y $s \leq r$ entonces $s \in \alpha$.
2. Si $r \in \alpha$ existe un $s \in \alpha$ tal que $r < s$.

Llamaremos $\overline{\mathbb{R}}$ al conjunto de todas las secciones iniciales abiertas en \mathbb{Q} .

Es fácil probar que los conjuntos α_P que hemos definido son secciones iniciales abiertas que cumplen dos propiedades más: no son vacías (pues hay puntos anteriores a P) y no contienen a todos los números racionales (pues hay puntos posteriores).

Definimos $\pm\infty \in \overline{\mathbb{R}}$ como $-\infty = \emptyset$ y $+\infty = \mathbb{Q}$. Llamaremos conjunto de los números reales a $\mathbb{R} = \overline{\mathbb{R}} \setminus \{\pm\infty\}$.

Teorema 2.7 Se cumplen las propiedades siguientes:

1. $\overline{\mathbb{R}}$ es un conjunto totalmente ordenado por la inclusión con mínimo $-\infty$ y máximo $+\infty$.
2. Todo subconjunto de $\overline{\mathbb{R}}$ tiene supremo e ínfimo.
3. \mathbb{R} es un conjunto totalmente ordenado sin máximo ni mínimo.
4. Si un subconjunto no vacío de \mathbb{R} está acotado superiormente tiene supremo, y si está acotado inferiormente tiene ínfimo.

DEMOSTRACIÓN: 1) La inclusión es claramente un orden parcial. Sólo falta ver que en este caso es total.

Sean $\alpha, \beta \in \overline{\mathbb{R}}$. Supongamos $\alpha \neq \beta$, por ejemplo, supongamos que existe $b \in \beta \setminus \alpha$.

Si $a \in \alpha$ entonces $a < b$ (o si no $b \in \alpha$), de donde $a \in \beta$.

Esto prueba que $\alpha \subset \beta$, es decir, $\alpha \leq \beta$. Por lo tanto $\overline{\mathbb{R}}$ está totalmente ordenado.

2) Sea S un subconjunto de $\overline{\mathbb{R}}$. Es inmediato comprobar que

$$\bigcup_{\alpha \in S} \alpha \in \overline{\mathbb{R}}$$

y es obviamente el supremo de S . El ínfimo de S no es sino el supremo del conjunto de sus cotas inferiores.

3) Si $\alpha \in \mathbb{R}$ entonces $\alpha \neq +\infty$, luego existe un número racional $r \in \mathbb{Q} \setminus \alpha$. Es fácil ver que si $r < s \in \mathbb{Q}$, el conjunto $\beta = \{t \in \mathbb{Q} \mid t < s\}$ es un número real tal que $\alpha < \beta$. Por lo tanto \mathbb{R} no tiene máximo.

Si $\alpha \in \mathbb{R}$, entonces $\alpha \neq -\infty$, luego existe un número racional $r \in \alpha$. Si $s \in \mathbb{Q}$ y $s < r$, entonces el conjunto $\beta = \{t \in \mathbb{Q} \mid t < s\}$ es un número real tal que $\beta < \alpha$. Por lo tanto \mathbb{R} no tiene mínimo.

4) Un subconjunto de \mathbb{R} no vacío y acotado superiormente tiene supremo en $\overline{\mathbb{R}}$, como no es vacío el supremo no es $-\infty$, como tiene una cota en \mathbb{R} tampoco es $+\infty$, luego tiene supremo en \mathbb{R} . Análogamente con ínfimos. ■

Ahora podemos probar:

Teorema 2.8 *La aplicación que a cada punto P de una recta graduada le asigna el número real α_P es una biyección entre la recta y el conjunto de los números reales. Además conserva el orden.*

DEMOSTRACIÓN: Veamos que α_P es un número real. En efecto: si $r \in \alpha_P$ y $s \leq r$ entonces $P_s \leq P_r < P$, luego $s \in \alpha_P$.

Si $r \in \alpha_P$ entonces $P_r < P$, luego hay un número s tal que $P_r < P_s < P$, luego $r < s$ y $s \in \alpha_P$.

Como existen puntos a la izquierda de P , también existen números racionales r tales que $P_r < P$, con lo que $\alpha_P \neq \emptyset$, similarmente existen números r tales que $P < P_r$, con lo que $r \notin \alpha_P$, luego $\alpha_P \neq \mathbb{Q}$.

Si $P < Q$ existe un número racional r tal que $P < P_r < Q$. Esto se traduce en que $r \in \alpha_Q$ y $r \notin \alpha_P$, luego $\alpha_P \neq \alpha_Q$ y es obvio que $\alpha_P \subset \alpha_Q$. Por lo tanto $\alpha_P < \alpha_Q$. Esto prueba en particular que la correspondencia es inyectiva.

Sea ahora un número real α . Vamos a probar que tiene un punto asociado en la recta. Sea X el conjunto de todos los puntos Q de la recta graduada tales que $P < P_r$ para algún $r \in \alpha$. Sea Y el conjunto de todos los puntos de la recta que no están en X . Es fácil ver que podemos aplicar el axioma D y obtener un punto P que claramente cumple $\alpha = \alpha_P$. ■

Consideremos una recta graduada y un número racional r . El punto P_r tiene asignado por una parte el número real r y por otra parte el número real formado por todos los números racionales menores que r . Podemos conciliar esta duplicidad identificando ambos números de acuerdo con el teorema siguiente:

Teorema 2.9 *La aplicación $i : \mathbb{Q} \longrightarrow \mathbb{R}$ dada por*

$$i(r) = \{s \in \mathbb{Q} \mid s < r\}$$

es inyectiva y conserva el orden. Si identificamos \mathbb{Q} con su imagen en \mathbb{R} , entre dos números reales hay siempre un número racional.

DEMOSTRACIÓN: Es inmediato que $i(r) \in \mathbb{R}$. Si $r, s \in \mathbb{Q}$ con $r < s$, entonces existe un $t \in \mathbb{Q}$ tal que $r < t < s$, de donde resulta que $t \in i(s) \setminus i(r)$. Como obviamente $i(r) \subset i(s)$, concluimos que $i(r) < i(s)$.

Esto prueba que i es inyectiva y creciente.

Si $\alpha < \beta$, existirá $b \in \beta \setminus \alpha$. Sean $r, s \in \beta$ tales que $b < r < s$. Entonces todo $a \in \alpha$ es menor que r , luego $\alpha \leq i(r)$. Así mismo todo $t < r$ está en β , luego $i(r) \leq \beta$.

Las desigualdades son estrictas pues $b \in i(r) \setminus \alpha$ y $s \in \beta \setminus i(r)$. ■

En lo sucesivo consideraremos a los números racionales como parte de los números reales a través de la aplicación que acabamos de definir. Ahora podemos generalizar la noción de razón entre dos segmentos de modo que sea aplicable a cualquier par de ellos.

Definición 2.10 Dado un segmento $u = \overline{AB}$ y un número real $\alpha > 0$, definimos el segmento αu como el segmento $\overline{AP_\alpha}$, donde P_α es el punto asociado a α en la prolongación de u cuando la graduamos tomando origen A , unidad u y la orientación de modo que B sea positivo.

Si en particular $v = ru$ para un número racional $r > 0$, tenemos que en una recta graduada $v \equiv \overline{P_0 P}$, donde $\alpha_P = \{s \in \mathbb{Q} \mid s < r\}$. Esto significa que $P_s < P$ si y sólo si $s < r$, lo que sólo es posible si $P = P_r$ (por el teorema 2.5). Por la definición de P_r tenemos que $v \equiv ru$ en el sentido de 2.4. Acabamos de probar que ambas definiciones coinciden sobre los números racionales.

Es inmediato que si $0 < \alpha < \beta$ entonces $\alpha u < \beta u$. Junto con lo anterior, esto permite probar fácilmente que la definición de αu depende sólo de la clase de congruencia de u . En efecto, si $u \equiv v$ pero $\alpha u < \alpha v$, existe un número racional r tal que $\alpha u < ru < \alpha v$, pero entonces $\alpha < r$, luego $\alpha v < rv \equiv ru$, lo cual es absurdo.

Teorema 2.11 Para todos los segmentos u, v y todos los números reales positivos α, β se cumple:

1. Si $\alpha < \beta$ entonces $\alpha u < \beta u$,
2. Si $u < v$ entonces $\alpha u < \alpha v$,
3. $\alpha(u + v) \equiv \alpha u + \alpha v$.

DEMOSTRACIÓN: La primera propiedad ya está probada (es otra forma de expresar que la correspondencia entre números reales y puntos de una recta conserva el orden). La segunda es una consecuencia de la tercera, pues si $u < v$ entonces $v \equiv u + w$, luego $\alpha u < \alpha u + \alpha w \equiv \alpha v$.

Para probar la tercera propiedad observemos en general que si $\alpha u < v$ entonces v será de la forma $v \equiv \beta u$, con $\beta > \alpha$, luego existirá un $r > \alpha$ (y menor que β) de modo que $ru < v$, e igualmente si las desigualdades son opuestas.

Supongamos que $\alpha(u + v) < \alpha u + \alpha v$. Entonces existe un número racional $r > \alpha$ tal que $r(u + v) < \alpha u + \alpha v$, pero $r(u + v) \equiv ru + rv > \alpha u + \alpha v$. Si por el contrario $\alpha(u + v) > \alpha u + \alpha v$ existirá un número racional $r < \alpha$ tal que $\alpha u + \alpha v < r(u + v)$, pero $r(u + v) \equiv ru + rv < \alpha u + \alpha v$. Así pues, se ha de dar la igualdad. ■

Si $v = \alpha u$, diremos también que la *proporción* que guardan v y u es α . Lo representaremos por

$$\frac{v}{u} = \alpha.$$

Un planteamiento alternativo es fijar un segmento u como unidad de longitud y decir que la *longitud* de v es α .

Es claro que dos segmentos son congruentes si y sólo si tienen la misma longitud medida con la misma unidad.

Definimos la *distancia* entre dos puntos P y Q como la longitud del segmento \overline{PQ} . La distancia es, pues, relativa a la unidad de longitud utilizada. Convenimos además que la distancia de un punto a sí mismo es 0.

En estos términos podemos decir, por ejemplo, que una circunferencia está formada por los puntos equidistantes de su centro.

Es muy importante notar que la construcción de los números reales a partir de los números racionales es puramente conjuntista, es decir, no se basa en los axiomas geométricos que estamos estudiando. Así pues, tenemos un modelo conjuntista del concepto geométrico de recta. Se trata del primer paso en la inmersión de la geometría en la teoría de conjuntos.

Vamos a definir una suma de números reales de manera que la longitud de una suma de segmentos será la suma de sus longitudes.

Definición 2.12 Dados $\alpha, \beta \in \mathbb{R}$, definimos

$$\alpha + \beta = \sup\{r + s \mid r, s \in \mathbb{Q}, r < \alpha, s < \beta\}.$$

Ciertamente tenemos que $\alpha + \beta \in \mathbb{R}$.

Teorema 2.13 *Se cumple:*

1. $(\mathbb{R}, +)$ es un grupo abeliano.
2. La suma $+$ restringida a \mathbb{Q} es la suma usual en \mathbb{Q} .
3. Si $\alpha \leq \beta$ entonces $\alpha + \gamma \leq \beta + \gamma$.
4. Se cumple $\alpha > 0$ si y sólo si $-\alpha < 0$.

DEMOSTRACIÓN: Notemos en general que para probar que dos números reales α y β son iguales es suficiente probar que todo número racional x que cumple $x < \alpha$ cumple también $x \leq \beta$ y viceversa.

Si $x \in \mathbb{Q}$, $x < (\alpha + \beta) + \gamma$, entonces existen $y, t \in \mathbb{Q}$ tales que $x < y + t$, $y < \alpha + \beta$, $t < \gamma$, luego existen $r, s \in \mathbb{Q}$, tales que $y < r + s$, $r < \alpha$, $s < \beta$, luego $s + t < \beta + \gamma$ y $x < r + (s + t) < \alpha + (\beta + \gamma)$. Similarmente se recorre el camino contrario, luego $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.

Es obvio que la suma es conmutativa.

Veamos que $\alpha + 0 = \alpha$. Si $x \in \mathbb{Q}$, $x < \alpha + 0$ existen $r, s \in \mathbb{Q}$ tales que $x < r + s$, $r < \alpha$, $s < 0$, luego $x < r + s < r < \alpha$. Así mismo, si $r \in \mathbb{Q}$, $r < \alpha$ entonces existe un $s \in \mathbb{Q}$ tal que $r < s < \alpha$, con lo que $r = s + (r - s) < \alpha + 0$.

Dado un número real α , definimos

$$-\alpha = \sup\{-r \mid r \in \mathbb{Q}, \alpha < r\}.$$

Una cota superior del conjunto es $-s$, donde $s \in \mathbb{Q}$, $s < \alpha$, luego $-\alpha$ es un número real. Veamos que $\alpha + (-\alpha) = 0$.

Si $x \in \mathbb{Q}$, $x < \alpha + (-\alpha)$ entonces existen $r, s \in \mathbb{Q}$ tales que $x < r + s$, $r < \alpha$, $-s > \alpha$, luego $x < r + s < 0$. Recíprocamente, si $x \in \mathbb{Q}$, $x < 0$, tomamos números racionales $x < -u < 0$ y $v < \alpha$. Entonces la sucesión $v + nu$ sobrepasará (un número racional mayor que) α para algún natural n , que podemos tomar mínimo. Así obtenemos un número $r < \alpha$ tal que $s = r + u > \alpha$ (si $r + u = \alpha$ cambiamos u por un número mayor que siga cumpliendo $x < -u < 0$). Entonces $x < -u = r - s < \alpha + (-\alpha)$.

Dados $u, v \in \mathbb{Q}$, si $x \in \mathbb{Q}$ cumple $x < i(u) + i(v)$ entonces existen $r, s \in \mathbb{Q}$ tales que $x < r + s$, $r < u$, $s < v$, luego $x < u + v$, luego $x < i(u + v)$. Si $x < i(u + v)$ tomamos $r \in \mathbb{Q}$ tal que $0 < r < (u + v - x)/2$, de modo que $x < (u - r) + (v - r) < i(u) + i(v)$. Por lo tanto $i(u + v) = i(u) + i(v)$.

Si $r < \alpha$ y $s < \gamma$ son números racionales y $\alpha \leq \beta$, entonces $r + s \leq \beta + \gamma$ por definición de suma, luego tomando el supremo, $\alpha + \gamma \leq \beta + \gamma$.

La propiedad 4 se sigue sin dificultad de la definición de $-\alpha$. ■

La interpretación geométrica de la suma de números reales está expresada en la relación siguiente, según la cual la longitud de una suma de segmentos es la suma de las longitudes.

Teorema 2.14 *Si α y β son números reales positivos y u es un segmento, entonces $(\alpha + \beta)u \equiv \alpha u + \beta u$.*

DEMOSTRACIÓN: Sea $\alpha u + \beta u = \gamma u$. Hay que ver que $\gamma = \alpha + \beta$. Si $r < \alpha$ y $s < \beta$ son números racionales positivos, entonces $(r + s)u \equiv ru + su < \alpha u + \beta u \equiv \gamma u$, luego $r + s < \gamma$, lo que prueba que $\alpha + \beta \leq \gamma$. Si fuera $\alpha + \beta < \gamma$ sea $\delta > 0$ tal que $\gamma = \alpha + \beta + \delta$ (existe por la propiedad 3 del teorema anterior). Sea $r \in \mathbb{Q}$ tal que $0 < r < \delta$. Tomemos números racionales s y t tales que $\alpha < s < \alpha + r/2$ y $\beta < t < \beta + r/2$. Entonces

$$\alpha u + \beta u < su + tu \equiv (s + t)u < (\alpha + \beta + r)u < (\alpha + \beta + \delta)u \equiv \alpha u + \beta u,$$

con lo que tenemos una contradicción. Por consiguiente se da la igualdad. ■

Los números reales negativos también tienen su interpretación geométrica en términos de las rectas graduadas.

Ejercicio: Probar que si P_α es el punto de una recta graduada que se corresponde con el número real α , entonces $\overline{P_0 P_\alpha} \equiv \overline{P_0 P_{-\alpha}}$, es decir, que los puntos asociados a números opuestos son los ‘simétricos’ respecto al punto P_0 .

Si tomamos dos unidades de longitud, por ejemplo el metro y el centímetro, para expresar en centímetros una longitud dada en metros hemos de multiplicar por 100, debido a que $1\text{m} = 100\text{ cm}$. En general, si tenemos dos unidades u y v , para expresar en términos de v la longitud de un segmento dada en términos de u necesitamos conocer la longitud de u en términos de v , es decir, el número β que cumple $u = \beta v$. Entonces, si $s = \alpha u$, tendremos también $s = \alpha\beta v$, de modo que la longitud de s en términos de v se obtiene multiplicando por β su longitud en términos de u . Todo esto es fácil de probar si las longitudes son todas racionales, pero si no es así ni siquiera tenemos definido el producto $\alpha\beta$. Vamos a definir un producto de números reales que dé sentido a estas fórmulas.

Definición 2.15 Si α y β son números reales positivos definimos

$$\alpha\beta = \sup\{rs \mid r, s \in \mathbb{Q}, 0 < r < \alpha, 0 < s < \beta\}.$$

El producto de dos números reales no nulos se define por las relaciones:

$$\alpha\beta = \begin{cases} -((- \alpha)\beta) & \text{si } \alpha < 0, \beta > 0, \\ -(\alpha(-\beta)) & \text{si } \alpha > 0, \beta < 0, \\ (-\alpha)(-\beta) & \text{si } \alpha < 0, \beta < 0. \end{cases}$$

Finalmente, si $\alpha = 0$ o $\beta = 0$ definimos $\alpha\beta = 0$.

Teorema 2.16 *Se cumple:*

1. $(\mathbb{R}, +, \cdot)$ es un cuerpo que contiene a \mathbb{Q} como subcuerpo.
2. Si $\alpha \geq 0$ y $\beta \geq 0$ entonces $\alpha\beta \geq 0$.

DEMOSTRACIÓN: La prueba de que los números reales positivos forman un grupo con el producto se obtiene reemplazando literalmente sumas por productos en la prueba que hemos visto de que \mathbb{R} es un grupo con la suma. Después las propiedades se trasladan formalmente a números reales arbitrarios a partir de la definición de producto. Del mismo modo se ve que el producto en \mathbb{R} extiende al producto en \mathbb{Q} , y la última propiedad es trivial. Sólo queda comprobar que la suma distribuye al producto.

Tomemos en primer lugar $\alpha, \beta, \gamma > 0$ y veamos que $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.

Si $r \in \mathbb{Q}$ cumple $0 < r < \alpha(\beta + \gamma)$, entonces existen $u, v \in \mathbb{Q}$ positivos y tales que $r < uv$, $u < \alpha$, $v < \beta + \gamma$, luego existen $x, y \in \mathbb{Q}$ positivos tales que $v < x + y$, $x < \beta$, $y < \gamma$. Entonces $r < u(x + y) = ux + uy \leq \alpha\beta + \alpha\gamma$.

Si $0 < r < \alpha\beta + \alpha\gamma$ entonces existen $u, v \in \mathbb{Q}$ positivos tales que $x < u + v$, $u < \alpha\beta$, $v < \alpha\gamma$. A su vez existen $a, b, c, d \in \mathbb{Q}$ positivos de modo que $x < ab + cd$, $a < \alpha$, $b < \beta$, $c < \alpha$, $d < \gamma$. Sea $e = \max\{a, c\}$. Entonces $e < \alpha$ y $x < eb + ed = e(b + d) < \alpha(\beta + \gamma)$. Esto prueba la igualdad.

Si $\beta + \gamma \geq 0$, $\beta \geq 0$, $\gamma < 0$, entonces $\alpha\beta = \alpha((\beta + \gamma) - \gamma) = \alpha(\beta + \gamma) + \alpha(-\gamma)$ (puesto que $\beta + \gamma \geq 0$ y $-\gamma \geq 0$), de donde $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$. Los demás casos se siguen formalmente de éstos dos. ■

Las propiedades siguientes son consecuencias inmediatas de los teoremas anteriores:

1. Si $\alpha \leq \beta$ entonces $-\beta \leq -\alpha$.
2. Para todo $\alpha \in \mathbb{R}$, $\alpha^2 \geq 0$.
3. Si $\alpha \leq \beta$ y $\gamma \geq 0$ entonces $\alpha\gamma \leq \beta\gamma$.
4. Si $0 < \alpha < \beta$, entonces $0 < \beta^{-1} < \gamma^{-1}$.

También podemos probar la propiedad geométrica que habíamos anunciado:

Si α y $\beta > 0$ y u es un segmento, entonces $\alpha(\beta u) \equiv (\alpha\beta)u$. Equivalentemente,

$$\frac{s}{u} = \frac{s}{v} \frac{v}{u}.$$

En efecto, un segmento s es menor que $\alpha(\beta u)$ si y sólo si $s < r(\beta u)$ para cierto $r < \alpha$, si y sólo si $(1/r)s < \beta u$ para cierto $r < \alpha$, si y sólo si $(1/r)s < r'u$ para ciertos $r < \alpha$ y $r' < \beta$, si y sólo si $s < rr'u$ para ciertos $r < \alpha$ y $r' < \beta$, si y sólo si $s < r''u$ para cierto $r'' < \alpha\beta$, si y sólo si $s < (\alpha\beta)u$. De aquí se sigue obviamente la igualdad. ■

Para terminar con los resultados básicos sobre longitud de segmentos demostramos una caracterización sencilla que nos será útil en el capítulo próximo.

Teorema 2.17 *Sea m una aplicación que a cada segmento le asigna un número real positivo y que cumpla las propiedades siguientes:*

1. Si $u \equiv v$ entonces $m(u) = m(v)$,
2. Para todo par de segmentos u, v se cumple $m(u + v) = m(u) + m(v)$.

Entonces, para todo par de segmentos u, v se cumple

$$\frac{m(v)}{m(u)} = \frac{v}{u}.$$

DEMOSTRACIÓN: Para todo número natural q se cumple por hipótesis que $m(qx) = qm(x)$, y si aplicamos esto a $x = u/q$ tenemos que $m(u) = qm(u/q)$, o equivalentemente

$$\frac{m(u/q)}{m(u)} = \frac{1}{q}.$$

Multiplicando esta igualdad por un número natural p obtenemos, para cualquier número racional positivo $r = p/q$, que

$$\frac{m(ru)}{m(u)} = r.$$

Si dos segmentos cumplen $x < y$, entonces existe un segmento z tal que $y = x + z$, luego $m(x) < m(x) + m(z) = m(y)$. Por lo tanto, si α es un número real arbitrario y r, s son números racionales tales que $r < \alpha < s$ tenemos

$$r = \frac{m(ru)}{m(u)} < \frac{m(\alpha u)}{m(u)} < \frac{m(su)}{m(u)} = s.$$

Como esto es válido para todo r y todo s , ha de ser

$$\frac{m(\alpha u)}{m(u)} = \alpha.$$

Expresando $v \equiv \alpha u$ tenemos la relación buscada. ■

Lo que afirma este teorema es que si una aplicación m cumple las propiedades indicadas, tomamos una unidad de longitud u y llamamos $k = m(u)$, entonces m asigna a cada segmento su longitud multiplicada por k .

2.2 Complementos sobre números reales

Recogemos aquí algunos conceptos y propiedades adicionales sobre los números reales que acabamos de construir y que a menudo resultan útiles.

Definición 2.18 Llamaremos \mathbb{R}^+ al conjunto de los números reales positivos (mayores que 0), llamaremos \mathbb{R}^- al conjunto de los números reales negativos (menores que 0). Definimos el *signo* de un número real α como

$$\text{sig } \alpha = \begin{cases} +1 & \text{si } \alpha > 0 \\ 0 & \text{si } \alpha = 0 \\ -1 & \text{si } \alpha < 0 \end{cases}$$

Llamaremos *valor absoluto* o *módulo* de un número real α al número

$$|\alpha| = \begin{cases} \alpha & \text{si } \alpha \geq 0 \\ -\alpha & \text{si } \alpha \leq 0 \end{cases}$$

La demostración de las propiedades siguientes no ofrece ninguna dificultad:

1. $|\alpha| \geq 0$,
2. $|\alpha| = |-\alpha|$,
3. $|\alpha| \leq \beta$ si y sólo si $-\beta \leq \alpha \leq \beta$,
4. $|\alpha + \beta| \leq |\alpha| + |\beta|$,
5. $||\alpha| - |\beta|| \leq |\alpha - \beta|$,
6. $|\alpha\beta| = |\alpha||\beta|$.

Todo número real α está comprendido entre dos números enteros. Llamaremos *parte entera* del número α al único número $E[\alpha]$ que cumple: $E[\alpha] \in \mathbb{Z}$ y $E[\alpha] \leq \alpha < E[\alpha] + 1$. Es fácil probar que siempre existe un único número entero en estas condiciones.

Definición 2.19 Si α y β son números reales definimos los conjuntos:

$$\begin{aligned}]\alpha, \beta[&= \{x \in \mathbb{R} \mid \alpha < x < \beta\} & [\alpha, \beta] &= \{x \in \mathbb{R} \mid \alpha \leq x \leq \beta\} \\]\alpha, \beta] &= \{x \in \mathbb{R} \mid \alpha < x \leq \beta\} & [\alpha, \beta[&= \{x \in \mathbb{R} \mid \alpha \leq x < \beta\}. \end{aligned}$$

Podemos considerar estas mismas definiciones en $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ y así tenemos en particular

$$\begin{aligned}]-\infty, \beta[&= \{x \in \mathbb{R} \mid x < \beta\} &]\alpha, +\infty[&= \{x \in \mathbb{R} \mid \alpha < x\} \\]-\infty, \beta] &= \{x \in \mathbb{R} \mid x \leq \beta\} & [\alpha, +\infty[&= \{x \in \mathbb{R} \mid \alpha \leq x\}. \end{aligned}$$

Los conjuntos de cualquiera de estos tipos reciben el nombre de *intervalos*.

La interpretación geométrica de los intervalos es clara. El teorema siguiente es una forma generalizada del axioma *D*, pero lo demostramos a partir de la definición de los números reales, es decir, sin hacer uso de los axiomas geométricos.

Teorema 2.20 *Los intervalos de \mathbb{R} son exactamente los subconjuntos I de \mathbb{R} que cumplen la propiedad siguiente: si $\alpha, \beta \in I$ y $\alpha < \gamma < \beta$, entonces $\gamma \in I$.*

DEMOSTRACIÓN: Es claro que todos los intervalos cumplen la propiedad indicada. Sea I un subconjunto de \mathbb{R} con dicha propiedad.

Si $I = \emptyset$, entonces $I =]\alpha, \alpha[$, que es un intervalo. Supongamos que I no es vacío.

Sean α y β el ínfimo y el supremo de I , respectivamente en $\overline{\mathbb{R}}$.

Si $x \in]\alpha, \beta[$, por definición de supremo e ínfimo, existen números $u, v \in I$ de manera que $\alpha \leq u < x < v \leq \beta$, luego $x \in I$, es decir, $] \alpha, \beta[\subset I$, y obviamente $I \subset [\alpha, \beta]$. Esto da lugar a cuatro casos según si α y β están o no en I , lo que nos lleva a una de las igualdades: $I =]\alpha, \beta[$, $I =]\alpha, \beta]$, $I = [\alpha, \beta[$, $I = [\alpha, \beta]$. ■

Veamos ahora que existen números irracionales. Para ello probamos lo siguiente:

Teorema 2.21 *Para todo número real $\alpha \geq 0$ existe un único número real $\beta \geq 0$ tal que $\alpha = \beta^2$. Diremos que β es la raíz cuadrada de α y lo representaremos por $\sqrt{\alpha}$.*

DEMOSTRACIÓN: Podemos suponer que $\alpha > 0$. Sea β el supremo del conjunto de números reales cuyo cuadrado es menor que α . Está acotado superiormente por cualquier número real mayor que α y que 1, luego β es ciertamente un número real. Supongamos que $\alpha < \beta^2$. Tomemos un número natural n que cumpla $n > 1/\beta$ y $n > 2\beta/(\beta^2 - \alpha)$. Así $2\beta/n < \beta^2 - \alpha$ y en consecuencia

$$\left(\beta - \frac{1}{n}\right)^2 = \beta^2 - 2\beta\frac{1}{n} + \frac{1}{n^2} > \beta^2 - \beta^2 + \alpha + \frac{1}{n^2} > \alpha.$$

Es claro que cualquier número cuyo cuadrado sea menor que α ha de ser menor que $\beta - 1/n$, luego este número es una cota superior del conjunto de todos ellos, en contra de que β sea su supremo.

Supongamos ahora que $\beta^2 < \alpha$. Entonces tomamos un número natural n que cumpla $n > 4\beta/(\alpha - \beta^2)$ y $n^2 > 2/(\alpha - \beta^2)$. Así

$$\left(\beta + \frac{1}{n}\right)^2 = \beta^2 + 2\beta\frac{1}{n} + \frac{1}{n^2} < \beta^2 + \frac{\alpha - \beta^2}{2} + \frac{\alpha - \beta^2}{2} = \beta^2 + \alpha - \beta^2 = \alpha,$$

luego $\beta + 1/n$ es un número mayor que β perteneciente al conjunto del que β es el supremo, lo cual es imposible. Por lo tanto ha de ser $\beta^2 = \alpha$.

La unicidad es clara, pues si γ es cualquier otro número real positivo entonces $\gamma^2 < \beta^2$ o $\gamma^2 > \beta^2$ según si $\gamma < \beta$ o $\gamma > \beta$. ■

Es conocido que un número natural no es un cuadrado en \mathbb{Q} salvo que sea un cuadrado en \mathbb{Z} , por lo que $\sqrt{2}$, $\sqrt{3}$, $\sqrt{5}$, etc. son ejemplos concretos de números irracionales.

Ejercicio: Probar que entre dos números reales cualesquiera existe un número irracional.

El teorema anterior admite una generalización que será esencial más adelante.

Teorema 2.22 *Todo polinomio de $\mathbb{R}[x]$ de grado impar tiene una raíz en \mathbb{R} .*

DEMOSTRACIÓN: Probemos en primer lugar que si $P(x)$ es un polinomio no constante, $c \in \mathbb{R}$ y $\epsilon > 0$, entonces existe un $\delta > 0$ tal que si $|h| < \delta$ entonces $|P(c+h) - P(c)| < \epsilon$.

En efecto, si $P(x) = \sum_{k=0}^n a_k x^k$, entonces

$$P(c+h) = \sum_{k=0}^n a_k \sum_{i=0}^k \binom{k}{i} c^i h^{k-i} = P(c) + \sum_{k=0}^n a_k \sum_{i=0}^{k-1} \binom{k}{i} c^i h^{k-1-i} h,$$

luego si $h \leq 1$ se cumple

$$|P(c+h) - P(c)| \leq \left(\sum_{k=0}^n \sum_{i=0}^{k-1} \binom{k}{i} a_k |c|^i \right) |h| = K|h|.$$

Basta tomar $\delta = \epsilon/K$.

Por otro lado, si suponemos que $P(x)$ es mónico (es decir, $a_n = 1$),

$$\frac{P(x)}{x^n} = 1 + \frac{a_{n-1}}{x} + \cdots + \frac{a_0}{x^n}.$$

Si tomamos $|x| > n \max_i |a_i|$ y $|x| > 1$, se cumple $|a_i/x^i| < 1/n$, luego

$$\left| \frac{a_{n-1}}{x} + \cdots + \frac{a_0}{x^n} \right| < 1,$$

con lo que

$$\frac{P(x)}{x^n} > 0,$$

es decir, $P(x)$ y x^n toman el mismo signo cuando $|x|$ es suficientemente grande. Si n es impar P toma valores positivos y negativos. Más aún, existe un número real x_0 tal que $P(t) < 0$ siempre que $t < x_0$. Esto prueba que el conjunto

$$A = \{x \in \mathbb{R} \mid P(t) < 0 \text{ para todo } t < x\}$$

no es vacío. Por otra parte cualquier x tal que $P(x) > 0$ es una cota superior de A , luego $\alpha = \sup A$ es un número real. Veamos que $P(\alpha) = 0$.

Notemos que si $t < \alpha$ entonces existe un $x \in A$ tal que $t < x < \alpha$, luego $P(t) < 0$. Si $P(\alpha) > 0$ tomamos $0 < \epsilon < P(\alpha)$ y sabemos que existe un $\delta > 0$ de modo que $|P(\alpha - \delta/2) - P(\alpha)| < \epsilon$, luego $P(\alpha - \delta/2) > P(\alpha) - \epsilon > 0$, y esto es imposible.

Supongamos ahora que $P(\alpha) < 0$. Entonces tomamos $P(\alpha) < -\epsilon < 0$, con lo que existe un $\delta > 0$ de modo que si $|h| < \delta$ entonces $|P(\alpha + h) - P(\alpha)| < \epsilon$, luego $P(\alpha + h) < P(\alpha) + \epsilon < 0$. Esto implica que $\alpha + \delta \in A$, lo cual es también imposible.

Concluimos que $P(\alpha) = 0$. ■

Si aplicamos el teorema anterior al polinomio $x^n - a$, con n impar, concluimos que todo número real tiene al menos una raíz n -sima, es decir, que existe un número b tal que $b^n = a$. Este número b es único, pues si hubiera otro b' tendríamos que $(b/b')^n = 1$ (podemos suponer $a \neq 0$), pero entonces $b/b' = 1$, ya que si $b/b' < 1$ entonces $(b/b')^n < 1$ y si $b/b' > 1$ entonces $(b/b')^n > 1$.

Llamaremos $\sqrt[n]{a}$ a la única raíz n -sima de a , donde n es impar. Notemos que el signo de $\sqrt[n]{a}$ es el mismo que el de a .

Si n no es impar pero $a > 0$ entonces a también tiene raíces n -simas. Concretamente, si $n = 2^r m'$, donde m' es impar, basta calcular r veces la raíz cuadrada de $\sqrt[m']{a}$. Sin embargo ahora tenemos dos raíces cuadradas, pues si b es una raíz n -sima de a con n par, es claro que $-b$ también lo es. El argumento anterior prueba, no obstante, que todo $a > 0$ tiene una única raíz n -sima positiva.

Si $a > 0$, llamaremos $\sqrt[n]{a}$ a su única raíz n -sima positiva.

Es inmediato comprobar que si $a < 1$ entonces $\sqrt[n]{a} < 1$, mientras que si $a > 1$ entonces $\sqrt[n]{a} > 1$.

Ahora observamos que si $m \geq 1$ es un número natural, n es un número entero y a un número real positivo entonces $(\sqrt[n]{a})^m$ depende sólo del número racional n/m , pues si k es un número natural no nulo, $(\sqrt[km]{a})^{km} = a$, luego $(\sqrt[km]{a})^k = \sqrt[n]{a}$, luego $(\sqrt[km]{a})^{kn} = (\sqrt[n]{a})^n$. Por consiguiente podemos definir

$$a^{n/m} = (\sqrt[n]{a})^m, \quad \text{para todo } \frac{n}{m} \in \mathbb{Q}.$$

Haciendo $m = 1$ vemos que esta exponenciación extiende a la usual (para exponentes enteros). Es muy simple comprobar que se conservan las propiedades básicas:

$$a^0 = 1, \quad a^1 = a, \quad 1^r = 1, \quad a^{r+s} = a^r a^s, \quad a^{rs} = (a^r)^s, \quad a^{-r} = \frac{1}{a^r}.$$

Además si $r < s$ entonces $a^r < a^s$ si $a > 1$, mientras que $a^s < a^r$ si $a < 1$. En efecto, basta comparar a^{s-r} con 1, luego basta probar que si $r > 0$ entonces a^r es mayor o menor que 1 según lo sea a , pero esto es inmediato.

Si $a > 1$ y $\alpha \in \mathbb{R}$ definimos $a^\alpha = \sup\{a^r \mid r \in \mathbb{Q}, r \leq \alpha\}$. Es claro que a^α así definido es un número real positivo. Además es inmediato que esta exponenciación con exponentes reales coincide con la que ya teníamos definida cuando los exponentes son racionales. Si $a < 1$ en la definición hemos de cambiar el supremo por un ínfimo. Por simplicidad supondremos siempre $a > 1$, pero el lector debe comprobar que todo vale igualmente cuando $a < 1$.

Se siguen cumpliendo todas las propiedades elementales:

$$a^0 = 1, \quad a^1 = a, \quad 1^\alpha = 1, \quad a^{\alpha+\beta} = a^\alpha a^\beta, \quad a^{\alpha\beta} = (a^\alpha)^\beta, \quad a^{-\alpha} = \frac{1}{a^\alpha}.$$

Comprobaremos la cuarta como ejemplo:

Si $r \leq \alpha$, $s \leq \beta$ son dos números racionales, entonces $r + s \leq \alpha + \beta$, luego por definición $a^{r+s} \leq a^{\alpha+\beta}$, luego $a^r a^s \leq a^{\alpha+\beta}$, luego $a^r \leq a^{\alpha+\beta}/a^s$. Tomando el supremo para $r \leq \alpha$ concluimos que $a^\alpha \leq a^{\alpha+\beta}/a^s$, luego $a^s \leq a^{\alpha+\beta}/a^\alpha$ y análogamente $a^\beta \leq a^{\alpha+\beta}/a^\alpha$, luego $a^\alpha a^\beta \leq a^{\alpha+\beta}$.

Si $r \leq \alpha + \beta$ es un número racional, de la definición de suma de números reales se sigue fácilmente que $r \leq s + t$, donde $s \leq \alpha$, $t \leq \beta$ son dos números racionales. Así $a^r \leq a^s a^t \leq a^\alpha a^\beta$, luego tomando el supremo para $r \leq \alpha + \beta$ queda $a^{\alpha+\beta} \leq a^\alpha a^\beta$.

Si $\alpha < \beta$, existen racionales $\alpha < r < s < \beta$, luego $a^\alpha \leq a^r < a^s \leq a^\beta$ (si $a < 1$ hay que cambiar el sentido de la desigualdad).

Con esto tenemos probado que la *función exponencial* $\exp_a : \mathbb{R} \rightarrow \mathbb{R}^+$ dada por $\exp_a(x) = a^x$ es un monomorfismo de grupos. Veamos que es suprayectivo, es decir, que para todo $\beta \in \mathbb{R}^+$ existe un $\alpha \in \mathbb{R}$ tal que $a^\alpha = \beta$.

En primer lugar probamos que si $a > 1$ y $x \in \mathbb{R}$ existe un natural n tal que $a^n > x$. En efecto, haciendo $a = 1 + b$, con $b > 0$ el teorema del binomio de Newton nos da que $a^n \geq 1 + nb$, luego basta tomar $n > (x - 1)/b$.

Tomamos $\alpha = \sup\{x \in \mathbb{R} \mid a^x \leq \beta\}$. La observación anterior prueba en particular que el conjunto está acotado superiormente. Si $r \leq \alpha$ entonces $a^r \leq \beta$ por definición, luego tomando supremos resulta que $a^\alpha \leq \beta$. No puede ser $a^\alpha < \beta$, pues entonces $1 < \beta/a^\alpha$ y existiría un natural n tal que $a < (\beta/a^\alpha)^n$, es decir, $a^{1/n} < \beta/a^\alpha$, o también $a^{\alpha+1/n} < \beta$, en contradicción con la definición de α . Así pues, $a^\alpha = \beta$.

Definimos la *función logarítmica* $\log_a : \mathbb{R}^+ \rightarrow \mathbb{R}$ como el isomorfismo inverso de la función exponencial \exp_a .

Las propiedades siguientes de la función logarítmica se demuestran inmediatamente a partir de las de la función exponencial:

$$\log_a(\alpha\beta) = \log_a(\alpha) + \log_a(\beta), \quad \log_a(\alpha^\beta) = \beta \log_a(\alpha), \quad \log_a(\alpha) = \frac{\log_b(\alpha)}{\log_b(a)}.$$

2.3 Amplitud de ángulos

Vamos a asignar un número real a cada ángulo tal y como hemos hecho con los segmentos. Dado un ángulo L , el ángulo nL , es decir, la suma de L consigo mismo n veces, no está definido para todos los números naturales n (puede incluso no estar definido para $n = 2$). En cambio el teorema 2.2 asegura que $(1/n)L$ está definido para todo n . Es claro entonces que rL está definido al menos para todo número racional menor o igual que 1. Más aún, si está definido para $r = p/q$, entonces está definido para n/q , para todo $n \leq p$, de donde es fácil ver que si rL está definido para un número racional $r > 0$, también lo está para todo número racional positivo menor. El teorema siguiente prueba que ciertamente hay números para los que rL no está definido junto con otro hecho importante.

Teorema 2.23 *Sea L un ángulo. Entonces*

1. *Existe un número real $\mu > 0$ tal que rL está definido exactamente para los números racionales que cumplen $0 < r \leq \mu$.*
2. *Dados dos ángulos $U < V$, hay un número racional r tal que $U < rL < V$.*

DEMOSTRACIÓN: Supongamos que nL está definido para todo número natural n . En particular los ángulos nL han de ser todos agudos, pues si nL fuera recto u obtuso entonces no estaría definido $3nL$. Sea \widehat{AOB} un ángulo recto y consideremos puntos A_n en \overline{AB} tales que $\widehat{AOA_n} \equiv nL$. Aplicando como de costumbre el axioma D probamos que el conjunto X de los puntos P en \overline{AB} tales que $\widehat{AOP} < \widehat{AOA_n}$ para algún n (más los puntos menores o iguales que A respecto al orden \leq_{AB}) es una semirrecta de origen un punto C en \overline{AB} .

Entonces \widehat{AOC} es mayor o igual que todos los ángulos nL . En particular podemos restarle un ángulo menor que L y obtener un punto D en \overline{AB} anterior a C tal que $\widehat{DOC} < L$. Entonces D está en X , luego existe un n tal que $\widehat{AOC} < \widehat{AOA_n}$. Pero entonces

$$\widehat{AOC} \equiv \widehat{AOD} + \widehat{DOC} < \widehat{AOA_n} + L \equiv \widehat{AOA_{n+1}},$$

mientras que ha de ser $\widehat{AOA_{n+1}} \leq \widehat{AOC}$. Por lo tanto nL no está definido para algún n (ni para ningún número racional posterior) y así el conjunto de los números racionales r para los que rL está definido tiene un supremo μ en \mathbb{R} , que es el número buscado.

De aquí se deduce que dados dos ángulos L y L' existe un número natural n tal que $(1/n)L < L'$. En efecto, si fuera $L' \leq (1/n)L$ para todo n , entonces es claro que nL' estaría definido para todo n , en contra de lo que acabamos de probar.

Sean ahora dos ángulos $U < V$. Existe un número natural n de manera que $(1/n)L < V - U < V$. Ahora observamos que si A es cualquier ángulo menor o igual que U , entonces $A + (1/n)L$ está definido y es un ángulo menor que

$U + (V - U) \equiv V$. Por lo tanto, o bien $U < (1/n)L < V$ o bien podemos ir calculando $(1/n)L$, $(2/n)L$, etc., definidos mientras sean menores o iguales que U . Como no puede haber definidos infinitos múltiplos de $(1/n)L$, ha de haber un k tal que $(k/n)L \leq U$ pero

$$U < \frac{k+1}{n}L < V,$$

como había que probar. ■

Ahora ya podemos definir la medida de ángulos:

Teorema 2.24 *Sea L un ángulo cualquiera y sea μ la constante del teorema anterior.*

1. *Para todo número real $0 < \alpha \leq \mu$ existe un único ángulo αL (salvo congruencia) tal que para todo número racional $0 < r < \alpha$ se cumple $rL < \alpha L$ y para todo número racional $\alpha < r \leq \mu$ se cumple $rL > \alpha L$.*
2. *Para todo ángulo L' existe un único número real $0 < \alpha \leq \mu$ tal que $L' \equiv \alpha L$. En particular μL es un ángulo llano.*

DEMOSTRACIÓN: Supongamos primero que $\alpha < \mu$. Sea un número racional $\alpha < s < \mu$. Sea $sL = \widehat{AOB}$. Sea X el conjunto de puntos P del intervalo \overline{AB} tales que $\widehat{OAP} < rL$ para algún número racional $r < \alpha$ (más los puntos anteriores a A). Aplicando el axioma D tal y como es habitual obtenemos que X es una semirrecta de origen un punto C que determina un ángulo \widehat{AOC} , que es al que vamos a llamar αL .

Si $r < \alpha$ existen un números racionales $r < r' < r'' < \alpha$. Si $r'L \equiv \widehat{AOP}$, entonces $P \in \overline{AB}$ (porque $r'L < sL$) y $\widehat{AOP} < r''L$, luego P está en X y por lo tanto es anterior a C , lo que implica que $rL < r'L \leq \alpha L$.

Si $\alpha < r$ entonces existen números racionales $\alpha < r' < r$ y $r' < s$. Si ponemos $r'L \equiv \widehat{AOP}$, entonces P no está en X , luego es posterior a C , luego $\alpha L \leq r'L < rL$.

El segundo apartado del teorema anterior garantiza que αL es único y que no depende más que de la clase de congruencia de L (si hubiera dos ángulos con la misma propiedad, uno sería menor que el otro y podríamos intercalar un ángulo rL entre ambos que incumpliría la propiedad que acabamos de probar).

Es claro que si definimos μL como un ángulo llano se cumple también el teorema en este caso (ningún ángulo no llano tiene la propiedad, pues sería posterior a un rL).

Dado un ángulo L' , sea α el supremo del conjunto de los números racionales tales que $rL < L'$. Necesariamente entonces $\alpha L \leq L'$, y si fuera estrictamente menor, existiría un número racional r tal que $\alpha L < rL < L'$, pero entonces sería $r \leq \alpha$ por definición de α y $r > \alpha$ por definición de αL . La unicidad de α es clara. ■

Es claro que si α es un número racional αL coincide con el definido algebraicamente mediante sumas y divisiones. Así mismo se demuestran sin dificultad

(y de modo totalmente análogo al caso de los segmentos) las propiedades siguientes:

Teorema 2.25 Sean α, β números reales y L, L' dos ángulos. Entonces se cumplen las propiedades siguientes (donde suponemos que las sumas y productos están definidos)

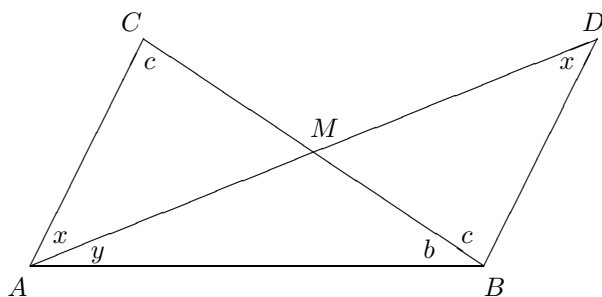
1. Si $\alpha < \beta$, entonces $\alpha L < \beta L$,
2. Si $L < L'$ entonces $\alpha L < \alpha L'$,
3. $(\alpha + \beta)L \equiv \alpha L + \beta L$,
4. $\alpha(L + L') \equiv \alpha L + \alpha L'$,
5. $(\alpha(\beta L)) \equiv (\alpha\beta)L$.

La interpretación geométrica de estas propiedades es la misma que en el caso de las longitudes de segmentos. En particular, si fijamos una unidad de ángulos arbitraria U , definimos la *amplitud* de un ángulo L con respecto a esta unidad como el número α que cumple $L = \alpha U$. Para que nuestra notación coincida con la habitual, vamos a adoptar provisionalmente un convenio no habitual: llamaremos π a la amplitud de un ángulo llano respecto a la unidad de ángulos que consideremos.

El teorema 1.49 puede reenunciarse ahora como que la suma de dos ángulos de un triángulo es siempre menor que π . En realidad podemos probar un poco más:

Teorema 2.26 (Lagrange) La suma de los tres ángulos de un triángulo es menor o igual que π .

DEMOSTRACIÓN: Dado un triángulo \widehat{ABC} cuyos ángulos midan a, b y c , consideramos el punto medio M del lado \overline{AB} y prolongamos el segmento \overline{AM} hasta un segmento $\overline{AD} \equiv 2\overline{AM}$. Es claro entonces que $\widehat{AMC} \equiv \widehat{DMB}$, con lo que el triángulo \widehat{ABD} resulta tener un ángulo igual a $b + c$ y otros dos ángulos x e y tales que $a = x + y$. Uno de estos dos será menor o igual que $a/2$.



De este modo, dado un triángulo arbitrario hemos obtenido otro cuyos ángulos suman la misma cantidad pero uno de ellos es menor o igual que la mitad de uno de los ángulos originales. Repitiendo el proceso podemos obtener triángulos con la misma suma de ángulos y donde uno de ellos sea menor o igual que $a/2^n$. Si la suma de los ángulos del triángulo original fuera $\pi + \epsilon$, entonces podemos llegar a un triángulo cuyos ángulos sumen $\pi + \epsilon$ con uno de ellos menor que ϵ . Por lo tanto, los otros dos sumarán al menos π , lo cual contradice al teorema 1.49. ■

Terminamos lo referente a la medida de ángulos con un resultado análogo al teorema 2.17. Omitimos la prueba por ser completamente análoga.

Teorema 2.27 *Sea m una aplicación que a cada ángulo le asigna un número real positivo y que cumpla las propiedades siguientes:*

1. Si $L \equiv L'$ entonces $m(L) = m(L')$,
2. Si L , y L' se pueden sumar, entonces $m(L + L') = m(L) + m(L')$.

Entonces, para todo par de ángulos L , L' se cumple

$$\frac{m(L)}{m(L')} = \frac{L}{L'}.$$

2.4 Arcos y sectores circulares

Fijemos una circunferencia ω de centro O y radio r . Si A y B son dos puntos de ω no diametralmente opuestos (es decir, no alineados con O), llamaremos *arco* menor de extremos A y B a la intersección de ω con \widehat{AOB} . El *arco* mayor de extremos A y B es el conjunto formado por los puntos de ω que no están en el arco menor, más los puntos A y B . Diremos que estos dos arcos son mutuamente *complementarios*.

Un arco determina a sus extremos, pues son los únicos puntos P con la propiedad de que todo ángulo de vértice O y que contenga a P en su interior (es decir, no en su frontera) contiene puntos de ω externos al arco.

Si A y B son puntos de ω diametralmente opuestos llamaremos *arcos* de extremos A y B a las intersecciones de ω con los dos semiplanos determinados por la recta AB . Así tenemos dos pares de arcos complementarios, sin que podamos decir que uno es mayor que otro. Estos arcos se llaman también *semicircunferencias*. Las semicircunferencias determinan sus extremos del mismo modo que los demás arcos. Usaremos la notación \widehat{AB} para referirnos a un arco de extremos A y B en una circunferencia dada, pero hemos de tener presente que \widehat{AB} no determina el arco, sino que éste depende de la circunferencia y además hay dos arcos distintos con los mismos extremos. Cuando no sean semicircunferencias los distinguiremos si conviene como \widehat{AB}_M (arco mayor) y \widehat{AB}_m (arco menor).

El segmento que une los extremos de un arco se llaman *cuerdas* del arco. Dos arcos complementarios comparten la misma cuerda.

Si en lugar de partir de una circunferencia partimos de un círculo obtenemos las definiciones de *sector circular* y *semicírculo*. La frontera de un sector circular está formada por dos radios \overline{OA} , \overline{OB} y un arco de extremos A y B , de modo que los arcos se corresponden biunívocamente con los sectores, por lo que es equivalente trabajar con unos u otros.

Los arcos admiten la caracterización siguiente:

Teorema 2.28 *Los arcos de extremos A y B en una circunferencia ω son las intersecciones con ω de los semiplanos determinados por la recta AB .*

DEMOSTRACIÓN: Si AB pasa por el centro O de ω los arcos de extremos A y B son semicircunferencias, y el resultado es cierto por definición. Supongamos el caso contrario. Si P es un punto del arco menor \widehat{AB}_m , entonces P está en el ángulo \widehat{AOB} , luego la semirrecta \overrightarrow{OP} corta al segmento \overline{AB} . Como este segmento está contenido en el círculo, el punto de corte ha de estar en \overline{OP} , y esto significa que O y P están en semiplanos distintos respecto a AB . Recíprocamente, si P es un punto de ω en el semiplano opuesto a O respecto a AB , entonces \overline{OP} ha de cortar a AB . Como \overline{OP} está contenido en el círculo, el punto de corte ha de estar en \overline{AB} , luego la semirrecta \overrightarrow{OP} está contenida en \widehat{AOB} , lo que se traduce en que P está en \widehat{AB}_m .

Con esto hemos probado que \widehat{AB}_m es la intersección de ω con el semiplano de AB que no contiene a O . Obviamente entonces el arco mayor ha de ser la intersección de ω con el semiplano que contiene a O . ■

Definición 2.29 La *amplitud* de un arco de extremos AB (en un círculo de centro O) es la del ángulo \widehat{AOB} si el arco es menor, 2π menos esta amplitud si es mayor y π si son semicircunferencias.

De este modo, los arcos menores son los arcos de amplitud menor que π y los mayores son los de amplitud mayor que π . Podemos pensar en las circunferencias como arcos de amplitud igual 2π .

Teorema 2.30 *Sean \widehat{AB} y \widehat{BC} dos arcos de una circunferencia con el extremo B como único punto común. Entonces $\widehat{AB} \cup \widehat{BC}$ es un arco de extremos A y C , y su amplitud es la suma de las amplitudes de los arcos dados.*

DEMOSTRACIÓN: Sea O el centro de la circunferencia que contiene a los arcos. Representemos por $m(X)$ la amplitud de un ángulo o arco X . Supongamos que la amplitud de uno de los arcos es mayor que π , por ejemplo $m(\widehat{AB}) > \pi$. Entonces \widehat{AB} es un arco mayor, luego está contenido en el complementario del ángulo \widehat{AOB} . Como \widehat{BC} no tiene puntos en común con \widehat{AB} (salvo B) ha de estar contenido en \widehat{AOB} . En particular C es un punto interior de este ángulo, luego $\widehat{AOB} = \widehat{AOC} + \widehat{COB}$. Es claro entonces que $\widehat{AB} \cup \widehat{BC} = \widehat{AC}_M$. Además

$$m(\widehat{AC}_M) = 2\pi - m(\widehat{AOC}) = 2\pi - m(\widehat{AOB}) + m(\widehat{COB}) = m(\widehat{AB}) + m(\widehat{BC}).$$

Si uno de los arcos tiene amplitud π la prueba es casi idéntica. Supongamos ahora que ambos arcos son menores que π . Esto significa que ambos arcos son menores, luego están contenidos en los ángulos \widehat{AOB} y \widehat{BOC} . Llamemos A' y B' a los puntos diametralmente opuestos a A y B .

Notemos que A y C han de estar en semiplanos opuestos respecto a BB' , pues en caso contrario los ángulos \widehat{AOB} y \widehat{BOC} estarían uno contenido en otro, y los arcos también. Por lo tanto C está en uno de los dos ángulos $\widehat{BOA'}$ o $\widehat{A'OB'}$.

Si C está en $\widehat{BOA'}$ (incluyendo el caso $C = A'$) entonces tenemos la relación $\widehat{AOC} = \widehat{AOB} + \widehat{BOC}$, luego $\widehat{AB} \cup \widehat{BC} = \widehat{AC}_m$ y la relación entre las amplitudes es clara.

Si C está en $\widehat{A'OB'}$ entonces $\widehat{BOC} = \widehat{BOA'} + \widehat{A'OC}$, luego

$$\widehat{AB} \cup \widehat{BC} = \widehat{AB} \cup \widehat{BA'} \cup \widehat{A'C} = \widehat{AA'} \cup \widehat{A'C}_m = \widehat{AC}_M,$$

donde $\widehat{AA'}$ es la semicircunferencia que contiene a B . Las amplitudes cumplen la relación indicada:

$$\begin{aligned} m(\widehat{AC}_M) &= 2\pi - m(\widehat{AOC}) = 2\pi - \pi + m(\widehat{A'OC}) = \pi + m(\widehat{A'OC}) \\ &= m(\widehat{AOB}) + m(\widehat{BOA'}) + m(\widehat{A'OC}) = m(\widehat{AOB}) + m(\widehat{BOC}) \\ &= m(\widehat{AB}) + m(\widehat{BC}). \end{aligned}$$

■

Capítulo III

La geometría euclídea

Los resultados de los capítulos anteriores eran intuitivamente evidentes, o por lo menos susceptibles de ser justificados sin más que observar una figura. En este capítulo nos ocuparemos de resultados más profundos, de los que ya no puede decirse lo mismo. Por una parte nos servirán para comprender que la geometría encierra resultados nada triviales de interés en sí mismo, y por otra una parte de ellos nos pondrán en condiciones de sumergir completamente la geometría en la teoría de conjuntos, de lo cual nos ocuparemos en el capítulo siguiente.

3.1 El axioma de las paralelas

Recordemos del capítulo anterior que dos rectas son paralelas si están contenidas en un mismo plano y no tienen puntos en común, así como que dos planos son paralelos si no tienen puntos en común. Muy poco es lo que podemos decir sobre rectas y planos paralelos sin más base que los axiomas que hemos dado hasta ahora. Probaremos tan sólo la existencia.

Teorema 3.1 *Por un punto exterior a una recta pasa una paralela.*

DEMOSTRACIÓN: Dada una recta r y un punto exterior P , sea s la perpendicular a r por P y sea t la perpendicular a s por P contenida en el plano de r y s . Las rectas r y t han de ser paralelas, pues si se cortaran formarían un triángulo con dos ángulos rectos. ■

Teorema 3.2 *Por un punto exterior a un plano pasa un plano paralelo.*

DEMOSTRACIÓN: Sea π un plano y P un punto exterior. Sea r la recta perpendicular a π por P , sea π_1 el plano perpendicular a r por P . Entonces π y π_1 son planos paralelos, pues si tuvieran un punto en común Q , el triángulo formado por P , Q y la intersección de r con π tendría dos ángulos rectos. ■

Introducimos ahora el último axioma de la geometría. Con ello estaremos en condiciones de formalizar cualquier razonamiento geométrico intuitivo.

Definición 3.3 Una *geometría (tridimensional) euclídea* es una geometría que satisfaga los axiomas A, B, C, D junto con el axioma siguiente:

Axioma E Por un punto exterior a una recta pasa una única paralela.

El hecho análogo para planos se demuestra fácilmente:

Teorema 3.4 *Por un punto exterior a un plano pasa un único plano paralelo.*

DEMOSTRACIÓN: Sea π un plano y P un punto exterior. Sea r la perpendicular a π por P y sea π_1 el plano perpendicular a r por P . Hemos probado que π_1 es paralelo a π . Veamos que es el único posible. Si π_2 es otro plano paralelo a π que pasa por P , entonces π_1 y π_2 se cortarán en una recta s (que contiene a P y es perpendicular a r). Sean t_1 y t_2 las rectas perpendiculares a s por P contenidas en π_1 y π_2 . Entonces r , t_1 y t_2 son coplanares (por ser perpendiculares a s por P), y el plano que forman corta a π en una recta t (que pasa por el punto de intersección de r y π). Por lo tanto las rectas t , t_1 y t_2 son coplanares, pero t_1 y t_2 son dos paralelas a t que pasan por P , luego han de coincidir. Así pues, π_1 y π_2 tienen en común las rectas perpendiculares s y t_1 , luego $\pi_1 = \pi_2$. ■

Teorema 3.5 *Dos rectas paralelas a una tercera son paralelas o coincidentes.*

DEMOSTRACIÓN: Sean r_1 y r_2 paralelas a r_3 . Supongamos que r_1 y r_2 no son paralelas ni coincidentes. Sea π_1 el plano que contiene a r_1 , r_3 y sea π_2 el plano que contiene a r_2 y r_3 . Si $\pi_1 = \pi_2$ entonces r_1 y r_2 están en el mismo plano, y como no son paralelas se cortan en un punto P , con lo que r_3 tiene dos paralelas distintas por P , contradicción.

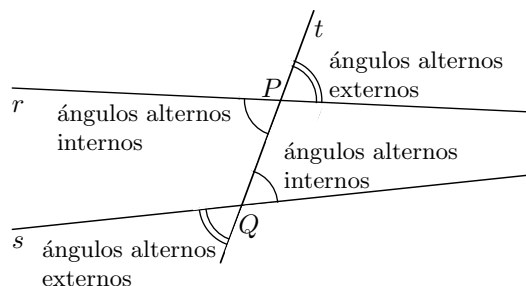
Si $\pi_1 \neq \pi_2$ entonces la intersección de ambos planos es la recta r_3 . Sea π el plano que contiene a r_1 y a un punto Q de r_2 . Entonces π corta a π_2 en una recta r' disjunta de r_3 (pues si R está en $r' \cap r_3$ entonces R no puede estar en r_1 , ya que entonces estaría en $\pi_1 \cap \pi_2 = r_3$ y por tanto en $r_1 \cap r_3$). De aquí que $\pi = \pi_1$, pues tienen a r_1 y a R en común, luego $r' = r_3$ y Q estaría en $r_2 \cap r_3$. Por lo tanto r' es paralela a r_3 y pasa por un punto de r_2 , luego $r' = r_2$. De nuevo tenemos que r_1 y r_2 están en un mismo plano, y concluimos como antes. ■

Más fácilmente se demuestra:

Teorema 3.6 *Dos planos paralelos a un tercero son paralelos o coincidentes.*

Consideremos una figura formada por dos rectas r y s que cortan a una recta t en puntos distintos P y Q . De los ocho ángulos determinados por los dos cortes, llamaremos *ángulos internos* de la figura a los dos que tienen vértice P y que están contenidos en el semiplano de r que contiene a Q , así como a los dos que tienen vértice Q y están contenidos en el semiplano de s que contiene a P . Hay, pues, cuatro ángulos internos, dos con vértice P , uno en cada semiplano respecto a t , y dos con vértice Q , también uno en cada semiplano respecto a t .

Los cuatro ángulos restantes son *externos*. También hay dos en cada vértice y cada uno en un semiplano distinto respecto a t . Diremos que dos ángulos son *alternos* si tienen vértices distintos y están contenidos en semiplanos distintos respecto a t .



Teorema 3.7 Sean r y s dos rectas que corten a una recta t en dos puntos P y Q . Entonces r y s son paralelas si y sólo si los ángulos alternos internos que determinan son iguales.

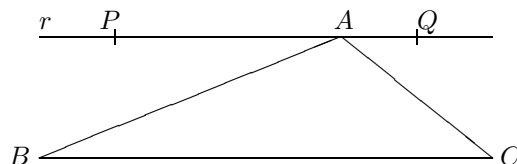
DEMOSTRACIÓN: Si los ángulos alternos internos son iguales, entonces dos ángulos internos no alternos son suplementarios, luego no pueden ser los ángulos de un triángulo, cosa que ocurriría si las rectas r y s se cortaran.

Supongamos ahora que r y s son paralelas. Tomamos uno de los ángulos internos de vértice P y lo transportamos a un ángulo de vértice Q , con un lado igual a \overrightarrow{QP} y con el otro lado en el semiplano opuesto al ángulo de partida respecto de t . La prolongación de dicho lado es una recta s' de modo que t determina ángulos alternos internos iguales en r y s' . Por la parte ya probada r y s' son rectas paralelas, y como s' pasa por Q , por el axioma E ha de ser $s = s'$, luego ciertamente los ángulos alternos internos son iguales. ■

Así pues cuando una recta t corta a dos rectas paralelas r y s se forman ocho ángulos divididos en dos grupos de cuatro que son iguales entre sí. Los de un grupo son los suplementarios de los del otro grupo. Como aplicación inmediata tenemos el hecho siguiente:

Teorema 3.8 La suma de los ángulos de un triángulo es igual a π .

DEMOSTRACIÓN: Dado un triángulo cualquiera \widehat{ABC} , trazamos la paralela r al lado BC por el punto A . Obviamente r deja a B y C en un mismo semiplano.



Las semirrectas \overrightarrow{AB} y \overrightarrow{AC} dividen a dicho semiplano en tres ángulos cuya suma es π . Uno de ellos es \hat{A} , y los otros dos son iguales a \hat{B} y \hat{C} por el teorema anterior.

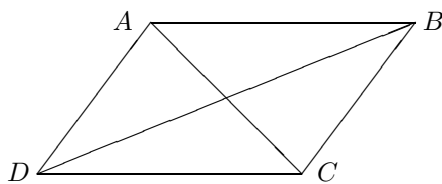
Con más detalle: tomemos puntos P y Q a ambos lados de A en r . Entonces los ángulos \widehat{PAB} y \widehat{PAC} están contenidos en el mismo semiplano respecto a r (pues r no corta a \overline{BC}) luego uno está contenido en el otro. Supongamos por ejemplo que \widehat{PAB} es el menor. Entonces \overrightarrow{AB} está contenida en \widehat{PAC} , luego corta a \overline{PC} , luego P y C están en semiplanos distintos respecto a AB , de donde se sigue que \widehat{PAB} y \widehat{ABC} son alternos internos. Así pues, $\widehat{PAB} \equiv \hat{B}$, luego $\widehat{PAC} \equiv \widehat{PAB} + \widehat{BAC} \equiv \hat{A} + \hat{B}$. Por otra parte, \widehat{QAC} es adyacente a \widehat{PAC} y de aquí que B y Q estén en semiplanos distintos respecto a AC , con lo que \widehat{QAC} y \widehat{ACB} son alternos internos. En total queda que $\hat{A} + \hat{B} + \hat{C} = \widehat{PAC} + \widehat{CAQ} = \pi$. ■

En particular, si dos triángulos tienen dos ángulos iguales, de hecho tienen los tres ángulos iguales, luego el criterio de igualdad ALA puede aplicarse con dos ángulos cualesquiera, no necesariamente los contiguos al lado.

Definición 3.9 Sean A, B, C, D cuatro puntos coplares tales que la recta AB sea paralela a CD y la recta AD sea paralela a BC . Entonces C y D están en el mismo semiplano respecto a AB , A y B están en el mismo semiplano respecto a CD , A y D están en el mismo semiplano respecto a BC y B y C están en el mismo semiplano respecto a AD . A la intersección de estos cuatro semiplanos se le llama *paralelogramo* (gr. ‘de líneas paralelas’) de vértices A, B, C y D .

Los segmentos \overline{AB} , \overline{BC} , \overline{CD} y \overline{DA} se llaman *lados* del paralelogramo. Dos lados son *contiguos* si tienen un vértice en común y son *opuestos* en caso contrario. La unión de todos ellos constituye su *frontera*.

Los ángulos \widehat{DAB} , \widehat{ABC} , \widehat{BCD} y \widehat{CDA} se llaman *ángulos* del paralelogramo. Los segmentos \overline{AC} y \overline{BD} se llaman *diagonales* del paralelogramo. Dos vértices son *contiguos* u *opuestos* según si son extremos de un lado o de una diagonal.



Los paralelogramos son intersecciones de semiplanos, luego son convexos. En particular las diagonales de un paralelogramo están contenidas en él. No es difícil probar que un paralelogramo determina sus vértices.

Teorema 3.10 *Los lados y los ángulos opuestos de un paralelogramo son iguales, los ángulos contiguos son suplementarios, las diagonales se cortan por su punto medio.*

DEMOSTRACIÓN: Con la notación de la definición anterior, tenemos que A y D están en el mismo semiplano respecto a BC , luego si D y B estuvieran en el mismo semiplano respecto a AC tendríamos que D pertenecería al ángulo \widehat{CAB} , luego AD cortaría al segmento BC , cuando en realidad son paralelos. Así pues, B y D están en semiplanos distintos respecto a AC , por lo que \overline{BD} corta a AC . Similarmente llegamos a que \overline{AC} ha de cortar a la recta BD , y como el punto de corte ha de ser el mismo en ambos casos, concluimos que las diagonales \overline{AC} y \overline{BD} se cortan en un punto P .

Es claro que dos ángulos contiguos, por ejemplo \hat{A} y \hat{B} son ángulos internos respecto a la figura formada por las rectas paralelas AD y BC cortadas por AB , pero no son alternos, sino que ambos están en el mismo semiplano respecto a AB . Por lo tanto \hat{A} y un ángulo adyacente a \hat{B} sí son ángulos alternos internos, luego \hat{A} y \hat{B} son suplementarios. De aquí que dos ángulos opuestos sean iguales, pues tienen un suplementario común.

Puesto que A y C están en semiplanos distintos respecto a BD , los ángulos \widehat{ABD} y \widehat{BDC} son alternos internos, luego son iguales. Igualmente se cumple que $\widehat{ADB} \equiv \widehat{DBC}$, luego $\widehat{ADB} \equiv \widehat{CBD}$, lo que en particular implica que los lados opuestos son iguales.

Es fácil ver también que $\widehat{PAB} \equiv \widehat{PCD}$, de donde se sigue que $\overline{PB} \equiv \overline{PD}$ y $\overline{PA} \equiv \overline{PC}$, luego P es el punto medio de las diagonales. ■

Definición 3.11 Un *rombo* (gr. ‘alternado’) es un paralelogramo cuyos cuatro lados son iguales. Un *rectángulo* es un paralelogramo cuyos cuatro ángulos son iguales (luego los cuatro son rectos). Un *cuadrado* es un paralelogramo con los cuatro lados y los cuatro ángulos iguales.

Ejercicio: Probar que las diagonales de un rombo son perpendiculares, y son las bisectrices de los ángulos que unen.

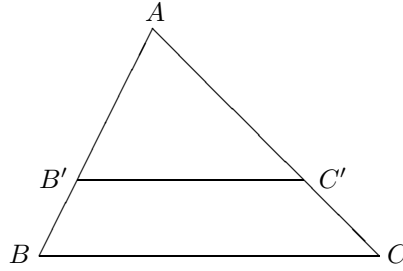
3.2 Semejanza de triángulos

Definición 3.12 Dos triángulos \widehat{ABC} y $\widehat{A'B'C'}$ son *semejantes* si sus lados son proporcionales dos a dos, es decir, si

$$\frac{\overline{AB}}{\overline{A'B'}} = \frac{\overline{AC}}{\overline{A'C'}} = \frac{\overline{BC}}{\overline{B'C'}}.$$

Este concepto de semejanza representa un papel muy importante en la geometría euclídea, y ello se debe fundamentalmente a que vamos a probar que dos triángulos son semejantes si y sólo si sus ángulos son iguales. Para ello necesitamos un par de resultados previos. Conviene introducir el concepto siguiente:

Definición 3.13 Diremos que dos triángulos \widehat{ABC} y $\widehat{A'B'C'}$ están en *posición de Tales* si, nombrando los vértices adecuadamente, $A = A'$, B' está en \overline{AB} y $B'C'$ es paralela a BC (con lo que también C' está en \overline{AC}).



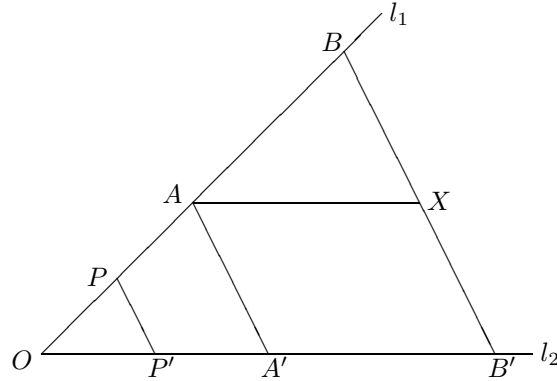
Es fácil probar el teorema siguiente:

Teorema 3.14 Si dos triángulos \widehat{ABC} y $\widehat{A'B'C'}$ están en posición de Tales entonces $\hat{A} = \hat{A}'$, $\hat{B} = \hat{B}'$ y $\hat{C} = \hat{C}'$. Recíprocamente, si \widehat{ABC} y $\widehat{A'B'C'}$ cumplen estas igualdades existe un triángulo $\widehat{A''B''C''}$ semejante a $\widehat{A'B'C'}$ de modo que \widehat{ABC} y $\widehat{A''B''C''}$ están en posición de Tales.

Con ayuda de este hecho ya podemos probar el resultado básico:

Teorema 3.15 Sean l_1 y l_2 dos semirrectas no alineadas de origen O . Sean P y P' puntos de l_1 y l_2 respectivamente distintos de O . Sean A y B dos puntos cualesquiera de l_1 distintos entre sí y distintos de O . Sean A' y B' los puntos donde las rectas paralelas a PP' por A y B , respectivamente, cortan a l_2 . Entonces la longitud de $\overline{A'B'}$ depende de la longitud de \overline{AB} , pero no de la elección de los puntos A y B en l_1 .

DEMOSTRACIÓN: Podemos suponer que A está más cerca de O que B . Notemos que la recta paralela a PP' por A ha de cortar efectivamente a l_2 , o de lo contrario serían paralelas, luego PP' también sería paralela a l_2 , lo cual es absurdo. Tracemos por A la paralela a l_2 , que cortará a $\overline{BB'}$ en un punto X (por el axioma B5).



Los triángulos $\widehat{BOB'}$ y \widehat{BAX} están en posición de Tales, luego se cumple $\widehat{BAX} = \widehat{BOB'} = \widehat{POP'}$. Lo mismo sucede con $\widehat{OPP'}$ y $\widehat{OBB'}$, luego también $\widehat{ABX} = \widehat{OBB'} = \widehat{OPP'}$.

Esto significa que los ángulos del triángulo \widehat{ABX} son independientes de la elección de A y B . Por lo tanto, si hacemos elecciones distintas conservando la longitud de \overline{AB} obtendremos triángulos iguales, luego los lados \overline{AX} también serán iguales, pero $\overline{AX} = \overline{A'B'}$ porque son los lados opuestos de un paralelogramo. Así pues, la longitud de $\overline{A'B'}$ sólo depende de la longitud de \overline{AB} . ■

Teorema 3.16 (Teorema de Tales) *Si dos triángulos están en posición de Tales entonces son semejantes.*

DEMOSTRACIÓN: Sean \widehat{ABC} y $\widehat{A'B'C'}$ triángulos en posición de Tales, digamos con el vértice A en común. Para cada segmento s podemos tomar dos puntos P y Q en \overline{AB} tales que $\overline{PQ} \equiv s$, transportarlos a \overline{AC} mediante rectas paralelas a BC de acuerdo con el teorema anterior y obtener así un segmento cuya longitud sólo depende de la longitud de s . Llamemos $m(s)$ a esta longitud. Es obvio que se cumplen las hipótesis del teorema 2.17, luego en particular, y teniendo en cuenta que \overline{AB} se transforma en \overline{AC} y $\overline{AB'}$ se transforma en $\overline{AC'}$, concluimos que

$$\frac{\overline{AB}}{\overline{A'B'}} = \frac{\overline{AC}}{\overline{AC'}}.$$

Ahora bien, los dos triángulos dados tienen los ángulos iguales, luego (pasando a otros congruentes) podemos ponerlos en posición de Tales con otro vértice en común, por ejemplo B , y entonces obtenemos la igualdad que nos falta para concluir que son semejantes. ■

Teorema 3.17 *Dos triángulos son semejantes si y sólo si tienen los ángulos iguales.*

DEMOSTRACIÓN: Si dos triángulos tienen los ángulos iguales, entonces pueden ponerse en posición de Tales, luego son semejantes. Supongamos ahora que \widehat{ABC} y $\widehat{A'B'C'}$ son dos triángulos semejantes. Es claro que podemos construir un triángulo $\widehat{A''B''C''}$ en posición de Tales respecto a \widehat{ABC} y de modo que $\overline{AA''} \equiv \overline{AA'}$. Entonces los triángulos $\widehat{A'B'C'}$ y $\widehat{A''B''C''}$ son semejantes y tienen un lado igual. De la definición de semejanza se sigue entonces que tienen los tres lados iguales, luego son iguales. ■

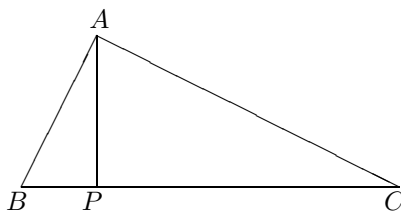
Teorema 3.18 *Dos triángulos son semejantes si sólo si tienen un ángulo igual y sus lados adyacentes proporcionales.*

DEMOSTRACIÓN: Sustituyendo uno de los triángulos por otro congruente (y por lo tanto semejante) podemos suponer que ambos tienen un ángulo coincidente. Sean, pues \widehat{ABC} y $\widehat{A'B'C'}$, donde las semirrectas \overrightarrow{AB} y \overrightarrow{AC} coinciden respectivamente con $\overrightarrow{AB'}$ y $\overrightarrow{AC'}$. Trazamos la paralela a BC que pasa por B' , la cual cortará a AC en un punto C'' . Por el teorema de Tales, el triángulo $\widehat{ABC''}$

es semejante a \widehat{ABC} luego a $\widehat{AB'C''}$, con lo que sus lados son proporcionales, pero tienen un lado coincidente, de hecho todos sus lados han de ser iguales. En particular $\overline{AC'} \equiv \overline{AC''}$, luego $C' = C''$ y así $\widehat{AB'C''}$ es semejante a \widehat{ABC} . El recíproco es obvio. ■

Teorema 3.19 (Teorema de Pitágoras) *En un triángulo rectángulo, el cuadrado de la hipotenusa es igual a la suma de los cuadrados de los catetos.*

DEMOSTRACIÓN: Sea \widehat{ABC} un triángulo rectángulo, donde \hat{A} es el ángulo recto. Sea P el punto donde la perpendicular a BC por A corta a BC . No puede ocurrir que P coincida con B o con C , o de lo contrario el triángulo tendría dos ángulos rectos. Además P ha de estar en el segmento \overline{BC} , pues si, por ejemplo, B estuviera entre P y C los ángulos \widehat{PBA} y \widehat{CBA} serían adyacentes, pero el primero forma parte del triángulo \widehat{PBA} , que tiene un ángulo recto en \hat{P} , luego \widehat{PBA} sería agudo, y \widehat{CBA} sería obtuso, lo cual es imposible. Por lo tanto P está ciertamente en \overline{BC} .



El triángulo rectángulo \widehat{ABP} comparte un ángulo agudo con el triángulo también rectángulo \widehat{CBA} , luego tienen dos—y por consiguiente tres—ángulos iguales, es decir, son semejantes. Así pues

$$\frac{\overline{AB}}{\overline{BP}} = \frac{\overline{BC}}{\overline{AB}},$$

y por consideraciones análogas sobre el triángulo \widehat{ACP} concluimos que

$$\frac{\overline{AC}}{\overline{CP}} = \frac{\overline{BC}}{\overline{AC}}.$$

Por lo tanto

$$\overline{AB}^2 + \overline{AC}^2 = \overline{BC} \overline{BP} + \overline{BC} \overline{CP} = \overline{BC} (\overline{BP} + \overline{CP}) = \overline{BC}^2.$$

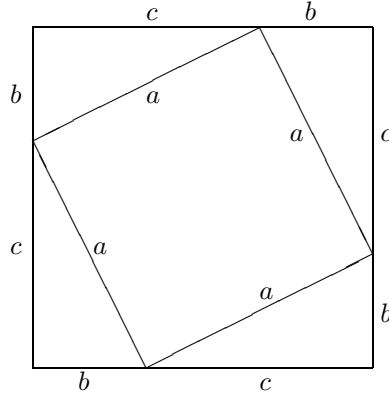
■

De aquí se sigue fácilmente que la diagonal de un cuadrado mide $\sqrt{2}$ veces su lado, luego no guarda una proporción racional éste. Así fue como los griegos descubrieron la existencia de números irracionales.

Otra prueba sencilla del teorema de Pitágoras se basa en el concepto de área. Si observamos la figura, vemos que el área del cuadrado se puede calcular de dos maneras:

$$a^2 + 4 \frac{bc}{2} = (b + c)^2.$$

Desarrollando se obtiene la igualdad $a^2 = b^2 + c^2$.



Esta prueba supone que tenemos definida el área de cualquier figura (o al menos de una familia de figuras que contenga a los triángulos y cuadrados), así como que el área de una unión disjunta de figuras es la suma de las áreas. Probar esto en general resultaría muy laborioso en estos momentos, sin embargo vamos a ver que con unos mínimos resultados sobre áreas se puede justificar el argumento.

Definición 3.20 Se llama *altura* de un triángulo a cualquiera de los segmentos que une perpendicularmente un vértice con la prolongación del lado opuesto. Dicho lado se llama *base* del triángulo correspondiente a la altura considerada. El punto donde la altura corta a la prolongación del lado se llama *pie* de la altura.

Teorema 3.21 El producto de una altura de un triángulo por su base correspondiente no depende de la elección de la altura.

DEMOSTRACIÓN: Sea \widehat{ABC} un triángulo. Sea P el pie de la altura que parte de A y sea Q el pie de la altura que parte de B . Hemos de probar que $\overline{AP} \overline{BC} = \overline{BQ} \overline{AC}$.

Los triángulos \widehat{APC} y \widehat{BQC} tienen los ángulos iguales, pues por una parte \hat{P} y \hat{Q} son rectos y por otra $\widehat{PCA} = \widehat{BCA} = \widehat{BCQ}$. Por lo tanto son semejantes y en consecuencia

$$\frac{\overline{AP}}{\overline{BQ}} = \frac{\overline{AC}}{\overline{BC}}.$$

Esto implica la igualdad buscada. ■

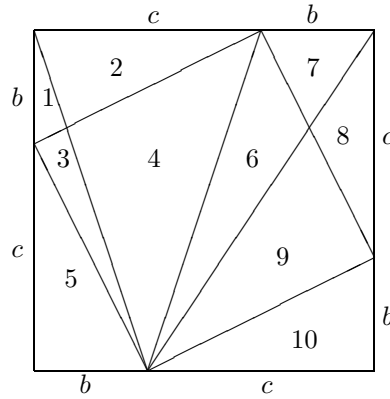
Definición 3.22 Llamaremos *área* de un triángulo \widehat{ABC} de altura h y base correspondiente b al número real

$$(ABC) = \frac{1}{2}bh.$$

Todos los razonamientos que haremos en lo sucesivo en los que intervenga el concepto de área se justificarán por el siguiente hecho evidente:

Si unimos un vértice de un triángulo con un punto del lado opuesto, obtenemos dos triángulos con la misma altura y cuyas bases suman la base del triángulo original. Por lo tanto el área del triángulo de partida es la suma de las áreas de los dos triángulos resultantes.

Por ejemplo, una demostración del teorema de Pitágoras se obtiene dividiendo el cuadrado anterior en diez triángulos de acuerdo con la figura siguiente:



Cada paréntesis en la igualdad siguiente significa una aplicación del principio de descomposición que hemos indicado.

$$\begin{aligned} & (1 + 2) + (7 + 8) + 5 + 10 + (3 + 4) + (6 + 9) \\ &= ((1 + 3) + 5) + ((2 + 4) + (6 + 7)) + ((8 + 9) + 10). \end{aligned}$$

El primer miembro es $2bc + a^2$, mientras que el segundo es

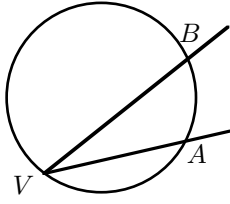
$$\frac{(b+c)b}{2} + \frac{(b+c)^2}{2} + \frac{(b+c)c}{2} = (b+c) \frac{b+b+c+c}{2} = (b+c)^2,$$

con lo que llegamos también a que $a^2 = b^2 + c^2$.

3.3 Relaciones entre ángulos y arcos

Definición 3.23 Diremos que un ángulo está *inscrito* en una circunferencia ω si su vértice está en ω y sus lados cortan a ω (en otros puntos distintos de su origen).

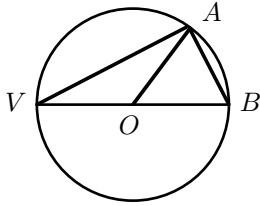
Cuando digamos que un ángulo \widehat{AVB} está inscrito en una circunferencia ω se sobreentenderá que A , V y B están en ω .



En estas condiciones, la intersección de \widehat{AVB} con ω es V y uno de los arcos de extremos A y B , concretamente el que no contiene a V . En efecto, si P es un punto de la intersección distinto de V , entonces la semirrecta \overrightarrow{VP} está contenida en el ángulo, luego corta a \overline{AB} , luego P y V están en semiplanos distintos respecto a AB , luego en arcos distintos. El recíproco se prueba igualmente.

Llamaremos *arco abarcado* por un ángulo inscrito en una circunferencia al arco formado por los puntos del ángulo contenidos en ella y distintos de su vértice.

Teorema 3.24 *Un ángulo inscrito en una circunferencia mide la mitad del arco que abarca.*



DEMOSTRACIÓN: Sea \widehat{AVB} un ángulo inscrito en una circunferencia ω de centro O . Supongamos primeramente que O está en uno de los lados del ángulo, por ejemplo en \overline{VB} . Sea α la amplitud de \widehat{AVB} . Como el triángulo $\triangle VOA$ es isósceles, tiene dos ángulos iguales a α , luego el ángulo $\angle VOA$ mide $\pi - 2\alpha$, luego \widehat{AOB} mide 2α .

Así pues, el arco menor \widehat{AB} tiene amplitud 2α , y éste es precisamente el arco abarcado, pues claramente no puede contener a V .

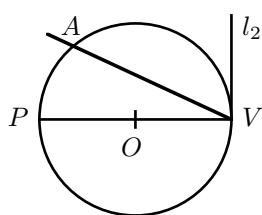
Supongamos ahora que O está contenido en \widehat{AVB} (pero no en su frontera). Al igual que antes, sea α su amplitud. Sea P el punto diametralmente opuesto a V . Entonces la semirrecta \overrightarrow{VP} divide el ángulo en suma de dos ángulos inscritos \widehat{AVP} y \widehat{PVB} . El arco abarcado \widehat{AB} será la unión de los arcos abarcados por los sumandos, \widehat{AP} y \widehat{PB} . Como el único punto en común entre estos dos es P , la amplitud de la unión es la suma de las amplitudes. Si llamamos α_1 y α_2 a las amplitudes de \widehat{AVP} y \widehat{PVB} , por el caso anterior tenemos que la amplitud de \widehat{AB} es $2\alpha_1 + 2\alpha_2 = 2\alpha$.

Por último supongamos que O no está contenido en \widehat{AVB} . Sea como antes P el punto diametralmente opuesto a V . Ahora A y B están en un mismo semiplano respecto a VP , luego uno de los ángulos \widehat{AVP} o \widehat{BVP} está contenido en el otro. Supongamos que el segundo es el menor. Entonces tenemos la

descomposición $\widehat{AVP} = \widehat{AVB} + \widehat{BVP}$, y la misma relación se tiene entre los arcos abarcados por los tres (igual que en el caso anterior). Aplicando el primer caso a los ángulos \widehat{AVP} y \widehat{BVP} , obtenemos la relación buscada. ■

El teorema anterior es válido también en el caso en que uno de los lados del ángulo sea tangente a la circunferencia. Entonces tenemos lo que se llama un ángulo semiinscrito:

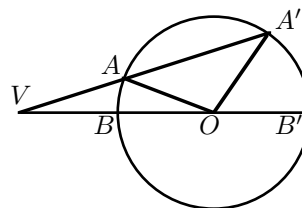
Definición 3.25 Diremos que un ángulo está *semiinscrito* en una circunferencia ω si su vértice está en ω , uno de sus lados es tangente a ω y el otro es secante.



Si $\widehat{l_1 l_2}$ está semiinscrito a ω , l_1 corta a ω en A y l_2 es tangente, es fácil ver que el arco abarcado por el ángulo, es decir, su intersección con ω es el arco \widehat{VA} contenido en el mismo semiplano que l_2 respecto a VA . Distinguiendo tres casos, según que el ángulo sea agudo, recto u obtuso y usando el teorema anterior, es fácil probar que la amplitud de un ángulo semiinscrito

es también la mitad de la del arco que abarca.

Supongamos ahora que un ángulo tiene su vértice V en el exterior de una circunferencia ω , y que sus lados cortan a ω en puntos A, A' y B, B' respectivamente. Supongamos que A y B están más cerca de V que A' y B' . Entonces la intersección del ángulo con ω está formada por dos arcos \widehat{AB} y $\widehat{A'B'}$ sin puntos comunes. Concretamente, \widehat{AB} es el arco contenido en el mismo semiplano que V respecto a su cuerda, mientras que $\widehat{A'B'}$ es el arco contenido en el semiplano opuesto a V respecto a su cuerda.



Teorema 3.26 La medida de un ángulo exterior a una circunferencia cuyos lados sean secantes a la misma es la semidiferencia de los arcos que abarca.

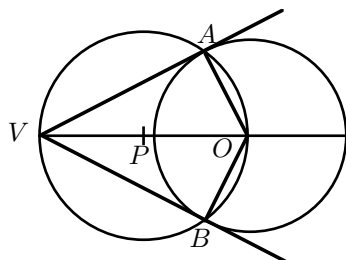
ESBOZO DE LA PRUEBA: El caso general se reduce al caso en que el centro O de la circunferencia está contenido en uno de los lados, por ejemplo en $\overline{BB'}$. Entonces

$$\begin{aligned} \widehat{AVB} &= \pi - \widehat{VBA} - \widehat{BAV} = \pi - (\pi - \widehat{ABO}) - (\pi - \widehat{BAO} - \widehat{OAA'}) \\ &= \widehat{ABO} + \widehat{BAO} + \widehat{OAA'} - \pi = \pi - \widehat{AB} + \pi/2 - \frac{\widehat{AA'}}{2} - \pi \\ &= \frac{\widehat{A'B'} - \widehat{AB}}{2}, \end{aligned}$$

usando que $\widehat{AB} + \widehat{AA'} + \widehat{A'B'} = \pi$. ■

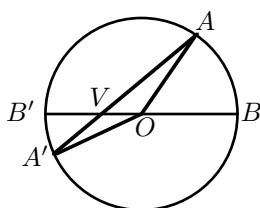
En el capítulo anterior comentamos que del axioma D se desprende que por un punto exterior a una circunferencia pasan dos tangentes a la misma. Veamos ahora una construcción explícita de dichas tangentes basada en los resultados que hemos probado.

Sea, pues ω una circunferencia de centro O y V un punto exterior. Sea P el punto medio de \overline{VO} . Entonces la circunferencia de centro P y radio \overline{VP} tiene un punto interior y un punto exterior a ω , luego corta a ésta en dos puntos distintos A y B .



Los ángulos \widehat{VAO} y \widehat{VBO} son rectos, pues abarcan arcos de amplitud π , luego las rectas VA y VB son tangentes a ω . Observar que la distancia de V a los dos puntos de tangencia es la misma.

Ejercicio: Probar que la amplitud de V es la semidiferencia de la de los arcos en que los puntos de tangencia dividen a ω . Generalizar este hecho al caso en que un lado sea tangente y el otro secante.



El último caso posible es aquel en que el vértice del ángulo está en el interior de la circunferencia. Aquí hemos de considerar tanto el arco \widehat{AB} abarcado por el ángulo dado como el arco $\widehat{A'B'}$ abarcado por su opuesto por el vértice. Dejamos a cargo del lector precisar las definiciones de estos arcos así como la prueba del teorema siguiente. Respecto a ella digamos tan sólo que se reduce como siempre al caso en que O está en uno de los lados (además podemos suponer que es distinto de V , o el resultado es trivial) y que entonces basta estudiar los triángulos \widehat{VOA} y $\widehat{AOA'}$.

Teorema 3.27 *La amplitud de un ángulo interior a una circunferencia es la semisuma de los arcos abarcados por él y por su opuesto por el vértice.*

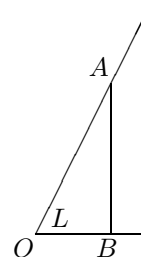
3.4 Las razones trigonométricas

Sabemos que los ángulos y los lados de un triángulo están sometidos a muchas relaciones. Por ejemplo, si conocemos los ángulos de un triángulo y uno de sus lados, los otros dos lados están completamente determinados, pero en principio no sabemos cómo calcularlos a partir de estos datos. Este tipo de problemas constituye el objeto de la *trigonometría* (gr. ‘medida de triángulos’), aunque ahora no profundizaremos mucho en ella, pues más adelante estaremos en condiciones de hacerlo con más fluidez.

Definición 3.28 Sea L un ángulo agudo de vértice O , sea A un punto arbitrario en uno de sus lados y sea B el punto donde la perpendicular al lado opuesto por A corta a dicho lado.

Entonces el triángulo \widehat{AOB} tiene un ángulo recto y otro igual a L , luego sus ángulos son independientes de la elección de A e incluso del lado en que lo tomamos (más aún, sólo dependen de la clase de congruencia de L). Si tomamos puntos distintos obtendremos triángulos semejantes, luego podemos definir el *seno* y el *coseno* de L como las razones

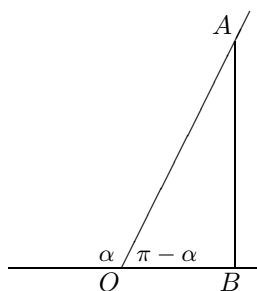
$$\operatorname{sen} L = \frac{\overline{AB}}{\overline{OA}}, \quad \cos L = \frac{\overline{OB}}{\overline{OA}}.$$



En particular podemos tomar el punto A a una unidad de distancia de O , y entonces el seno y el coseno de L son simplemente \overline{AB} y \overline{OB} . El teorema de Pitágoras nos da entonces la que llamaremos *relación fundamental*:

$$\operatorname{sen}^2 L + \cos^2 L = 1.$$

Puesto que el seno y el coseno dependen sólo de las clases de congruencia de los ángulos, podemos definir el seno y el coseno de un número real positivo menor que π como el seno y el coseno de los ángulos de amplitud correspondiente.



Las definiciones de seno y coseno valen tal cual para ángulos obtusos, pero en este caso convendremos en que el coseno tiene signo negativo. Es claro entonces que se cumplen las relaciones:

$$\operatorname{sen} \alpha = \operatorname{sen}(\pi - \alpha), \quad \cos \alpha = -\cos(\pi - \alpha).$$

Se sigue cumpliendo la relación fundamental.

Observemos que el seno y el coseno pueden caracterizarse del modo siguiente: dado un ángulo L , si tomamos un punto en uno de sus lados a una distancia r del vértice, su proyección sobre el otro se encuentra a una distancia $r \cos L$ del vértice (donde el signo es negativo si el punto de corte está fuera del lado opuesto) y la distancia entre el punto y su proyección es $r \operatorname{sen} L$. Ahora bien, todo esto tiene sentido también si el ángulo es recto, en cuyo caso la proyección del punto escogido es el propio vértice del ángulo, de modo que todo lo dicho se sigue cumpliendo si definimos

$$\operatorname{sen} \frac{\pi}{2} = 1, \quad \cos \frac{\pi}{2} = 0.$$

Para cada número real α en el intervalo $] -1, 1[$ es fácil construir un ángulo con coseno igual a α , que claramente será único. Así mismo, si α está en el intervalo $] 0, 1]$ existen exactamente dos ángulos (suplementarios) con seno igual a α .

La relación entre las razones trigonométricas y los problemas sobre triángulos viene dada por el teorema del coseno y el teorema de los senos, que demostramos a continuación. El teorema del coseno es una generalización del teorema de Pitágoras:

Teorema 3.29 (Teorema del coseno) *Todo triángulo \widehat{ABC} verifica la relación*

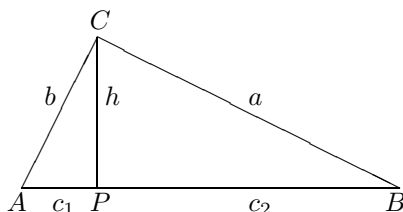
$$a^2 = b^2 + c^2 - 2bc \cos \hat{A}.$$

DEMOSTRACIÓN: Sea P el pie de la altura del triángulo desde C (que puede o no estar en \overline{AB}). Podemos suponer que el triángulo no es rectángulo, pues si no el teorema se reduce al de Pitágoras. Entonces P es distinto de B y C . Sean c_1 , c_2 y h las longitudes indicadas en la figura. Por el teorema de Pitágoras se cumple

$$a^2 = h^2 + c_2^2, \quad b^2 = h^2 + c_1^2,$$

luego $a^2 = b^2 - c_1^2 + c_2^2$.

Por otra parte, $c = c_2 \pm c_1$, donde el signo es positivo si A es agudo y negativo si es obtuso. Así, $c^2 = (c \mp c_1)^2 = c^2 + c_1^2 \mp 2cc_1$, con lo que $a^2 = b^2 + c^2 \mp 2cc_1$. Ahora bien, $c_1 = \pm b \cos \hat{A}$, luego concluimos que $a^2 = b^2 + c^2 - 2bc \cos \hat{A}$. ■



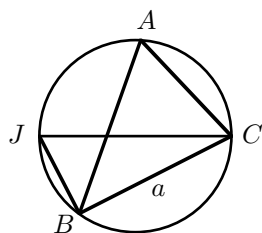
Para probar el teorema de los senos recordamos que la mediatriz de un segmento es la perpendicular por su punto medio, así como que está formada por los puntos que equidistan de sus extremos. Las mediatrices de dos lados de un triángulo \widehat{ABC} , por ejemplo, \overline{AB} y \overline{BC} se cortan en un punto O que equidista de A y B por una parte y de B y C por otra, luego equidista de los tres vértices, lo que implica que también se encuentra sobre la tercera mediatriz.

Definición 3.30 Se llama *circuncentro* de un triángulo a la intersección O de las mediatrices de sus lados. La distancia R del punto O a cualquiera de los vértices se llama *circunradio* del triángulo, la circunferencia de centro O y radio R se llama *circunferencia circunscrita* al triángulo.

Claramente la circunferencia circunscrita a un triángulo es la única que pasa por sus vértices.

Teorema 3.31 (Teorema de los senos) *Si R es el circunradio de un triángulo \widehat{ABC} , entonces se cumple*

$$\frac{a}{\sin \hat{A}} = \frac{b}{\sin \hat{B}} = \frac{c}{\sin \hat{C}} = 2R.$$



DEMOSTRACIÓN: Sea ω la circunferencia circunscrita del triángulo. Sea J el punto diametralmente opuesto a C . Si $J = B$ entonces el ángulo \hat{A} abarca un semicírculo, luego es un ángulo recto, y trivialmente se cumple

$$\frac{a}{\sin \hat{A}} = \frac{2R}{1} = 2R.$$

Si $B \neq J$ entonces los ángulos \hat{J} y \hat{A} abarcan arcos que comparten la cuerda \overline{BC} . Esto implica que los ángulos son iguales o suplementarios, luego tienen el mismo seno. Por otra parte el ángulo \widehat{JBC} es recto, pues abarca un semicírculo. Así pues

$$\frac{a}{\sin \hat{A}} = \frac{a}{\sin \hat{J}} = \frac{a}{a/2R} = 2R.$$

■

Si tenemos datos suficientes sobre un triángulo (los tres lados, dos lados y el ángulo que forman, dos ángulos y un lado) los teoremas anteriores nos permiten calcular los lados y ángulos restantes en términos de senos y cosenos. Naturalmente esto no es de mucha utilidad práctica si no sabemos calcular efectivamente las razones trigonométricas de un ángulo, cosa que de momento no estamos en condiciones de hacer. Pese a ello, la trigonometría es de gran utilidad teórica. Un ejemplo sencillo nos lo proporciona el teorema siguiente, que proporciona una expresión para el área de un triángulo donde se ve explícitamente su carácter invariante.

Teorema 3.32 *El área de un triángulo \widehat{ABC} viene dada por*

$$(ABC) = \frac{abc}{4R},$$

donde R es el circunradio del triángulo.

DEMOSTRACIÓN: La altura de \widehat{ABC} que parte de A es claramente $b \sin \hat{C}$, y por el teorema de los senos $\sin \hat{C} = c/2R$. La conclusión es obvia. ■

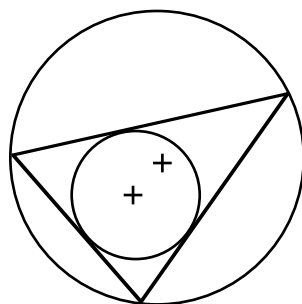
3.5 Propiedades de los triángulos

Terminaremos el capítulo introduciendo algunos conceptos adicionales relacionados con los triángulos. Con la teoría que hemos desarrollado estamos en condiciones de probar resultados interesantes, aunque en los próximos capítulos profundizaremos más en ellos. Cada triángulo tiene asociados varios puntos de interés. Ya conocemos el circuncentro, que es el punto de intersección de sus mediatrices. Similarmente podemos considerar el punto de intersección de sus bisectrices.

El incentro Recordemos que la bisectriz de un ángulo es la recta que lo divide en dos ángulos iguales, y es fácil probar que está formada por los puntos que equidistan de sus lados (la distancia de un punto a una recta es la longitud del segmento que lo une perpendicularmente con ella). Si dos bisectrices de un triángulo se cortan en un punto I , entonces I equidista de los tres lados del triángulo, luego está también sobre la tercera bisectriz.

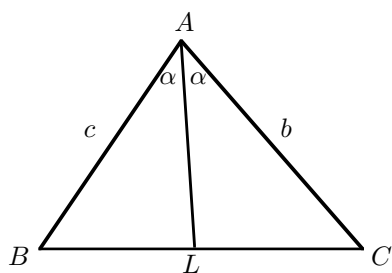
Definición 3.33 Se llama *incentro* de un triángulo al punto I en el que se cortan sus bisectrices. La distancia r de I a cualquiera de los lados se llama *inradio* del triángulo. La circunferencia de centro I y radio r se llama *circunferencia inscrita* al triángulo.

El punto donde la perpendicular por el incentro a uno de los lados corta al lado pertenece a la circunferencia inscrita, luego los lados del triángulo son tangentes a la circunferencia.



Circunferencias inscrita y circunscrita de un triángulo

En principio no conocemos los puntos de tangencia, pero podemos determinarlos. Para ello consideramos el punto L donde una bisectriz corta al lado opuesto.



Notemos que los dos ángulos con vértice L son suplementarios, luego tienen el mismo seno, al cual podemos llamar sin ambigüedad $\sin \hat{L}$. El teorema de los senos implica entonces que

$$\frac{\overline{BL}}{\sin \alpha} = \frac{c}{\sin \hat{L}}, \quad \frac{\overline{LC}}{\sin \alpha} = \frac{b}{\sin \hat{L}},$$

con lo que

$$\frac{\overline{BL}}{\overline{LC}} = \frac{c}{b}.$$

Por lo tanto hemos demostrado:

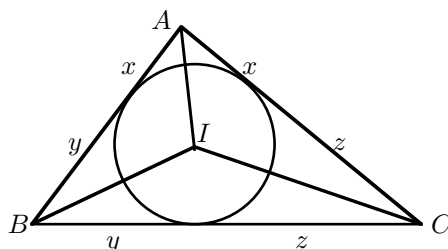
Teorema 3.34 Una bisectriz de un triángulo divide a su lado opuesto en segmentos proporcionales a los lados adyacentes.

Todavía podemos decir más. Llamemos x a la distancia entre A y los puntos donde la circunferencia inscrita toca a los lados \overline{AB} y \overline{AC} (hemos visto que las tangentes por un punto exterior tienen la misma longitud). Similarmente, sea y la distancia de B a los puntos de tangencia en \overline{BA} y \overline{BC} y z la distancia de C a los puntos de tangencia en \overline{CA} y \overline{CB} . Entonces

$$a = y + z, \quad b = x + z, \quad c = x + y.$$

Euler introdujo la costumbre de llamar s al semiperímetro de un triángulo, es decir, $s = (a + b + c)/2$. En estos términos $x + y + z = s$, y

$$x = s - a, \quad y = s - b, \quad z = s - c.$$



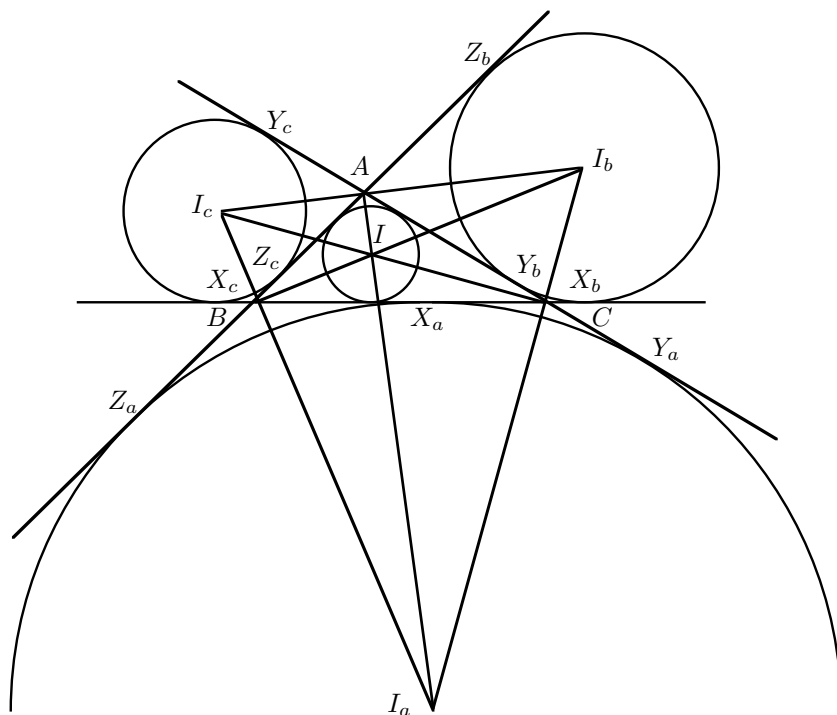
Además, observamos que el triángulo \widehat{AIB} tiene base c y altura r (el inradio) luego su área es $cr/2$. Calculando del mismo modo el área de los otros dos triángulos \widehat{AIC} y \widehat{BIC} concluimos que

$$(ABC) = sr.$$

(Es fácil justificar que (ABC) es la suma de las áreas de los tres triángulos.)

Los excentros Se llaman *ángulos exteriores* de un triángulo los ángulos adyacentes a sus ángulos (interiores). Cada vértice tiene, pues, dos ángulos exteriores opuestos por el vértice y, por lo tanto, iguales. Los dos ángulos exteriores por un vértice comparten la bisectriz. Consideremos el punto I_a donde se cortan las bisectrices exteriores de los ángulos B y C (es fácil probar que no pueden ser paralelas). Entonces I_a equidista de las prolongaciones de los tres lados, luego está también sobre la bisectriz del ángulo (interior) A . Hemos probado:

Las bisectrices exteriores de dos ángulos de un triángulo son concurrentes con la bisectriz interna del tercer ángulo.



Definición 3.35 Llamaremos *excentros* de un triángulo a los puntos de intersección de cada par de bisectrices exteriores con la bisectriz interior del tercer ángulo. Los representaremos por I_a, I_b, I_c . Las distancias de cada excentro a las prolongaciones de los lados se llaman *exradios*, r_a, r_b, r_c . Las circunferencias determinadas por los excentros y los exradios se llaman *circunferencias excritas* al triángulo y son tangentes a un lado y a las prolongaciones de los otros dos. Las circunferencias excritas y la circunferencia inscrita se llaman *circunferencias tritangentes* al triángulo.

Es fácil determinar los puntos donde los excírculos tocan a los lados. Notemos que $\overline{BX_b} = \overline{BZ_b}$ y por otro lado

$$\begin{aligned}\overline{BX_b} + \overline{BZ_b} &= \overline{BC} + \overline{CX_b} + \overline{BA} + \overline{AZ_b} \\ &= \overline{BC} + \overline{CY_b} + \overline{BA} + \overline{AY_b} = a + b + c = 2s.\end{aligned}$$

Por consiguiente $\overline{BX_b} = s$, luego $\overline{CY_b} = \overline{CX_b} = \overline{BX_b} - \overline{BC} = s - a$. Similarmente podemos calcular la distancia de cualquier vértice a cualquier punto de tangencia.

El teorema de Ceva Los triángulos tienen asociadas varias ternas de rectas concurrentes, como, por ejemplo, las bisectrices. El matemático italiano Giovanni Ceva (1647–1734) obtuvo un resultado general sobre este tipo de rectas, a raíz del cual se llaman *cevianas* de un triángulo a las rectas que pasan por un vértice y cortan a la prolongación del lado opuesto en puntos distintos de

los vértices, de modo que cuando hablamos de tres cevianas en un triángulo se sobrentiende que cada una pasa por uno de los vértices.

Para enunciar el teorema de Ceva conviene introducir un convenio: si A , B , C y D son cuatro puntos colineales, consideraremos que la razón

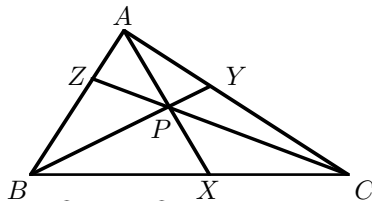
$$\frac{\overline{AB}}{\overline{CD}}$$

es positiva si las relaciones $<_{AB}$ y $<_{CD}$ son iguales, mientras que será negativa si las relaciones son mutuamente inversas. De este modo, dados tres puntos colineales A , B y C con $A \neq B$, existe un único punto D colineal con ellos y para el que la proporción anterior tome un valor dado (en principio habría dos puntos posibles, uno en cada semirrecta de origen C , pero el ajuste del signo descarta a uno de ellos).

Teorema 3.36 (Ceva) *En un triángulo arbitrario, tres cevianas \overline{AX} , \overline{BY} , \overline{CZ} son concurrentes si y sólo si*

$$\frac{\overline{BX}}{\overline{XC}} \frac{\overline{CY}}{\overline{YA}} \frac{\overline{AZ}}{\overline{ZB}} = 1.$$

DEMOSTRACIÓN: Supongamos en primer lugar que las cevianas se cortan en un punto P . Por definición de ceviana el punto X es distinto de B y C . Así mismo, el punto P no puede estar en la recta BC , o de lo contrario sería $Y = C$. Por lo demás X puede estar o no entre B y C . En cualquier caso los triángulos \widehat{ABX} y \widehat{AXC} tienen la misma altura, así como \widehat{PBX} y \widehat{PXC} , luego



$$\left| \frac{\overline{BX}}{\overline{XC}} \right| = \frac{(ABX)}{(AXC)} = \frac{(PBX)}{(PXC)} = \frac{\pm(ABX) \pm (PBX)}{\pm(AXC) \pm (PXC)} = \frac{(ABP)}{(ACP)}.$$

(Notar que la elección del signo es la misma en el numerador y el denominador, según la posición de P respecto a A y X .)

Del mismo modo obtenemos

$$\left| \frac{\overline{CY}}{\overline{YA}} \right| = \frac{(BCP)}{(BAP)}, \quad \left| \frac{\overline{AZ}}{\overline{ZB}} \right| = \frac{(CAP)}{(CBP)}.$$

Multiplicando las tres igualdades resulta

$$\frac{\overline{BX}}{\overline{XC}} \frac{\overline{CY}}{\overline{YA}} \frac{\overline{AZ}}{\overline{ZB}} = \pm 1.$$

Estudiemos el signo. Si dos de los factores son positivos, por ejemplo los dos primeros, eso significa que X está entre B y C y que Y está entre A y C . Entonces la semirrecta \overrightarrow{AX} está contenida en \hat{A} , luego ha de cortar al segmento

\overrightarrow{BY} , y la intersección es P , luego P está en el triángulo, luego está en \hat{C} , luego \overrightarrow{CZ} ha de cortar a \overline{AB} y la intersección es Z , luego Z está entre A y B y el tercer factor es también positivo.

Sólo falta ver que los tres factores no pueden ser negativos a la vez. Si así fuera, entonces X no está entre B y C . Podemos suponer que C está entre B y X . Por otra parte Y no está entre A y C , luego \overrightarrow{BY} no está contenida en \hat{B} , luego no corta a \overline{AX} , luego P no está entre A y X . Distinguimos dos casos: si es A quien está entre P y X , entonces P está en el mismo semiplano que A respecto a BC y en semiplano opuesto a X respecto a AC , luego en el mismo que B . Por consiguiente P está en \hat{C} , luego \overrightarrow{CP} corta a \overline{AB} , y la intersección es Z , con lo que el tercer factor es positivo. Si por el contrario es X quien está entre A y P entonces P está en el mismo semiplano que X respecto a AC , luego en el opuesto a B , y también en el semiplano opuesto a A respecto a BC , luego está en el ángulo opuesto por el vértice a \hat{C} . Por consiguiente la semirrecta complementaria de \overrightarrow{CP} está contenida en \hat{C} , luego corta a \overline{AB} y de nuevo la intersección ha de ser Z .

Probemos el recíproco. Si se cumple la relación indicada, sea P el punto donde AX corta a BY y sea CZ' la ceviana que pasa por P . Notar que Z' ha de ser distinto de A y B , pues en caso contrario P estaría en uno de los lados y X o Y coincidiría con un vértice. Podemos aplicar la parte ya probada a las cevianas \overline{AX} , \overline{BY} , $\overline{CZ'}$, con lo que obtenemos la relación

$$\frac{\overline{BX}}{\overline{XC}} \cdot \frac{\overline{CY}}{\overline{YA}} \cdot \frac{\overline{AZ'}}{\overline{Z'B}} = 1.$$

Comparando con la hipótesis concluimos que

$$\frac{\overline{AZ}}{\overline{ZB}} = \frac{\overline{AZ'}}{\overline{Z'B}},$$

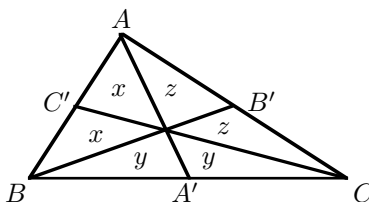
lo cual sólo es posible si $Z = Z'$, luego las cevianas dadas se cortan en P . ■

El baricentro Una aplicación inmediata del teorema de Ceva nos da la intersección de las medianas:

Definición 3.37 Se llaman *medianas* de un triángulo a los segmentos que unen cada vértice con el punto medio del lado opuesto. El punto de intersección de las medianas de un triángulo se llama *baricentro* (gr./lat. ‘centro de peso’), y lo representaremos por G .

El baricentro debe su nombre a que la física demuestra que es el centro de gravedad del triángulo.

Los triángulos marcados con la misma letra en la figura tienen la misma área, pues comparten una altura y sus bases son iguales.



Por la misma razón $(BCC') = (ACC')$, luego $2y + x = 2z + x$ (es claro que podemos sumar las áreas). Así pues $y = z$, y del mismo modo se concluye que las seis áreas son iguales. Esto a su vez implica que $(GAB) = 2(GBA')$, y como los dos triángulos comparten una altura, se ha de cumplir que $\overline{AG} = 2\overline{GA'}$. En otros términos: las medianas de un triángulo se cortan en razón 2 : 1.

El ortocentro Veamos que el teorema de Ceva es aplicable a las alturas de un triángulo no rectángulo (notemos que las alturas de un triángulo rectángulo se cortan trivialmente en el vértice correspondiente al ángulo recto).

Para ello basta observar que

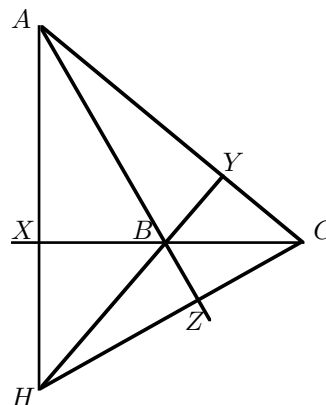
$$\frac{\overline{BX}}{\overline{XC}} = \frac{c \cos \hat{B}}{b \cos \hat{C}},$$

donde el signo del coseno ajusta correctamente el signo de la proporción (tener presente que a lo sumo uno de los ángulos es obtuso).

Las otras dos proporciones se calculan del mismo modo:

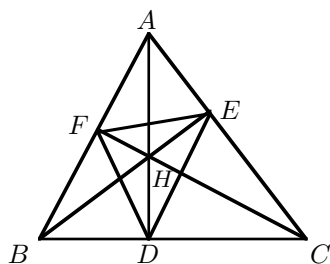
$$\frac{\overline{CY}}{\overline{YA}} = \frac{a \cos \hat{C}}{c \cos \hat{A}}, \quad \frac{\overline{AZ}}{\overline{ZB}} = \frac{b \cos \hat{A}}{a \cos \hat{B}},$$

y es claro que su producto es 1.



Definición 3.38 La intersección de las alturas de un triángulo se llama *ortocentro* (gr./lat. ‘centro de perpendiculares’), y lo representaremos por H . Si el triángulo no es rectángulo, los pies de las alturas determinan un triángulo que recibe el nombre de *triángulo órtico*.

Es fácil comprobar que las bisectrices de dos ángulos adyacentes son perpendiculares, de donde se sigue que todo triángulo es el triángulo órtico del triángulo formado por sus excentros. Veamos que se cumple el recíproco. Consideremos un triángulo acutángulo.



Es claro que el circuncentro de un triángulo rectángulo es el punto medio de su hipotenusa. Por lo tanto, los triángulos \widehat{BFH} y \widehat{BHD} tienen el mismo circuncentro, luego los cuatro puntos B , F , E y D están sobre un mismo círculo. Por lo tanto los ángulos \widehat{FBH} y \widehat{FDH} son iguales (abarcan el mismo arco). El primero es claramente $\pi - \hat{A}$, luego lo mismo vale para el segundo. Por el mismo argumento $\widehat{EDH} = \pi - \hat{A}$. Vemos, pues, que la altura \overline{AD} divide el ángulo \hat{D} del triángulo órtico en dos ángulos iguales, o sea, es la bisectriz de \hat{D} . Además entonces el lado \overline{BC} es la bisectriz de los ángulos adyacentes de \hat{D} . Como esto es válido por igual para todos los vértices, hemos probado lo siguiente:

El ortocentro y los vértices de un triángulo acutángulo coinciden respectivamente con el incentro y los excentros de su triángulo órtico.

La situación es similar en el caso de triángulos obtusángulos. De hecho, si en la figura anterior consideramos que \widehat{AHC} es un triángulo obtusángulo arbitrario, entonces \widehat{DEF} sigue siendo el triángulo órtico y B es el ortocentro, de modo que en un triángulo obtusángulo, el vértice correspondiente al ángulo obtuso es el incentro del triángulo órtico y el ortocentro junto con los otros dos vértices son los excentros. Dejamos las comprobaciones a cargo del lector.

Para terminar señalamos que los triángulos \widehat{AEF} , \widehat{DFB} y \widehat{DEC} son todos semejantes a \widehat{ABC} .

Capítulo IV

La geometría analítica

En el siglo XVII, Descartes revolucionó la geometría al descubrir la *geometría analítica*, una potente técnica capaz de convertir los problemas geométricos en problemas algebraicos equivalentes y, a menudo, más fáciles de tratar.¹ Por contraposición, el tratamiento de la geometría que hemos seguido hasta ahora recibe el nombre de *geometría sintética*. La primera sección de este capítulo contiene los resultados técnicos necesarios para que el tránsito de una a la otra se produzca sin salto lógico alguno.

4.1 Vectores

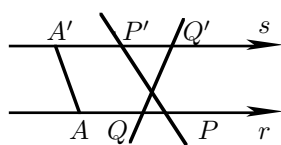
Definición 4.1 Una *recta orientada* es una recta en la que hemos fijado una de las ordenaciones $<_{AB}$ determinadas por dos cualesquiera de sus puntos.

Una misma recta da lugar a dos rectas orientadas distintas. Si fijamos un punto P en una recta orientada r , entonces P divide a r en dos semirrectas, una formada por los puntos menores o iguales que P y otra por los puntos mayores o iguales que P . Las llamaremos, respectivamente, semirrecta menor y semirrecta mayor.

Diremos que dos rectas paralelas orientadas r y s tienen *la misma orientación* si existe una recta t que las cruza en puntos P y P' de modo que las semirrectas menores determinadas por estos puntos estén contenidas en el mismo semiplano respecto de t .

Veamos que esta definición no depende de la recta t escogida. En efecto, supongamos que esto se cumple con una recta t y que t' es cualquier otra recta que corte a r y a s en dos puntos Q y Q' . Tomemos puntos A y A' en r y s que sean menores que P , Q , P' , Q' respectivamente. Basta probar que t' no corta al segmento $\overline{AA'}$, pues entonces A y A' están en el mismo semiplano respecto a t' , y como están ambos en las semirrectas menores respecto a P' y Q' , concluimos que ambas están en el mismo semiplano respecto a t' .

¹Al parecer, Descartes partió de una sugerencia que le hizo Fermat en una carta.



Supongamos que $\overline{AA'}$ corta a t' . Es fácil ver que entonces de hecho corta al segmento $\overline{QQ'}$ (El primer segmento está contenido en la *banda* determinada por r y s , esto es, en la intersección del semiplano de frontera r que contiene a s con el semiplano de frontera s que contiene a r . Por otra parte la intersección de t' con esta banda es $\overline{QQ'}$, luego si la recta corta a $\overline{AA'}$ el punto de corte ha de estar en dicho segmento.)

Si $\overline{AA'}$ corta a $\overline{QQ'}$ y $P \neq Q$, entonces el axioma B5 aplicado al triángulo $\overline{PQQ'}$ obtenemos que $\overline{AA'}$ corta a $\overline{PQ'}$ (no puede cortar al otro lado \overline{PQ} porque entonces A estaría entre P y Q). Si $P = Q$ tenemos directamente que $\overline{AA'}$ corta a $\overline{PQ'}$. Por el mismo argumento aplicado al triángulo $\overline{PQ'P'}$ llegamos a que $\overline{AA'}$ corta a $\overline{P'P'}$, pero el punto de corte ha de estar en $\overline{AA'}$, con lo que A y A' están en semiplanos distintos respecto de t , en contradicción con la hipótesis. ■

Así pues, siempre que una recta t corta a dos rectas paralelas igualmente orientadas r y s , las semirrectas menores quedan en el mismo semiplano (y las mayores también).

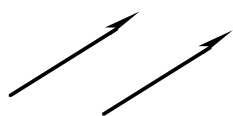
A partir de aquí se sigue inmediatamente que el estar igualmente orientado es una relación de equivalencia, es decir, que si dos rectas paralelas están orientadas igual que una tercera, entonces están igualmente orientadas. Por lo tanto, cada haz de rectas paralelas admite dos orientaciones posibles.

Definición 4.2 Un *vector fijo* (lat ‘transportador’) es un par ordenado de puntos del plano. Lo representaremos \overrightarrow{PQ} (no confundir con la notación para las semirrectas). El punto P se llama *origen* del vector \overrightarrow{PQ} , mientras que el punto Q es el *extremo*. Si el origen es igual al extremo diremos que el vector es *nulo*.

Llamaremos *norma* de un vector fijo \overrightarrow{PQ} a la longitud $\|\overrightarrow{PQ}\|$ del segmento \overline{PQ} (respecto a una unidad de longitud prefijada). Convenimos que la longitud de los vectores nulos es 0.

Llamaremos *dirección* de un vector fijo no nulo \overrightarrow{PQ} al haz de todas las rectas paralelas a PQ , de modo que dos vectores \overrightarrow{PQ} y $\overrightarrow{P'Q'}$ tienen la misma dirección si y sólo si las rectas PQ y $P'Q'$ son paralelas o iguales.

Llamaremos *sentido* de un vector fijo no nulo \overrightarrow{PQ} al conjunto de las relaciones sobre las rectas paralelas a PQ que inducen la misma orientación que \angle_{PQ} . De este modo, los vectores fijos no nulos con una misma dirección quedan divididos en dos clases de vectores de sentidos opuestos. Los vectores \overrightarrow{PQ} y \overrightarrow{QP} tienen sentidos opuestos.



Vectores equipolentes

Diremos que dos vectores fijos \overrightarrow{PQ} y $\overrightarrow{P'Q'}$ son *equipolentes* (lat. de igual fuerza) si ambos son nulos o si ambos son no nulos y tienen la misma norma, la misma dirección y el mismo sentido.

Es claro que la equipolencia es una relación de equivalencia.

Un *vector libre* es una clase de equipolencia de vectores fijos. Si \vec{v} es un vector libre, a los vectores fijos que lo componen los llamaremos *trasladados* de \vec{v} . Si \overrightarrow{PQ} es un trasladado de \vec{v} escribiremos simplemente $\vec{v} = \overrightarrow{PQ}$. Por definición de equipolencia, los vectores nulos forman un mismo vector libre al que representaremos por $\vec{0}$. Llamaremos norma, dirección y sentido de un vector libre al de cualquiera de sus trasladados.

Dado un punto P y un vector libre \vec{v} , existe un único punto Q tal que $\vec{v} = \overrightarrow{PQ}$. En efecto, si $\vec{v} = \vec{0}$ es claro. Si $\vec{v} = \overrightarrow{AB}$, un vector con origen P y la misma dirección que \vec{v} ha de tener el extremo en la recta paralela (o igual) a AB por P . En dicha recta hay dos puntos Q tales que \overrightarrow{PQ} tiene la misma norma que \vec{v} , y cada uno de ellos determina una ordenación $<_{PQ}$ diferente, luego sólo uno de ellos hace que \overrightarrow{PQ} tenga el mismo sentido que \vec{v} .

Al único punto Q que cumple $\overrightarrow{PQ} = \vec{v}$ lo llamaremos *trasladado* de P por \vec{v} y lo representaremos por $Q = P + \vec{v}$. También diremos que \overrightarrow{PQ} es el *trasladado* de \vec{v} de origen P . En particular tenemos $P + \overrightarrow{PQ} = Q$.

Es claro que dos vectores libres no nulos \vec{v} y \vec{w} tienen la misma dirección si y sólo si al tomar trasladados $\vec{v} = \overrightarrow{PQ}$ y $\vec{w} = \overrightarrow{P'Q'}$ con origen común se cumple que P, Q y Q' son colineales. Además tendrán el mismo sentido si y sólo si Q y Q' están en la misma semirrecta respecto a P .

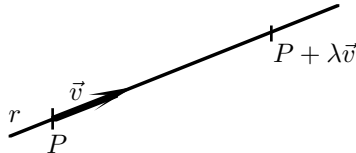
Dado un vector libre $\vec{v} \neq \vec{0}$ y un número real $\alpha \neq 0$, definimos el vector libre $\alpha\vec{v}$ como el único vector con norma $|\alpha| \|\vec{v}\|$, la misma dirección que \vec{v} y el mismo sentido que \vec{v} o sentido opuesto según si $\alpha > 0$ o $\alpha < 0$. Definimos también $0\vec{v} = \alpha\vec{0} = \vec{0}$. Es fácil comprobar que $(\alpha\beta)\vec{v} = \alpha(\beta\vec{v})$, así como que $1\vec{v} = \vec{v}$.

Dada una recta r , un *vector director* de r es cualquier vector $\vec{v} = \overrightarrow{PQ}$, donde P y Q son dos puntos distintos de r .

Teorema 4.3 Si P es un punto de una recta r y \vec{v} es un vector director, entonces cada punto de r se expresa de forma única como $Q = P + \lambda\vec{v}$, con $\lambda \in \mathbb{R}$.

DEMOSTRACIÓN: Tenemos $\vec{v} = \overrightarrow{PP'}$, para un cierto P' en r . Por lo tanto los vectores \vec{v} y \overrightarrow{PQ} tienen la misma dirección. Por lo tanto existe un $\lambda \in \mathbb{R}$ tal que $\overrightarrow{PQ} = \lambda\vec{v}$, luego $Q = P + \lambda\vec{v}$.

Si $Q = P + \lambda\vec{v} = P + \lambda'\vec{v}$, entonces $\lambda\vec{v} = \lambda'\vec{v}$. Obviamente $\lambda = 0$ si y sólo si $\lambda' = 0$. Si no son nulos, comparando las normas y los sentidos concluimos que $\lambda = \lambda'$. ■



Es fácil ver que $P_\lambda = P + \lambda \vec{v}$ es una graduación de la recta r .

Más en general, si r es una recta graduada y \vec{v} es un vector director, entonces $\vec{v} = \overrightarrow{P_0 P_\beta}$, para algún β . Entonces $Q = P_\alpha + \vec{v}$ cumple que $\overrightarrow{P_\alpha Q} = \vec{v} = \overrightarrow{P_0 P_\beta}$, luego Q está en r y dista $|\beta|$ unidades de P_α , por lo que claramente $Q = P_{\alpha \pm \beta}$.

Si $\beta > 0$ entonces $P_0 < P_\beta$, luego $P_\alpha < Q$ (pues ambos pares de puntos han de inducir el mismo orden en r), luego $Q = P_{\alpha+\beta}$. Similarmente, si $\beta < 0$ ha de ser $Q < P_\alpha$, luego también $Q = P_{\alpha+\beta}$. Así pues, en toda recta graduada

$$P_\alpha + \overrightarrow{P_0 P_\beta} = P_{\alpha+\beta}.$$

En particular

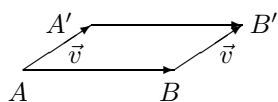
$$P + (\alpha + \beta)\vec{v} = (P + \alpha\vec{v}) + \beta\vec{v}$$

Las propiedades básicas de los vectores se siguen del teorema siguiente.

Teorema 4.4 *Para todo par de puntos A y B y todo vector libre \vec{v} , si $A' = A + \vec{v}$ y $B' = B + \vec{v}$, entonces $\overrightarrow{AB} = \overrightarrow{A'B'}$.*

DEMOSTRACIÓN: Podemos suponer que $\vec{v} \neq \vec{0}$. Supongamos ahora que A , B y A' están alineados. Graduemos la recta que los contiene tomando $P_0 = A$ y $P_1 = B$. Sea $\vec{v} = \overrightarrow{P_0 P_\alpha}$. Entonces $A + \vec{v} = P_\alpha$ y $B + \vec{v} = P_{1+\alpha}$, y claramente $\overrightarrow{P_0 P_1} = \overrightarrow{P_\alpha P_{\alpha+1}}$ (ambos miden una unidad de longitud e inducen la ordenación de la graduación).

Supongamos ahora que A' no está en AB . Entonces las rectas AA' y BB' son paralelas. Si las ordenamos con las relaciones $<_{AA'}$ y $<_{BB'}$ la orientación es la misma, luego A' y B' están en el mismo semiplano respecto de AB .



La paralela a AB que pasa por A' corta a BB' en un punto X que forma un paralelogramo con A , A' , B , luego X dista de B lo mismo que A' dista de A , luego lo mismo que B' dista de B . Como además B' y X están en el mismo semiplano respecto de AB (el que contiene a B), de hecho están en la misma semirrecta de origen B , luego $X = B'$. Así pues, $\overrightarrow{A'B'}$ y \overrightarrow{AB} tienen la misma dirección y la misma norma. Como la recta AA' deja a B y a B' en el mismo semiplano, el sentido también es el mismo (B' y B están en las semirrectas mayores para las ordenaciones que inducen $\overrightarrow{A'B'}$ y \overrightarrow{AB}). Esto prueba que los dos vectores son iguales. ■

Teorema 4.5 *Para todos los puntos A y B y todos los vectores \vec{v} y \vec{w} se cumple:*

1. Si $A + \vec{v} = B + \vec{v}$ entonces $A = B$.
2. Si $A + \vec{v} = A + \vec{w}$ entonces $\vec{v} = \vec{w}$.
3. $(A + \vec{v}) + \vec{w} = (A + \vec{w}) + \vec{v}$.
4. Si $A' = (A + \vec{v}) + \vec{w}$ y $B' = (B + \vec{v}) + \vec{w}$, entonces $\overrightarrow{AA'} = \overrightarrow{BB'}$.

DEMOSTRACIÓN: 1) Si llamamos $C = A + \vec{v} = B + \vec{v}$, el teorema anterior nos da que \overrightarrow{AB} es equipolente a \overrightarrow{CC} , luego ha de ser $A = B$.

2) Si $C = A + \vec{v} = A + \vec{w}$, obviamente $\vec{v} = \overrightarrow{AC} = \vec{w}$.

3) Sean $B = A + \vec{v}$, $A' = A + \vec{w}$ y $B' = B + \vec{w}$. Por el teorema anterior $\overrightarrow{A'B'} = \overrightarrow{AB} = \vec{v}$. Así

$$(A + \vec{v}) + \vec{w} = B + \vec{w} = B', \quad (A + \vec{w}) + \vec{v} = A' + \overrightarrow{A'B'} = B'.$$

4) Sea $\vec{x} = \overrightarrow{AB}$. Entonces

$$B' = (B + \vec{v}) + \vec{w} = ((A + \vec{x}) + \vec{v}) + \vec{w} = ((A + \vec{v}) + \vec{w}) + \vec{x} = A' + \vec{x},$$

y como también $B = A + \vec{x}$, el teorema anterior implica que $\overrightarrow{AB} = \overrightarrow{A'B'}$. ■

La propiedad 4) del teorema anterior nos permite definir la suma de dos vectores \vec{v} y \vec{w} como el vector $\vec{v} + \vec{w}$ que cumple

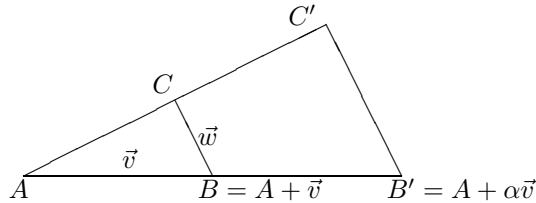
$$A + (\vec{v} + \vec{w}) = (A + \vec{v}) + \vec{w},$$

para todo punto A .

Teorema 4.6 *El conjunto V de los vectores libres con la suma y el producto que hemos definido forma un espacio vectorial sobre \mathbb{R} en el sentido algebraico del término, es decir, se cumplen las propiedades siguientes:*

1. $(\vec{v} + \vec{w}) + \vec{x} = \vec{v} + (\vec{w} + \vec{x})$,
2. $\vec{v} + \vec{w} = \vec{w} + \vec{v}$,
3. $\vec{v} + \vec{0} = \vec{v}$,
4. Para cada vector \vec{v} existe un vector $-\vec{v}$ tal que $\vec{v} + (-\vec{v}) = \vec{0}$,
5. $\alpha(\beta\vec{v}) = (\alpha\beta)\vec{v}$,
6. $(\alpha + \beta)\vec{v} = \alpha\vec{v} + \beta\vec{v}$,
7. $\alpha(\vec{v} + \vec{w}) = \alpha\vec{v} + \alpha\vec{w}$,
8. $1\vec{v} = \vec{v}$.

DEMOSTRACIÓN: Todas las propiedades son inmediatas a partir de lo ya probado salvo la séptima. Las propiedades restantes permiten reducir la prueba al caso en que $\alpha > 0$ y los vectores \vec{v} y \vec{w} tienen direcciones distintas.



Sea A un punto cualquiera, $B = A + \vec{v}$, $C = A + \vec{v} + \vec{w}$ y $B' = A + \alpha\vec{v}$. Sea C' el punto donde la paralela a BC por B' corta a AC .

Entonces $\overrightarrow{B'C'}$ tiene la misma dirección y sentido que \vec{w} , y por el teorema de Tales su norma es $\alpha\|\vec{w}\|$, luego $\overrightarrow{B'C'} = \alpha\vec{w}$ y por consiguiente $\overrightarrow{AC'} = \alpha\vec{v} + \alpha\vec{w}$.

Por otra parte $\overrightarrow{AC'}$ tiene la misma dirección y sentido que $\vec{v} + \vec{w}$ y también por el teorema de Tales, su norma es $\alpha\|\vec{v} + \vec{w}\|$, luego $\overrightarrow{AC'} = \alpha(\vec{v} + \vec{w})$. ■

El teorema 4.3 se interpreta ahora como que una recta está formada por los trasladados de uno cualquiera de sus puntos mediante los vectores de un subespacio vectorial de V de dimensión 1. Vamos a probar un resultado análogo para los planos:

Teorema 4.7 *Sea π un plano, sean P , Q y R tres puntos de π no colineales. Sean $\vec{u} = \overrightarrow{PQ}$ y $\vec{v} = \overrightarrow{PR}$. Entonces cada punto de π se expresa de forma única como*

$$X = P + \lambda\vec{u} + \mu\vec{v}, \quad \lambda, \mu \in \mathbb{R}.$$

DEMOSTRACIÓN: Dado un punto X en π , consideramos la recta paralela a PR que pasa por X . Ésta cortará a PQ en un punto de la forma $Y = P + \lambda\vec{u}$. El vector \overrightarrow{YX} tiene la misma dirección que $\overrightarrow{PR} = \vec{v}$, luego será de la forma $\overrightarrow{YX} = \mu\vec{v}$. Así

$$X = P + \overrightarrow{PY} + \overrightarrow{YX} = P + \lambda\vec{u} + \mu\vec{v}.$$

La unicidad equivale a que los vectores \vec{u} y \vec{v} sean linealmente independientes, pero esto se sigue de que P , Q y R no son colineales. Es fácil ver de modo similar que todo punto de la forma indicada está en π . ■

Del mismo modo tenemos:

Teorema 4.8 *Sean P , Q , R , S cuatro puntos no coplanares. Sean $\vec{u} = \overrightarrow{PQ}$, $\vec{v} = \overrightarrow{PR}$ y $\vec{w} = \overrightarrow{PS}$. Entonces todo punto X se expresa de forma única como*

$$X = P + \lambda\vec{u} + \mu\vec{v} + \nu\vec{w}, \quad \lambda, \mu, \nu \in \mathbb{R}.$$

DEMOSTRACIÓN: La recta paralela a PS que pasa por X cortará al plano PQR en un punto que por el teorema anterior será de la forma

$$Y = P + \lambda\vec{u} + \mu\vec{v}.$$

El vector \overrightarrow{YP} tiene la misma dirección que $\overrightarrow{PS} = \vec{w}$, luego $\overrightarrow{YP} = \nu\vec{w}$. Como en el teorema anterior concluimos que X tiene la forma indicada. La expresión es única porque los vectores \vec{u} , \vec{v} y \vec{w} son linealmente independientes (o de lo contrario todo el espacio estaría contenido en un plano). ■

El teorema anterior implica que el espacio de los vectores libres tiene dimensión 3.

4.2 Espacios afines

Los resultados algebraicos que hemos obtenido en la sección anterior se expresan más adecuadamente en términos del concepto siguiente:

Definición 4.9 Un *espacio afín* de dimensión n es una terna $(E, \vec{E}, +)$, donde E es un conjunto no vacío, a cuyos elementos se les llama *puntos*, \vec{E} es un espacio vectorial de dimensión n sobre un cuerpo K y $+$ es una aplicación $E \times \vec{E} \rightarrow E$ que cumple las propiedades siguientes:

1. Para cada par de puntos P, Q , existe un único vector \vec{v} tal que $Q = P + \vec{v}$, y se le representa por $\vec{v} = \overrightarrow{PQ}$.
2. $P + \vec{0} = P$,
3. $(P + \vec{v}) + \vec{w} = P + (\vec{v} + \vec{w})$.

En la sección anterior hemos probado que el espacio de la geometría (tridimensional) euclídea es un espacio afín tridimensional con \mathbb{R} como cuerpo de escalares. Es fácil demostrar que todas las propiedades de vectores que hemos probado allí son válidas en cualquier espacio afín. Por ejemplo,

$$P + \overrightarrow{PQ} + \overrightarrow{QP} = Q + \overrightarrow{QP} = P = P + \vec{0}, \quad \text{luego} \quad \overrightarrow{QP} = -\overrightarrow{PQ}.$$

Del mismo modo se prueba que $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$.

Una *variedad lineal*² de dimensión m en un espacio afín es un conjunto de puntos de la forma

$$L = P + \vec{L} = \{P + \vec{w} \mid \vec{w} \in \vec{L}\},$$

donde \vec{L} es un subespacio vectorial de \vec{E} de dimensión m , llamado *espacio director* de la variedad.

Notemos que $\vec{L} = \{\vec{v} \in \vec{E} \mid P + \vec{v} \in L\}$, luego \vec{L} está determinado por L . Recíprocamente, una variedad lineal L está determinada por su espacio director \vec{L} y uno cualquiera de sus puntos, pues si $Q \in P + \vec{L}$, entonces $\overrightarrow{PQ} \in \vec{L}$, luego $Q + \vec{L} = P + \overrightarrow{PQ} + \vec{L} = P + \vec{L} = L$.

En la sección anterior hemos probado que las rectas son las variedades lineales de dimensión 1 del espacio de la geometría euclídea, mientras que los planos son las variedades de dimensión 2. Por ello las variedades lineales de dimensión 1 y 2 de cualquier espacio afín reciben el nombre de *rectas* y *planos*. Las variedades lineales de dimensión 0 son los puntos (con rigor, los conjuntos con un solo punto). En un espacio afín de dimensión n , las variedades de dimensión $n - 1$ se llaman *hiperplanos*.

Toda variedad lineal $L = P + \vec{L}$ de dimensión m se puede considerar en sí misma como un espacio afín de dimensión m con \vec{L} como espacio vectorial asociado.

²El nombre hace referencia a que los elementos de una variedad están determinados por varias coordenadas.

Dado un conjunto de puntos $A \neq \emptyset$, existe una mínima variedad lineal que lo contiene, la llamaremos la variedad lineal generada por A y la representaremos por $\langle A \rangle$. En efecto, si $P \in A$ es fácil ver que

$$\langle A \rangle = P + \langle \overrightarrow{PQ} \mid Q \in A \rangle.$$

es una variedad que contiene a A y está contenida en cualquier otra que cumpla lo mismo.

Diremos que $n + 1$ puntos A_0, \dots, A_n son *afínmente independientes* si la variedad lineal que generan tiene dimensión n (en caso contrario se dice que son afínmente dependientes). Según lo que acabamos de ver, esto equivale a que los vectores $\overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_n}$ sean linealmente independientes.

En particular tres puntos P, Q, R son colineales si y sólo si son afínmente dependientes, si y sólo si los vectores \overrightarrow{PQ} y \overrightarrow{PR} son linealmente dependientes. Análogamente, cuatro puntos P, Q, R, S son coplanares si y sólo si $\overrightarrow{PQ}, \overrightarrow{PR}$ y \overrightarrow{PS} son linealmente dependientes, etc.

Es claro que en un espacio afín de dimensión n hay conjuntos afínmente independientes con $n + 1$ puntos, pero no con más.

Diremos que dos variedades lineales son *paralelas* si no tienen puntos en común y el espacio director de una está contenido en el de la otra.

El teorema siguiente prueba que esta noción general de paralelismo coincide con la que conocemos para rectas y planos en los espacios tridimensionales.

Teorema 4.10 *Dos variedades lineales de dimensiones k y l con $k \leq l$ son paralelas si y sólo si no tienen puntos en común y están contenidas en una variedad de dimensión $l + 1$.*

DEMOSTRACIÓN: Sean $P + \langle \vec{v}_1, \dots, \vec{v}_k \rangle$ y $Q + \langle \vec{w}_1, \dots, \vec{w}_l \rangle$ las dos variedades. Supongamos que no se cortan y están contenidas en una variedad de dimensión $l + 1$. Veamos en primer lugar que \overrightarrow{PQ} no es combinación lineal de $\vec{v}_1, \dots, \vec{v}_k, \vec{w}_1, \dots, \vec{w}_l$. Si lo fuera tendríamos

$$\alpha_1 \vec{v}_1 + \dots + \alpha_k \vec{v}_k = \overrightarrow{PQ} + \beta_1 \vec{w}_1 + \dots + \beta_l \vec{w}_l,$$

con lo que el punto

$$P + \alpha_1 \vec{v}_1 + \dots + \alpha_k \vec{v}_k = Q + \beta_1 \vec{w}_1 + \dots + \beta_l \vec{w}_l$$

estaría en las dos variedades.

Una variedad de dimensión $l + 1$ que contenga a ambas variedades contiene a los puntos $P, Q, P + \vec{v}_i, Q + \vec{w}_i$, luego su espacio director contiene a los vectores $\overrightarrow{PQ}, \vec{v}_i, \vec{w}_i$. Como la dimensión es $l + 1$ dicho espacio ha de ser $\langle \overrightarrow{PQ}, \vec{w}_1, \dots, \vec{w}_l \rangle$, luego

$$\langle \vec{v}_1, \dots, \vec{v}_k \rangle \subset \langle \overrightarrow{PQ}, \vec{w}_1, \dots, \vec{w}_l \rangle,$$

y como \overrightarrow{PQ} no es combinación lineal de ninguno de los demás vectores,

$$\langle \vec{v}_1, \dots, \vec{v}_k \rangle \subset \langle \vec{w}_1, \dots, \vec{w}_l \rangle,$$

lo que prueba que las variedades son paralelas.

Recíprocamente, si se da la inclusión anterior entonces las dos variedades están contenidas en la variedad $P + \langle \overrightarrow{PQ}, \vec{w}_1, \dots, \vec{w}_l \rangle$. ■

Así dos rectas son paralelas si y sólo si no se cortan y están contenidas en un plano, dos hiperplanos (en particular dos planos en un espacio tridimensional) son paralelos si y sólo si no se cortan.

Ejercicio: Probar que si dos variedades lineales tienen intersección no vacía entonces ésta es una variedad lineal y que su espacio director es la intersección de los espacios directores.

Es fácil comprobar que cualquier espacio afín de dimensión 3 cumple el grupo de axiomas A del capítulo I, así como el axioma E . De hecho, si la dimensión es 2 se siguen cumpliendo los axiomas que no hablan de planos, y si la dimensión es mayor que 3 se cumplen todos los axiomas del grupo excepto $A6$.

Por ejemplo, si dos planos en un espacio tridimensional tienen un punto en común P , la intersección es una variedad lineal cuyo espacio director es la intersección de dos subespacios de dimensión 2, luego la relación

$$\dim(V + W) = \dim V + \dim W - \dim(V \cap W)$$

implica que la dimensión de la intersección ha de ser al menos 1, luego la intersección es un plano o una recta.

Si el cuerpo K del espacio afín está ordenado (como es el caso de \mathbb{R}) podemos definir el orden $<_{AB}$ en la recta AB como el dado por³

$$X <_{AB} Y \quad \text{si y sólo si} \quad X = A + \lambda \overrightarrow{AB}, \quad Y = A + \mu \overrightarrow{AB} \quad \text{y} \quad \lambda < \mu.$$

Con este orden se satisface también el grupo de axiomas B . Vamos a demostrar como ejemplo el axioma $B5$, que es el más técnico. Notemos en general que los puntos situados entre dos puntos A y B son los de la forma

$$A + \lambda \overrightarrow{AB}, \quad \lambda \in [0, 1].$$

Sean A , B y C tres puntos no colineales y una recta r en su mismo plano que no pasa por ninguno de ellos, pero sí por un punto X entre A y B . Sea \vec{v} un vector director de r , es decir, $r = X + \langle \vec{v} \rangle$.

Los vectores \vec{v} y \overrightarrow{XB} han de ser linealmente independientes, pues en caso contrario B estaría en r . El punto A está en la recta XB , luego $A = X + \alpha \overrightarrow{XB}$,

³Notemos que en el caso del espacio euclídeo hemos probado que $X_\lambda = A + \lambda \overrightarrow{AB}$ es una graduación de AB , luego esta definición da el orden usual.

para cierto escalar α . Como ha de ser $A <_{XB} X$ y $X = X + 0\overrightarrow{XB}$, la definición de orden implica que $\alpha < 0$.

Es claro que el plano $X + \langle \overrightarrow{XB}, \vec{v} \rangle$ es ABC . Por lo tanto

$$C = X + \beta \overrightarrow{XB} + \gamma \vec{v},$$

para ciertos escalares β, γ . No puede ser $\beta = 0$ o de lo contrario C estaría en r . De hecho, el argumento que sigue va a demostrar que el signo de β será positivo o negativo según que C esté en el mismo semiplano que B o que A respecto a r .

Supongamos por ejemplo que $\beta > 0$ y vamos a ver que r corta a \overrightarrow{AC} . El punto de corte entre r y AC (si existe) ha de cumplir

$$X + \lambda \vec{v} = A + \mu \overrightarrow{AC}$$

para ciertos escalares λ y μ , luego

$$\begin{aligned} X + \lambda \vec{v} &= X + \overrightarrow{XA} + \mu(\overrightarrow{XC} - \overrightarrow{XA}) = X + \alpha \overrightarrow{XB} + \mu(\beta \overrightarrow{XB} + \gamma \vec{v} - \alpha \overrightarrow{XB}) \\ &= X + (\alpha + \mu(\beta - \alpha)) \overrightarrow{XB} + \mu\beta \vec{v}. \end{aligned}$$

Esto sucede exactamente cuando

$$\mu = \frac{-\alpha}{\beta - \alpha},$$

que, teniendo en cuenta los signos, es un número entre 0 y 1, luego el punto de corte $A + \mu \overrightarrow{AC}$ está ciertamente entre A y C . Similarmente se prueba que r no corta al segmento \overrightarrow{BC} . ■

Como ya hemos comentado, en la prueba se ha visto que si r es una recta que contiene a un punto X y B es un punto exterior a r , entonces un punto C está en el semiplano⁴ de frontera r que contiene a B si y sólo si es de la forma

$$C = X + \beta \overrightarrow{XB} + \gamma \vec{v}, \quad \text{con } \beta \geq 0.$$

Equivalentemente, dada una recta $r = P + \langle \vec{v} \rangle$, todo plano que la contiene es de la forma $\pi = P + \langle \vec{v}, \vec{w} \rangle$, para un cierto vector \vec{w} independiente de \vec{v} , y entonces los semiplanos que r determina en π son

$$\{P + \lambda \vec{v} + \mu \vec{w} \mid \mu \geq 0\} \quad \text{y} \quad \{P + \lambda \vec{v} + \mu \vec{w} \mid \mu \leq 0\}.$$

Dada una recta $r = O + \langle \vec{v} \rangle$, es fácil ver que las semirrectas que O determina en r son los conjuntos

$$\{O + \lambda \vec{v} \mid \lambda \geq 0\} \quad \text{y} \quad \{O + \lambda \vec{v} \mid \lambda \leq 0\}.$$

⁴Desde el momento en que hemos probado que todo espacio afín (sobre un cuerpo ordenado) satisface los grupos de axiomas A y B , todos los conceptos definidos a partir de ellos tienen sentido en cualquier espacio afín en estas condiciones (segmentos, semirrectas, etc.).

Por lo tanto, el ángulo de vértice O y lados las semirrectas

$$\{O + \lambda \vec{v} \mid \lambda \geq 0\} \quad \text{y} \quad \{O + \lambda \vec{w} \mid \lambda \geq 0\}$$

es el conjunto

$$A(O; \vec{v}, \vec{w}) = \{O + \lambda \vec{v} + \mu \vec{w} \mid \lambda \geq 0 \text{ y } \mu \geq 0\}.$$

Es fácil comprobar que todo espacio afín sobre \mathbb{R} cumple también el axioma D , luego en total tenemos que la geometría afín satisface todos los grupos de axiomas a excepción del grupo C . No tiene sentido plantearse si los espacios afines cumplen estos axiomas porque en ellos no tenemos definida la noción de congruencia.

Terminamos la sección estudiando brevemente las aplicaciones que conservan la estructura afín.

Definición 4.11 Sean E y F dos espacios afines. Una *aplicación afín* o *afinidad* entre E y F es una aplicación $f : E \longrightarrow F$ tal que para todo punto P de E se cumple

$$f(P) = f(O) + \vec{f}(\overrightarrow{OP}),$$

donde O es un punto prefijado en E , y $\vec{f} : \vec{E} \longrightarrow \vec{F}$ es una aplicación lineal, llamada *aplicación lineal asociada* a f .

En primer lugar hemos de notar que si f admite una expresión como la indicada para un punto O y una cierta aplicación lineal \vec{f} , entonces f admite una expresión análoga para cualquier otro punto O' y con la misma aplicación \vec{f} . En efecto, puesto que $f(O') = f(O) + \vec{f}(\overrightarrow{OO'})$, resulta que

$$f(P) = f(O) + \vec{f}(\overrightarrow{OP}) = f(O') + \vec{f}(\overrightarrow{OP}) - \vec{f}(\overrightarrow{OO'}) = f(O') + \vec{f}(\overrightarrow{O'P}).$$

En particular una afinidad determina su aplicación lineal asociada.

Notemos que la relación entre f y \vec{f} se puede escribir también en la forma $\vec{f}(\overrightarrow{PQ}) = \overrightarrow{f(P)f(Q)}$, para todo par de puntos P y Q . Teniendo esto en cuenta es fácil comprobar que la composición de aplicaciones afines es una aplicación afín, así como que la inversa de una biyección afín es una biyección afín. Además $\overrightarrow{f \circ g} = \vec{f} \circ \vec{g}$ y $\overrightarrow{f^{-1}} = (\vec{f})^{-1}$.

Se prueba sin dificultad que las biyecciones afines conservan todas las propiedades definibles a partir de la estructura de espacio afín: aplican variedades en variedades, paralelas en paralelas, semiplanos en semiplanos, ángulos en ángulos, triángulos en triángulos, conservan las relaciones de orden de las rectas, etc.

4.3 Coordenadas cartesianas y baricéntricas

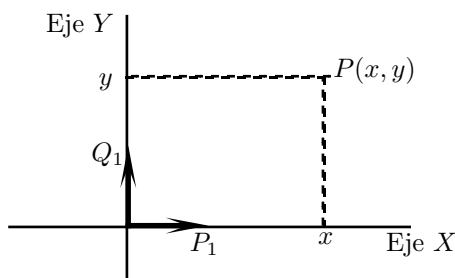
Tras todas estas consideraciones podemos exponer el núcleo de la geometría analítica, en virtud de la cual los conceptos geométricos se caracterizan en términos de ecuaciones y desigualdades. Para ello necesitamos los conceptos de sistema de referencia y coordenadas de un punto.

Definición 4.12 Un *sistema de referencia* en un espacio afín E está formado por un punto O y una base $(\vec{e}_1, \dots, \vec{e}_n)$ del espacio \vec{E} .

Fijado un sistema de referencia, podemos identificar cada vector \vec{v} con sus coordenadas en la base del sistema. Así, $\vec{v} = (x_1, \dots, x_n)$ se interpreta como que $\vec{v} = x_1\vec{e}_1 + \dots + x_n\vec{e}_n$. Llamaremos *vector de posición* de un punto P (siempre respecto al sistema de referencia fijado) al vector \vec{OP} . Las *coordenadas cartesianas* (lat. ‘de Descartes’) de un punto P serán las coordenadas de su vector de posición. Escribiremos $P(x_1, \dots, x_n)$ para indicar que las coordenadas de P en un sistema de referencia dado son (x_1, \dots, x_n) . Según lo dicho, esto equivale a que

$$P = O + x_1\vec{e}_1 + \dots + x_n\vec{e}_n.$$

Por ejemplo, si tomamos dos rectas que se corten perpendicularmente en un punto O , las graduamos con la misma unidad y consideramos los vectores $\vec{v} = \vec{OP_1}$, $\vec{w} = \vec{OQ_1}$ (donde P_1 y Q_1 son los respectivos puntos unitarios de las rectas) los resultados de la sección anterior muestran que tenemos un sistema de referencia cartesiano, y las coordenadas (x, y) de un punto P se interpretan como los números asociados por las graduaciones a las proyecciones de P por rectas paralelas a las rectas OP_1 y OQ_1 , es decir, sus distancias al origen más un signo que indica la semirrecta en la que se encuentran. La recta $O + \lambda\vec{v}$, cuyos puntos tienen coordenadas $(x, 0)$, se llama simplemente ‘Eje X ’ o eje de abscisas, mientras que la recta $O + \lambda\vec{w}$ recibe el nombre de ‘Eje Y ’, o eje de ordenadas.



Igualmente podemos interpretar las coordenadas (x, y, z) de un punto del espacio respecto a un sistema de referencia determinado por tres rectas perpendiculares graduadas con la misma unidad (ejes X , Y y Z). Conviene tener presente que la definición de sistema de referencia no exige que los ejes se tomen perpendiculares. De hecho la noción de perpendicularidad no está definida en un espacio afín arbitrario.

Consideremos un hiperplano $H = P + \langle \vec{v}_1, \dots, \vec{v}_{n-1} \rangle$ en un espacio afín en el que hemos fijado un sistema de referencia de origen O . Un punto $Q(x_1, \dots, x_n)$ está en H si y sólo si $\overrightarrow{PQ} \in \langle \vec{v}_1, \dots, \vec{v}_{n-1} \rangle$. Si (p_1, \dots, p_n) son las coordenadas de P , entonces las coordenadas de $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}$ son $(x_1 - p_1, \dots, x_n - p_n)$, luego la condición anterior equivale a que el determinante formado por este vector y las coordenadas de los \vec{v}_i sea igual a 0. Esto se traduce en una ecuación de la forma

$$a_1x_1 + \dots + a_nx_n = b,$$

donde necesariamente alguno de los coeficientes es no nulo.

Recíprocamente, el conjunto de puntos cuyas coordenadas satisfacen una ecuación de esta forma es un hiperplano. En efecto, el miembro izquierdo de la ecuación define una aplicación lineal no nula de K^n en K , que necesariamente será suprayectiva y su núcleo tendrá dimensión $n - 1$. La aplicación que a cada vector le asigna sus coordenadas en una base dada es un isomorfismo de V en K^n , luego los vectores cuyas coordenadas anulan el miembro izquierdo de la ecuación forman un subespacio vectorial de base $\vec{v}_1, \dots, \vec{v}_{n-1}$. Así mismo ha de existir un vector \vec{v} cuyas coordenadas satisfagan la ecuación. Si llamamos $P = O + \vec{v}$, entonces los puntos del hiperplano $P + \langle \vec{v}_1, \dots, \vec{v}_{n-1} \rangle$ son exactamente los puntos cuyas coordenadas satisfacen la ecuación dada.

En resumen:

Los puntos de un hiperplano en un espacio afín están caracterizados por que sus coordenadas satisfacen una determinada ecuación lineal no nula. Toda ecuación lineal no nula es la ecuación de un hiperplano.

Dos hiperplanos son paralelos si y sólo si no tienen puntos comunes, lo que fácilmente se traduce en que sus ecuaciones tienen los miembros izquierdos proporcionales pero el término derecho no respeta la proporción.

De este modo, una recta en el plano está formada por los puntos cuyas coordenadas en un sistema de referencia dado satisfacen una ecuación de la forma $ax + by = c$. Una recta en el espacio se puede expresar como intersección de dos planos, luego sus elementos son los puntos cuyas coordenadas satisfacen un sistema de dos ecuaciones lineales independientes y compatibles

$$\left. \begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \end{aligned} \right\}$$

En general, los puntos de una variedad de dimensión m en un espacio afín de dimensión k están caracterizados por que sus coordenadas satisfacen un sistema de $n - m$ ecuaciones lineales independientes.

Veamos ahora el modo en que una afinidad transforma las coordenadas afines de los puntos. Sea $f : E \rightarrow F$ una afinidad, sea $(O; \vec{v}_1, \dots, \vec{v}_n)$ un sistema de referencia en E y $(O'; \vec{w}_1, \dots, \vec{w}_m)$ un sistema de referencia en F . Sean $A = (a_1, \dots, a_m)$ las coordenadas de $f(O)$ en O' , sea M la matriz asociada a la aplicación lineal \vec{f} en las bases de los sistemas de referencia. Sea P un punto en E con coordenadas $X = (x_1, \dots, x_n)$. Éstas son por definición las coordenadas

de \overrightarrow{OP} en la base \vec{v}_i , luego las coordenadas de $\vec{f}(\overrightarrow{OP})$ en la base \vec{w}_i son XM . Puesto que

$$f(P) = f(O) + \vec{f}(\overrightarrow{OP}) = O' + \overrightarrow{O'f(O)} + \vec{f}(\overrightarrow{OP}),$$

vemos que las coordenadas de $\overrightarrow{O'f(P)}$ en la base \vec{w}_i son

$$Y = A + XM.$$

En resumen:

La relación entre las coordenadas X de un punto en un sistema de referencia O y las coordenadas Y de su imagen por una afinidad f es $Y = A + XM$, donde A son las coordenadas de $f(O)$ en el sistema O' y M es la matriz de \vec{f} .

Recíprocamente, es fácil ver que toda ecuación matricial $Y = A + XM$ está asociada a una única afinidad en unos sistemas de referencia prefijados.

El sumando A puede suprimirse si en el segundo espacio tomamos como origen de coordenadas el punto $O' = f(O)$. Sin embargo, cuando f es una afinidad de un espacio en sí mismo, resulta más conveniente considerar un único sistema de referencia, y entonces sólo podemos eliminar el sumando A si f tiene un punto fijo, es decir, un punto O tal que $f(O) = O$.

Si K es un cuerpo, podemos dotar al conjunto K^n de estructura de espacio afín n -dimensional sin más que tomar sus elementos como puntos y como vectores a un tiempo, de modo que la suma de puntos y vectores sea la misma suma vectorial de K^n . Las variedades lineales de K^n son de la forma $P + V$, donde V es un subespacio de K^n , es decir, son las clases de congruencia módulo los subespacios de K^n . El *sistema de referencia canónico* en K^n es el formado por $(0, \dots, 0)$ como origen y la base canónica de K^n . Respecto a este sistema, cada punto coincide con sus propias coordenadas.

Si E es cualquier espacio afín n -dimensional sobre K , la aplicación que a cada punto le asigna sus coordenadas respecto a un sistema de referencia prefijado es una biyección afín (cuya expresión coordenada respecto a este sistema y al sistema canónico de K^n es simplemente $Y = X$).

Así pues, todo espacio afín n -dimensional sobre K es isomorfo a K^n , luego dichos espacios tienen todos las mismas propiedades afines (las mismas propiedades que sean invariantes por biyecciones afines).

En particular podemos identificar el espacio de la geometría tridimensional euclídea que venimos estudiando con el conjunto \mathbb{R}^3 .

Si $(O; \vec{e}_1, \dots, \vec{e}_n)$ es un sistema de referencia de un espacio afín E , entonces los puntos $A_0 = O$, $A_1 = O + \vec{e}_1, \dots, A_n = O + \vec{e}_n$ son afinmente independientes y, recíprocamente, un conjunto de $n + 1$ puntos afinmente independientes A_0, \dots, A_n determina el sistema de referencia $(A_0; \overrightarrow{A_0A_1}, \dots, \overrightarrow{A_0A_n})$. Por lo tanto todo punto P de E se expresa de forma única como

$$P = A_0 + x_1 \overrightarrow{A_0A_1} + \dots + \overrightarrow{A_0A_n}.$$

En algunas ocasiones esta expresión no es satisfactoria, pues sitúa al punto A_0 en una situación asimétrica respecto de los otros, cuando en realidad la noción de puntos afinmente independientes es completamente simétrica. Por ello, a veces conviene considerar las coordenadas baricéntricas que definimos a continuación.

Notemos en primer lugar, que para puntos cualesquiera A_1, \dots, A_n y escalares $\lambda_1 + \dots + \lambda_n = 0$, el vector

$$\sum_{i=1}^n \lambda_i \overrightarrow{OA_i}$$

es independiente de la elección del punto O , pues si tomamos otro punto O' vemos que

$$\sum_{i=1}^n \lambda_i \overrightarrow{O'A_i} - \sum_{i=1}^n \lambda_i \overrightarrow{OA_i} = \sum_{i=1}^n \lambda_i \overrightarrow{OO_i} = \vec{0}.$$

De aquí se desprende que si $\lambda_1 + \dots + \lambda_n = \lambda \neq 0$ el punto

$$B = O + \frac{1}{\lambda} \sum_{i=1}^n \lambda_i \overrightarrow{OA_i}$$

es independiente de O . Al punto B se le llama *baricentro* de (A_i, λ_i) y lo representaremos por

$$\frac{1}{\lambda} (\lambda_1 A_1 + \dots + \lambda_n A_n) = O + \frac{1}{\lambda} \sum_{i=1}^n \lambda_i \overrightarrow{OA_i}.$$

Las sumas y productos del miembro izquierdo son meramente formales, sin un significado intrínseco. Tan sólo sugieren que las coordenadas del baricentro en el sistema de referencia $(O; \overrightarrow{OA_i})$ se obtienen multiplicando por λ_i las coordenadas de A_i , sumando y dividiendo entre λ . El baricentro recibe este nombre porque se corresponde con el centro de gravedad de un sistema de n partículas puntuales situadas en los puntos A_i con masas λ_i .

Retomando ahora el conjunto afinmente independiente A_0, \dots, A_n , ahora notamos que un punto P se puede expresar en la forma

$$P = A_0 + \sum_{i=1}^n \lambda_i \overrightarrow{A_0 A_i} = \sum_{i=0}^n \lambda_i A_i,$$

donde

$$\lambda_0 = 1 - \sum_{i=1}^n \lambda_i.$$

Los números $(\lambda_0, \dots, \lambda_n)$ (sujetos a la condición $\lambda_0 + \dots + \lambda_n = 1$) están unívocamente determinados por P y se llaman *coordenadas baricéntricas* de P respecto al sistema (A_i) .

Teorema 4.13 *Los puntos A_0, \dots, A_n son afínmente independientes si y sólo si ninguno de ellos es un baricentro de los demás.*

DEMOSTRACIÓN: Si uno de los puntos, por ejemplo A_0 es un baricentro de los demás, entonces

$$A_0 = \sum_{i=1}^n \lambda_i A_i, \quad \text{con} \quad \sum_{i=1}^n \lambda_i = 1.$$

Alguno de los coeficientes ha de ser no nulo, digamos que $\lambda_n \neq 0$. Así

$$A_0 = A_n + \sum_{i=1}^{n-1} \lambda_i \overrightarrow{A_n A_i},$$

donde no todos los coeficientes son nulos, luego $\overrightarrow{A_n A_0}$ es combinación lineal de los restantes. Invirtiendo el razonamiento tenemos la otra implicación. ■

4.4 Espacios euclídeos

Nos ocupamos ahora de extender la estructura de espacio afín para recoger el concepto de congruencia, el último concepto geométrico primitivo que nos queda por representar analíticamente. Es importante notar que todo lo que desde el punto de vista de la teoría de conjuntos hemos tomado como definiciones, desde el punto de vista de la geometría euclídea son teoremas. Por ejemplo, hemos definido una recta en \mathbb{R}^3 como una variedad lineal de dimensión 1, pero hemos demostrado que las rectas en la geometría euclídea son las variedades de dimensión 1. Esto se interpreta como que las definiciones dadas son las únicas posibles para que se satisfagan los axiomas de la geometría, de modo que si tomamos un espacio euclídeo cualquiera y a cada punto le asignamos una terna de coordenadas respecto a un sistema de referencia arbitrario, entonces esta correspondencia transformará necesariamente las rectas y planos del espacio en las variedades lineales de \mathbb{R}^3 , las relaciones de orden entre los puntos de una recta se corresponderán necesariamente con las que hemos definido en \mathbb{R}^3 , etc.

Si probamos que no hay más que una definición analítica de congruencia que sea consistente con los axiomas de la geometría, habremos demostrado que sólo existe un espacio tridimensional euclídeo, en el sentido de que cualquiera de ellos es identificable con \mathbb{R}^3 a través de un sistema de coordenadas. Por ello vamos a desarrollar un poco más nuestro concepto sintético de congruencia y los relacionados con él para comprobar que en la definición de congruencia no tenemos ningún grado de libertad.

En primer lugar notamos que si \vec{v} y \vec{w} son dos vectores no nulos cualesquiera, O es un punto, $A = O + \vec{v}$ y $B = O + \vec{w}$, entonces el ángulo \widehat{AOB} (si está definido) es independiente de O , pues el triángulo \widehat{AOB} tiene lados $\|\vec{v}\|$, $\|\vec{w}\|$ y $\|\vec{v} - \vec{w}\|$, luego los triángulos contruidos a partir de puntos O distintos son congruentes, y en particular el ángulo indicado es el mismo.

El ángulo \widehat{AOB} no está definido cuando los vectores \vec{v} y \vec{w} son linealmente dependientes (lo cual tampoco depende de O). Todo esto nos permite definir el *ángulo* entre dos vectores no nulos como el ángulo \widehat{AOB} si los puntos no están alineados, si $\vec{w} = \alpha\vec{v}$ con $\alpha < 0$ consideraremos que el ángulo es llano y si $\alpha > 0$ diremos que el ángulo que forman es nulo y tiene medida 0. Cuando dos vectores forman un ángulo recto se dice que son *ortogonales* (gr. ‘en ángulo recto’).

Consideremos un sistema de referencia de origen O y con una base $\vec{e}_1, \vec{e}_2, \vec{e}_3$ formada por vectores ortogonales y de norma 1. Entonces, un punto arbitrario P cumplirá que

$$P = O + x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3,$$

para ciertos números reales (x, y, z) (sus coordenadas en el sistema de referencia indicado). Si $Q = O + x\vec{e}_1 + y\vec{e}_2$, vemos que la recta OQ está contenida en el plano XY , mientras que la recta QP es paralela al eje Z , que es perpendicular al plano XY (suponiendo que $Q \neq O$ y $Q \neq P$). Por lo tanto el triángulo \widehat{OQP} es rectángulo, y el teorema de Pitágoras nos da que

$$\|\vec{OP}\|^2 = \|\vec{OQ}\|^2 + \|\vec{QP}\|^2 = \|x\vec{e}_1 + y\vec{e}_2\|^2 + \|z\vec{e}_3\|^2.$$

Si $Q = O$ o $Q = P$ la conclusión es trivialmente cierta. Del mismo modo, si llamamos $R = O + x\vec{e}_1$ tenemos que el triángulo \widehat{OQR} es rectángulo (o bien $Q = O$ o $Q = R$), con lo que

$$\|\vec{OP}\|^2 = \|x\vec{e}_1\|^2 + \|y\vec{e}_2\|^2 + \|z\vec{e}_3\|^2 = x^2 + y^2 + z^2,$$

luego $\|\vec{OP}\| = \sqrt{x^2 + y^2 + z^2}$. Puesto que P es un punto arbitrario, tenemos que si un vector \vec{v} tiene coordenadas (x, y, z) respecto a una base formada por vectores ortogonales y unitarios, entonces

$$\|\vec{v}\| = \sqrt{x^2 + y^2 + z^2}.$$

Si \vec{v} y \vec{w} son dos vectores arbitrarios cuyas coordenadas en una base en las condiciones anteriores son respectivamente (x, y, z) y (x', y', z') , definimos su *producto escalar* como

$$\vec{v}\vec{w} = xx' + yy' + zz'.$$

En estos términos hemos probado que $\|\vec{v}\| = \sqrt{\vec{v}\vec{v}}$. Es claro además que se cumplen las propiedades siguientes:

1. $\vec{u}(\vec{v} + \vec{w}) = \vec{u}\vec{v} + \vec{u}\vec{w}$,
2. $(\vec{u} + \vec{v})\vec{w} = \vec{u}\vec{w} + \vec{v}\vec{w}$,
3. $(\alpha\vec{u})\vec{v} = \alpha(\vec{u}\vec{v}) = \vec{u}(\alpha\vec{v})$,
4. $\vec{u}\vec{v} = \vec{v}\vec{u}$.

Teniendo esto en cuenta vamos a probar unos hechos muy importantes:

Consideremos dos vectores no nulos \vec{v} y \vec{w} . Supongamos en primer lugar que son linealmente independientes. Sean $A = O + \vec{v}$, $B = O + \vec{w}$. Entonces los puntos O , A , B no son colineales, luego forman un triángulo cuyos lados miden $\|\vec{OA}\| = \|\vec{v}\|$, $\|\vec{OB}\| = \|\vec{w}\|$, $\|\vec{AB}\| = \|\vec{w} - \vec{v}\|$. Por otra parte

$$\|\vec{AB}\|^2 = (\vec{w} - \vec{v})(\vec{w} - \vec{v}) = \vec{w}\vec{w} + \vec{v}\vec{v} - 2\vec{v}\vec{w} = \|\vec{OA}\|^2 + \|\vec{OB}\|^2 - 2\vec{v}\vec{w}.$$

Si comparamos con el teorema del coseno concluimos que

$$\vec{v}\vec{w} = \|\vec{v}\| \|\vec{w}\| \cos \widehat{\vec{v}\vec{w}}.$$

Esta expresión sigue siendo válida cuando $\vec{w} = \alpha\vec{v}$ si convenimos en que el coseno de un ángulo nulo es 1 y el coseno de un ángulo llano es -1 . En efecto:

$$\vec{v}\vec{w} = \alpha\vec{v}\vec{v} = \pm|\alpha| \|\vec{v}\|^2 = \pm\|\vec{v}\| \|\vec{w}\|.$$

En particular esto prueba que el producto escalar de dos vectores no depende del sistema de referencia que elegimos para calcularlo (siempre y cuando los vectores de su base sean ortogonales y de norma 1).

Estas consideraciones condicionan ya la estructura algebraica que hemos de imponer a un espacio afín para definir en él un concepto de ortogonalidad consistente con los axiomas de la geometría euclídea.

Definición 4.14 Un *espacio vectorial euclídeo* es un espacio vectorial real V de dimensión finita sobre el que hay definido un *producto escalar*, que es una aplicación $V \times V \longrightarrow \mathbb{R}$ que cumple las propiedades siguientes:

1. $\vec{v}\vec{v} \geq 0$ y $\vec{v}\vec{v} = 0$ si y sólo si $\vec{v} = 0$,
2. $\vec{u}(\vec{v} + \vec{w}) = \vec{u}\vec{v} + \vec{u}\vec{w}$,
3. $(\vec{u} + \vec{v})\vec{w} = \vec{u}\vec{w} + \vec{v}\vec{w}$,
4. $(\alpha\vec{u})\vec{v} = \alpha(\vec{u}\vec{v}) = \vec{u}(\alpha\vec{v})$,
5. $\vec{u}\vec{v} = \vec{v}\vec{u}$.

Un *espacio afín euclídeo* es un espacio afín real E tal que en \vec{E} hay definido un producto escalar que lo dota de estructura de espacio euclídeo.

Definimos la *norma* de un vector en un espacio euclídeo como $\|\vec{v}\| = \sqrt{\vec{v}\vec{v}}$. La *distancia* entre dos puntos A y B de un espacio afín euclídeo será $\|\vec{AB}\|$. Diremos que dos vectores \vec{v} y \vec{w} son *ortogonales* si $\vec{v}\vec{w} = 0$. Lo representaremos por $\vec{v} \perp \vec{w}$. Un conjunto de vectores es *ortogonal* si no contiene al 0 y sus elementos son ortogonales dos a dos. Un conjunto de vectores es *ortonormal* si es ortogonal y todos sus vectores tienen norma 1.

Hemos probado que el espacio euclídeo en el sentido de la geometría sintética es un espacio afín euclídeo tridimensional en el sentido analítico. Probaremos

que el recíproco también es cierto (lo que justifica el nombre de espacio euclídeo en el segundo caso). En primer lugar sucede que esta definición algebraica no muestra directamente la existencia de bases ortonormales. Probar esto será nuestro primer objetivo. Para ello conviene notar algunos hechos elementales.

Se cumple que $\vec{0}\vec{v} = (\vec{0}\vec{v})\vec{v} = 0(\vec{0}\vec{v}) = 0$.

La norma verifica la relación $\|\alpha\vec{v}\| = |\alpha| \|\vec{v}\|$. Por lo tanto, si dividimos un vector no nulo por su norma obtenemos un vector de norma 1, y si los vectores de un conjunto ortogonal los dividimos por sus normas respectivas obtenemos un conjunto ortonormal. Así pues, basta probar la existencia de bases ortogonales.

Observemos también que si un conjunto de vectores $\{\vec{v}_1, \dots, \vec{v}_n\}$ es ortogonal, entonces es linealmente independiente. En efecto, si $\vec{\alpha}_1\vec{v}_1 + \dots + \vec{\alpha}_m\vec{v}_m = \vec{0}$ entonces $0 = \vec{v}_j\vec{0} = \alpha_1\vec{v}_j\vec{v}_1 + \dots + \alpha_m\vec{v}_j\vec{v}_m = \alpha_j\vec{v}_j\vec{v}_j$, luego $\alpha_j = 0$.

Teorema 4.15 (Teorema de ortogonalización de Gram-Schmidt) *Si V es un espacio vectorial euclídeo, todo conjunto ortogonal de vectores de V se extiende hasta una base ortogonal de V .*

DEMOSTRACIÓN: Sea $A = \{\vec{v}_1, \dots, \vec{v}_r\}$ un conjunto de vectores ortogonales (y por lo tanto linealmente independientes). Si A es ya una base de V no hay nada que probar. En caso contrario tomamos un vector \vec{v} de V que no esté en $\langle A \rangle$. Consideremos un vector de la forma

$$\vec{w} = \vec{v} - \sum_{i=1}^r \alpha_i \vec{v}_i.$$

Claramente $\vec{v}_i\vec{w} = \vec{v}_i\vec{v} - \alpha_i\vec{v}_i\vec{v}_i$, luego si tomamos cada α_i como el único escalar que cumple esta ecuación tenemos que $\vec{w} \perp A$. Así obtenemos un conjunto ortogonal con un vector más. Repitiendo el proceso llegamos hasta un conjunto ortogonal con tantos vectores como la dimensión de V , que será una base ortogonal. ■

En particular todo espacio euclídeo V tiene una base ortonormal $\{\vec{e}_1, \dots, \vec{e}_n\}$. Si \vec{v} y \vec{w} son vectores cualesquiera y $(x_1, \dots, x_n), (y_1, \dots, y_n)$ son sus coordenadas en dicha base, tenemos que

$$\vec{v}\vec{w} = (x_1\vec{e}_1 + \dots + x_n\vec{e}_n)(y_1\vec{e}_1 + \dots + y_n\vec{e}_n) = x_1y_1 + \dots + x_ny_n.$$

Esto significa que un producto escalar está completamente determinado en cuanto conocemos una base ortonormal. En particular podemos dotar a \mathbb{R}^n de estructura de espacio euclídeo definiendo el producto escalar como

$$(x_1, \dots, x_n)(y_1, \dots, y_n) = x_1y_1 + \dots + x_ny_n.$$

Es claro que se trata ciertamente de un producto escalar y, respecto a él, la base canónica es ortonormal. Los resultados del comienzo de la sección prueban que si en un espacio euclídeo (en el sentido sintético) fijamos un sistema de referencia con base ortonormal y a cada punto le asignamos sus coordenadas respecto a

dicho sistema, entonces el producto escalar se corresponde con el producto en \mathbb{R}^3 que acabamos de definir, luego la ortogonalidad y la distancia entre puntos se corresponden con los conceptos homónimos en \mathbb{R}^3 definidos analíticamente.

El teorema siguiente es obvio desde un punto de vista sintético, pero es lo único necesario para probar analíticamente las propiedades de la distancia entre puntos y la ortogonalidad de vectores.

Teorema 4.16 (Desigualdad de Cauchy-Schwarz) *Si \vec{v} y \vec{w} son dos vectores de un espacio euclídeo, entonces $|\vec{v}\vec{w}| \leq \|\vec{v}\| \|\vec{w}\|$ y se da la igualdad si y sólo si \vec{v} y \vec{w} son linealmente dependientes.*

DEMOSTRACIÓN: Si $\vec{w} = \vec{0}$ la igualdad se da trivialmente. Supongamos que $\vec{w} \neq \vec{0}$. Llamando $\vec{u} = \vec{w}/\|\vec{w}\|$, entonces $\|\vec{u}\| = 1$ y lo que hemos de probar es que $|\vec{v}\vec{u}| \leq \|\vec{v}\|$, y que se da la igualdad si y sólo si \vec{u} y \vec{v} son linealmente dependientes.

Tomemos $\lambda \in \mathbb{R}$. Entonces

$$0 \leq \|\vec{v} + \lambda\vec{u}\|^2 = (\vec{v} + \lambda\vec{u})(\vec{v} + \lambda\vec{u}) = \vec{v}\vec{v} + \lambda^2\vec{u}\vec{u} + 2\lambda\vec{v}\vec{u} = \|\vec{v}\|^2 + \lambda^2 + 2\lambda\vec{v}\vec{u},$$

En particular, si tomamos $\lambda = -\vec{v}\vec{u}$ queda que $\|\vec{v}\|^2 + (\vec{v}\vec{u})^2 - 2(\vec{v}\vec{u})^2 \geq 0$, o sea, $(\vec{v}\vec{u})^2 \leq \|\vec{v}\|^2$, luego $|\vec{v}\vec{u}| \leq \|\vec{v}\|$.

Además la igualdad se da si y sólo si $\|\vec{v} - (\vec{v}\vec{u})\vec{u}\| = 0$, o sea, si y sólo si $\vec{v} = (\vec{v}\vec{u})\vec{u}$. Por lo tanto, si se da la igualdad \vec{u} y \vec{v} son linealmente dependientes. Recíprocamente, si $\vec{v} = \lambda\vec{u}$, entonces $|\vec{v}\vec{u}| = |\lambda\vec{u}\vec{u}| = |\lambda| = \|\vec{v}\|$. ■

Ahora podemos probar las propiedades esenciales de la norma euclídea:

Teorema 4.17 *En todo espacio euclídeo se cumple:*

1. $\|\vec{v}\| \geq 0$ y $\|\vec{v}\| = 0$ si y sólo si $\vec{v} = \vec{0}$.
2. $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$.
3. $\|\alpha\vec{v}\| = |\alpha| \|\vec{v}\|$.

DEMOSTRACIÓN: La única propiedad que no es inmediata es la segunda, pero

$$\begin{aligned} \|\vec{v} + \vec{w}\|^2 &= (\vec{v} + \vec{w})(\vec{v} + \vec{w}) = \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\vec{v}\vec{w} \leq \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2|\vec{v}\vec{w}| \\ &\leq \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\|\vec{v}\| \|\vec{w}\| = (\|\vec{v}\| + \|\vec{w}\|)^2. \end{aligned}$$

Por lo tanto $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$. ■

La propiedad 2 del teorema anterior es la versión analítica de la desigualdad triangular. Es fácil ver que es una igualdad exactamente cuando uno de los vectores es un múltiplo positivo del otro.

Ahora nos ocupamos de las aplicaciones que conservan la estructura euclídea. Vamos a dar una definición que aparentemente es mucho más débil que la que

cabría esperar, pues vamos a pedir que se conserven las distancias, pero no la estructura lineal ni mucho menos el producto escalar. Sin embargo probaremos seguidamente que la conservación de las distancias implica todo lo demás. Representaremos por $d(P, Q) = \|\vec{PQ}\|$ la distancia entre dos puntos P y Q de un espacio afín euclídeo.

Definición 4.18 Sean A y B dos subconjuntos de dos espacios afines euclídeos. Una *isometría* entre A y B es una aplicación biyectiva $f : A \longrightarrow B$ tal que para todo par de puntos $P, Q \in A$ se cumple $d(P, Q) = d(f(P), f(Q))$.

Tomemos un sistema de referencia del primer espacio con origen en un punto $O \in A$ y un sistema de referencia en el segundo espacio con origen en $O' = f(O)$.

Por simplificar la notación usaremos un apóstrofo para representar las imágenes por f , es decir, si $P \in A$, entonces P' representará a $f(P)$. Llamemos \tilde{A} al conjunto de los vectores \vec{OP} , con $P \in A$. Similarmente, sea \tilde{B} el conjunto de los vectores $\vec{O'P'}$ con $P' \in B$. Si $\vec{v} = \vec{OP} \in \tilde{A}$, llamaremos $\vec{v}' = \vec{O'P'} \in \tilde{B}$.

En estos términos $\vec{0} \in \tilde{A}$, $\vec{0}' = \vec{0} \in \tilde{B}$ y $\|\vec{v} - \vec{w}\| = \|\vec{v}' - \vec{w}'\|$, para todo par de vectores \vec{v} y \vec{w} en \tilde{A} . En particular, si hacemos $\vec{v} = \vec{0}$ tenemos $\|\vec{v}\| = \|\vec{v}'\|$.

La relación

$$\|\vec{v} - \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2 - 2\vec{v}\vec{w}$$

implica que $\vec{v}\vec{w} = \vec{v}'\vec{w}'$, para todo par de vectores $\vec{v}, \vec{w} \in \tilde{A}$.

Veamos que si $\vec{v} \in \tilde{A}$ y $\alpha\vec{v} \in \tilde{A}$ entonces $(\alpha\vec{v})' = \alpha\vec{v}'$. Si $\vec{v} = \vec{0}$ es trivial. En caso contrario

$$|\vec{v}'(\alpha\vec{v})'| = |\vec{v}(\alpha\vec{v})| = |\alpha| \|\vec{v}\|^2 = \|\vec{v}\| \|\alpha\vec{v}\| = \|\vec{v}'\| \|(\alpha\vec{v})'\|.$$

Como la desigualdad de Cauchy-Schwarz es en este caso una igualdad, tenemos que $(\alpha\vec{v})' = \lambda\vec{v}'$, para cierto escalar λ . Multiplicando ambos miembros por \vec{v}' obtenemos $\alpha\|\vec{v}\|^2 = \lambda\|\vec{v}\|^2$, luego $\lambda = \alpha$, y se cumple lo afirmado.

Ahora probamos que si $\vec{u}, \vec{v}, \vec{u} + \vec{v} \in \tilde{A}$, entonces $(\vec{u} + \vec{v})' = \vec{u}' + \vec{v}'$. Tenemos

$$\|\vec{u}' + \vec{v}'\|^2 = \|\vec{u}'\|^2 + \|\vec{v}'\|^2 + 2\vec{u}'\vec{v}' = \|\vec{u}\|^2 + \|\vec{v}\|^2 + 2\vec{u}\vec{v} = \|\vec{u} + \vec{v}\|^2,$$

luego $\|\vec{u}' + \vec{v}'\| = \|\vec{u} + \vec{v}\|$. Por otra parte

$$(\vec{u} + \vec{v})'(\vec{u}' + \vec{v}') = (\vec{u} + \vec{v})'\vec{u}' + (\vec{u} + \vec{v})'\vec{v}' = (\vec{u} + \vec{v})\vec{u} + (\vec{u} + \vec{v})\vec{v} = \|\vec{u} + \vec{v}\|^2$$

Reuniendo ambas igualdades $(\vec{u} + \vec{v})'(\vec{u}' + \vec{v}') = \|(\vec{u} + \vec{v})'\| \|\vec{u}' + \vec{v}'\|$. De nuevo la desigualdad de Cauchy-Schwarz es una igualdad, lo que implica que $\vec{u}' + \vec{v}' = \lambda(\vec{u} + \vec{v})'$. Consecuentemente,

$$\|\vec{u} + \vec{v}\|^2 = (\vec{u} + \vec{v})'(\vec{u}' + \vec{v}') = \lambda\|\vec{u}' + \vec{v}'\|^2 = \lambda\|\vec{u} + \vec{v}\|^2.$$

Esto implica $\lambda = 1$ (y por lo tanto la relación que queremos probar) salvo si $\vec{u} + \vec{v} = \vec{0}$, pero en este caso tenemos $\|\vec{u}' + \vec{v}'\| = 0$, luego $\vec{u}' + \vec{v}' = \vec{0} = (\vec{u} + \vec{v})'$.

Sea $W = \langle \tilde{A} \rangle$ y $W' = \langle \tilde{B} \rangle$. Consideramos una base de W contenida en \tilde{A} y definimos $\vec{f} : W \longrightarrow W'$ como la aplicación lineal que sobre la base es

$\vec{f}(\vec{v}) = \vec{v}'$. Los resultados que acabamos de probar justifican que esta relación vale para todos los vectores de \tilde{A} . Ahora observamos que

$$\vec{f}(\vec{u})\vec{f}(\vec{v}) = \vec{u}\vec{v}, \quad \text{para todo } \vec{u}, \vec{v} \in W.$$

En efecto, basta tener en cuenta que esto es cierto para vectores de \tilde{A} y usar la linealidad de \vec{f} y del producto escalar. Haciendo $\vec{u} = \vec{v}$ obtenemos que \vec{f} es inyectiva. La suprayectividad es clara, por construcción. También vemos que $\|\vec{f}(\vec{u})\| = \|\vec{u}\|$. Los isomorfismos entre espacios vectoriales euclídeos que conservan el producto escalar en este sentido se llaman *isometrías lineales*.

Consideramos ahora las variedades lineales $L = O + W$ y $L' = O' + W'$. Es claro que se trata de las menores variedades lineales que contienen a A y B respectivamente. Hemos probado que tienen la misma dimensión. Más aún, la afinidad $\tilde{f} : L \rightarrow L'$ dada por $\tilde{f}(P) = O' + \vec{f}(\overrightarrow{OP})$ es una isometría que extiende a f . Esta isometría es única, pues si g fuera otra, le aplicamos todo el razonamiento anterior, tomando ahora $A = L$, con lo que $\tilde{A} = W$. Con ello probamos que la aplicación $\vec{u} \mapsto \vec{u}'$ es lineal y coincide con \vec{f} en el conjunto \tilde{A} original (el generador de L), luego coincide con \vec{f} en L , luego g coincide con \tilde{f} .

El teorema siguiente recoge lo que hemos obtenido hasta ahora:

Teorema 4.19 *Toda isometría entre dos subconjuntos de dos espacios euclídeos se extiende a una única isometría entre las variedades lineales que generan. Además la extensión es una biyección afín cuya aplicación lineal asociada \vec{f} es una isometría lineal entre los espacios directores de las variedades.*

Ahora veamos que es posible extender la isometría a todo el espacio, aunque perdemos la unicidad. Para ello conviene introducir el concepto siguiente:

Definición 4.20 Dado un subespacio W de un espacio vectorial euclídeo V , se llama *complemento ortogonal* de W al espacio

$$W^\perp = \{\vec{v} \in V \mid \vec{v} \perp \vec{w} \text{ para todo } \vec{w} \in W\}.$$

Es claro que W^\perp es un subespacio vectorial, así como que $W \cap W^\perp = 0$. Tomemos una base ortogonal $\vec{v}_1, \dots, \vec{v}_r$ de W y extendámosla hasta una base ortogonal $\vec{v}_1, \dots, \vec{v}_n$ de V . Entonces es claro que $W^\perp = \langle \vec{v}_{r+1}, \dots, \vec{v}_n \rangle$, luego $V = W \oplus W^\perp$.

En general, si V_1, \dots, V_r son espacios vectoriales cuyos elementos son ortogonales entre sí, es fácil ver que su suma es directa, y el tal caso diremos que tienen *suma ortogonal* y la representaremos por $V_1 \perp \dots \perp V_r$. Hemos probado que $V = W \perp W^\perp$. El teorema siguiente es inmediato:

Teorema 4.21 *Un isomorfismo $f : V \rightarrow V'$ entre dos espacios vectoriales euclídeos es una isometría lineal si y sólo si existe una base $\{\vec{v}_1, \dots, \vec{v}_n\}$ de V tal que $f(\vec{v}_i \vec{v}_j) = f(\vec{v}_i)f(\vec{v}_j)$ para todo i, j . En particular si aplica una base ortonormal a una base ortonormal.*

Con todo esto ya podemos concluir:

Teorema 4.22 Sean E y F dos espacios euclídeos de la misma dimensión. Entonces

1. Las isometrías de E en F son las biyecciones afines f tales que \vec{f} es una isometría lineal.
2. Toda isometría entre un subconjunto A de E y un subconjunto B de F se extiende a una isometría de E en F .
3. Si A contiene un conjunto de $n + 1$ puntos afínmente independientes, entonces la extensión es única.

DEMOSTRACIÓN: Sean $\langle A \rangle = O + \vec{W}$ y $\langle B \rangle = O' + \vec{W}'$, de modo que f se extienda a una isometría afín sobre $O + \vec{W}$. En particular W y W' tienen la misma dimensión, luego sus complementos ortogonales también. Extendamos \vec{f} a V (el espacio vectorial de E) asignando a una base ortonormal de W^\perp una base ortonormal de W'^\perp . Por el teorema anterior la extensión (que seguiremos llamando \vec{f}) es una isometría lineal, y la afinidad que induce en E es una isometría y extiende a f . ■

A través de las isometrías podemos definir una noción general de congruencia de conjuntos arbitrarios que coincide con la que ya conocemos en el caso de segmentos, ángulos y triángulos y con la cual todo espacio euclídeo cumplirá el grupo de axiomas C de la geometría euclídea.

Definición 4.23 Diremos que dos conjuntos A y B en un espacio euclídeo E son *congruentes* si existe una isometría $f : E \longrightarrow E$ tal que $f[A] = B$. Los puntos de A y B correspondientes por (una isometría en particular) f se llaman puntos *homólogos*.

En el sentido de la geometría axiomática, dos segmentos son equivalentes si y sólo si tienen la misma longitud pero, si esto es así, una biyección entre los conjuntos de sus extremos es una isometría, que se extiende a todo el espacio, y es fácil ver que transforma uno de los segmentos en el otro, luego son semejantes en el sentido de la definición anterior. El recíproco es análogo.

Similarmente, si un ángulo \widehat{AOB} es congruente con otro de vértice O' , podemos tomar puntos en sus lados de modo que éste sea $\widehat{A'O'B'}$ y además $\overline{OA} \equiv \overline{O'A'}$, $\overline{OB} \equiv \overline{O'B'}$. Entonces se cumplirá también que $\overline{AB} \equiv \overline{A'B'}$, luego la correspondencia entre los conjuntos $\{A, O, B\}$ y $\{A', O', B'\}$ es una isometría, que se extiende a todo el espacio, y de nuevo es fácil ver que hace corresponder los ángulos, luego son semejantes en el sentido de la definición anterior. De igual modo se prueba el recíproco y la equivalencia correspondiente a triángulos.

Ahora veamos que todo espacio euclídeo cumple el grupo de axiomas C. La comprobación de los axiomas sobre segmentos es muy sencilla, y el axioma C4 sobre ángulos se comprueba de modo similar al axioma C5 sobre triángulos. Por ello probaremos sólo C4.

En primer lugar consideremos dos semiplanos cualesquiera, que serán conjuntos de la forma

$$\pi = \{O' + \lambda \vec{v}_1 + \mu \vec{v}_2 \mid \mu \geq 0\} \quad \text{y} \quad \pi' = \{O' + \lambda \vec{w}_1 + \mu \vec{w}_2 \mid \mu \geq 0\},$$

donde podemos suponer que las bases (\vec{v}_1, \vec{v}_2) y (\vec{w}_1, \vec{w}_2) son ortonormales. Si llamamos \vec{f} a la aplicación lineal determinada por $\vec{f}(\vec{v}_i) = \vec{w}_i$, entonces \vec{f} es una isometría lineal, y es claro que la aplicación dada por $f(P) = O' + \vec{f}(\overrightarrow{OP})$ es una isometría entre los planos que contienen a π y a π' , tal que $f[\pi] = \pi'$ y que además transforma la semirrecta $O + \{\lambda \vec{v}_1 \mid \lambda \geq 0\}$, en la semirrecta $O' + \{\lambda \vec{w}_1 \mid \lambda \geq 0\}$.

De aquí se sigue que dados dos semiplanos y en sus fronteras sendas semirrectas, existe una isometría que transforma uno en el otro haciendo corresponder las semirrectas. Por lo tanto, si tenemos un ángulo, un semiplano y en su frontera una semirrecta (según el axioma C4) existe una isometría entre el semiplano que contiene al ángulo con un lado en su frontera y el semiplano dado, de modo que el lado del ángulo se corresponda con la semirrecta dada. La imagen del ángulo por la isometría indicada será un ángulo congruente con el dado, contenido en el semiplano dado y con un lado igual a la semirrecta dada, tal y como exige el axioma C4.

Para probar la unicidad suponemos que tenemos dos ángulos con un lado en común y contenidos en el mismo semiplano respecto a éste. Serán de la forma $A(O; \vec{v}_1, \vec{v}_2)$ y $A(O; \vec{v}_1, \vec{v}_2')$. Además ambos están contenidos en el semiplano

$$\{O + \lambda \vec{v}_1 + \mu \vec{w} \mid \mu \geq 0\},$$

lo que se traduce en que $\vec{v}_2 = x\vec{v}_1 + y\vec{w}$ con $y > 0$ y $\vec{v}_2' = x'\vec{v}_1 + y'\vec{w}$ con $y' > 0$. Podemos suponer que \vec{v}_1 , \vec{v}_2 y \vec{v}_2' tienen todos norma 1.

Una isometría f entre los ángulos ha de cumplir⁵ $\vec{f}(\vec{v}_1) = \vec{v}_1$ y $\vec{f}(\vec{v}_2) = \vec{v}_2'$ o bien $\vec{f}(\vec{v}_1) = \vec{v}_2'$ y $\vec{f}(\vec{v}_2) = \vec{v}_1$. En cualquier caso tendremos $\vec{v}_1 \vec{v}_2 = \vec{v}_1 \vec{v}_2'$, luego $x^2 = x'^2$. Por otro lado,

$$x^2 + y^2 = \|\vec{v}_2\|^2 = \|\vec{v}_2'\|^2 = x'^2 + y'^2,$$

y como $y, y' > 0$, de hecho $y = y'$. Desarrollando igualmente la condición $\|\vec{v}_2 - \vec{v}_1\|^2 = \|\vec{v}_2' - \vec{v}_1\|^2$ se concluye $x = x'$, con lo que los ángulos son iguales. ■

Así pues, tenemos que todo espacio tridimensional euclídeo en el sentido analítico cumple los axiomas de la geometría euclídea, y que todo espacio que cumpla dichos axiomas puede ser dotado de estructura de espacio tridimensional euclídeo. La aplicación que a cada punto de un espacio euclídeo de dimensión n le hace corresponder sus coordenadas en \mathbb{R}^n respecto a un sistema de referencia ortonormal es una isometría (porque la aplicación lineal asociada envía la base

⁵El punto más delicado es comprobar que si $\widehat{AOB} = \widehat{A'O'B'}$ entonces $OA = O'A'$ y $OB = O'B'$. Por ejemplo, puede probarse que si una recta corta a un ángulo en un segmento, entonces los extremos del segmento están sobre los lados. De aquí se sigue fácilmente la propiedad indicada.

ortonormal en la base canónica), y las isometrías conservan todas las propiedades geométricas, luego resulta que todos los espacios euclídeos de la misma dimensión tienen las mismas propiedades, luego cualquier resultado geométrico que se obtenga en uno de ellos es aplicable a todos los demás. En este sentido se dice que la geometría euclídea es *completa* o, mejor, *categorica*.

Por último señalamos que, desde el momento en que hemos probado que los espacios afines euclídeos verifican los axiomas geométricos de congruencia, todos los hechos que hemos probado al comienzo de la sección a partir de tales axiomas son válidas en general.⁶ En particular tenemos la relación entre el coseno de un ángulo y el producto escalar:

$$\vec{v}\vec{w} = \|\vec{v}\| \|\vec{w}\| \cos \widehat{\vec{v}\vec{w}},$$

de la cual se deduce a su vez la versión analítica del teorema del coseno:

$$\|\overrightarrow{PQ}\|^2 = \|\overrightarrow{OP}\|^2 + \|\overrightarrow{OQ}\|^2 - 2\|\overrightarrow{OP}\| \|\overrightarrow{OQ}\| \cos A(O; \overrightarrow{OP} \overrightarrow{OQ}),$$

válida para puntos cualesquiera O, P, Q (sin exigir que sean distintos o no colineales) si convenimos que el ángulo es nulo si uno de los vectores que lo definen es nulo. A su vez esta fórmula contiene como caso particular el teorema de Pitágoras.

4.5 Los giros y la medida de ángulos

Dedicamos esta sección a definir el concepto de giro, el último concepto euclídeo que nos obligará a resolver algunos problemas técnicos, concretamente al respecto de la medida de ángulos.

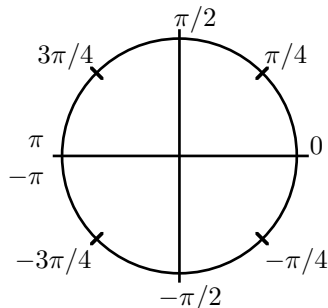
Consideremos un sistema de referencia ortonormal $(O; \vec{v}_1, \vec{v}_2)$ en un plano afín euclídeo. Vamos a adoptar por primera vez el punto de vista habitual de la geometría analítica. Identificaremos a cada punto P con su vector de posición \overrightarrow{OP} y éste a su vez con sus coordenadas (x, y) . Notemos que, en estos términos, $\overrightarrow{PQ} = Q - P$. Es costumbre llamar semiplano derecho al semiplano $x \geq 0$, semiplano izquierdo al semiplano $x \leq 0$, semiplano superior a $y \geq 0$ y semiplano inferior a $y \leq 0$ (pero hay que tener presente que la diferencia entre derecha e izquierda no es intrínseca, en el sentido de que depende del sistema de referencia y es imposible establecerla de forma absoluta sin hacer referencia a la anatomía humana, o algo similar).

Consideramos también la circunferencia ω de centro $(0, 0)$ y radio 1, cuyos puntos se caracterizan por la relación $x^2 + y^2 = 1$.

Vamos a asignar a cada punto de ω un número real al que llamaremos su argumento. Fijamos una unidad de ángulos. Si un punto P de ω está en el semiplano superior, su argumento será la medida del arco menor de extremos

⁶Aunque los axiomas que hemos dado describen la geometría tridimensional, los hechos citados son bidimensionales, y todos los planos euclídeos bidimensionales son isométricos.

$(1, 0)$ y P (entendiendo que el argumento de $(1, 0)$ es 0). Si el punto P está en el semiplano inferior, su argumento se define igualmente pero con signo negativo.



Con esto tenemos bien definido el argumento sobre todos los puntos de ω excepto $(-1, 0)$, diametralmente opuesto a $(1, 0)$, cuyo argumento debería ser por una parte π y por otra parte $-\pi$. Resolvemos esto asignando a cada punto, no un argumento, sino una clase de argumentos módulo 2π , es decir, si α es el argumento de P según lo acabamos de definir, ahora definimos $\arg P = \{\alpha + 2k\pi \mid k \in \mathbb{Z}\}$. De este modo, $\arg(-1, 0) = \{\pi + 2k\pi \mid k \in \mathbb{Z}\}$, que contiene tanto a π como a $-\pi$, y la definición resulta consistente. Tenemos así una biyección entre la circunferencia ω y el grupo cociente $\mathbb{R}/2\pi\mathbb{R}$.

Intuitivamente, el argumento de un punto indica el ángulo que hay que girar desde $(1, 0)$ para llegar hasta él. Así, el hecho de que un mismo punto tenga argumentos $\pi/2$, $-3\pi/2$ y $5\pi/2$ significa que se llega hasta él girando un recto en un sentido, o tres rectos en sentido opuesto, o dando una vuelta entera de 4 rectos más otro recto. El resultado básico en torno a los argumentos es el siguiente:

Teorema 4.24 Sean P y Q dos puntos en ω , sean $\arg P = [\alpha]$, $\arg Q = [\beta]$ de modo que $|\beta - \alpha| \leq \pi$. Entonces el ángulo \widehat{POQ} tiene amplitud $|\beta - \alpha|$.

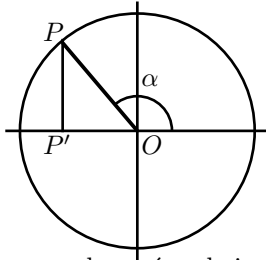
DEMOSTRACIÓN: Notemos en general que cada clase módulo 2π tiene un único representante en el intervalo $[-\pi, \pi]$ (excepto π , que tiene dos pero con el mismo valor absoluto). Esto implica que el valor de $|\beta - \alpha|$ menor o igual que π está unívocamente determinado por P y Q y no importa con qué representantes concretos lo calculamos. En principio tomamos α y β en $[-\pi, \pi]$.

Sean $I = (1, 0)$, $I' = (-1, 0)$. Si P y Q están ambos en el semiplano superior, podemos suponer que $0 \leq \alpha \leq \beta \leq \pi$. Descartando casos triviales tenemos que \widehat{IOP} está contenido en \widehat{IOQ} y sus amplitudes son α y β , luego es claro que la de \widehat{POQ} es $\beta - \alpha$. El mismo argumento vale si P y Q están ambos en el semiplano inferior.

Supongamos que P está en el semiplano superior y Q está en el inferior. Si Q es diametralmente opuesto a P , entonces \widehat{IOQ} es adyacente a \widehat{IOP} , luego sus amplitudes α y $-\beta$ suman π , como había que probar.

Sea P' el punto diametralmente opuesto a P . Podemos suponer que P es distinto de I, I' . Entonces la recta PP' deja a I y a I' en semiplanos distintos. Supongamos que Q e I están en el mismo semiplano. Entonces es claro que $\widehat{POQ} = \widehat{POI} + \widehat{IOQ}$ y de nuevo la suma de las amplitudes es $\alpha - \beta < \pi$.

Finalmente consideramos el caso en que Q e I' están en el mismo semiplano. Entonces $\widehat{POQ} = \widehat{POI'} + \widehat{QOI'}$. Las amplitudes de los sumandos son $\pi - \alpha$ y $\pi + \beta$, luego la amplitud de \widehat{POQ} es $2\pi + \beta - \alpha$. Si tomamos $2\pi + \beta$ como argumento de Q , se cumple lo pedido. ■



Consideremos un punto $P(x, y)$ en ω , sea α su argumento en $]-\pi, \pi]$, sea $P' = (x, 0)$. Si x, y son no nulos, entonces el triángulo $\widehat{OPP'}$ es rectángulo, pues $P - P' = (0, y)$ es ortogonal a P' . Por otra parte, si $I = (1, 0)$, la definición de argumento nos da que $\widehat{IOP} = \pm\alpha$, donde el signo es el de la coordenada y .

Puesto que $\|P\| = 1$, la definición del seno y el coseno de un ángulo implican entonces que $(x, y) = (\cos(\pm\alpha), \pm\sin(\pm\alpha))$. Esta fórmula sigue siendo cierta para los puntos $(\pm 1, 0)$ y $(0, \pm 1)$ (cuyos argumentos son $0, \pm\pi/2, \pm\pi$) por la definición del seno y del coseno en estos casos particulares. Recordemos que al comienzo de la sección anterior hemos convenido en que el coseno de un ángulo nulo es 1 y el coseno de un ángulo llano es -1 (y en ambos casos definimos el seno como 0).

Hasta aquí tenemos definidas las funciones seno y coseno en el intervalo $[0, \pi]$. Ahora extendemos las definiciones al intervalo $[-\pi, \pi]$ mediante las relaciones

$$\sin(-\alpha) = -\sin \alpha, \quad \cos(-\alpha) = \cos \alpha.$$

Así la relación entre α y las coordenadas de P se reduce a

$$(x, y) = (\cos \alpha, \sin \alpha).$$

Puesto que todos los números $\alpha + 2k\pi$ son argumentos de P , podemos hacer que esta fórmula valga para cualquiera de ellos si extendemos las funciones seno y coseno a todo el conjunto \mathbb{R} mediante

$$\sin(\alpha + 2k\pi) = \sin \alpha, \quad \cos(\alpha + 2k\pi) = \cos \alpha, \quad \text{para todo } k \in \mathbb{Z}.$$

A continuación extendemos el concepto de argumento a cualquier punto P del plano distinto de O . Para ello consideramos la intersección con ω de la semirrecta de origen O que pasa por P , esto es, el punto $P^* = P/\|P\|$. Definimos el *argumento* de P como $\arg P = \arg P^*$. Si este argumento es α , las coordenadas de P son

$$(x, y) = \|P\|(\cos \alpha, \sin \alpha).$$

Definición 4.25 Llamaremos *coordenadas polares* de un punto $P \neq 0$ (respecto al sistema de referencia dado) a los números $\rho = \|P\|$ (el *módulo* de P) y $\theta = \arg P$ (el argumento de P , determinado módulo 2π). Representaremos por $P = \rho_\theta$ al único punto P del plano de módulo ρ y argumento θ .

En estos términos, la relación entre las coordenadas cartesianas y las coordenadas polares de un punto P es

$$(x, y) = (\rho \cos \theta, \rho \sin \theta).$$

Definición 4.26 Sea π un plano afín y $(O; \vec{v}_1, \vec{v}_2)$ un sistema de referencia. Sea α un número real. Llamaremos *giro* de centro O y ángulo α a la aplicación determinada por $G(O) = O$ y $G_\alpha(\rho_\theta) = \rho_{\theta+\alpha}$.

Los giros son isometrías. En efecto, dados dos puntos P y Q , hemos de probar que $\|Q - P\| = \|G_\alpha(Q) - G_\alpha(P)\|$. Si $P = O$ el término izquierdo es la coordenada ρ de Q , mientras que el término derecho es la coordenada ρ de $G_\alpha(Q)$, y ésta se conserva por definición. Si $P \neq O \neq Q$, entonces el teorema 4.24 prueba que $\widehat{POQ} = G_\alpha(\widehat{P}O\widehat{Q})$ (incluyendo los valores posibles 0 y π si los puntos están alineados). Como también tenemos que $\|P\| = \|G_\alpha(P)\|$ y $\|Q\| = \|G_\alpha(Q)\|$, por el teorema del coseno $\|Q - P\| = \|G_\alpha(Q) - G_\alpha(P)\|$.

Puesto que $G_\alpha(O) = O$, la expresión en coordenadas de G_α es lineal (coincide con la de \vec{G}_α). Vamos a calcular su matriz en la base del sistema de referencia que estamos considerando. Se trata de la matriz que tiene por filas las imágenes de los puntos $(1, 0)$ y $(0, 1)$. El punto $P(1, 0) = 1_0$ se transforma en $1_\alpha = (\cos \alpha, \sin \alpha)$, mientras que el punto $P(0, 1) = 1_{\pi/2}$ se transforma en $1_{\pi/2+\alpha} = (\cos(\pi/2 + \alpha), \sin(\pi/2 + \alpha))$. Por otra parte este vector tiene norma 1 y es ortogonal a $(\cos \alpha, \sin \alpha)$. Sólo hay dos vectores en estas condiciones (el complemento ortogonal del subespacio generado por este último vector tiene dimensión 1). Así pues,

$$G_\alpha(0, 1) = (\cos(\pi/2 + \alpha), \sin(\pi/2 + \alpha)) = \pm(-\sin \alpha, \cos \alpha)$$

(pues estos dos vectores cumplen las condiciones y no puede haber más). Es fácil ver que el signo ha de ser positivo. Por ejemplo, si $0 < \alpha < \pi/2$ entonces $\cos \alpha > 0$ y por otra parte $\pi/2 < \pi/2 + \alpha < \pi$, luego $\sin(\pi/2 + \alpha) > 0$. Similarmente se discuten las demás posibilidades. Así pues, la matriz del giro de ángulo α en un sistema de referencia dado es

$$M_\alpha = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}.$$

Teniendo en cuenta que $G_\alpha(O) = O$, la expresión en coordenadas de G_α resulta ser

$$G_\alpha(x, y) = (x, y) \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}.$$

Es fácil ver que, salvo que $\alpha = 2k\pi$ con $k \in \mathbb{Z}$, el único punto fijo de G_α es el origen O , luego un giro determina su centro. Ahora vamos a comparar los giros definidos en distintos sistemas de referencia ortonormales con un mismo origen O . Para ello hemos de notar que de la definición de giro se sigue inmediatamente que $G_\alpha \circ G_\beta = G_{\alpha+\beta}$, luego $M_\alpha M_\beta = M_{\alpha+\beta}$. En particular $M_\alpha^{-1} = M_{-\alpha}$.

Recordemos que habíamos fijado el sistema $(O; \vec{v}_1, \vec{v}_2)$ y sea ahora $(O; \vec{w}_1, \vec{w}_2)$ otro sistema de referencia ortonormal. Sea $O + \vec{w}_1 = 1_\theta$. Entonces $\vec{w}_1 = \vec{G}_\theta(\vec{v}_1)$. Los vectores $\vec{G}_\theta(\vec{v}_2)$ y \vec{w}_2 son ambos ortogonales a \vec{w}_1 y tienen módulo 1, luego $\vec{w}_2 = \pm \vec{G}_\theta(\vec{v}_2)$. Teniendo esto en cuenta es fácil ver que la matriz de cambio de base de (\vec{v}_1, \vec{v}_2) a (\vec{w}_1, \vec{w}_2) es

$$M_\theta \begin{pmatrix} 1 & 0 \\ 0 & \pm 1 \end{pmatrix},$$

luego su inversa es

$$\begin{pmatrix} 1 & 0 \\ 0 & \pm 1 \end{pmatrix} M_{-\theta}.$$

Así pues, la matriz del giro de ángulo α (definido respecto al primer sistema de referencia) en el segundo sistema de referencia es

$$M_{\theta} \begin{pmatrix} 1 & 0 \\ 0 & \pm 1 \end{pmatrix} M_{\alpha} \begin{pmatrix} 1 & 0 \\ 0 & \pm 1 \end{pmatrix} M_{-\theta} = M_{\theta} M_{\pm\alpha} M_{-\theta} = M_{\pm\alpha}.$$

Esto significa que el giro de ángulo α respecto a un sistema de referencia es el giro de ángulo $\pm\alpha$ respecto a cualquier otro sistema de referencia con el mismo origen, y además el posible cambio de signo depende sólo de los sistemas de referencia, es decir, o bien todos los giros coinciden, o bien todos cambian de signo. Vemos, pues, que un giro (distinto de la identidad) determina su centro y su ángulo salvo signo. Dicho signo depende del sistema de referencia.

Como aplicación vamos a obtener unas importantes relaciones trigonométricas. Basta desarrollar la igualdad $M_{\alpha+\beta} = M_{\alpha} M_{\beta}$ para concluir

$$\begin{aligned} \operatorname{sen}(\alpha + \beta) &= \operatorname{sen} \alpha \cos \beta + \cos \alpha \operatorname{sen} \beta, \\ \cos(\alpha + \beta) &= \cos \alpha \cos \beta - \operatorname{sen} \alpha \operatorname{sen} \beta. \end{aligned}$$

Enseguida veremos que con estas fórmulas tenemos completamente determinadas (salvo un factor constante) las funciones seno y coseno. Primero enunciaremos un teorema que recoja las propiedades esenciales de ambas funciones. Todas están ya demostradas.

Teorema 4.27 *Las funciones seno y coseno (respecto a una medida de ángulos prefijada en la que los ángulos llanos tengan amplitud π) verifican las propiedades siguientes:*

1. $\operatorname{sen}(\alpha + 2k\pi) = \operatorname{sen} \alpha$, $\cos(\alpha + 2k\pi) = \cos \alpha$, para todo $k \in \mathbb{Z}$.
2. $\operatorname{sen} 0 = 0$, $\cos 0 = 1$, $\operatorname{sen}(\pi/2) = 1$, $\cos(\pi/2) = 0$.
3. $\operatorname{sen}^2 \alpha + \cos^2 \alpha = 1$, $\operatorname{sen}(-\alpha) = -\operatorname{sen} \alpha$, $\cos(-\alpha) = \cos \alpha$.
4. $\operatorname{sen} \alpha \geq 0$ si $\alpha \in [0, \pi]$, $\cos \alpha \geq 0$ si $\alpha \in [-\pi/2, \pi/2]$.
5. Para cada par $(x, y) \in \mathbb{R}^2$ tal que $x^2 + y^2 = 1$ existe un único $\alpha \in]-\pi, \pi]$ tal que $(x, y) = (\cos \alpha, \operatorname{sen} \alpha)$.
6. $\operatorname{sen}(\alpha + \beta) = \operatorname{sen} \alpha \cos \beta + \cos \alpha \operatorname{sen} \beta$,
7. $\cos(\alpha + \beta) = \cos \alpha \cos \beta - \operatorname{sen} \alpha \operatorname{sen} \beta$.

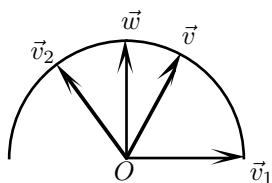
Nuestra definición de las funciones seno y coseno se basa en la delicada construcción de la medida de ángulos, que no admite una formulación algebraica sencilla. Estas funciones pueden ser definidas mucho más fácilmente mediante técnicas analíticas. Vamos a probar que cualquier par de funciones que cumplan

las propiedades del teorema anterior son de hecho el seno y el coseno que hemos definido (respecto a una unidad de ángulos adecuada). Así tendremos justificado que nuestra definición geométrica equivale a cualquiera de las varias definiciones analíticas posibles.

Supongamos, pues, que tenemos dadas dos funciones sen y cos que satisfagan las propiedades del teorema anterior. Entonces, dado $x \in [0, 1]$, existe un único número $y \geq 0$ tal que $x^2 + y^2 = 1$, y así $(x, y) = (\cos \alpha, \sin \alpha)$ para un único $\alpha \in [0, \pi]$. El teorema de Cauchy-Schwarz nos permite definir la *amplitud* del ángulo entre dos vectores \vec{v}_1 y \vec{v}_2 como el único número $\alpha \in [0, \pi]$ que cumple

$$\cos \alpha = \frac{\vec{v}_1 \vec{v}_2}{\|\vec{v}_1\| \|\vec{v}_2\|}.$$

La amplitud de un ángulo \widehat{ABC} es la de los vectores \overrightarrow{BA} y \overrightarrow{BC} . Es fácil ver que esta definición no depende de la elección de A y C en los lados del ángulo, así como que se conserva por isometrías, es decir, que ángulos congruentes tienen la misma amplitud. Vamos a comprobar que esta definición de amplitud satisface las condiciones del teorema 2.27. Acabamos de probar la primera.



Para probar la segunda consideramos un ángulo $\hat{A} = A(O; \vec{v}_1, \vec{v}_2)$, donde podemos suponer que \vec{v}_1 y \vec{v}_2 tienen norma 1, y una semirrecta $O + \{\lambda \vec{v} \mid \lambda \geq 0\}$ contenida en el ángulo, lo cual equivale a que

$$\vec{v} = \lambda \vec{v}_1 + \mu \vec{v}_2, \quad \lambda, \mu \geq 0.$$

(También podemos suponer que \vec{v} tiene norma 1.)

Sea \vec{w} un vector ortogonal a \vec{v}_1 en el plano del ángulo. Cambiando \vec{w} por $-\vec{w}$ si es preciso podemos suponer que $\vec{v}_2 = x\vec{v}_1 + y\vec{w}$, con $y > 0$. Sean α , β y γ las amplitudes de los ángulos $\widehat{\vec{v}_1 \vec{v}_2}$, $\widehat{\vec{v}_1 \vec{v}}$ y $\widehat{\vec{v} \vec{v}_2}$ respectivamente. Hemos de probar que $\alpha = \beta + \gamma$.

Tenemos $\cos \alpha = \vec{v}_1 \vec{v}_2 = x$, luego $y = \sin \alpha$ (pues $x^2 + y^2 = 1$ e $y > 0$). Así pues, $\vec{v}_2 = \cos \alpha \vec{v}_1 + \sin \alpha \vec{w}$. De aquí que

$$\vec{v} = (\lambda + \mu \cos \alpha) \vec{v}_1 + \mu \sin \alpha \vec{w}.$$

A su vez esto implica que $\cos \beta = \vec{v}_1 \vec{v} = \lambda + \mu \cos \alpha$, luego $\sin \beta = \mu \sin \alpha$. Usando las fórmulas 6 y 7 del teorema 4.27 vemos que

$$\begin{aligned} \cos(\alpha - \beta) &= \cos \alpha \cos \beta + \sin \alpha \sin \beta = \lambda \cos \alpha + \mu = \vec{v} \vec{v}_2 = \cos \gamma, \\ \sin(\alpha - \beta) &= \sin \alpha \cos \beta - \sin \beta \cos \alpha = \lambda \sin \alpha \geq 0. \end{aligned}$$

Puesto que $|\alpha - \beta| \leq \pi$, la última desigualdad implica que $0 \leq \alpha - \beta \leq \pi$, luego de $\cos(\alpha - \beta) = \cos \gamma$ podemos concluir $\alpha - \beta = \gamma$, como queríamos probar.

Llamemos $m(\hat{A})$ a la amplitud de un ángulo \hat{A} tal y como la hemos definido a partir de las funciones sen y cos. Sea $m'(\hat{A})$ la amplitud de \hat{A} en el sentido usual, digamos que tomando como unidad de ángulos el ángulo recto. Claramente los

ángulos rectos tienen medida m igual a $\pi/2$. El teorema 2.27 nos da entonces que $m(\hat{A}) = (\pi/2)m'(\hat{A})$, para todo ángulo \hat{A} , dicho de otro modo, que la amplitud de un ángulo cualquiera en el sentido usual es $(2/\pi)m(\hat{A})$ ángulos rectos.

Si suponemos además que $\pi > 1$, entonces podemos tomar como unidad de ángulos $2/\pi$ veces un ángulo recto y, respecto a esta unidad, la amplitud de un ángulo cualquiera es precisamente $m(\hat{A})$.

Finalmente, dado $\alpha \in]0, \pi[$, fijamos un sistema de referencia ortonormal y consideramos los puntos $P(\cos \alpha, \sin \alpha)$ y $Q(1, 0)$. Entonces el triángulo \widehat{POQ} tiene un ángulo recto \hat{Q} y la perpendicular a OQ por P corta a OQ en el punto $(\cos \alpha, 0)$. Como $\|P\| = 1$, concluimos que $\sin \alpha$ y $\cos \alpha$ son el seno y el coseno de \hat{O} en el sentido geométrico, es decir, hemos probado:

Teorema 4.28 *Si unas funciones \sin y \cos satisfacen las propiedades del teorema 4.27 con $\pi > 1$, entonces existe una unidad de ángulos respecto a la cual $\sin \alpha$ y $\cos \alpha$ son el seno y el coseno de los ángulos de amplitud α (para todo $\alpha \in [0, \pi]$).*

Es importante insistir en que nosotros hemos demostrado la existencia de las funciones seno y coseno. No obstante, si se supone conocida su existencia junto con las propiedades del teorema 4.27, la fórmula $\vec{v}\vec{w} = \|\vec{v}\| \|\vec{w}\| \cos \widehat{vw}$ (usada como definición de la amplitud de \widehat{vw}) permite introducir rápidamente la medida de ángulos en los espacios euclídeos, y de ella se siguen inmediatamente el teorema del coseno y el teorema de Pitágoras. Por completitud daremos la prueba analítica de otros dos resultados fundamentales: probaremos que los ángulos de un triángulo suman dos rectos y también el teorema de los senos.

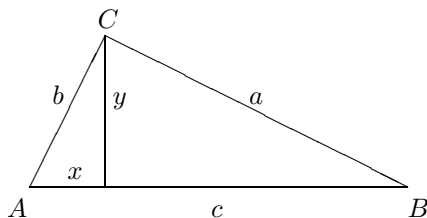
Dado un triángulo \widehat{ABC} , tomando un sistema de referencia ortonormal con origen en A y un vector en la dirección de \overrightarrow{AB} podemos suponer que $A = (0, 0)$, $B = (c, 0)$, $C = (x, y)$, con $c, y \geq 0$. Se comprueba fácilmente que

$$\cos \hat{A} = \frac{x}{b}, \quad \sin \hat{A} = \frac{y}{b}, \quad \cos \hat{B} = \frac{c-x}{a}, \quad \sin \hat{B} = \frac{y}{a}.$$

Las fórmulas de la suma nos dan que

$$\begin{aligned} \sin(\hat{A} + \hat{B}) &= \frac{yc}{ab} \geq 0, \\ \cos(\hat{A} + \hat{B}) &= \frac{cx - x^2 - y^2}{ab} = -\frac{\overrightarrow{AC}\overrightarrow{BC}}{ab} = -\cos \hat{C}. \end{aligned}$$

Estas relaciones implican que $0 \leq \hat{A} + \hat{B} \leq \pi$ y $\hat{A} + \hat{B} = \pi - \hat{C}$. ■



Conservando la notación, tenemos

$$C = (b \cos \hat{A}, b \sin \hat{A}), \quad B - C = (a \cos \hat{B}, -a \sin \hat{B}),$$

y como $B = (c, 0)$, al sumar las segundas coordenadas queda

$$0 = b \sin \hat{A} - a \sin \hat{B},$$

que es una de las igualdades del teorema de los senos. La otra se prueba intercambiando los papeles los vértices (no entraremos en la relación con el radio de la circunferencia circunscrita). ■

4.6 Complementos sobre trigonometría

Ahora que tenemos caracterizadas completamente las funciones seno y coseno podemos probar unos pocos resultados adicionales sobre trigonometría que nos serán útiles más adelante. En primer lugar destacamos las fórmulas del seno y el coseno del ángulo doble, que son consecuencias inmediatas de las fórmulas de la suma:

$$\sin 2\alpha = 2 \sin \alpha \cos \alpha, \quad \cos 2\alpha = \cos^2 \alpha - \sin^2 \alpha.$$

De ellas se deducen sin esfuerzo las fórmulas para el ángulo mitad:

$$\sin \frac{\alpha}{2} = \sqrt{\frac{1 - \cos \alpha}{2}}, \quad \cos \frac{\alpha}{2} = \sqrt{\frac{1 + \cos \alpha}{2}}.$$

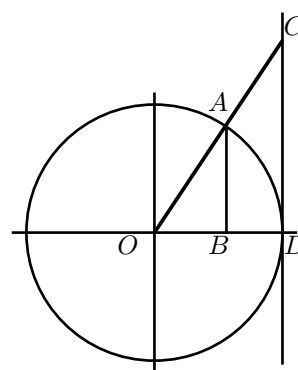
Introducimos ahora una tercera función trigonométrica de interés:

Definición 4.29 Llamaremos *tangente* de un ángulo α a

$$\tan \alpha = \frac{\sin \alpha}{\cos \alpha}.$$

La tangente está definida para todos los números reales α excepto aquellos cuyo coseno es nulo, es decir, $\alpha = \pm\pi/2 + 2k\pi$, con $k \in \mathbb{Z}$. Su interpretación geométrica es fácil de obtener. Tomemos $0 < \alpha < \pi/2$. Si la circunferencia de la figura tiene radio 1, sabemos que las longitudes de los segmentos \overline{OA} y \overline{OB} son respectivamente $\sin \alpha$ y $\cos \alpha$. Por lo tanto la $\tan \alpha$ es el cociente de ambas longitudes, y por el teorema de Tales este cociente es independiente de la recta vertical AB que escojamos para calcularlo. En particular si tomamos la tangente CD vemos que $\tan \alpha = \overline{CD}$.

Para los demás valores posibles de α la tangente se interpreta del mismo modo, y la única diferencia es el signo.



En general, vemos que la tangente de uno de los ángulos agudos de un triángulo rectángulo es la razón entre el cateto opuesto y el cateto contiguo. De aquí se sigue que existen ángulos agudos cuya tangente es igual a cualquier número real positivo prefijado. Las fórmulas siguientes se prueban sin dificultad a partir de las correspondientes para senos y cosenos:

$$\tan(\alpha + 2\pi) = \tan \alpha, \quad \tan(-\alpha) = -\tan \alpha,$$

$$\tan(\alpha + \beta) = \frac{\tan \alpha + \tan \beta}{1 - \tan \alpha \tan \beta}, \quad \tan 2\alpha = \frac{2 \tan \alpha}{1 - \tan^2 \alpha}.$$

Dejamos al lector la comprobación de la tabla siguiente. Es fácil obtener pruebas tanto geométricas como algebraicas.

α	sen	cos	tan
0	0	1	0
$\pi/6$	1/2	$\sqrt{3}/2$	$1/\sqrt{3}$
$\pi/4$	$\sqrt{2}/2$	$\sqrt{2}/2$	1
$\pi/3$	$\sqrt{3}/2$	1/2	$\sqrt{3}$
$\pi/2$	1	0	—

4.7 Circunferencias

Veamos la expresión analítica de las circunferencias. Dado un plano afín euclídeo y fijado en él un sistema de referencia ortonormal, la circunferencia de centro un punto (a, b) y radio $r > 0$ está formada por los puntos (x, y) cuya distancia a (a, b) es igual a r , es decir, $\|(x - a, y - b)\| = r$, o equivalentemente:

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

Desarrollando la ecuación, vemos que las circunferencias están formadas por los puntos cuyas coordenadas satisfacen una ecuación de la forma

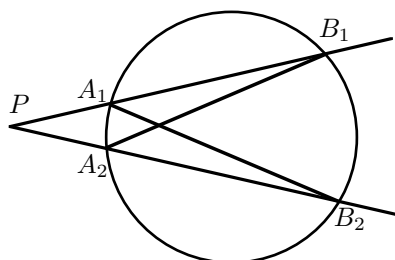
$$x^2 + y^2 + Ax + By + C = 0, \quad (4.1)$$

donde $A = -2a$, $B = -2b$, $C = a^2 + b^2 - r^2$.

Demostramos ahora un resultado clásico sobre circunferencias cuya prueba sintética es inmediata. A continuación lo relacionaremos con la geometría analítica.

Teorema 4.30 (Teorema de la potencia) *Sea P un punto que no esté en una circunferencia ω . Sean L_1 y L_2 dos rectas secantes a ω que pasen por P . Sean A_1 y B_1 los puntos en que L_1 corta a ω y sean A_2 y B_2 los puntos en que L_2 corta a ω . Entonces*

$$\overline{PA_1} \cdot \overline{PB_1} = \overline{PA_2} \cdot \overline{PB_2}.$$



DEMOSTRACIÓN: Basta observar que los triángulos $\widehat{PA_2B_1}$ y $\widehat{PA_1B_2}$ tienen dos ángulos iguales ($\widehat{B_1}$ y $\widehat{B_2}$ abarcan arcos iguales), luego son semejantes. Si el punto P es interior a la circunferencia el razonamiento es similar. ■

Definición 4.31 La *potencia* de un punto P respecto a una circunferencia ω se define como 0 si P está en ω y en caso contrario como la cantidad constante considerada en el teorema anterior, con signo positivo si P es exterior a ω y con signo negativo si P es interior.

Si ω tiene centro O y radio r y la distancia de P a r es d , podemos calcular la potencia de P mediante la recta OP (tomamos una recta cualquiera si $P = O$). Es fácil ver que si P es exterior la potencia es $(d - r)(d + r) = d^2 - r^2$, y si P es interior obtenemos $-(d + r)(r - d) = d^2 - r^2$. Esta expresión vale trivialmente si P está en ω , luego en general tenemos que la potencia de P respecto a ω viene dada por la fórmula $d^2 - r^2$.

Teniendo en cuenta cómo hemos obtenido la ecuación (4.1) es claro que la potencia de un punto (x, y) respecto a la circunferencia que ésta determina es precisamente el valor del miembro izquierdo de la ecuación.

Ejercicio: Probar que si P es un punto exterior a una circunferencia ω , entonces la potencia de P es el cuadrado de la distancia de P a los puntos donde las tangentes a ω por P tocan a ω .

Ejercicio: Probar que el conjunto de los puntos con la misma potencia respecto a dos circunferencias no concéntricas es una recta perpendicular a la recta que pasa por sus centros (esta recta recibe el nombre de *eje radical* de las circunferencias).

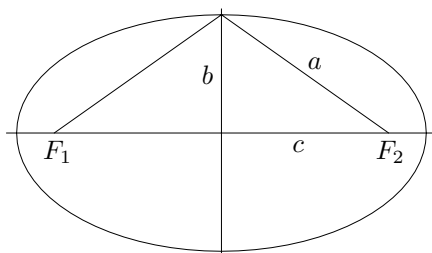
4.8 Cónicas

Como ilustración de las técnicas analíticas que hemos introducido en este capítulo vamos a exponer la teoría básica sobre las secciones cónicas, que incluyen tres clases de curvas: elipses, hipérbolas y parábolas. Empezaremos estudiándolas individualmente y después veremos las relaciones entre ellas. En toda esta sección E será un plano afín euclídeo.

La elipse Es posible definir una elipse de muchas formas alternativas, pero quizá la más intuitiva sea la siguiente:

Definición 4.32 Una *elipse* es el conjunto de puntos del plano tales que la suma de sus distancias a dos puntos fijos llamados *focos* es constante.

La figura siguiente muestra una elipse de focos F_1 y F_2 . Podemos pensar que unimos F_1 y F_2 con una cuerda de longitud l , la tensamos estirando de un punto y movemos dicho punto de modo que la cuerda nunca deje de estar tensada. La trayectoria del punto será la elipse correspondiente a los focos dados y a la suma de distancias l .



El punto medio del segmento que une los focos se llama *centro* de la elipse, los segmentos que unen dos puntos de la elipse y que pasan por su centro se llaman *diámetros*. El diámetro que contiene a los focos se llama *eje mayor* y el diámetro perpendicular se llama *eje menor*. Los extremos del diámetro mayor se llaman *vértices mayores* y los extremos del eje menor *vértices menores*. La distancia c del centro a cada foco se llama *distancia focal*.

Es claro que los vértices menores están a la misma distancia de los dos focos, luego si llamamos a a dicha distancia, entonces el valor constante de la suma de las distancias a los focos desde cualquier punto de la elipse es $l = 2a$. Notemos que para que pueda existir una elipse con distancia focal c y suma de longitudes $2a$ es necesario (y suficiente) que $c < a$.

La distancia de los vértices mayores/menores al centro de la elipse se llama *semieje mayor/menor* de la elipse. Al semieje menor lo representaremos por b , mientras que el semieje mayor no es sino la distancia a que acabamos de definir.

En efecto, si llamamos a' al semieje mayor, la distancia de un vértice mayor a su foco más próximo es $a' - c$, mientras que la distancia al foco más alejado es $c + a'$, luego la suma de ambas distancias es $a' - c + c + a' = 2a'$ y, como dicha suma ha de ser $2a$, resulta que $a' = a$.

La figura muestra entonces la relación $a^2 = b^2 + c^2$. En particular vemos que el semieje mayor es siempre mayor que el semieje menor, por lo que los nombres son adecuados.

Conviene observar que la definición de elipse sigue siendo válida si reducimos los dos focos a un mismo punto C . Lo que obtenemos entonces es el conjunto de puntos P tales que el doble de la distancia de P a C es una constante $2a$, pero esto no es sino la circunferencia de centro C y radio a . En suma, podemos ver a las circunferencias como elipses con distancia focal nula.

En general, una elipse tiene el aspecto de una circunferencia achatada, y el “grado de achatamiento” puede medirse a través de su *excentricidad* $e = c/a$.

Para una elipse propiamente dicha tenemos que $0 < e < 1$, y las circunferencias pueden verse como elipses de excentricidad $e = 0$.

Observemos que $b/a = \sqrt{1 - e^2}$, por lo que la excentricidad determina (y está determinada por) la proporción entre los semiejes de la elipse. Así se ve más claramente que es una medida del “achataamiento” de la elipse.

Fijemos ahora un sistema de referencia ortonormal cuyo origen sea el centro de la elipse y su eje X contenga al eje mayor. Entonces los focos tienen coordenadas $(\pm c, 0)$ y un punto $P = (x, y)$ está en la elipse si cumple que

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

Despejando la primera raíz cuadrada, elevando al cuadrado y simplificando llegamos a que

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

Elevando de nuevo al cuadrado queda

$$x^2 + c^2 + y^2 = a^2 + c^2 x^2 / a^2$$

o, equivalentemente,

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

Por último, dividiendo entre b^2 llegamos a la ecuación

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \quad (4.2)$$

Recíprocamente, vamos a ver que toda ecuación de esta forma con $a > b$ corresponde a una elipse de semiejes a y b y focos en los puntos $(\pm c, 0)$, donde $c = \sqrt{a^2 - b^2}$. Para ello calculamos la distancia d_1 de un punto $P = (x, y)$ que cumpla la ecuación al punto $(-c, 0)$:

$$d_1^2 = (x + c)^2 + y^2 = x^2 + 2cx + c^2 + b^2 - \frac{b^2}{a^2} x^2 = \frac{c^2}{a^2} x^2 + 2cx + a^2 = (a + ex)^2,$$

donde hemos llamado $e = c/a$. De (4.2) se sigue que $|x| \leq a$ y $0 < e < 1$, luego $a + ex > 0$, luego $d_1 = a + ex$. Similarmente se prueba que la distancia de P al vértice $(c, 0)$ es $d_2 = a - ex$, luego, ciertamente, la suma de ambas distancias es constante: $d_1 + d_2 = 2a$. Esto prueba que la ecuación (4.2) corresponde a una elipse con los parámetros indicados.

Del argumento anterior se extrae una consecuencia relevante: llamemos

$$d = a/e = a^2/c > a$$

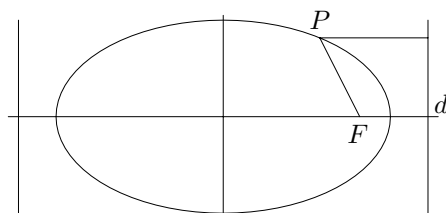
y consideremos las *directrices* de la elipse, que son las rectas perpendiculares a su eje mayor y que distan de su centro una distancia d . El cociente de la distancia de un punto P al foco $(c, 0)$ y a la directriz que pasa por $(d, 0)$ es

$$\frac{a - ex}{d - x} = e \frac{a - ex}{a - ex} = e,$$

e igualmente sucede si consideramos el otro foco y la otra directriz. En resumen:

Teorema 4.33 *El cociente entre la distancia de un punto de una elipse (que no sea una circunferencia) a uno de sus focos sobre la distancia a su directriz correspondiente es igual a la excentricidad de la elipse.*

La figura siguiente muestra las directrices de la elipse anterior junto con las distancias de un punto arbitrario P a un foco y a su directriz. Vemos así que una elipse está completamente determinada por uno de sus focos, su directriz correspondiente y su excentricidad.



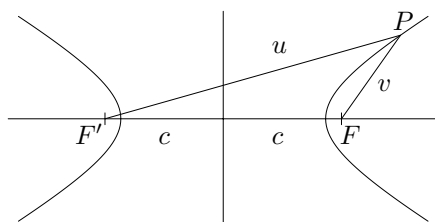
La hipérbola Un mínimo cambio en la definición de elipse nos da la definición de hipérbola:

Definición 4.34 Una *hipérbola* es el conjunto de puntos del plano tales que la diferencia de sus distancias a dos puntos fijos llamados *focos* es constante.

Aquí hemos de entender que la diferencia se toma en valor absoluto o, equivalentemente, que la ecuación de una hipérbola respecto de un sistema de referencia en el que los focos tengan coordenadas $(\pm c, 0)$ es de la forma

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

con $a > 0$. Observemos que si llamamos u y v a las dos distancias y, por ejemplo $u \geq v$, entonces $u - v = 2a$ y, por otro lado, la diferencia entre las longitudes de dos lados de un triángulo ha de ser menor que el tercer lado, luego $2a \leq 2c$.



Si se da la igualdad $a = c$ entonces la “hipérbola” ha de estar contenida en la recta $y = 0$ (y es fácil ver que, de hecho, es toda la recta), así que vamos considerar que la definición de hipérbola excluye este caso trivial. Suponemos, pues, que $a < c$.

Las mismas manipulaciones que hemos hecho con la ecuación de la elipse nos llevan ahora a

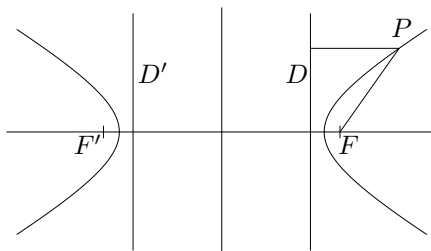
$$\frac{a^2 - c^2}{a^2} x^2 + y^2 = a^2 - c^2.$$

Como $a < c$, podemos llamar $b^2 = c^2 - a^2$, con lo que la ecuación se reduce a la forma

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \quad (4.3)$$

Como en el caso de la elipse, podemos definir el *centro* de una hipérbola como el punto medio de sus focos, sus *diámetros* como los segmentos que unen dos puntos de la hipérbola y pasan por su centro, su *eje mayor* como el diámetro que pasa por sus focos y sus *vértices* como los extremos del eje mayor. Respecto al sistema de referencia correspondiente a la ecuación (4.3) el centro es el punto $(0,0)$ y los vértices son los puntos $(\pm a, 0)$. La hipérbola sólo tiene dos vértices, porque no corta al eje $x = 0$. El número b no tiene una interpretación geométrica, pero determina la *distancia focal* c a través de la relación $c^2 = a^2 + b^2$. Definimos la *excentricidad* como $e = c/a > 1$.

Definimos las *directrices* de una hipérbola igual que para una elipse, es decir, como las rectas perpendiculares a su eje que distan $\pm d$ de su centro, donde $d = a/e$. Dejamos a cargo del lector la demostración de que el teorema 4.33 vale igualmente para elipses. La figura siguiente muestra una hipérbola con sus focos y sus directrices. La distancia de P a F dividida entre la distancia de P a D es igual a la excentricidad e .



Cónicas euclídeas El hecho de que tanto las elipses como las hipérbolas satisfagan el teorema 4.33 nos lleva a la siguiente definición de cónica en el plano euclídeo:

Definición 4.35 Sea D una recta en el plano euclídeo, sea F un punto no contenido en D y sea $e > 0$ un número real. Llamaremos *cónica* de foco F , directriz D y excentricidad e al conjunto de los puntos P tales que la distancia de P a F sobre la distancia de P a D sea igual a e . Una *cónica* será una curva que cumpla esta definición o bien una circunferencia, de modo que, por definición, las circunferencias son las cónicas de excentricidad 0.

Vamos a probar que las cónicas en este sentido son elipses si su excentricidad es $0 < e < 1$ y son hipérbolas si $e > 1$. Para ello tomamos D , F y e en las

condiciones de esta definición con $e \neq 1$. Llamamos $k > 0$ a la distancia de F a D , y definimos $a > 0$ mediante la relación

$$a \left| \frac{1}{e} - e \right| = k.$$

De este modo, si llamamos $c = ae$ y $d = a/e$, se cumple que $|c - d| = a$, luego podemos tomar un sistema de referencia en el que el punto F tenga coordenadas $(c, 0)$ y la recta D tenga ecuación $x = d$. Si $0 < e < 1$ el punto F estará a la izquierda de D , pero si $e > 1$ será al revés.

Los puntos $P = (x, y)$ de la cónica definida por D , F y e son los que cumplen

$$\frac{\sqrt{(x - c)^2 + y^2}}{|d - x|} = e. \quad (4.4)$$

Esta ecuación equivale a la que resulta de elevarla al cuadrado, que es:

$$x^2 - 2cx + c^2 + y^2 = e^2 d^2 + e^2 x^2 - 2e^2 dx,$$

la cual es equivalente a

$$(1 - e^2)x^2 + y^2 = a^2(1 - e^2).$$

Si llamamos $\epsilon = \pm 1$ al signo de $1 - e^2$ y

$$b = a\sqrt{|1 - e^2|} = \sqrt{|a^2 - c^2|},$$

tenemos que

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

lo que equivale a que

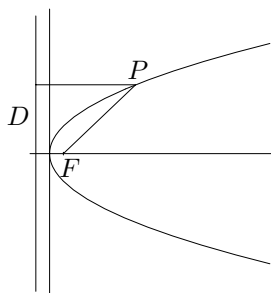
$$\frac{x^2}{a^2} + \epsilon \frac{y^2}{b^2} = 1.$$

Así, si $0 < e < 1$ tenemos que $\epsilon = 1$ y hemos llegado a la ecuación de una elipse, mientras que para $e > 1$ se cumple que $\epsilon = -1$ y la ecuación corresponde a una hipérbola. En ambos casos es claro que e , F y D son respectivamente la excentricidad, un foco y una directriz en el sentido definido específicamente para elipses e hipérbolas, luego la definición 4.35 es coherente con las definiciones previas.

La parábola Las cónicas de excentricidad $e = 1$ no son ni elipses ni hipérbolas, sino que reciben el nombre de *parábolas*. La ecuación más simple de una parábola se tiene cuando se toma un sistema de referencia en el que el foco tenga coordenadas $(p, 0)$ y la directriz tenga ecuación $x = -p$. Entonces un punto $P = (x, y)$ está en la parábola si cumple

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

lo cual equivale a $y^2 = 4px$.



Podemos definir el *eje* de la parábola como la recta perpendicular a D que contiene a F , y su *vértice* es el punto donde la parábola corta al eje. En el sistema de referencia en el que la parábola tiene la ecuación canónica $y^2 = 4px$ tenemos que el eje es $y = 0$ y el vértice es $(0, 0)$, por lo que p es la distancia del foco al vértice (o de la directriz al vértice) y se llama *parámetro* de la parábola.

El teorema siguiente resume la clasificación de las cónicas euclídeas que hemos obtenido:

Teorema 4.36 *Fijado un sistema de referencia ortonormal en un plano euclídeo real, toda cónica es isométrica a una sola de las cónicas siguientes:*

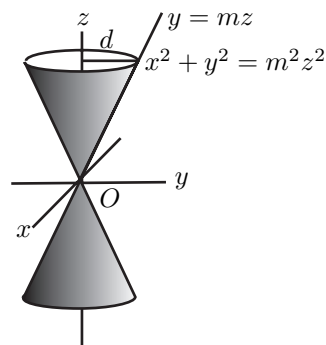
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (\text{con } a \geq b > 0), \quad \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad (\text{con } a, b > 0), \quad y^2 = 4px, \\ (\text{con } p > 0).$$

La unicidad se debe a que las isometrías han de conservar los parámetros intrínsecos de las cónicas, como los semiejes, la distancia focal, etc.

Secciones cónicas Finalmente vamos a probar que las cónicas pueden caracterizarse por su definición clásica, es decir, por que son las curvas que se obtienen al cortar un cono con un plano que no pase por su vértice.

Dados un punto O y una recta r que pase por O , un *cono* (gr. ‘piña’) de *vértice* O y *eje* r es el conjunto de los puntos del espacio que pertenecen a alguna de las rectas que pasan por O y forman un ángulo fijo α con r . Dichas rectas se llaman *generatrices* del cono. En realidad esta definición corresponde a lo que más propiamente se llama un *cono doble*, formado por dos *conos simples* opuestos por el vértice. Las generatrices de los conos simples son las semirrectas de origen en O y contenidas en un mismo semiespacio respecto del plano perpendicular a r que pasa por O , pero aquí vamos a considerar conos dobles.

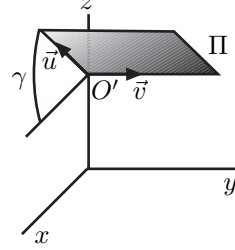
Dado un cono en un espacio afín euclídeo E , podemos tomar un sistema de referencia ortonormal cuyo origen se encuentre en el vértice del cono y cuyo eje Z sea el eje



del cono. Si llamamos d a la distancia al eje Z de un punto del cono situado a altura z , tenemos que $m = \tan \alpha = d/z$, luego los puntos del cono cumplen la ecuación

$$x^2 + y^2 = m^2 z^2$$

Consideremos ahora un plano Π que corte al eje Z en un punto $O' = (0, 0, l)$, con $l > 0$. Supongamos primeramente que Π no es vertical. Entonces su intersección con el plano XY es una recta y , como podemos elegir los ejes X e Y sin alterar la ecuación del cono, no perdemos generalidad si suponemos que dicha intersección es el eje Y . Esto significa que el plano es de la forma



$$\Pi = O' + \langle (1, 0, k), (0, 1, 0) \rangle,$$

para cierto $k \geq 0$. Concretamente, notemos que $k = \tan \gamma$, donde γ es el ángulo que forma Π con el plano XY . Por lo tanto, un punto está en Π si y sólo si cumple la ecuación

$$z = l + kx$$

Podemos considerar a Π como plano afín euclídeo con el producto escalar en $\vec{\Pi}$ inducido desde \vec{E} por restricción. Un sistema de referencia ortonormal de Π es entonces $(O'; \vec{u}, \vec{v})$, donde

$$\vec{u} = \left(\frac{1}{\sqrt{k^2 + 1}}, 0, \frac{k}{\sqrt{k^2 + 1}} \right), \quad \vec{v} = (0, 1, 0).$$

Claramente

$$(x, y, l + kx) = O' + \sqrt{k^2 + 1} x \vec{u} + y \vec{v},$$

lo que se interpreta como que las coordenadas en Π del punto que en E tiene coordenadas (x, y, z) son

$$(x', y') = (\sqrt{k^2 + 1} x, y).$$

Los puntos de la intersección de Π con el cono son los que están en Π y además cumplen la ecuación

$$x^2 + y^2 = m^2 (kx + l)^2 \quad \Leftrightarrow \quad (1 - m^2 k^2) x^2 + y^2 = 2m^2 k l x + m^2 l^2.$$

En términos de las coordenadas en Π esto equivale a

$$\frac{1 - m^2 k^2}{k^2 + 1} x'^2 + y'^2 = \frac{2m^2 k l}{\sqrt{k^2 + 1}} x' + m^2 l^2. \quad (4.5)$$

Vamos a probar que esta ecuación corresponde a una cónica en Π . Supongamos en primer lugar que $mk \neq 1$, con lo que el coeficiente de x'^2 no es nulo. Realizamos entonces la traslación

$$x' = x + \frac{\sqrt{k^2 + 1} k m^2 l}{1 - m^2 k^2}, \quad y' = y,$$

con lo que la ecuación se transforma en

$$\frac{1 - m^2 k^2}{k^2 + 1} x^2 + y^2 = \frac{m^2 l^2}{1 - m^2 k^2}$$

Si $mk < 1$ todos los coeficientes son positivos, por lo que tenemos una elipse. Concretamente, dividiendo entre el término independiente, vemos que sus semiejes son

$$a = \frac{m\sqrt{k^2 + 1}}{1 - m^2 k^2} l, \quad b = \frac{m}{\sqrt{1 - m^2 k^2}} l,$$

luego

$$c = \sqrt{a^2 - b^2} = \frac{\sqrt{1 + m^2} mk}{1 - m^2 k^2} l.$$

La excentricidad resulta ser

$$e = \frac{\sqrt{1 + m^2} k}{\sqrt{k^2 + 1}}. \quad (4.6)$$

Si $mk > 1$ tenemos una hipérbola, y cálculos análogos llevan a expresiones similares para a , b , c (el único cambio es que aparece $m^2 k^2 - 1$ donde antes teníamos $1 - m^2 k^2$) que dan lugar a la misma expresión para e . Por último, si $mk = 1$ la ecuación (4.5) se reduce a

$$x'^2 - \frac{2m^2 l}{\sqrt{m^2 + 1}} y' = m^2 l^2,$$

y una traslación en y la reduce a

$$x'^2 = \frac{2m^2 l}{\sqrt{m^2 + 1}} y',$$

que corresponde a una parábola. Notemos que la fórmula (4.6) da $e = 1$ cuando $k = 1/m$, luego es válida en todos los casos.

Notemos a continuación que $\sqrt{m^2 + 1} = \sqrt{\tan^2 \alpha + 1} = \sec \alpha$ e, igualmente⁷ $\sqrt{k^2 + 1} = \sec \gamma$, por lo que (4.6) equivale a

$$e = \frac{\cos \gamma \tan \gamma}{\cos \alpha} = \frac{\sin \gamma}{\cos \alpha}.$$

Si llamamos β al ángulo que Π forma con el eje Z se cumple que $\sin \gamma = \cos \beta$, y así obtenemos una expresión más simple para la excentricidad:

$$e = \frac{\cos \beta}{\cos \alpha}. \quad (4.7)$$

Así, la cónica es una elipse (gr. ‘defecto’), una parábola (gr. ‘comparación’) o una hipérbola (gr. ‘exceso’) según si $\beta < \alpha$, $\beta = \alpha$ o $\beta > \alpha$.

Ejercicio: Comprobar que si el plano Π es vertical (y no pasa por O) corta al cono en una hipérbola de excentricidad dada igualmente por (4.7) con $\beta = 0$.

■

⁷Esta relación surge de dividir entre $\cos^2 \alpha$ la identidad $\sin^2 \alpha + \cos^2 \alpha = 1$.

Capítulo V

Números complejos

Es bien conocido que las raíces de una ecuación polinómica de segundo grado con coeficientes reales $ax^2 + bx + c = 0$ vienen dadas por la fórmula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

supuesto que el discriminante $b^2 - 4ac$ tenga raíz cuadrada, esto es, no sea negativo. Sin embargo, los matemáticos renacentistas italianos observaron que aun en el caso de que el discriminante fuera negativo se podía trabajar consistentemente con unas hipotéticas raíces “imaginarias” sin más que admitir la existencia de una cantidad $i = \sqrt{-1}$. Más concretamente, los algoritmos de resolución de ecuaciones cúbicas pasaban en determinados casos por soluciones imaginarias de ecuaciones cuadráticas asociadas que después daban lugar a soluciones reales de la cúbica de partida. La teoría de extensiones algebraicas proporciona un fundamento riguroso de esta introducción directa de los números imaginarios a partir de la ecuación $i^2 = -1$:

5.1 Definición y propiedades básicas

El polinomio $x^2 + 1$ no tiene raíces reales, y al tener grado 2 es irreducible en el anillo $\mathbb{R}[x]$. Por lo tanto existe una extensión algebraica de \mathbb{R} en la cual tiene una raíz:

Definición 5.1 Llamaremos *cuerpo de los números complejos* $\mathbb{C} = \mathbb{R}[i]$ a la adjunción a \mathbb{R} de una raíz i del polinomio $x^2 + 1$.

La teoría de cuerpos nos asegura que \mathbb{C} está unívocamente determinado salvo isomorfismos que fijan a \mathbb{R} . El cuerpo \mathbb{C} es un espacio vectorial de dimensión 2 sobre \mathbb{R} y tiene por base a $\{1, i\}$. Por lo tanto todo número complejo z se expresa de forma única como $z = a + bi$, donde $a, b \in \mathbb{R}$. Así pues, podemos identificar el número complejo $a + bi$ con el par $(a, b) \in \mathbb{R}^2$.

El nombre de “números complejos” hace referencia precisamente a que cada número complejo $z = a + bi$ está compuesto o determinado por dos números reales a y b , llamados respectivamente *parte real* y *parte imaginaria* de z . Las representaremos por $\operatorname{Re} z$ e $\operatorname{Im} z$.

El adjetivo “imaginario” es herencia de las vacilaciones originales sobre la naturaleza de los números complejos. En general, los números $a + bi$ con $b \neq 0$ se llaman *imaginarios* (y de aquí viene, por oposición, el calificativo de números reales para los números no imaginarios) los números de la forma bi con $b \neq 0$ se llaman *imaginarios puros*. Al número i se le llama *unidad imaginaria*.

Teniendo en cuenta que $i^2 = -1$ las operaciones en \mathbb{C} son

$$(a + bi) + (c + di) = (a + c) + (b + d)i, \quad (a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

Vemos, pues, que la suma de números complejos es la suma usual en \mathbb{R}^2 , y el producto de un número real por un número complejo es también el producto usual. Queda pendiente interpretar geométricamente el producto de dos números complejos cualesquiera.

Puesto que $x^2 + 1 = (x + i)(x - i)$, la extensión \mathbb{C}/\mathbb{R} es finita de Galois, luego tiene exactamente dos \mathbb{R} -automorfismos. Uno es la identidad, y el otro envía i a $-i$. Lo llamaremos *conjugación*. Representaremos por \bar{z} al conjugado de z . Según lo dicho,

$$\overline{a + bi} = a - bi.$$

La norma euclídea de \mathbb{R}^2 se corresponde con lo que en la teoría de números complejos se llama el *módulo* de un número complejo $z = a + bi$, definido como

$$|z| = \sqrt{a^2 + b^2} = \sqrt{z\bar{z}}.$$

Observemos que $|z|^2$ es la norma de z en el sentido de la teoría de Galois, luego el módulo es multiplicativo, es decir, $|z_1 z_2| = |z_1| |z_2|$.

La relación $|z|^2 = z\bar{z}$ nos da una expresión sencilla para el inverso de un número complejo no nulo, a saber,

$$z^{-1} = \frac{\bar{z}}{|z|^2}.$$

También conviene notar que $|\operatorname{Re} z| \leq |z|$, $|\operatorname{Im} z| \leq |z|$. En efecto, si $z = a + bi$, tenemos que $|\operatorname{Re} z|^2 = a^2 \leq a^2 + b^2 = |z|^2$, luego $|\operatorname{Re} z| \leq |z|$, e igualmente con la parte imaginaria.

Las partes real e imaginaria de un número z son las coordenadas cartesianas de z respecto al sistema de referencia canónico determinado por el 0 y la base $\{1, i\}$. También podemos expresar todo número complejo no nulo en términos de sus coordenadas polares, es decir,

$$z = |z|(\cos \theta + i \operatorname{sen} \theta),$$

con el argumento θ determinado módulo 2π .

Retomemos ahora el problema de interpretar el producto de números complejos. Consideremos en primer lugar un número complejo de módulo 1, digamos $1_\theta = \cos \theta + i \sin \theta$. Consideremos otro número complejo $x + iy$. Entonces

$$1_\theta(x + iy) = (x \cos \theta - y \sin \theta) + (x \sin \theta + y \cos \theta)i.$$

Si lo interpretamos en \mathbb{R}^2 tenemos

$$1_\theta(x, y) = (x, y) \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} = G_\theta(x, y),$$

donde G_θ es el giro de ángulo θ respecto al sistema de referencia canónico. Así pues, si expresamos $x + yi$ en coordenadas polares, $\rho'_{\theta'}$, acabamos de probar que $1_\theta \rho'_{\theta'} = \rho'_{\theta+\theta'}$.

Teniendo en cuenta que el módulo del producto es el producto de los módulos, ahora es claro que

$$\rho_\theta \rho'_{\theta'} = \rho \cdot 1_\theta \rho'_{\theta'} = \rho \rho'_{\theta+\theta'} = (\rho \rho')_{\theta+\theta'}.$$

Así pues, el producto de dos números complejos no nulos es el número complejo cuyo módulo es el producto de los módulos y cuyo argumento es la suma de los argumentos.

Ahora es evidente que todo número complejo tiene una raíz cuadrada. Concretamente, la raíz cuadrada de ρ_θ es $\sqrt{\rho_{\theta/2}}$. Vamos a dar una prueba de este hecho que no dependa de consideraciones geométricas sobre argumentos.

Partamos de un número complejo z que podemos suponer imaginario. Sea $z/|z| = a + bi$. La bisectriz del ángulo formado por 1 y $a + bi$ pasará por el punto medio del segmento de extremos 1 y $a + bi$. Éste es $(a + 1)/2 + bi/2$. Tenemos así un número cuyo argumento es la mitad del argumento de z . Ahora hemos de ajustar su módulo, que es

$$\sqrt{\frac{a+1}{2}}$$

Dividiéndolo entre este número obtenemos un número de módulo 1 y multiplicando por $\sqrt{|z|}$ obtenemos, según lo dicho anteriormente, una raíz cuadrada de z . Ésta es

$$w = \sqrt{\frac{2|z|}{a+1}} \left(\frac{a+1}{2} + \frac{b}{2}i \right). \quad (5.1)$$

Elevando al cuadrado y usando que $a^2 + b^2 = 1$ obtenemos $|z|(a + bi) = z$, luego se trata ciertamente de una raíz cuadrada de z .

5.2 La clausura algebraica de \mathbb{C}

El hecho de que todo número complejo tenga raíces cuadradas se traduce en que toda ecuación de segundo grado tiene raíces en \mathbb{C} . Los algebristas sabían bien que lo mismo sucede con las ecuaciones de grado tres y cuatro, así como

con muchos casos particulares de ecuaciones de grados superiores. Esto sustentó durante siglos la conjetura de que de hecho toda ecuación polinómica podía resolverse en \mathbb{C} , sin embargo, la primera prueba tuvo que esperar hasta principios del siglo XIX. Fue en este período cuando Argrand descubrió la representación geométrica de los números complejos, aunque fue Gauss el primero en usarla y difundirla. Gauss descubrió así mismo la interpretación geométrica del producto de números complejos y en su tesis doctoral probó el que hoy se conoce como teorema fundamental del álgebra. La prueba que veremos aquí se basa en la teoría de Galois y sólo requiere dos hechos concretos sobre números reales y complejos, ambos ya probados:

A Todo polinomio de grado impar con coeficientes reales tiene al menos una raíz real.

B Todo número complejo tiene una raíz cuadrada.

El último ingrediente que necesitamos es un resultado fundamental de la teoría de grupos finitos:

Teorema de Sylow Sea G un grupo finito, p un número primo y n un número natural tal que $p^n \mid |G|$. Entonces G tiene un subgrupo de orden p^n .

Para comodidad del lector incluimos al final del capítulo un apéndice con la demostración del teorema de Sylow. Para probar el teorema fundamental del álgebra conviene reformular A y B en términos de la teoría de cuerpos. En primer lugar, A equivale a que todo polinomio irreducible en $\mathbb{R}[x]$ de grado mayor que 1 es de grado par. Como toda extensión finita de \mathbb{R} es simple, su grado es el grado del polinomio mínimo en $\mathbb{R}[x]$ de su elemento primitivo, luego

A' Todas las extensiones propias de \mathbb{R} tienen grado par.

Por otro lado B equivale a que toda ecuación de segundo grado sobre \mathbb{C} tiene raíces en \mathbb{C} , es decir, que no hay polinomios irreducibles de grado 2 en $\mathbb{C}[x]$, así pues:

B' No existen extensiones de \mathbb{C} de grado 2.

Teorema 5.2 (Teorema fundamental del álgebra) \mathbb{C} es algebraicamente cerrado.

DEMOSTRACIÓN: Hay que probar que \mathbb{C} no tiene extensiones algebraicas propias. Supongamos que K es una tal extensión. Adjuntando a \mathbb{C} un elemento cualquiera de K que no esté en \mathbb{C} obtenemos una extensión propia y finita de \mathbb{C} . Podemos suponer, pues, que K/\mathbb{C} es una extensión finita. También será finita la extensión K/\mathbb{R} . Existe una extensión finita normal de \mathbb{R} que contiene a K . Por lo tanto podemos suponer que K/\mathbb{R} es una extensión finita de Galois (de grado mayor que 1).

Por A sabemos que $|K : \mathbb{R}|$ es un número par. Digamos que $|K : \mathbb{R}| = 2^n \cdot m$, con $n \geq 1$ y m impar.

Sea G el grupo de Galois de la extensión. Como $|G| = 2^n \cdot m$, el teorema de Sylow nos da la existencia de un subgrupo H de orden 2^n . El cuerpo F fijado por H cumple $|K : F| = 2^n$, luego $|F : \mathbb{R}| = m$. Por A' resulta que m ha de ser igual a 1. Así pues, $|K : \mathbb{R}| = 2^n$ y, por lo tanto, $|K : \mathbb{C}| = 2^{n-1}$. Sea $r = n - 1$. Estamos suponiendo que K es una extensión propia de \mathbb{C} , luego $r \geq 1$.

Sea ahora L el grupo de Galois de la extensión K/\mathbb{C} . Entonces $|L| = 2^r$ y, de nuevo por el teorema de Sylow, L tiene un subgrupo H de orden 2^{r-1} . El cuerpo fijado por este subgrupo es una extensión de \mathbb{C} de grado 2, en contradicción con la afirmación B' . ■

Como consecuencia, el conjunto \mathbb{A} de los números complejos algebraicos sobre \mathbb{Q} es algebraicamente cerrado (pues las raíces complejas de cualquier polinomio de $\mathbb{A}[x]$ serán elementos algebraicos sobre \mathbb{A} , luego sobre \mathbb{Q} , luego estarán en \mathbb{A}). Así pues, \mathbb{A} es la clausura algebraica de \mathbb{Q} , luego podemos considerar contenidas en \mathbb{C} a todas las extensiones algebraicas de \mathbb{Q} . Otra consecuencia es que \mathbb{C} es la clausura algebraica de \mathbb{R} , por lo que los polinomios irreducibles en $\mathbb{R}[x]$ tienen todos grado 1 o 2. Esto será importante en la clasificación de las isometrías en \mathbb{R}^n .

5.3 Construcciones con regla y compás

Veamos cómo la estructura de cuerpo que le hemos dado al plano euclídeo resulta de gran ayuda para estudiar la constructibilidad con regla y compás. Los antiguos griegos se interesaron especialmente por los problemas geométricos asociados a figuras que pueden ser trazadas con los instrumentos más simples: una regla para trazar rectas y un compás para trazar circunferencias. Con más precisión, una construcción con regla y compás en sentido estricto no admite que se tracen rectas o circunferencias aleatorias. Para que una recta se pueda considerar ‘construida’ es necesario tener contruidos dos de sus puntos, sobre los que apoyar la regla; un punto está construido cuando lo hemos determinado como intersección de dos rectas, dos circunferencias, o una recta y una circunferencia; finalmente, para construir una circunferencia tenemos que tener construido su centro, sobre el que clavar el compás, y su radio ha de ser la distancia entre dos puntos contruidos, con los que fijar la apertura del compás.

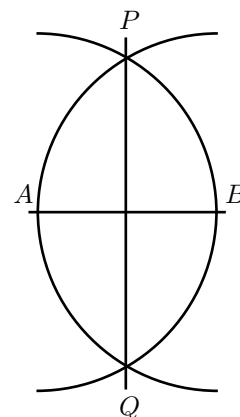
Para aplicar estos criterios necesitamos tener unos datos, por ejemplo, dado un triángulo arbitrario (trazado al azar, si se quiere) podemos plantearnos la construcción de su circunferencia circunscrita. Esto supone tomar como ‘construidos’ los vértices del triángulo y a partir de ellos realizar una cadena de construcciones en el sentido anterior que acaben con el trazado de la circunferencia buscada.

Es claro que a partir de un único punto es imposible realizar construcción alguna. El menor número de puntos para iniciar una construcción es 2. Una construcción a partir de dos puntos es una construcción *absoluta*. Una construcción que acepte como datos más de dos puntos es una construcción *relativa* a los datos. Veamos algunos ejemplos de construcciones elementales:

Mediatriz de un segmento dado Tomando como datos los extremos de un segmento \overline{AB} , para construir su mediatriz procedemos del modo siguiente: trazamos las circunferencias de centros A y B y radio \overline{AB} . Éstas se cortarán en dos puntos P y Q equidistantes de A y B . La recta que pasa por P y Q es la mediatriz buscada. ■

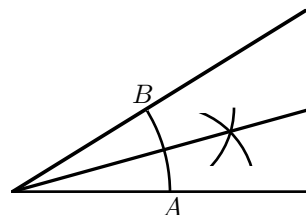
Perpendicular a una recta Para trazar la perpendicular a una recta por un punto O dado sobre ella se traza una circunferencia de centro O y cualquier radio constructible, que cortará a la recta en dos puntos A y B . La perpendicular buscada es la mediatriz del segmento \overline{AB} .

Para obtener la perpendicular a una recta por un punto exterior P se toma un punto A en la recta y se traza la circunferencia de centro P y radio PA , que cortará a la recta en otro punto B . La perpendicular buscada es la mediatriz del segmento \overline{AB} . ■

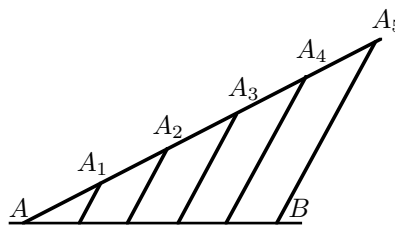


Paralela a una recta Para trazar la paralela a una recta r por un punto exterior P se traza la perpendicular a r por P , llamémosla s , y luego la perpendicular a s por P .

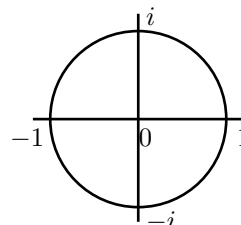
Bisectriz de un ángulo Para trazar la bisectriz de un ángulo dado se traza una circunferencia con centro en su vértice y cualquier radio constructible. Ésta cortará a los lados en dos puntos A y B . La bisectriz buscada es la mediatriz de \overline{AB} . ■



División de un segmento Para dividir un segmento \overline{AB} en n partes iguales trazamos cualquier recta constructible que pase por A y que sea no contenga a B . Con el compás marcamos n puntos equidistantes $A = A_0, A_1, \dots, A_n$. Unimos A_n con B y trazamos paralelas a esta recta que pasen por los puntos A_i . Por el teorema de Tales, los cortes de estas rectas con \overline{AB} dividen el segmento en partes iguales. ■



Consideremos un sistema de referencia ortonormal en un plano euclídeo. Cada punto del plano se corresponde a través de este sistema con un único número complejo. Podemos elegir arbitrariamente los puntos que se corresponden con 0 y 1. Trazamos la recta que contiene a ambos y la perpendicular a ella por 0. La circunferencia



de centro 0 y radio 1 corta a estas rectas en los puntos ± 1 y $\pm i$. Tenemos libertad para decidir quién es i y quién $-i$. Una vez hecha esta elección, cada punto del plano tiene asignado un único número complejo. Llamaremos *números constructibles* a los números complejos asociados a puntos del plano constructibles con regla y compás (únicamente a partir de 0 y 1). Llamaremos \mathbb{C} al conjunto de todos los números complejos constructibles.

Notemos que si un número real a es constructible, también lo es la circunferencia de centro 0 y radio a , y ésta corta al eje imaginario en ai , luego también éste es constructible. En general, si $z = a + bi$ es constructible, entonces también lo son las rectas que pasan por z perpendiculares a los ejes, luego también los puntos de intersección con los ejes, luego los números a y bi son constructibles, luego a y b son constructibles. El recíproco se prueba igualmente. Por consiguiente tenemos:

Un número complejo es constructible si y sólo si lo son su parte real y su parte imaginaria.

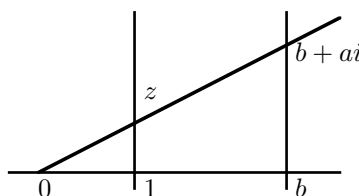
Ahora podemos probar el resultado básico sobre constructibilidad:

Teorema 5.3 *El conjunto \mathbb{C} de los números complejos constructibles es un subcuerpo de \mathbb{C} .*

DEMOSTRACIÓN: Veamos en primer lugar que $\mathbb{C} \cap \mathbb{R}$ es un subcuerpo de \mathbb{R} .

Sean $a, b \in \mathbb{C} \cap \mathbb{R}$. Si $b = 0$, entonces $a - b = a \in \mathbb{C} \cap \mathbb{R}$. En otro caso, la circunferencia de centro a y radio $|b|$ es constructible y corta al eje real en $a + b$ y $a - b$, luego $a - b \in \mathbb{C} \cap \mathbb{R}$. Esto prueba que $\mathbb{C} \cap \mathbb{R}$ es un subgrupo de \mathbb{R} .

Ahora supongamos que $b \neq 0$ y veamos que $ab^{-1} \in \mathbb{C} \cap \mathbb{R}$, lo que probará que es un subcuerpo. Podemos suponer que $a, b > 0$. Usando las construcciones que hemos visto antes es claro que todos los puntos marcados en la figura son constructibles. El teorema de Tales nos da que la distancia de z a 1 es ab^{-1} .



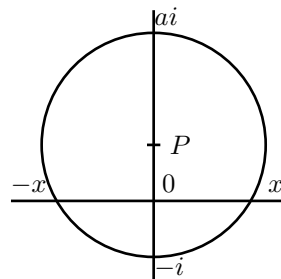
Dados z_1 y z_2 en \mathbb{C} , tenemos que sus partes reales e imaginarias son constructibles, de donde se sigue fácilmente, aplicando la parte ya probada, que las partes reales e imaginarias de $z_1 - z_2$ y $z_1 z_2^{-1}$ (si $z_2 \neq 0$) son constructibles, luego también lo son estos números y el teorema queda probado. ■

En particular todos los números racionales son constructibles, pues \mathbb{Q} es el menor cuerpo contenido en \mathbb{C} . Es fácil dar una prueba directa de la constructibilidad de los números racionales. A su vez esto legitima el uso de reglas graduadas. En principio hemos considerado a una regla como un instrumento para trazar rectas por dos puntos dados. Ahora sabemos que si marcamos una graduación (racional) sobre la regla y la usamos para construir segmentos de una longitud racional determinada, esto no aumenta nuestra capacidad de construcción.

Vamos a dar una caracterización puramente algebraica de los números complejos constructibles. Para ello necesitamos un último resultado auxiliar:

Las raíces cuadradas de los números constructibles son constructibles.

Probamos primero que si $a \in \mathbb{C} \cap \mathbb{R}$ y $a > 0$, entonces $\sqrt{a} \in \mathbb{C} \cap \mathbb{R}$. Sea P el punto medio entre ai y $-i$ (se obtiene de la bisectriz del segmento). La circunferencia de centro P que pasa por $-i$ corta al eje real en dos puntos $\pm x$. La potencia de 0 respecto a esta circunferencia (en módulo) es por una parte x^2 y por otra parte a , luego $x = \sqrt{a}$.



La raíz cuadrada de un número constructible arbitrario no nulo viene dada por la fórmula 5.1 y es claramente constructible. ■

Teorema 5.4 *Un número complejo z es constructible si y sólo si existe una cadena de cuerpos*

$$\mathbb{Q} = F_0 \subset F_1 \subset \cdots \subset F_n \subset \mathbb{C}$$

de modo que $|F_{j+1} : F_j| = 2$ y $z \in F_n$.

DEMOSTRACIÓN: Si z es constructible podemos considerar la sucesión de puntos intermedios que se trazan en una construcción de z :

$$z_0 = 0, z_1 = 1, z_2, \dots, z_n = z.$$

Así, cada z_j se construye directamente a partir de los puntos anteriores, bien como intersección de dos rectas que pasan por dos de los puntos anteriores, bien como intersección de dos circunferencias con centros en dos de los puntos anteriores y radios iguales a las distancias entre dos puntos anteriores o bien como intersección de una recta y una circunferencia en estas mismas condiciones.

Vamos a probar que existe una sucesión de cuerpos

$$\mathbb{Q} = K_0 \subset K_1 \subset \cdots \subset K_n \subset \mathbb{R}$$

tales que $|K_{i+1} : K_j| \leq 2$ y las partes reales e imaginarias de cada z_j están en K_j . Lo demostramos por inducción. Todo se reduce a probar que si un cuerpo K contiene las partes reales e imaginarias de los puntos anteriores a z_j entonces las partes real e imaginaria de z_j están en K o en una extensión cuadrática de K .

Sea $z_j = (x, y)$. Supongamos que z_j es la intersección de dos rectas que pasan por dos puntos de entre los anteriores, en particular dos puntos con coordenadas en K . La recta que pasa por dos puntos (a, b) y (c, d) en K^2 tiene vector director $\vec{v} = (c - a, d - b)$, luego está formada por los puntos (x, y) tales que $(x - a, y - b)$ es múltiplo de \vec{v} o, equivalentemente, tales que

$$\begin{vmatrix} x - a & y - b \\ c - a & d - b \end{vmatrix} = 0.$$

Al desarrollar obtenemos una ecuación $Ax + By = C$ con coeficientes en K . Tenemos, pues, que $z_j = (x, y)$ es la solución única de dos ecuaciones de este

tipo, y al resolver explícitamente el sistema obtenemos una expresión de x e y en función de los coeficientes, que prueba que ambos están en K .

Supongamos ahora que z_j está en la intersección de una recta que pasa por dos puntos de K^2 y una circunferencia con centro en K^2 y radio r igual a la distancia entre dos puntos de K^2 . Entonces $r^2 \in K$, luego tenemos que (x, y) satisface las ecuaciones

$$\left. \begin{aligned} Ax + By &= C \\ x^2 + y^2 + ax + by + c &= 0 \end{aligned} \right\}$$

donde todos los coeficientes están en K . Uno de los coeficientes A o B ha de ser no nulo. Al despejar la variable correspondiente en la primera ecuación y sustituirla en la segunda obtenemos una ecuación de segundo grado con coeficientes en K que ha de tener entre sus soluciones una de las coordenadas de z_j , luego ésta se encuentra en una extensión de $K(\sqrt{\alpha})$ de K , donde α es el discriminante de la ecuación, que está en K . La otra coordenada se despeja de la primera ecuación, luego está también en $K(\sqrt{\alpha})$. Claramente $|K(\sqrt{\alpha}) : K| \leq 2$.

Por último suponemos que z_j está en la intersección de dos circunferencias constructibles a partir de los puntos anteriores. Por el mismo razonamiento de antes, sus coordenadas cumplen dos ecuaciones de la forma

$$\left. \begin{aligned} x^2 + y^2 + ax + by + c &= 0 \\ x^2 + y^2 + a'x + b'y + c' &= 0 \end{aligned} \right\}$$

con los coeficientes en K . Restando ambas ecuaciones obtenemos la ecuación

$$(a - a')x + (b - b')y + c - c' = 0,$$

que también es satisfecha por las coordenadas de z_j y cuyos coeficientes no pueden ser todos nulos, o si no las dos circunferencias serían la misma. En definitiva tenemos que z_j está también en una recta cuya ecuación tiene sus coeficientes en K y este caso se reduce al anterior.

Tenemos, pues, la sucesión K_j en las condiciones pedidas. Ahora observamos que $z_j \in K_j(i)$ y la sucesión

$$\mathbb{Q} \subset \mathbb{Q}(i) \subset K_1(i) \subset \cdots \subset K_n(i)$$

cumple las condiciones del teorema.

Recíprocamente, si z cumple las condiciones del enunciado, entonces es claro que $F_{j+1} = F_j(\sqrt{\alpha_j})$ para un cierto $\alpha_j \in F_j$, y si $F_j \subset \mathbb{C}$, entonces $\sqrt{\alpha_j} \in \mathbb{C}$, luego $F_{j+1} \subset \mathbb{C}$. Así pues, $z \in F_n$ es constructible. ■

Esto nos da una condición necesaria sencilla para que un número complejo sea constructible. Observar que es muy restrictiva.

Teorema 5.5 *Si z es un número complejo constructible entonces es algebraico y $|\mathbb{Q}(z) : \mathbb{Q}|$ es potencia de 2.*

Esta condición no es en general necesaria, pero tenemos un recíproco parcial:

Teorema 5.6 *Si K es una extensión finita de Galois de \mathbb{Q} , entonces $K \subset \mathbb{C}$ si y sólo si $|K : \mathbb{Q}|$ es potencia de 2.*

Ejemplo: El problema delíaco La ciudad de Delos era una de las más importantes de la Grecia clásica. Destacaba en ella el templo de Apolo, con la colosal estatua del dios, ofrenda de Naxos, que los griegos tenían por el más excelso de sus santuarios. El oráculo de Delos era uno de los más reputados, considerado poco menos que infalible. Cuenta la leyenda que una gran plaga azotó la ciudad, y sus habitantes consultaron el oráculo, el cual les ordenó duplicar el altar de Apolo. Éste tenía forma cúbica, y así construyeron otro altar más lujoso con el doble de arista. Como la plaga no arreciaba, volvieron a consultar el oráculo, que esta vez especificó que debían construir un altar con el doble de volumen. Si tomamos como unidad la arista del altar original, el problema consistía en construir un cubo de volumen 2, luego con arista $\sqrt[3]{2}$. Los arquitectos de Delos no supieron cómo hacer los cálculos y llamaron en su ayuda a los geómetras más famosos. Desde entonces muchos geómetras han trabajado en el problema. Euclides probó que $\sqrt[3]{2}$ es un número irracional, Apolonio, Herón, Papos, Huygens, Fermat y Descartes propusieron soluciones, y en 1837 Wantzel probó que $\sqrt[3]{2}$ no es constructible con regla y compás. Para nosotros es evidente, pues el cuerpo $\mathbb{Q}(\sqrt[3]{2})$ tiene grado 3 sobre \mathbb{Q} .

5.4 Polígonos regulares

Es especialmente interesante la teoría en torno a la construcción de polígonos regulares. Para evitar una digresión sobre polígonos en general daremos una definición fuerte.

Definición 5.7 Diremos que n puntos P_1, \dots, P_n , con $n \geq 3$ son los vértices de un *polígono regular* (gr. ‘muchos ángulos’) de n lados si están contenidos en una circunferencia y los arcos menores $\widehat{P_1P_2}, \widehat{P_2P_3}, \dots, \widehat{P_nP_1}$ son todos disjuntos y tienen amplitud $2\pi/n$. Los lados del polígono son los segmentos $\overline{P_1P_2}, \overline{P_2P_3}, \dots, \overline{P_nP_1}$, el polígono con los vértices indicados es la intersección de todos los semiplanos respecto a las rectas P_1P_2, \dots, P_nP_1 que contienen a los vértices restantes. La circunferencia que contiene a los vértices de un polígono regular se llama *circunferencia circunscrita* al polígono (también se dice que el polígono está *inscrito* en la circunferencia). El centro y el radio de la circunferencia se llaman también centro y radio del polígono.

Es fácil probar que fijado un punto en una circunferencia existe un único polígono regular inscrito en ella con un vértice en el punto dado. Los polígonos regulares de tres lados son simplemente los triángulos equiláteros, los de cuatro lados son los cuadrados, los siguientes reciben el nombre de pentágono regular, hexágono regular, heptágono regular, etc. También es fácil ver que dos polígonos regulares son congruentes si y sólo si tienen el mismo radio y el mismo número de lados.

Fijada una circunferencia ω y un punto en ella, podemos tomar como 0 el centro y como 1 el punto dado, de modo que los vértices del n -ágono regular serán los números complejos de módulo 1 y argumento $2k\pi/n$, para $n = 0, \dots, n-1$.

Construir dicho polígono es, por definición, construir estos vértices, pero en realidad basta construir $\zeta_n = 1_{2\pi/n}$ pues, una vez construido éste, la circunferencia de centro ζ_n que pasa por 1 corta a ω en $1_{4\pi/n}$, la circunferencia de centro $1_{4\pi/n}$ y que pasa por $1_{2\pi/n}$ corta a ω en $1_{6\pi/n}$, y de este modo se construyen los vértices restantes.

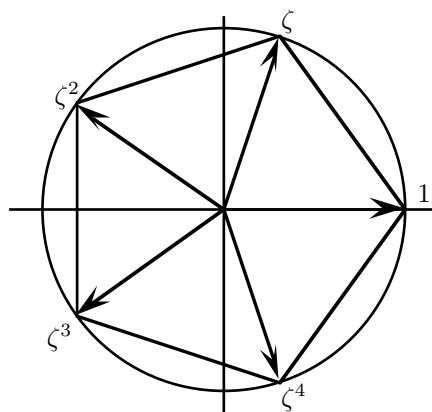
Así pues, un polígono regular de n vértices es constructible si y sólo si lo es el número complejo

$$\zeta_n = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}.$$

Estos números tienen una caracterización algebraica muy simple. Notemos que

$$\zeta_n^k = 1_{2k\pi/n} = 1_{2k\pi/n} = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n},$$

luego los números ζ_n^k para $k = 0, \dots, n-1$ son los vértices del n -ágono regular y $\zeta_n^n = 1$.



En términos algebraicos, ζ_n es una raíz n -sima primitiva de la unidad, el cuerpo $\mathbb{Q}(\zeta_n)$ es el cuerpo ciclotómico n -simo (gr. = ‘que divide el círculo’), que es una extensión de Galois de grado $\phi(n)$, donde ϕ es la función de Euler.¹ El teorema 5.6 nos da la caracterización siguiente:

Teorema 5.8 *Los polígonos regulares de n vértices son constructibles con regla y compás si y sólo si $\phi(n)$ es potencia de 2.*

Notemos que la constructibilidad de los polígonos regulares de n vértices equivale a la constructibilidad de los ángulos de amplitud $2\pi/n$. Esto nos lleva a un interesante resultado:

¹ $\phi(n)$ es en número de números naturales menores que n y primos con n .

Ejemplo: la trisección del ángulo Otro problema clásico es la división de un ángulo dado en tres partes iguales. Ahora es fácil ver que no existe una solución general de este problema pues, por ejemplo, el ángulo de amplitud $2\pi/6$ es constructible, ya que $\phi(6) = 2$, pero su tercera parte, es decir, el ángulo de amplitud $2\pi/18$ no lo es, ya que $\phi(18) = 6$.

Ejercicio: Describir las construcciones de los polígonos regulares de 3, 4 y 6 lados.

Vamos a dar una caracterización aritmética de cuándo $\phi(n)$ es potencia de 2. Factoricemos $n = 2^{r_0} p_1^{r_1} \cdots p_k^{r_k}$, donde p_1, \dots, p_k son primos impares distintos. Entonces $\phi(n) = (p_1 - 1)p_1^{r_1-1} \cdots (p_k - 1)p_k^{r_k-1}$ y la única posibilidad para que este número sea una potencia de dos es que todos los exponentes sean iguales a 1 y que $p_i - 1$ sea potencia de 2 para todo i .

Definición 5.9 Los números primos de la forma $2^n + 1$ se llaman *primos de Fermat*.

En estos términos:

Teorema 5.10 *El polígono regular de n vértices es constructible con regla y compás si y sólo si n es producto de una potencia de 2 y de primos de Fermat con exponente 1.*

Sólo nos queda estudiar qué números de la forma $2^n + 1$ son primos. Construyamos una tabla:

n	1	2	3	4	5	6	7	8	9	10
$2^n + 1$	3	5	9	17	33	65	129	257	513	1.025

Resultan primos los números correspondientes a los valores $n = 1, 2, 4, 8, \dots$. Fermat conjeturó basándose en esto que los primos que hoy llevan su nombre son exactamente los números de la forma $2^{2^n} + 1$.

Ciertamente todo primo de Fermat es de esta forma: sea $p = 2^n + 1$ un primo de Fermat. Tomamos clases módulo p . Entonces el orden de $[2]$ divide a $p - 1 = 2^n$, luego es potencia de 2. Sea $n = 2^u v$, con v impar. Entonces $2^{2^u v} \equiv -1 \pmod{p}$, luego $2^{2^{u+1} v} \equiv 1 \pmod{p}$. Así, el orden de $[2]$ divide a $2^{u+1} v$ y es potencia de 2, luego de hecho divide a 2^{u+1} . Sea, pues $d \leq u + 1$ tal que $[2]^{2^d} = [1]$ pero $[2]^{2^{d-1}} \neq 1$. Como el cuadrado de esta clase es $[1]$, ha de ser $[2]^{2^{d-1}} = [-1]$ y así $2^n + 1 \mid 2^{2^{d-1}} + 1$. Puesto que $2^{d-1} \leq n$, necesariamente $2^n + 1 = 2^{2^{d-1}} + 1$ y $n = 2^{d-1}$.

Queda por ver si todo número de la forma $2^{2^n} + 1$ es primo. Volvamos a la tabla y ampliémosla:

n	0	1	2	3	4	5
$2^{2^n} + 1$	3	5	17	257	65.537	4.294.967.297

Ahora chocamos con un problema computacional. No es fácil decidir si los números que aparecen son primos o no. De hecho el lector ha aceptado en la

tabla anterior que el número 257 es primo, lo que supone probar que no es divisible entre primos menores o iguales que 13.

Aceptemos este hecho y enfrentémonos al caso más complejo de determinar si el número 65.537 es o no primo. Sabemos que se trata de $2^{2^4} + 1$. Supongamos que es divisible entre un primo p . Entonces, módulo p se cumple que $[2]^{2^4} = [-1]$, y elevando al cuadrado $[2]^{2^5} = [1]$. Consecuentemente $o([2]) \mid 2^5$, pero $o([2]) \nmid 2^4$ (o sería $[2]^{2^4} = [1]$). Esto significa que $o([2]) = 2^5$.

Por otra parte $o([2]) \mid p - 1$, o sea, $2^5 \mid p - 1$ y por tanto p ha de ser de la forma $p = 32k + 1$.

La raíz cuadrada de $2^{2^4} + 1$ es apenas mayor que $2^{2^3} = 256$, luego si nuestro número no es primo tiene un divisor primo $p < 256$. Los primeros valores de la sucesión $32k + 1$ son:

$$33, 65, 97, 129, 161, 193, 225, 257, \dots$$

y los únicos primos a considerar son 97 y 193 (161 es múltiplo de 7). Por lo tanto, 65.537 será primo si y sólo si no es divisible entre 97 ni 193. Comprobar que esto es cierto no supone ninguna dificultad, luego, en efecto, 65.537 es primo.

El siguiente número a comprobar es muchísimo mayor. Si aplicamos el mismo razonamiento concluimos que si $2^{2^5} + 1$ no es primo, entonces tiene un divisor primo de la forma $p = 64k + 1$ y $p \leq 65.536$. Esto nos deja todavía un gran número de candidatos, pero el argumento puede ser refinado como sigue:

Si $p = 64k + 1$ en particular $p \equiv 1 \pmod{8}$, lo que implica que 2 es un resto cuadrático módulo p , es decir, existe un natural x tal que $x^2 \equiv 2 \pmod{p}$, o sea, $p \mid x^2 - 2$, y así, módulo p , sabemos que $[x]^{p-1} = [1]$, luego $[2]^{(p-1)/2} = [1]$. Por lo tanto $o([2]) = 64 \mid (p-1)/2$ y en consecuencia $p = 128k + 1$.

Aún así la lista de candidatos es enorme. Los diez primeros son

$$129, \textbf{257}, 385, 513, \textbf{641}, \textbf{769}, 897, 1.025, \textbf{1.153}, 1.281, \dots$$

(los primos están en negrita).

El primer primo es fácil de descartar por su forma, ya que se trata de $2^8 + 1$ y módulo $2^8 + 1$ se cumple $[2^{2^5} + 1] = [2^8]^4 + [1] = [-1]^4 + [1] = [2] \neq [0]$.

Respecto al siguiente, las potencias de 2 módulo 641 son las siguientes:

$$\begin{array}{cccccccccc} 2, & 4, & 8, & 16, & 32, & 64, & 128, & 256, & -129, & -258, \\ 125, & 250, & -141, & -282, & 77, & 154, & 308, & -25, & -50, & -100, \\ -200, & -400, & -159, & -318, & 5, & 10, & 20, & 40, & 80, & 160, \\ 320, & 640. & & & & & & & & \end{array}$$

Así pues, $[2^{2^5} + 1] = [0]$, es decir, $641 \mid 2^{2^5} + 1$ que no es, por tanto, primo.

Posiblemente así fue cómo Gauss llegó a descubrir este hecho. Por supuesto una calculadora nos da más rápidamente que $2^{2^5} + 1 = 641 \cdot 6.700.417$ (el segundo factor resulta ser primo).

Hoy en día no se conoce ningún otro primo de Fermat distinto de los cinco que ya hemos encontrado, a saber:

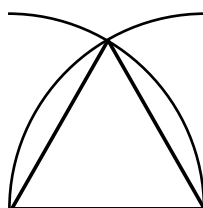
$$3, \quad 5, \quad 17, \quad 257, \quad 65.537.$$

A partir de un polígono de n lados es muy fácil construir un polígono de $2n$ lados (basta bisecar sus ángulos o sus lados), luego sólo tiene interés la construcción de polígonos con un número impar de lados. A partir de los primos de Fermat que conocemos sólo puede construirse un número finito de tales polígonos, concretamente 31. Los primeros tienen los siguientes números de lados:

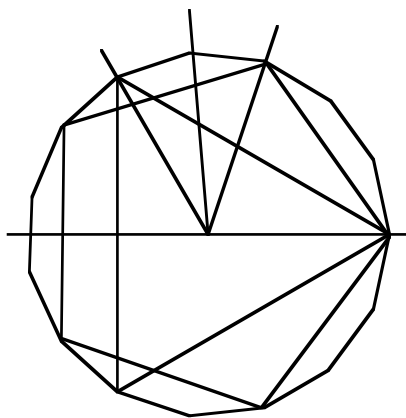
$$3, \quad 5, \quad 15, \quad 17, \quad 51, \quad 85, \quad 255, \quad 257, \quad \dots$$

La construcción de los tres primeros era conocida por los griegos, mientras que la del heptadecágono regular fue descubierta por Gauss. Vamos a mostrar construcciones explícitas de estos cuatro polígonos.

En primer lugar notemos que es muy fácil construir un triángulo equilátero:



Por otro lado, si sabemos construir un pentágono regular, también sabemos construir un pentadecágono regular: basta restar de un ángulo de amplitud $2\pi/3$ otro de amplitud $2\pi/5$ para obtener uno de amplitud $4\pi/15$, y después bisecar el resultado:



Sólo necesitamos, pues, encontrar una construcción para el pentágono y el heptadecágono regular. La construcción del pentágono es muy simple, y nos servirá de guía para obtener la del heptadecágono.

5.4.1 Construcción del pentágono regular

Sea $K = \mathbb{Q}(\zeta)$, donde $\zeta = 1_{2\pi/5}$, el cuerpo ciclotómico quinto. Su grado es 4 y su grupo de Galois G es cíclico. Un generador es el automorfismo σ que cumple $\sigma(\zeta) = \zeta^2$. Hay un único cuerpo intermedio L , el cuerpo fijado por σ^2 ,

al cual pertenecen los períodos de longitud 2: $\eta_1 = \zeta + \zeta^4$ y $\eta_2 = \zeta^2 + \zeta^3$, que constituyen una \mathbb{Q} -base de L .

La conjugación compleja induce un automorfismo en K de orden 2, luego ha de ser σ^2 . Por lo tanto $\zeta^4 = \bar{\zeta}$, de donde $\eta_1 = 2 \operatorname{Re} \zeta = 2 \cos 2\pi/5 > 0$.

Teniendo en cuenta que ζ es una raíz de $x^4 + x^3 + x^2 + x + 1$, es inmediato que $\eta_1 + \eta_2 = \eta_1\eta_2 = -1$, luego η_1 y η_2 son las raíces del polinomio $x^2 + x - 1$. Estas raíces son

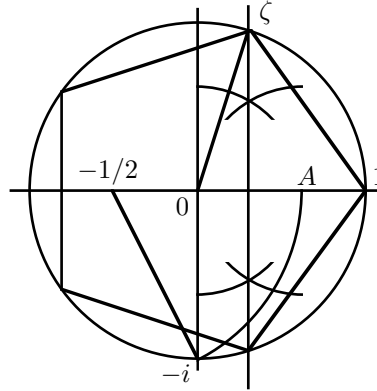
$$\eta_1 = \frac{-1 + \sqrt{5}}{2} \quad \text{y} \quad \eta_2 = \frac{-1 - \sqrt{5}}{2}$$

(sabemos que η_1 es la raíz positiva). Así pues,

$$\cos \frac{2\pi}{5} = \frac{-1 + \sqrt{5}}{4}.$$

Aunque no lo vamos a necesitar, hemos obtenido que $\sqrt{5} \in L$, luego $L = \mathbb{Q}(\sqrt{5})$.

Una forma rápida de construir el pentágono es trazar un arco con el compás apoyado en $-1/2$ y el otro extremo en $-i$, hasta cortar el eje real en el punto A . La distancia de $-1/2$ a $-i$ es $\sqrt{5}/2$, luego $A = \eta_1$. Después trazamos la mediatriz del segmento $\overline{0A}$, que pasa por $\cos 2\pi/5$ y cortará al círculo de centro 0 y radio 1 en el punto ζ :



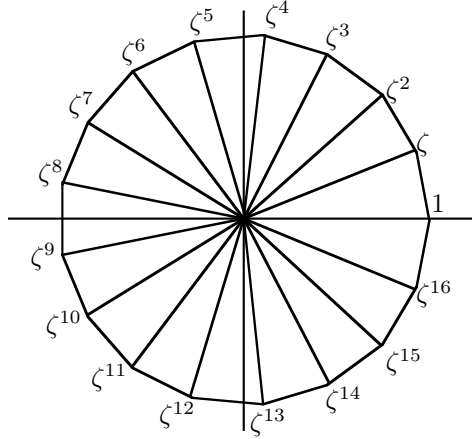
5.4.2 Construcción del heptadecágono regular

Sea $\zeta = 1_{2\pi/17}$ y $K = \mathbb{Q}(\zeta)$. Ahora el grupo de Galois G tiene orden 16. Una raíz primitiva módulo 17 es 3, luego el automorfismo σ determinado por $\sigma(\zeta) = \zeta^3$ tiene orden 16 y por lo tanto es un generador de G . El grupo G tiene tres subgrupos, a saber, $\langle \sigma^8 \rangle$, $\langle \sigma^4 \rangle$ y $\langle \sigma^2 \rangle$, que por el teorema de Galois se corresponden con tres cuerpos intermedios: $\mathbb{Q} \subset L_1 \subset L_2 \subset L_3 \subset K$, de grados 2, 4 y 8.

Para llegar desde los números racionales hasta ζ (y obtener una expresión radical de $\cos 2\pi/17$) hemos de dar cuatro pasos, lo que significa que la expresión de ζ será demasiado complicada para manejarla con comodidad. En lugar de eso nos limitaremos a describir ζ en función de su polinomio mínimo respecto

a L_3 , y los coeficientes de este polinomio en función de sus polinomios mínimos en L_2 y así sucesivamente.

En el caso del pentágono, los números η_1 y η_2 estaban determinados por que sabíamos que eran las raíces de $x^2 + x - 1$ y porque además sabíamos que η_1 era la raíz positiva. Aquí también tendremos que hacer estimaciones de signos con el mismo fin, para lo cual deberemos tener presente la distribución de las raíces de la unidad:



Como en el caso del pentágono, el automorfismo de orden 2, o sea, σ^8 es la conjugación compleja y así, ζ y ζ^{16} son conjugados, al igual que ζ^2 y ζ^{15} , etc. (como muestra la figura).

El cuerpo L_3 (fijado por σ^8) contiene a los períodos de longitud 2, es decir, a los pares de raíces conjugadas:

$$\begin{aligned} \lambda_1 &= \zeta + \zeta^{16}, & \lambda_2 &= \zeta^2 + \zeta^{15}, & \lambda_3 &= \zeta^3 + \zeta^{14}, & \lambda_4 &= \zeta^4 + \zeta^{13}, \\ \lambda_5 &= \zeta^5 + \zeta^{12}, & \lambda_6 &= \zeta^6 + \zeta^{11}, & \lambda_7 &= \zeta^7 + \zeta^{10}, & \lambda_8 &= \zeta^8 + \zeta^9, \end{aligned}$$

y como $\zeta\zeta^{16} = 1$, es claro que ζ y ζ^{16} son las raíces del polinomio $x^2 - \lambda_1x + 1$. En general tenemos que

ζ	y	ζ^{16}	son las raíces de	$x^2 - \lambda_1x + 1$
ζ^2	y	ζ^{15}	son las raíces de	$x^2 - \lambda_2x + 1$
ζ^3	y	ζ^{14}	son las raíces de	$x^2 - \lambda_3x + 1$
ζ^4	y	ζ^{13}	son las raíces de	$x^2 - \lambda_4x + 1$
ζ^5	y	ζ^{12}	son las raíces de	$x^2 - \lambda_5x + 1$
ζ^6	y	ζ^{11}	son las raíces de	$x^2 - \lambda_6x + 1$
ζ^7	y	ζ^{10}	son las raíces de	$x^2 - \lambda_7x + 1$
ζ^8	y	ζ^9	son las raíces de	$x^2 - \lambda_8x + 1$

En realidad para la construcción del heptadecágono nos interesan más los λ_k que las raíces de la unidad, pues claramente $\lambda_k = 2 \operatorname{Re} \zeta^k = 2 \cos 2k\pi/17$. De aquí se sigue claramente (ver la figura anterior) que

$$\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4 > 0 > \lambda_5 > \lambda_6 > \lambda_7 > \lambda_8.$$

Una vez construido un λ_k , basta dividirlo entre 2 y levantar la perpendicular para obtener ζ^k y su conjugado como las intersecciones de dicha recta y la circunferencia unidad. Para construir los λ_k vamos a expresarlos en función del cuerpo L_2 , que es el cuerpo fijado por σ^4 y que, por tanto, tiene por base a los períodos de longitud 4:

$$\begin{aligned}\xi_1 &= \zeta + \zeta^{16} + \zeta^4 + \zeta^{13}, & \xi_2 &= \zeta^2 + \zeta^{15} + \zeta^8 + \zeta^9, \\ \xi_3 &= \zeta^3 + \zeta^{14} + \zeta^5 + \zeta^{12}, & \xi_4 &= \zeta^6 + \zeta^{11} + \zeta^7 + \zeta^{10}.\end{aligned}$$

(Si aplicamos a ζ las potencias de σ^4 obtenemos ζ , ζ^{16} , ζ^4 y ζ^{13} , por lo que σ^4 permuta los sumandos de ξ_1 , que es, por lo tanto, invariante).

Es obvio que

$$\xi_1 = \lambda_1 + \lambda_4, \quad \xi_2 = \lambda_2 + \lambda_8, \quad \xi_3 = \lambda_3 + \lambda_5, \quad \xi_4 = \lambda_6 + \lambda_7.$$

Por otro lado

$$\begin{aligned}\lambda_1 \cdot \lambda_4 &= (\zeta + \zeta^{16})(\zeta^4 + \zeta^{13}) = \zeta^5 + \zeta^{14} + \zeta^3 + \zeta^{12} = \xi_3 \\ \lambda_2 \cdot \lambda_8 &= (\zeta^2 + \zeta^{15})(\zeta^8 + \zeta^9) = \zeta^{10} + \zeta^{11} + \zeta^6 + \zeta^7 = \xi_4 \\ \lambda_3 \cdot \lambda_5 &= (\zeta^3 + \zeta^{14})(\zeta^5 + \zeta^{12}) = \zeta^8 + \zeta^{15} + \zeta^2 + \zeta^9 = \xi_2 \\ \lambda_6 \cdot \lambda_7 &= (\zeta^6 + \zeta^{11})(\zeta^7 + \zeta^{10}) = \zeta^{13} + \zeta^{16} + \zeta + \zeta^4 = \xi_1\end{aligned}$$

de donde se desprende que

$$\begin{array}{llll} \lambda_1 & \text{y} & \lambda_4 & \text{son las raíces de } x^2 - \xi_1 x + \xi_3 \\ \lambda_2 & \text{y} & \lambda_8 & \text{son las raíces de } x^2 - \xi_2 x + \xi_4 \\ \lambda_3 & \text{y} & \lambda_5 & \text{son las raíces de } x^2 - \xi_3 x + \xi_2 \\ \lambda_4 & \text{y} & \lambda_7 & \text{son las raíces de } x^2 - \xi_4 x + \xi_1 \end{array}$$

Finalmente, L_1 es el cuerpo fijado por σ^2 , y su base son los períodos de longitud 8:

$$\begin{aligned}\eta_1 &= \zeta + \zeta^9 + \zeta^{13} + \zeta^{15} + \zeta^{16} + \zeta^8 + \zeta^4 + \zeta^2 \\ \eta_2 &= \zeta^3 + \zeta^{10} + \zeta^5 + \zeta^{11} + \zeta^{14} + \zeta^7 + \zeta^{12} + \zeta^6.\end{aligned}$$

Un simple cálculo nos da $\eta_1 + \eta_2 = -1$, $\eta_1 \eta_2 = -4$, luego son las raíces del polinomio $x^2 + x - 4$.

Por otro lado se comprueba que ξ_1 y ξ_2 son las raíces de $x^2 - \eta_1 x - 1$, mientras que ξ_3 y ξ_4 son las raíces de $x^2 - \eta_2 x - 1$.

Estos datos nos permiten ascender gradualmente hasta construir ζ . Para agilizar la construcción nos valdremos de un truco.

Consideremos el ángulo θ que cumple $0 < 4\theta < \pi/2$ y $\tan 4\theta = 4$.

Las raíces del polinomio $x^2 + x - 4$ son $2 \tan 2\theta$ y $-2/\tan 2\theta$, pues su producto es -4 y su suma vale

$$2 \tan 2\theta - \frac{2}{\tan 2\theta} = 2 \left(\frac{\tan^2 2\theta - 1}{\tan 2\theta} \right) = -\frac{4}{\tan 4\theta} = -1$$

(por la fórmula de la tangente del ángulo doble).

Por otra parte sabemos que las raíces son η_1 y η_2 , luego

$$\eta_1 = 2 \tan 2\theta \quad \text{y} \quad \eta_2 = -2 \tan 2\theta,$$

(hay que comprobar que $\eta_1 > 0$, pero eso se sigue de que

$$\eta_1 = \lambda_1 + \lambda_2 + \lambda_4 + \lambda_8 = 2 \left(\cos \frac{2\pi}{17} + \cos \left(2\frac{2\pi}{17} \right) + \cos \left(4\frac{2\pi}{17} \right) + \cos \left(8\frac{2\pi}{17} \right) \right)$$

y los tres primeros sumandos son positivos, más aún, los dos primeros suman más de 1, luego compensan al último.)

Las raíces de $x^2 - \eta_1 x - 1$ son $\tan(\theta + \pi/4)$ y $\tan(\theta - \pi/4)$, pues su producto es -1 (los ángulos se diferencian en $\pi/2$, luego la tangente de uno es la inversa de la tangente del otro, y además tienen signos opuestos) y su suma es

$$\begin{aligned} \tan \left(\theta + \frac{\pi}{4} \right) + \tan \left(\theta - \frac{\pi}{4} \right) &= \frac{\tan \theta + 1}{1 - \tan \theta} + \frac{\tan \theta - 1}{1 + \tan \theta} \\ &= \frac{4 \tan \theta}{1 - \tan^2 \theta} = 2 \tan 2\theta = \eta_1. \end{aligned}$$

Por lo tanto $\xi_1 = \tan(\theta + \pi/4)$ y $\xi_2 = \tan(\theta - \pi/4)$ (pues $\xi_1 = \lambda_1 + \lambda_4 > 0$).

Con el mismo razonamiento se concluye que $\xi_3 = \tan \theta$ y $\xi_4 = -\cot \theta$. Ya tenemos suficiente para realizar la construcción:

Consideramos los puntos $A = 1$ y $B = i$. Dividimos el segmento \overline{OB} en cuatro partes, con lo que obtenemos el punto $I = (1/4)i$. El ángulo \widehat{OIA} , tiene tangente 4, luego es 4θ . Lo biseamos dos veces y obtenemos θ igual al ángulo \widehat{OIE} (figura 5.1).

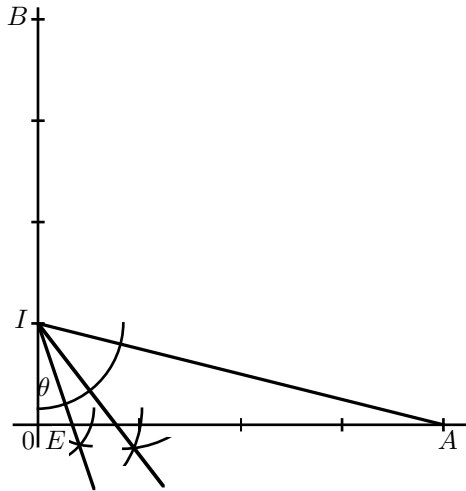


Figura 5.1: Construcción de θ .

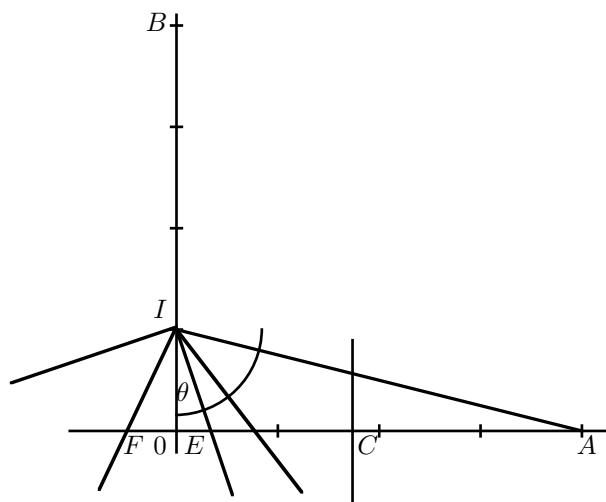


Figura 5.2: Construcción de C .

Según nuestros cálculos $E = \xi_3/4$, $F = -\xi_2/4$ y la distancia de A a F es $\overline{AF} = (4 - \xi_2)/4$. Por lo tanto la distancia de F a C (que es igual a la distancia de C a A) es $(4 - \xi_2)/8$. Consecuentemente

$$C = \overline{OC} = \frac{4 - \xi_2}{8} + \frac{\xi_2}{4} = \frac{\xi_2 + 4}{8}.$$

Trazamos el círculo de centro C y radio \overline{AC} , que cortará al eje imaginario en un punto K . Por el teorema de Pitágoras

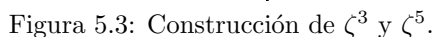
$$\begin{aligned} K &= \overline{OK} = \sqrt{\overline{OK}^2 - \overline{OC}^2} = \sqrt{\overline{CA}^2 - \overline{OC}^2} = \frac{\sqrt{-\xi_2}}{2}, \\ \overline{EK} &= \sqrt{\overline{OK}^2 + \overline{OE}^2} = \frac{\sqrt{\xi_3 - 4\xi_2}}{4}. \end{aligned}$$

A continuación trazamos el círculo de centro E y radio \overline{EK} , que cortan al eje horizontal en los puntos N_5 y N_3 . Entonces

$$N_3 = E + \overline{EK} = \frac{\xi_3 + \sqrt{\xi_3^2 - 4\xi_2}}{2} = \frac{\lambda_3}{2},$$

pues λ_3 es la raíz positiva del polinomio $x^2 - \xi_3 x + \xi_2$. Así pues, $N_3 = \cos 6\pi/17$. Por otra parte,

$$N_5 = \overline{EN}_5 - E = \overline{EK} - E = \frac{\xi_3^2 - 2\xi_2}{4} - \frac{\xi_3}{4} = \frac{-\xi_3 + \sqrt{\xi_3^2 - 4\xi_2}}{2} = -\frac{\lambda_5}{2},$$



Levantando perpendiculares por N_3 y N_5 y trazando el círculo de centro 0 y radio 1 obtenemos ζ_3 y ζ_5 (figura 5.3).

Para acabar, bisecamos el ángulo $\widehat{\zeta_3 O \zeta_5}$ con lo que obtenemos ζ_4 y, uniendo ζ_3 con ζ_4 obtenemos el lado del heptadecágono regular (figura 5.4).

Aunque en la definición de espacio afín euclídeo hemos exigido que el cuerpo base sea \mathbb{R} , es claro que todos los resultados en torno a ellos son válidos si trabajamos con cualquier subcuerpo K de \mathbb{R} en el que todo número positivo tenga una raíz cuadrada (para definir la norma). Un ejemplo es tomar como K el cuerpo de los números reales constructibles. Así K^3 verifica todos los axiomas de la geometría euclídea salvo el axioma de constructibilidad. Puesto que \mathbb{C} es isomorfo a K^2 como espacio vectorial, los planos euclídeos sobre K son identificables con \mathbb{C} . Las propiedades sobre circunferencias que hemos probado con el axioma D siguen siendo válidas, así como la propiedad de Arquímedes; en cambio el comportamiento de los ángulos es peculiar, pues en K^2 no hay heptágonos regulares, ni se puede trisecar un ángulo arbitrario.

Una geometría euclídea discontinua menos atípica es la que se obtiene al tomar como K el cuerpo de los números reales algebraicos, de modo que los planos se pueden identificar con la clausura algebraica de \mathbb{Q} . De nuevo K^2 satisface todos los axiomas excepto el axioma D pero ahora existen polígonos regulares de cualquier número de lados (pues los números $1_{2\pi/n}$ son algebraicos).

Ejercicio: Probar que un número complejo $a + bi$ es algebraico sobre \mathbb{Q} si y sólo si lo son a y b .

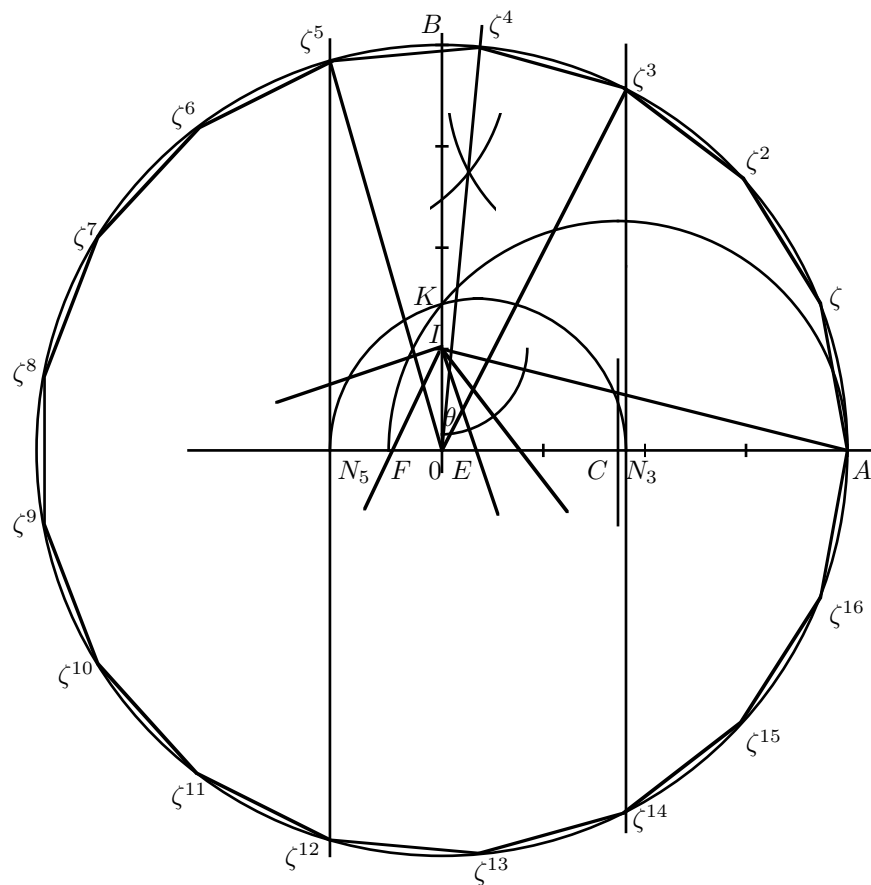


Figura 5.4: Construcción del heptadecágono regular

Esto muestra que los resultados esenciales de la geometría no dependen del axioma D, sino que éste puede sustituirse por requisitos específicos más débiles.

5.6 Apéndice: El teorema de Sylow

Probamos aquí el teorema de Sylow que hemos usado en la demostración del teorema fundamental del álgebra. Emplearemos un argumento inductivo que se apoyará en el hecho siguiente:

Si G es un grupo abeliano finito y p un número primo tal que $p \mid |G|$, entonces G tiene un elemento de orden p .

Esto es cierto porque todo grupo abeliano finito es producto directo de grupos cíclicos, uno de los cuales debe tener orden múltiplo de p , y en un grupo cíclico hay elementos de todos los órdenes posibles (divisores del orden del grupo).

Si G es un grupo finito, en G tenemos definida la relación de conjugación, según la cual dos elementos $x, y \in G$ son *conjugados* si existe un $g \in G$ tal que $y = g^{-1}xg = x^g$. Se trata de una relación de equivalencia. La clase de equivalencia de un elemento x se llama *clase de conjugación* de x , y la representaremos por $\text{cl}(x)$.

Necesitamos calcular el cardinal de $\text{cl}(x)$. Para ello definimos el *centralizador* de x en G como el conjunto

$$C_G(x) = \{g \in G \mid x^g = x\} = \{g \in G \mid xg = gx\},$$

que claramente es un subgrupo de G . Tenemos el teorema siguiente:

Teorema 1 *Sea G un grupo finito y $x \in G$. Entonces $|\text{cl}(x)| = |G : C_G(x)|$.*

DEMOSTRACIÓN: Observamos que $x^g = x^h$ si y sólo si $hg^{-1}x = xhg^{-1}$, si y sólo si $hg^{-1} \in C_G(x)$, si y sólo si $C_G(x)h = C_G(x)g$.

El subgrupo $C_G(x)$ no tiene por qué ser un subgrupo normal, luego no podemos hablar de grupo cociente, pero sí del conjunto $G/C_G(x)$ de las clases de equivalencia $C_G(x)g$ para la relación de la congruencia por la derecha módulo $C_G(x)$.

Lo que hemos obtenido es que la aplicación $f : G/C_G(x) \rightarrow \text{cl}(x)$ definida por $f(C_G(x)g) = x^g$ es biyectiva. ■

En particular, un $x \in G$ cumplirá $\text{cl}(x) = \{x\}$ si y sólo si $C_G(x) = G$, es decir, si y sólo si $gx = xg$ para todo $g \in G$, si x conmuta con todo elemento de G . Definimos el *centro* de G como el subgrupo

$$Z(G) = \{x \in G \mid gx = xg \text{ para todo } g \in G\}.$$

Así, $Z(G)$ está formado por los elementos cuya clase de conjugación tiene cardinal 1, por los elementos que conmutan con todos los elementos de G . En

particular los elementos de $Z(G)$ conmutan entre sí, es decir, $Z(G)$ es un grupo abeliano. También es claro que $Z(G)$ es normal en G .

Dado un grupo finito G , sabemos que tiene $|Z(G)|$ clases de conjugación con un elemento y, digamos, n clases con más de un elemento. Sean x_1, \dots, x_n representantes de estas clases. El orden de G es la suma de los órdenes de sus clases de conjugación, luego teniendo en cuenta el teorema 1 concluimos:

Teorema 2 (ecuación de clases) *Si G es un grupo finito, existen elementos $x_1, \dots, x_n \in G$ tales que $C_G(x_i) < G$ y*

$$|G| = |Z(G)| + \sum_{i=1}^n |G : C_G(x_i)|.$$

Con esto tenemos suficiente para probar la parte principal del teorema de Sylow. Conviene dar la definición siguiente:

Definición Sea G un grupo finito y p un número primo. Sea $|G| = p^n \cdot m$, con $(p, m) = 1$ (quizá con $n = 0$). Un subgrupo H de G de orden p^n se llama *p-subgrupo de Sylow* de G .

Teorema 3 *Si G es un grupo finito y p un número primo, entonces G tiene un p-subgrupo de Sylow.*

DEMOSTRACIÓN: Por inducción sobre el orden de G . Si G tiene orden 1 es obvio. Supongamos que todos los grupos de orden menor que $|G|$ tienen p -subgrupos de Sylow y demostremos que G también los tiene.

Si $p \nmid |G|$, entonces el subgrupo trivial es un p -subgrupo de Sylow de G . Supongamos, pues, que $p \mid |G|$. Sea $|G| = p^n \cdot m$, con $(p, m) = 1$.

Distinguimos dos casos:

CASO 1 Existe un subgrupo $H < G$ tal que $p \nmid |G : H|$.

Entonces $p^n \mid |H|$ y por hipótesis de inducción H tiene un p -subgrupo de Sylow P de orden p^n , y así, P es también un p -subgrupo de Sylow de G .

CASO 2 Para todo subgrupo $H \leq G$, se cumple que $p \mid |G : H|$.

Entonces la ecuación de clases nos da que $p \mid |Z(G)|$. Como se trata de un grupo abeliano, tiene un elemento de orden p o, lo que es lo mismo, tiene un subgrupo $H \leq Z(G)$ de orden p . Como los elementos de H conmutan con todos los elementos de G , es evidente que $H^g = H$ para todo $g \in G$, o sea, H es normal en G .

Se cumple que $|G/H| = p^{n-1} \cdot m$ y tiene un subgrupo de Sylow P/H que cumplirá $|P/H| = p^{n-1}$, luego $|P| = p^n$, es decir, P es un subgrupo de Sylow de G . ■

Ahora falta probar que todo grupo de orden potencia de primo tiene subgrupos de todos los órdenes posibles, lo que concluirá la prueba del teorema de Sylow. Los grupos de orden p^n se llaman p -grupos. El teorema siguiente basta:

Teorema 4 *Si p es un primo, todo p -grupo no trivial tiene un subgrupo de índice p .*

DEMOSTRACIÓN: Sea P un p -grupo. Lo probaremos por inducción sobre el orden de P . Si $|P| = p$ el subgrupo trivial tiene índice p . Supongamos que los p -grupos de orden menor que $|P|$ tienen subgrupos de índice p .

Si P es abeliano, entonces P es producto de grupos cíclicos. Los grupos cíclicos tienen subgrupos de todos los órdenes posibles, luego podemos tomar un subgrupo de índice p en uno cualquiera de los factores. El producto de este subgrupo por los factores restantes es un subgrupo de índice p en P .

Supongamos que P no es abeliano. Entonces $Z(P) < P$. De nuevo la ecuación de clases nos da que $p \mid |Z(P)|$, luego $|Z(P)| \neq 1$. Como $|P/Z(P)| < |P|$ y es un p -grupo no trivial, por hipótesis de inducción tenemos que existe un grupo $H/Z(P) < P/Z(P)$ de índice p , luego $|P : H| = p$. ■

Conviene observar que hemos demostrado lo siguiente:

Teorema 5 *Si p es un primo y P es un p -grupo no trivial, entonces $Z(P) \neq 1$.*

De aquí se sigue, por ejemplo:

Teorema 6 *Todo p -grupo es resoluble.*

DEMOSTRACIÓN: Sea P un p -grupo no trivial. Por inducción sobre $|P|$. Si P es abeliano es resoluble. Si no es abeliano, entonces $P/Z(P)$ es un p -grupo no trivial y de orden menor. Por hipótesis de inducción $P/Z(P)$ es resoluble y $Z(P)$ es abeliano, luego resoluble. De aquí se sigue que P es resoluble. ■

Capítulo VI

Bijecciones afines

La geometría sintética suele ser la más indicada para tratar problemas “locales”, en torno a figuras que involucran pocos puntos y rectas, mientras que la geometría analítica resulta más apta, por no decir imprescindible, para cuestiones “globales”, como son las relacionadas con las transformaciones del espacio en sí mismo. Dedicamos este capítulo a profundizar en las propiedades de estas transformaciones. Conviene distinguir siempre entre los resultados afines, válidos en todo espacio afín, de los genuinamente euclídeos, pues la geometría afín es una herramienta algebraica susceptible de ser aplicada al estudio de cuerpos arbitrarios, mientras que la estructura euclídea es muy particular, válida para espacios reales y sólo parcialmente generalizable a espacios sobre el cuerpo de los números complejos.

6.1 El grupo afín y el grupo lineal

Definición 6.1 Dado un espacio vectorial V sobre un cuerpo K , llamaremos *grupo lineal* de V , representado por $GL(V)$, al conjunto de todos los automorfismos de V , es decir, el conjunto de todas las aplicaciones lineales biyectivas de V en sí mismo, que es un grupo con la composición de aplicaciones. Claramente, si V tiene dimensión n , entonces $GL(V)$ es isomorfo al grupo $GL(n, K)$ de las matrices inversibles $n \times n$ con coeficientes en K .

Dado un espacio afín E , llamaremos *grupo afín* de E al grupo $GA(E)$ formado por las biyecciones afines de E en sí mismo, también con la composición de aplicaciones.

Recordemos del capítulo IV que una afinidad $f \in GA(E)$ es una aplicación que actúa sobre cada punto P en la forma

$$f(P) = f(O) + \vec{f}(\overrightarrow{OP}),$$

donde $\vec{f} \in GL(\vec{E})$ y O es un punto de E . Allí probamos que una afinidad admite una expresión de esta forma para cualquier punto O prefijado y siempre con la misma aplicación \vec{f} .

Es inmediato comprobar que la aplicación

$$\begin{array}{ccc} \text{GA}(E) & \longrightarrow & \text{GL}(\vec{E}) \\ f & \mapsto & \vec{f} \end{array}$$

es un epimorfismo de grupos. Su núcleo está formado por las aplicaciones de la forma

$$f(P) = f(O) + \overrightarrow{OP} = P + \overrightarrow{Pf(O)} + \overrightarrow{OP} = P + \overrightarrow{Of(O)},$$

donde $\overrightarrow{Of(O)}$ es un vector arbitrario de V . Así pues, el núcleo del epimorfismo anterior es el grupo $T(E)$ de las *traslaciones* de E , esto es, el grupo formado por las afinidades $T_{\vec{v}}$ dadas por $T_{\vec{v}}(P) = P + \vec{v}$.

Así pues, tenemos que $\text{GA}(E)/T(E) \cong \text{GL}(V)$. Por otro lado, es obvio que la aplicación $\vec{v} \mapsto T_{\vec{v}}$ es un isomorfismo $\vec{E} \cong T(E)$. Diremos que dos conjuntos son *trasladados* si uno es la imagen del otro por una traslación.

También podemos sumergir $\text{GL}(\vec{E})$ en $\text{GA}(E)$. Dado un punto $O \in E$, el conjunto $\text{GA}_O(E)$ formado por los elementos $f \in \text{GA}(E)$ tales que $f(O) = O$ es claramente un subgrupo de $\text{GA}(E)$ isomorfo a $\text{GL}(\vec{E})$. Concretamente, cada $\vec{f} \in \text{GL}(\vec{E})$ se corresponde con la afinidad \vec{f}_O dada por

$$\vec{f}_O(P) = O + \vec{f}(\overrightarrow{OP}).$$

El isomorfismo inverso es simplemente $f \mapsto \vec{f}$.

Fijado O , toda biyección afín f se expresa como

$$f(P) = f(O) + \vec{f}(\overrightarrow{OP}) = O + \vec{f}(\overrightarrow{OP}) + \overrightarrow{Of(O)} = \vec{f}_O(P) + \overrightarrow{Of(O)},$$

o sea, como composición de $\vec{f}_O \in \text{GA}_O(E)$ con la traslación de dirección $\overrightarrow{Of(O)}$. Esto significa que $\text{GA}(E) = \text{GA}_O(E)T(E)$. Por otro lado es claro que las traslaciones no tienen puntos fijos, luego $\text{GA}_O(E) \cap T(E) = 1$.

En términos de la teoría de grupos esto se interpreta como que $\text{GA}(E)$ es isomorfo al producto semidirecto $\text{GL}(\vec{E})[\vec{E}]$.

Si una afinidad tiene un punto fijo O , entonces su expresión respecto a un sistema de referencia con origen O es particularmente simple, por ello es interesante el teorema siguiente, que describe el conjunto de puntos fijos de una afinidad y en un caso particular garantiza incluso su existencia:

Teorema 6.2 *Sea $f : E \longrightarrow E$ una afinidad en un espacio afín E . Entonces*

1. *El conjunto $\vec{I}(\vec{f})$ de puntos fijos (o invariantes) de \vec{f} es un subespacio vectorial de \vec{E} .*
2. *Si f tiene un punto fijo O , entonces el conjunto $I(f)$ de puntos fijos de f es la variedad lineal $O + \vec{I}(\vec{f})$.*
3. *Si $\vec{I}(\vec{f}) = 0$, entonces f tiene un único punto fijo.*

DEMOSTRACIÓN: 1 es evidente. En 2 es claro que todos los puntos de $O + \vec{I}(\vec{f})$ son puntos fijos de f . Recíprocamente, si $f(P) = P$ entonces $P = O + \vec{f}(\vec{OP})$, luego $\vec{f}(\vec{OP}) = \vec{OP}$ y por lo tanto $\vec{OP} \in \vec{I}(\vec{f})$.

Veamos 3. Sea O un origen arbitrario. Estamos buscando un punto P tal que $P = f(O) + \vec{f}(\vec{OP}) = O + \vec{Of}(\vec{O}) + \vec{f}(\vec{OP})$, es decir, tal que

$$\vec{OP} - \vec{f}(\vec{OP}) = \vec{Of}(\vec{O}).$$

La aplicación $1 - \vec{f}$ (donde 1 representa a la identidad) es lineal y por hipótesis su núcleo es 0, luego es inyectiva, y por consiguiente suprayectiva, luego ciertamente existe un vector \vec{OP} tal que $(1 - \vec{f})(\vec{OP}) = \vec{Of}(\vec{O})$, como queríamos probar. La unicidad del punto fijo se sigue del apartado anterior. ■

6.2 Homotecias

Las homotecias son una familia importante de aplicaciones afines, pues contienen esencialmente el concepto de semejanza de figuras, que hasta ahora hemos estudiado únicamente para el caso de los triángulos. Comenzamos definiéndolas algebraicamente y después las caracterizaremos geométricamente.

Definición 6.3 Sea E un espacio afín sobre un cuerpo K . La *homotecia* de centro $O \in E$ y razón $k \in K^* = K \setminus \{0\}$ es la aplicación $H(O, k) \in \text{GA}(E)$ dada por

$$H(O, k)(P) = O + k \vec{OP}.$$

Llamaremos $H(O, E)$ al grupo de todas las homotecias de centro O en E . Dos conjuntos son *homotéticos* (gr. ‘con el mismo aspecto’) si uno es la imagen del otro por una homotecia.¹

Es inmediato comprobar que la aplicación $k \mapsto H(O, k)$ es un isomorfismo $K^* \cong H(O, E)$.

Las homotecias se pueden caracterizar por sus aplicaciones lineales asociadas. Llamaremos *homotecia lineal* de razón $k \in K^*$ en un espacio vectorial V sobre un cuerpo K a la aplicación dada por $f(\vec{v}) = k\vec{v}$. Claramente las homotecias lineales forman un grupo $H(V)$ isomorfo a K^* .

Teorema 6.4 Una afinidad f de un espacio afín E en sí mismo es una homotecia de razón $k \neq 1$ si y sólo si \vec{f} es una homotecia lineal de razón k .

DEMOSTRACIÓN: Si $f = H(O, k)$, entonces $f(P) = O + \vec{f}(\vec{OP}) = O + k \vec{OP}$, luego $\vec{f}(\vec{OP}) = k \vec{OP}$ para todo vector \vec{OP} de \vec{E} , es decir, \vec{f} es la homotecia lineal de razón k .

¹Notar que esta relación es reflexiva y simétrica, pero enseguida veremos que no es transitiva.

Recíprocamente, si \vec{f} es la homotecia de razón $k \neq 1$, entonces es claro que $\vec{I}(\vec{f}) = O$, luego por 6.2 sabemos que f tiene un punto fijo O . Entonces es claro que $f = H(O, k)$. ■

Observar que la única homotecia de razón 1 es la identidad. Ahora es claro que la composición de dos homotecias de centros distintos no es necesariamente una homotecia. Concretamente, la composición de una homotecia de razón k con otra de razón $1/k$ da lugar a una afinidad con $\vec{f} = 1$, pero si los centros no son el mismo es fácil ver que $f \neq 1$, y entonces f es una traslación. Pronto veremos que la composición de dos homotecias es siempre una homotecia o una traslación. Antes necesitamos algunos conceptos auxiliares:

Definición 6.5 Diremos que dos variedades lineales L y M de un espacio afín E son *suplementarias* si $\vec{E} = \vec{L} \oplus \vec{M}$.

Notemos que si L y M son suplementarias entonces existe un punto O tal que $L \cap M = \{O\}$. En efecto, dado $P \in L$ y $Q \in M$, entonces $\overrightarrow{PQ} = \vec{v} + \vec{w}$, donde $\vec{v} \in \vec{L}$ y $\vec{w} \in \vec{M}$, luego $O = P + \vec{v} = Q - \vec{w}$ cumple $O \in L \cap M$. Si $O' \in L \cap M$, entonces $\overrightarrow{OO'} \in \vec{L} \cap \vec{M} = 0$, luego $O = O'$.

Dadas dos variedades lineales suplementarias L y M en un espacio afín E , todo punto de E se expresa de forma única como $O + \vec{v} + \vec{w}$, donde $L \cap M = \{O\}$, $\vec{v} \in \vec{L}$ y $\vec{w} \in \vec{M}$. Llamaremos *proyección* de E en L *paralela* a M a la aplicación $p : E \rightarrow L$ dada por $p(O + \vec{v} + \vec{w}) = O + \vec{v}$.

Alternativamente, $p(P)$ es el único punto de intersección entre las variedades suplementarias L y $P + \vec{M}$. También es claro que p es una afinidad, pues

$$p(P) = O + \vec{p}(\overrightarrow{OP}),$$

donde \vec{p} es la proyección de $\vec{L} \oplus \vec{M}$ sobre \vec{L} .

Ahora estamos en condiciones de probar una versión general del teorema de Tales. Podría pensarse que el teorema de Tales es genuinamente euclídeo, por lo que no puede enunciarse en espacios afines. No es así, y ello se debe a la observación siguiente: si P, Q, R y S son cuatro puntos colineales en un espacio afín sobre un cuerpo K y $P \neq Q$, entonces $\overrightarrow{RS} = \alpha \overrightarrow{PQ}$, para un único $\alpha \in K$. Por lo tanto podemos definir

$$\frac{\overrightarrow{RS}}{\overrightarrow{PQ}} = \alpha.$$

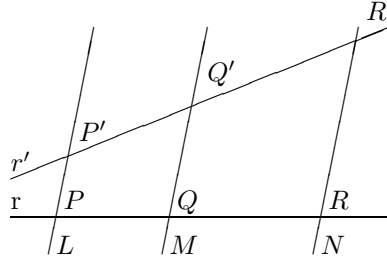
Notar que esta definición generaliza a la que dimos en el capítulo III antes del teorema de Ceva. En un espacio afín no podemos comparar vectores que no correspondan a puntos colineales,² pero para formular el teorema de Tales no necesitamos más.

²Podríamos haber definido más en general la razón entre dos vectores con la misma dirección, aunque los puntos que los definen no sean colineales, pero esto ya no sería un invariante afín (las afinidades no lo conservan) mientras que sobre puntos colineales sí lo es, como es fácil comprobar.

Teorema 6.6 (Teorema de Tales) Sean L, M, N tres hiperplanos paralelos en un espacio afín E . Si r es cualquier recta no paralela a estos hiperplanos, entonces la razón

$$\frac{\overrightarrow{PQ}}{\overrightarrow{PR}}, \quad \text{donde } r \cap L = \{P\}, \quad r \cap M = \{Q\}, \quad r \cap N = \{R\},$$

es independiente de r .

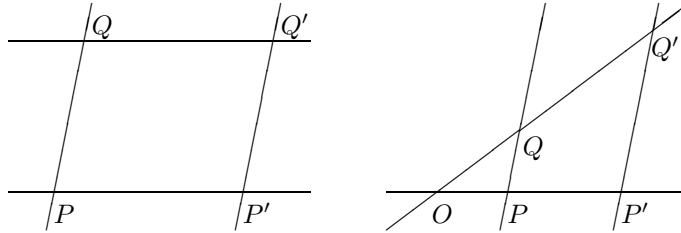


DEMOSTRACIÓN: Sea W el espacio director de los tres hiperplanos. Sea r' otra recta no paralela a L, M, N que los corte en los puntos P', Q', R' . Sea $p : E \rightarrow r'$ la proyección paralela a W . Entonces es claro que $p(P) = P'$, $p(Q) = Q'$, $p(R) = R'$. Por lo tanto $\overrightarrow{P'R'} = \vec{p}(\overrightarrow{PR})$ y $\overrightarrow{P'Q'} = \vec{p}(\overrightarrow{PQ})$. Si se cumple $\overrightarrow{PQ} = \alpha \overrightarrow{PR}$, aplicando \vec{p} obtenemos $\overrightarrow{P'Q'} = \alpha \overrightarrow{P'R'}$. ■

Como consecuencia obtenemos:

Teorema 6.7 Sean P, Q, P', Q' puntos distintos de un espacio afín tales que las rectas PQ y $P'Q'$ sean paralelas.

1. Si las rectas PP' y QQ' son paralelas, entonces $Q' = T_{\vec{v}}(Q)$, con $\vec{v} = \overrightarrow{PP'}$.
2. Si las rectas PP' y QQ' se cortan en O , entonces $Q' = H(O, k)(Q)$, donde $k = \overrightarrow{OP'} / \overrightarrow{OP}$.



DEMOSTRACIÓN: 1) Hemos de probar que $Q' = Q + \overrightarrow{PP'}$. Para ello notamos que $Q + \overrightarrow{PP'}$ está en la recta QQ' y por otro lado

$$Q + \overrightarrow{PP'} = P + \overrightarrow{PQ} + \overrightarrow{PP'} = P' + \overrightarrow{PQ}$$

está en la recta $P'Q'$, luego se trata del punto de corte entre ambas, Q' , como queríamos probar.

2) Hemos de probar que $Q' = O + k\overrightarrow{OQ}$, o sea, $\overrightarrow{OQ'}/\overrightarrow{OQ} = \overrightarrow{OP'}/\overrightarrow{OP}$, pero esto es consecuencia inmediata del teorema de Tales (aplicado al plano que contiene las rectas consideradas y teniendo en cuenta además la recta paralela a PQ que pasa por O). ■

Esto nos lleva a la caracterización geométrica de las homotecias que habíamos anunciado:

Teorema 6.8 *Sea E un espacio afín de dimensión mayor o igual que 2. El conjunto de las biyecciones de E en sí mismo que transforman cada recta en una recta paralela (o igual) es el subgrupo de $\text{GA}(E)$ formado por las homotecias y las traslaciones de E . Lo representaremos por $\text{HT}(E)$.*

DEMOSTRACIÓN: Las biyecciones indicadas forman obviamente un grupo. Es fácil ver que contiene a las homotecias y a las traslaciones. Por ejemplo, si $f = H(O, k)$ y $r = P + \langle \vec{v} \rangle$ es una recta, la imagen de un punto $Q = P + \lambda \vec{v}$ es $O + k\overrightarrow{OQ}$, pero $k\overrightarrow{OQ} = k\overrightarrow{OP} + k\overrightarrow{PQ} = k\overrightarrow{OP} + \lambda k, \vec{v}$, luego concluimos que $f[r] = (O + k\overrightarrow{OP}) + \langle \vec{v} \rangle$, que es una recta paralela a r .

Veamos que cualquier elemento f del grupo es una homotecia o una traslación. Distinguimos varios casos.

Si f tiene dos puntos fijos, Q y Q' , entonces, dado cualquier punto P que no esté en la recta QQ' , tenemos que $f[PQ]$ es una recta paralela o igual a PQ que contiene a P , luego es PQ . Igualmente $f[PQ'] = PQ'$. Como P es el único punto en común entre PQ y PQ' , necesariamente $f(P) = P$. Así pues, f fija a todos los puntos fuera de la recta QQ' . Si R está en la recta QQ' y P es un punto exterior a esta recta, entonces todos los puntos de PR salvo quizá R son fijados, luego R también. Por consiguiente f es la identidad.

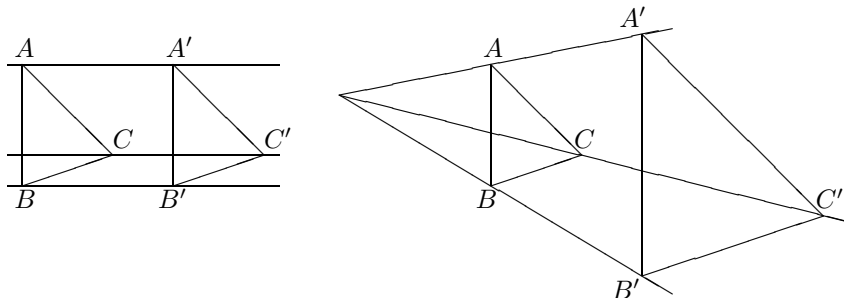
Si f tiene un único punto fijo O , tomemos un punto cualquiera $P \neq O$. Sea $P' = f(P)$, que es un punto distinto de O y P y colineal con ellos. Si Q es cualquier punto exterior a la recta OP , entonces $Q' = f(Q)$ cumple que las rectas QQ' y PP' se cortan en O y las rectas PQ y $P'Q'$ son paralelas. Por el teorema anterior $f(Q) = H(O, k)(Q)$, donde $k = \overrightarrow{OP'}/\overrightarrow{OP}$. Esto vale en principio para todo Q fuera de la recta OP , pero razonando con otro punto P fuera de esta recta, también vale para ella.

Si f no tiene puntos fijos, consideremos un punto cualquiera P y su imagen $P' = f(P)$. Para cualquier punto Q exterior a la recta PP' el punto $Q' = f(Q)$ cumple que PQ es paralela a $P'Q'$. Además PP' es paralela a QQ' , pues si ambas rectas tuvieran un punto en común, éste sería un punto fijo de f . Por el teorema anterior $f(Q) = T_{\vec{v}}(Q)$, donde $\vec{v} = \overrightarrow{PP'}$. En principio esto vale para todo punto Q fuera de la recta PP' , pero partiendo de otra recta, también vale para ella. ■

Ejercicio: Probar que los grupos $H(O, E)$ y $H(O', E)$ son conjugados en $\text{HT}(E)$ mediante la traslación de vector $\overrightarrow{OO'}$.

Veamos ahora una consecuencia muy importante del teorema de Tales:

Teorema 6.9 (Teorema de Desargues) *En un espacio afín, sean \widehat{ABC} y $\widehat{A'B'C'}$ dos triángulos con $A \neq A'$, $B \neq B'$, $C \neq C'$ y tales que las rectas AB , AC , BC sean paralelas (o iguales) respectivamente a $A'B'$, $A'C'$ y $B'C'$. Entonces las rectas AA' , BB' y CC' son paralelas o concurrentes.*



DEMOSTRACIÓN: Notar que no se exige que los triángulos estén contenidos en el mismo plano. Basta probar que si dos de las tres rectas AA' , BB' , CC' se cortan en un punto, la tercera pasa por dicho punto (si dos de las rectas coinciden el teorema se cumple trivialmente, las tres rectas no pueden coincidir). Supongamos que AA' y CC' concurren en O . Si $A = O$ las rectas AC y $A'C'$ no serían paralelas. En general, O ha de ser distinto de los seis vértices. Por el teorema de Tales $\overrightarrow{OA}/\overrightarrow{OA'} = \overrightarrow{OC}/\overrightarrow{OC'} = k$. Como $A \neq A'$ ha de ser $k \neq 1$. Sea $f = H(O, k)$ y $B'' = f(B)$. Por el teorema 6.7 tenemos que B'' está en la paralela a AB que pasa por A' , es decir, en $A'B'$, y por el mismo motivo está en $B'C'$, luego ha de ser $B'' = B'$. Esto prueba que O , B y B' están alineados (de hecho los triángulos son homotéticos). ■

Ejercicio: Refinar el teorema anterior (analizando los casos que no ha hecho falta analizar) para concluir que bajo las hipótesis del teorema de Desargues los triángulos son trasladados u homotéticos.

El teorema de Desargues es equivalente a su recíproco, por lo que también éste es conocido como teorema de Desargues:

Teorema 6.10 (Teorema de Desargues) *En un espacio afín, sean \widehat{ABC} y $\widehat{A'B'C'}$ dos triángulos con $A \neq A'$, $B \neq B'$, $C \neq C'$ y tales que las rectas AA' , BB' y CC' sean paralelas (o iguales) o concurrentes en un punto O distinto de todos los vértices. Si los triángulos tienen dos pares de lados paralelos, también el tercer par lo es.*

DEMOSTRACIÓN: Supongamos que AB es paralela a $A'B'$ y que AC es paralela a $A'C'$. No puede ser que ambos pares de rectas sean coincidentes (o sería $A = A'$) supongamos que las rectas AB y $A'B'$ son distintas y sea B'' el punto de corte entre $A'B'$ y la recta paralela a BC por C' . Entonces $B \neq B''$ y podemos aplicar el teorema anterior para concluir que las rectas AA' , BB'' , CC' son paralelas o concurrentes, lo que conjuntamente con la hipótesis implica que $BB'' = BB'$, luego $B' = B''$, pues ambos son el punto de corte de BB'' con $A'B'$. ■

6.3 El teorema fundamental de la geometría afín

Vamos a probar un teorema que muestra claramente el papel exacto que juegan las biyecciones afines en el marco de la geometría afín, al menos en el caso de espacios reales. Se trata del siguiente:

Teorema 6.11 (Teorema fundamental de la geometría afín) *Si E y E' son dos espacios afines reales de la misma dimensión $n \geq 2$ y $f : E \rightarrow E'$ es una biyección que transforma puntos colineales en puntos colineales, entonces f es una biyección afín.*

Así pues, toda aplicación que conserve la colinealidad conserva de hecho todas las propiedades afines, es decir, todas las propiedades que se conservan por biyecciones afines. En el capítulo siguiente interpretaremos este hecho mostrando que toda la geometría afín puede construirse a partir del mero concepto de colinealidad.

DEMOSTRACIÓN: Dividiremos la prueba en varios pasos.

1. Si $A = \{P_0, \dots, P_m\}$ es un conjunto de puntos de E afínmente independientes, entonces $f[\langle A \rangle] \subset \langle f[A] \rangle$.

Lo probamos por inducción sobre m . Para $m = 1$ se trata de la hipótesis. Supongámoslo cierto para conjuntos con m puntos y tomemos $P \in \langle A \rangle$. Usando coordenadas baricéntricas tenemos

$$P = \lambda_0 P_0 + \dots + \lambda_m P_m, \quad \text{con } \lambda_0 + \dots + \lambda_m = 1.$$

Puesto que $m \geq 1$ hay al menos un $\lambda_i \neq 1$. Supongamos por ejemplo que se trata de λ_m . Entonces

$$P = (1 - \lambda_m)P' + \lambda_m P_m,$$

donde

$$P' = \frac{1}{1 - \lambda_m}(\lambda_0 P_0 + \dots + \lambda_{m-1} P_{m-1}) \in \langle P_0, \dots, P_{m-1} \rangle.$$

Entonces $f(P)$ está en la recta $f(P')f(P_m)$ y $f(P')$ está en la variedad $\langle f(P_0), \dots, f(P_{m-1}) \rangle$ por hipótesis de inducción. Así pues, $f(P) \in \langle f[A] \rangle$.

2. Las imágenes por f de $m + 1$ puntos afínmente independientes son afínmente independientes

Podemos completar el conjunto hasta un sistema de referencia afín y suponer, por lo tanto, que $m = n$. Si A es un conjunto de $n + 1$ puntos afínmente independientes, por el apartado anterior $E' = f[\langle A \rangle] \subset \langle f[A] \rangle$, luego los puntos de $f[A]$ han de ser independientes.

3. f transforma cada recta PQ en la recta $f(P)f(Q)$.

La hipótesis del teorema nos da una inclusión. Sea R' un punto de $f(P)f(Q)$. Como f es biyectiva $R' = f(R)$ para un cierto R . Si R no

estuviera en PQ entonces los puntos P, Q, R serían afinmente libres, luego por el punto anterior lo mismo ocurriría con $f(P), f(Q), f(R)$, lo cual es una contradicción.

4. *f transforma rectas paralelas en rectas paralelas.*

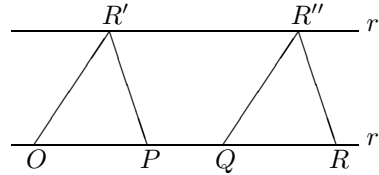
Sean PQ y $P'Q'$ rectas paralelas distintas en E . Entonces los puntos P, Q, P' son afinmente independientes y generan el plano que contiene a las dos rectas. Por el paso 1 tenemos que $f(P), f(Q), f(P')$ son afinmente independientes, luego generan un plano que contiene a $f(Q')$ por 2, luego de hecho contiene a las rectas $f[PQ]$ y $f[P'Q']$. Como f es biyectiva estas rectas no tienen puntos comunes, luego son paralelas.

5. *Para todos los puntos $O, P, Q \in E$, se cumple*

$$f(P + \overrightarrow{OQ}) = f(P) + \overrightarrow{f(O)f(Q)}.$$

Sea $R = P + \overrightarrow{OQ}$. Si O, P, Q no están alineados, lo mismo sucede con $f(P), f(Q), f(R)$. Por el teorema 6.7 tenemos que R es la intersección de la paralela a OP por Q y la paralela a OQ por P . Por consiguiente $f(R)$ es la intersección de la paralela a $f(O)f(P)$ por $f(Q)$ y la paralela a $f(O)f(Q)$ por $f(P)$. De nuevo por el teorema 6.7 podemos concluir que $f(R) = f(O) + \overrightarrow{f(O)f(Q)}$.

Si O, P, Q están contenidos en una recta r , tomamos una paralela cualquiera r' y en ella un punto R' . La paralela a OR' por Q corta a r' en un punto $R'' = R' + \overrightarrow{OQ}$ y la paralela a PR' por R'' corta a r en $P + \overrightarrow{OQ} = R$. Al aplicar f a todos estos puntos y rectas se conservan las relaciones de incidencia y paralelismo, luego tenemos que $f(R'') = f(R') + \overrightarrow{f(O)f(Q)}$ y $f(R) = f(P) + \overrightarrow{f(O)f(Q)}$.



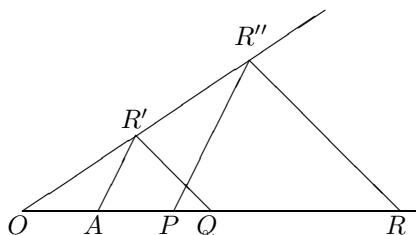
6. *Fijemos una recta r y sean O, A dos de sus puntos. Para cada $\alpha \in \mathbb{R}$ sea $P = O + \alpha \overrightarrow{OA}$ y sea $\sigma(\alpha) = \overrightarrow{f(O)f(P)} / \overrightarrow{f(O)f(A)}$, Entonces $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ es un automorfismo de cuerpos.*

La figura anterior ilustra que σ conserva la suma. En efecto, si suponemos $P = O + \alpha \overrightarrow{OA}$, $Q = O + \beta \overrightarrow{OA}$, entonces $R'' = R' + \beta \overrightarrow{OA}$, y de aquí que $R = P + \beta \overrightarrow{OA} = O + (\alpha + \beta) \overrightarrow{OA}$.

Por definición de σ tenemos que $f(P) = f(O) + \sigma(\alpha) \overrightarrow{f(O)f(A)}$, $f(Q) = f(O) + \sigma(\beta) \overrightarrow{f(O)f(A)}$, $f(R) = f(O) + \sigma(\alpha + \beta) \overrightarrow{f(O)f(A)}$, y como f

conserva la figura, también tenemos que $f(R) = (\sigma(\alpha) + \sigma(\beta)) \overrightarrow{f(O)f(A)}$, luego $\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta)$.

Para el producto empleamos un argumento similar basado en la construcción que muestra la figura siguiente:



Se prueba sin dificultad que si $P = O + \alpha \overrightarrow{OA}$ y $Q = O + \beta \overrightarrow{OB}$, entonces $R = O + (\alpha\beta) \overrightarrow{OA}$. Por definición de σ tenemos que $f(R) = \sigma(\alpha\beta) \overrightarrow{f(O)f(A)}$, y como f conserva la figura $f(R) = \sigma(\alpha)\sigma(\beta) \overrightarrow{f(O)f(A)}$, luego $\sigma(\alpha\beta) = \sigma(\alpha)\sigma(\beta)$. El carácter biyectivo de σ es inmediato.

7. *El único automorfismo de \mathbb{R} es la identidad.*

En efecto, si σ es un automorfismo de \mathbb{R} , entonces σ envía cuadrados a cuadrados, luego $\alpha > 0$ si y sólo si $\sigma(\alpha) > 0$. A su vez esto implica que $\alpha < \beta$ si y sólo si $\sigma(\alpha) < \sigma(\beta)$. Por otro lado σ ha de fijar a \mathbb{Q} , y una biyección de \mathbb{R} en \mathbb{R} que conserve el orden y fije a los números racionales es necesariamente la identidad.

8. *f es una afinidad.*

Fijemos un punto $O \in E$. Sea $\vec{f}(\vec{v}) = \overrightarrow{f(O)f(O+\vec{v})}$. Por 5 se cumple que $f(O+\overrightarrow{OP}+\overrightarrow{OQ}) = f(O+\overrightarrow{OP}) + \overrightarrow{f(O)f(Q)} = f(O) + \overrightarrow{f(O)f(P)} + \overrightarrow{f(O)f(Q)}$, luego

$$\vec{f}(\overrightarrow{OP} + \overrightarrow{OQ}) = \vec{f}(\overrightarrow{OP}) + \vec{f}(\overrightarrow{OQ}).$$

Por 7 tenemos que σ es la identidad, luego por definición de σ

$$\vec{f}(\alpha \overrightarrow{OA}) = \alpha \vec{f}(\overrightarrow{OA}).$$

Por consiguiente \vec{f} es una aplicación lineal, y $f(P) = f(O) + \vec{f}(\overrightarrow{OP})$, luego f es afín. ■

Observar que la prueba es válida para espacios afines sobre cualquier subcuerpo K de \mathbb{R} en el que todo número positivo tenga una raíz cuadrada. En particular es válido para los modelos de geometría discontinua que comentamos en el capítulo anterior.

Ejercicio: Probar que el teorema anterior es falso para espacios afines complejos.

6.4 Isometrías y semejanzas

Nos ocupamos ahora de los espacios euclídeos. Recordemos del capítulo IV que una isometría de un espacio afín euclídeo E es una afinidad biyectiva cuya aplicación lineal asociada es una isometría lineal. Recordemos a su vez que las isometrías lineales son los automorfismos \vec{f} de \vec{E} tales que

$$\vec{f}(\vec{v})\vec{f}(\vec{w}) = \vec{v}\vec{w}, \quad \text{para todo } \vec{v}, \vec{w} \in \vec{E},$$

lo cual a su vez equivale a que $\|\vec{f}(\vec{v})\| = \|\vec{v}\|$ para todo $\vec{v} \in \vec{E}$, debido a la relación

$$\|\vec{v} - \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2 - 2\vec{v}\vec{w}.$$

Otra caracterización importante es que una aplicación lineal $\vec{f}: \vec{E} \rightarrow \vec{E}$ es una isometría si y sólo si transforma una base ortonormal en una base ortonormal, pues si conserva el producto escalar sobre una base, por linealidad lo conserva sobre vectores arbitrarios. De aquí se sigue a su vez otra caracterización importante que no vimos en su momento.

Definición 6.12 Diremos que una matriz cuadrada A en un cuerpo K es *ortogonal* si cumple $AA^t = I_n$, donde A^t es la matriz traspuesta de A y I_n es la matriz identidad. Así, A es ortogonal si su inversa es su traspuesta.

Veamos la caracterización de las isometrías lineales a la que hacíamos referencia:

Teorema 6.13 Sea E un espacio vectorial euclídeo y $f: E \rightarrow E$ una aplicación lineal. Entonces f es una isometría lineal si y sólo si la matriz de f en una base ortonormal es ortogonal (y en tal caso la matriz de f en cualquier base ortonormal es ortogonal).

DEMOSTRACIÓN: Fijada una base ortonormal de E , la aplicación que a cada vector le asigna sus coordenadas respecto a esta base es una isometría entre E y \mathbb{R}^n . Sea A la matriz de f en la base fijada. Entonces f será una isometría si y sólo si las imágenes de los vectores de la base forman una base ortonormal, si y sólo si sus coordenadas (esto es, las filas de A) son una base ortonormal en \mathbb{R}^n . Ahora observamos que el elemento (i, j) de AA^t es el producto escalar de la fila i -ésima de A por la fila j -ésima de A , luego las filas de A son una base ortonormal de \mathbb{R}^n si y sólo si $AA^t = I_n$. ■

Definición 6.14 Dado un espacio vectorial euclídeo V sobre un cuerpo K , llamaremos *grupo ortogonal* de V , representado por $O(V)$, al conjunto de todas las isometrías lineales de V , que es un grupo con la composición de aplicaciones. Si V tiene dimensión n , entonces $O(V)$ es isomorfo al grupo $O(n, K)$ de las matrices ortogonales $n \times n$ con coeficientes en K .

Dado un espacio afín euclídeo E , llamaremos $\text{Is}(E)$ al grupo de las isometrías de E , también con la composición de aplicaciones.

Es claro que la aplicación

$$\begin{array}{ccc} \text{Is}(E) & \longrightarrow & \text{O}(\vec{E}) \\ f & \mapsto & \vec{f} \end{array}$$

es un epimorfismo de grupos cuyo núcleo es el grupo de las traslaciones $T(E)$.

Definición 6.15 Sea E un espacio afín euclídeo. Una biyección $f : A \longrightarrow B$ entre dos subconjuntos de E es una *semejanza* de razón $k > 0$ si para todo par de puntos $P, Q \in A$ se cumple

$$\|\overrightarrow{f(P)f(Q)}\| = k\|\overrightarrow{PQ}\|.$$

Si existe una semejanza entre A y B se dice que son *semejantes*, y cada par de puntos $P, f(P)$ correspondientes por una semejanza f se llaman puntos *homólogos* (respecto a f).

Es claro que la inversa de una semejanza de razón k es una semejanza de razón $1/k$ y que la composición de semejanzas es una semejanza de razón igual al producto de las razones. Las semejanzas de razón 1 son las isometrías

Notemos que dos triángulos son semejantes en el sentido usual si y sólo si lo son en el sentido general que acabamos de definir.

Es obvio que una homotecia de razón k es una semejanza de razón $|k|$. Por consiguiente, si f es una semejanza de razón k entre dos subconjuntos A y B de un espacio afín euclídeo E , al componerla con una homotecia g de razón $1/k$ obtenemos una semejanza de razón 1, es decir, una isometría. Por el teorema 4.22, $f \circ g$ se extiende a una isometría de E en sí mismo, llamémosla h . Entonces $h \circ g^{-1}$ es una semejanza de E en sí mismo que extiende a f . Esto prueba el teorema siguiente:

Teorema 6.16 *Toda semejanza de razón k entre dos subconjuntos de un espacio afín euclídeo E se extiende a una semejanza de razón k de E en sí mismo. Toda semejanza de razón k de E en sí mismo es una afinidad y se puede expresar como composición de una isometría y una homotecia de razón k .*

Definición 6.17 Llamaremos $\text{Sem}(E)$ al grupo de las semejanzas de un espacio afín euclídeo E en sí mismo.

Hemos probado que $\text{Sem}(E) \leq \text{GA}(E)$. La aplicación que a cada semejanza le asigna su razón es un epimorfismo de $\text{Sem}(E)$ en $]0, +\infty[$ cuyo núcleo es claramente $\text{Is}(E)$.

Vamos a caracterizar las semejanzas en términos de sus aplicaciones lineales asociadas. Si V es un espacio vectorial euclídeo, diremos que $f \in \text{GL}(V)$ es una *semejanza lineal* de razón $k > 0$ si $\|f(\vec{v})\| = k\|\vec{v}\|$ para todo $\vec{v} \in V$. Llamaremos $\text{Sem}(V)$ al grupo de las semejanzas lineales de V .

Es claro que una afinidad f es una semejanza si y sólo si \vec{f} es una semejanza lineal. Tenemos, pues, un epimorfismo $\text{Sem}(E) \longrightarrow \text{Sem}(\vec{E})$ cuyo núcleo es el

grupo $T(E)$ de las traslaciones de E . Por otra parte, el hecho de que toda semejanza de E se descomponga en producto de una isometría y una homotecia se traduce, a través de este epimorfismo, en que toda semejanza lineal es producto de una isometría lineal y una homotecia lineal. Más aún, es fácil ver que si llamamos $H_+(\vec{E})$ al grupo de las homotecias lineales de \vec{E} de razón positiva, se cumple

$$\text{Sem}(\vec{E}) = O(\vec{E}) \times H_+(\vec{E}).$$

Para ello hay que probar que los dos subgrupos de la derecha son normales, pero ello se sigue inmediatamente del hecho de que las homotecias lineales conmutan con todos los elementos de $GL(\vec{E})$.

Ejercicio: Probar que $HT(E)$ es un subgrupo normal de $\text{Sem}(E)$ y

$$\text{Sem}(E)/HT(E) \cong \text{Sem}(\vec{E})/H(\vec{E}).$$

Una propiedad interesante de las semejanzas propiamente dichas, es decir, las que no son isometrías, es que siempre tienen un único punto fijo.

Teorema 6.18 *Toda semejanza de razón $k \neq 1$ en un espacio afín euclídeo tiene un único punto fijo llamado centro de la semejanza.*

DEMOSTRACIÓN: Es consecuencia inmediata del teorema 6.2, pues si f cumple las hipótesis entonces \vec{f} es una semejanza lineal de razón $k \neq 1$, y es claro que su único punto fijo es 0. ■

En general, las semejanzas lineales no conservan las normas, luego tampoco el producto escalar. Sin embargo, el hecho de que toda semejanza lineal f de razón k se exprese como composición de una isometría lineal por una homotecia lineal de razón k implica la relación $f(\vec{v})f(\vec{w}) = k^2\vec{v}\vec{w}$. Esto implica que las semejanzas lineales conservan ángulos, pues

$$\cos \widehat{f(\vec{v})f(\vec{w})} = \frac{f(\vec{v})f(\vec{w})}{\|f(\vec{v})\|\|f(\vec{w})\|} = \frac{k^2\vec{v}\vec{w}}{k^2\|\vec{v}\|\|\vec{w}\|} = \cos \widehat{\vec{v}\vec{w}}.$$

En particular las semejanzas lineales conservan la ortogonalidad. Vamos a probar el recíproco, es decir, que las únicas aplicaciones lineales que conservan la ortogonalidad (o *a fortiori* los ángulos) son las semejanzas lineales. De aquí deduciremos el análogo para las semejanzas afines.

Teorema 6.19 *Sea V un espacio vectorial euclídeo y $f : V \rightarrow V$ una aplicación lineal no nula. Entonces son equivalentes:*

1. f es una semejanza lineal.
2. f conserva la ortogonalidad de vectores.
3. f conserva los ángulos entre vectores.

DEMOSTRACIÓN: Es suficiente probar que 2 implica 1. Suponemos, pues, que si $\vec{v}\vec{w} = 0$ entonces $f(\vec{v})f(\vec{w}) = \vec{0}$ (notar que no suponemos la doble implicación).

Fijado un vector $\vec{v} \in V$, la aplicación $u(\vec{w}) = f(\vec{v})f(\vec{w})$ es lineal. Supongamos que existe un \vec{w} tal que $u(\vec{w}) \neq 0$. Entonces el núcleo de u tiene dimensión $n - 1$ (donde n es la dimensión de V), y por hipótesis contiene a $\langle \vec{v} \rangle^\perp$, cuya dimensión es la misma. Por consiguiente u se anula exactamente sobre los vectores ortogonales a \vec{v} .

Definimos $k(\vec{v})$ como el único escalar que cumple

$$u(\vec{w}) = f(\vec{v})f(\vec{w}) = k(\vec{v})\vec{v}\vec{w}.$$

Todo vector de V se expresa como $\vec{t} + \alpha\vec{w}$, donde $\vec{t} \perp \vec{v}$, con lo que la relación anterior vale en realidad no sólo para el \vec{w} que hemos tomado sino de hecho para todo $\vec{w} \in V$. Si u es nula la relación sigue siendo válida sin más que tomar $k(\vec{v}) = 0$.

Por simetría, si $\vec{v}\vec{w} \neq 0$ se ha de cumplir $k(\vec{v}) = k(\vec{w})$. Si $\vec{v}\vec{w} = 0$ con ambos factores no nulos, entonces $\vec{y} = \vec{v} + \vec{w}$ cumple que $\vec{v}\vec{y} \neq 0 \neq \vec{w}\vec{y}$, luego $k(\vec{v}) = k(\vec{y}) = k(\vec{w})$. Esto prueba que $k = k(\vec{v})$ es independiente de $\vec{v} \neq \vec{0}$. Además

$$\|f(\vec{v})\|^2 = k\|\vec{v}\|^2, \quad \text{para todo } \vec{v} \in V.$$

Como existen vectores con imagen no nula, ha de ser $k > 0$, de donde f es una semejanza de razón \sqrt{k} . ■

Teorema 6.20 *Sea E un espacio afín euclídeo y $f : E \longrightarrow E$ una aplicación biyectiva. Las afirmaciones siguientes son equivalentes:*

1. f es una semejanza.
2. Para toda terna de puntos distintos se cumple $\widehat{ABC} = f(A)\widehat{f(B)f(C)}$.
3. Para todos los puntos $A \neq B$, $C \neq D$, si $AB \perp CD$, entonces

$$f(A)f(B) \perp f(C)f(D).$$

DEMOSTRACIÓN: Usando el teorema anterior es fácil ver que $1 \rightarrow 2 \rightarrow 3$. Para probar $3 \rightarrow 1$ basta demostrar que f es afín, y por el teorema fundamental esto equivale a probar que f transforma puntos alineados en puntos alineados.

Sean A, B, C tres puntos alineados. Podemos formar una base ortogonal de \vec{E} de la forma $\overrightarrow{AA_i}$, para $i = 1, \dots, n$ con $A_1 = B$. Por hipótesis los vectores $\overrightarrow{f(A)f(A_i)}$ son ortogonales dos a dos, luego son una base ortogonal de \vec{E} . Podemos expresar

$$\overrightarrow{f(A)f(C)} = \lambda_1 \overrightarrow{f(A)f(A_1)} + \dots + \lambda_n \overrightarrow{f(A)f(A_n)}. \quad (6.1)$$

Ahora bien, si $i \neq 1$ entonces $AA_i \perp AC$, luego también $f(A)f(A_i) \perp f(A)f(C)$. Multiplicando (6.1) por $\overrightarrow{f(A)f(A_i)}$ obtenemos $\lambda_i = 0$, luego (6.1) se reduce a $\overrightarrow{f(A)f(C)} = \lambda_1 \overrightarrow{f(A)f(A_1)}$, lo que significa que $f(A), f(A_1) = f(B)$ y $f(C)$ están alineados. ■

6.5 Clasificación de endomorfismos

En esta sección desarrollaremos una potente teoría algebraica que nos dará un gran control sobre los endomorfismos de un espacio vectorial y, de aquí, sobre las afinidades e isometrías de los espacios afines. El problema principal que nos proponemos resolver consiste en que a veces es fácil reconocer que una isometría lineal es, por ejemplo, un giro en \mathbb{R}^3 , si conocemos su matriz en una base determinada, mientras que es irreconocible si se nos da su matriz en otra base. Vamos a buscar criterios que nos permitan determinar cuándo dos matrices pueden corresponder a la misma aplicación escogiendo bases adecuadas así como técnicas que nos permitan encontrar la expresión más simple posible para la matriz de una aplicación lineal.

Por razones que pronto se comprenderán, conviene que comencemos nuestro estudio en un contexto muy general. Consideraremos homomorfismos entre módulos libres sobre un dominio euclídeo. Haremos uso de los teoremas de estructura de módulos: todo módulo M finitamente generado sobre un dominio de ideales principales³ A se descompone como $M = L \oplus M_t$, donde L es un submódulo libre (cuyo rango sólo depende de M) y M_t es el *submódulo de torsión*, formado por todos los elementos m tales que $am = 0$ para cierto $a \in A$ no nulo. Cada elemento m de M_t tiene definido (salvo unidades) un orden $o(m) \in A$ de modo que $am = 0$ si y sólo si $o(m) \mid a$. A su vez es posible descomponer $M_t = \langle v_1 \rangle \oplus \cdots \oplus \langle v_n \rangle$, donde podemos exigir que cada $o(v_i)$ sea potencia de primo (y entonces se llaman *divisores elementales* de M) o bien que $o(v_i) \mid o(v_{i+1})$ (y entonces se llaman *factores invariantes* de M). Los divisores elementales y los factores invariantes están unívocamente determinados por M salvo unidades.

Supongamos que $f : M \longrightarrow N$ es un homomorfismo entre A -módulos libres de rangos m y n respectivamente y que B y B' son bases respectivas. Sea S la matriz de f en estas bases. Supongamos que f tiene otra matriz T respecto a otras bases C y C' . ¿Cuál es entonces la relación entre S y T ? Es sencilla: Sea P la matriz de la identidad en M respecto de las bases C y B (la llamada matriz del cambio de base). Sea Q la matriz de la identidad en N respecto de las bases C' y B' . Entonces si x es la m -tupla de coordenadas en C de un elemento $m \in M$, tendremos que xP es la m -tupla de coordenadas de m en la base B , luego xPS es la n -tupla de coordenadas de $f(m)$ en B' , luego $xPSQ$ es la n -tupla de coordenadas de $f(m)$ en C' . Por la unicidad de la matriz, se ha de cumplir $T = PSQ$. Además como P y Q son matrices de isomorfismos, ambas son regulares (tienen inversa). En vista de esto definimos:

Definición 6.21 Sea A un anillo conmutativo y unitario. Diremos que dos matrices $S, T \in \text{Mat}_{m \times n}(A)$ son *equivalentes* si existen matrices $P \in \text{GL}(m, A)$ y $Q \in \text{GL}(n, A)$ tales que $T = PSQ$.

Es evidente que la equivalencia de matrices es una relación de equivalencia en el conjunto $\text{Mat}_{m \times n}(A)$. Acabamos de demostrar que si S y T son matrices

³En realidad nos bastará trabajar con dominios euclídeos.

de un mismo homomorfismo entre A -módulos libres $f : M \longrightarrow N$, entonces S y T son equivalentes. Recíprocamente, si S es la matriz de f en ciertas bases y T es equivalente a S , es fácil ver que T es la matriz de f en otras bases adecuadas (las que convierten a P y Q en las matrices del cambio de base).

Ahora vamos a dar un criterio sencillo para determinar cuándo dos matrices dadas son o no equivalentes, al tiempo que encontraremos para cada matriz una equivalente lo más sencilla posible. Para ello usaremos el resultado siguiente:

Teorema 6.22 *Sea A un anillo conmutativo y unitario y sean S y T dos matrices $m \times n$ en A . Si T resulta de realizar sobre S una de las siguientes operaciones, entonces T es equivalente a S :*

1. *Permutar dos filas (o columnas).*
2. *Cambiar la fila (o columna) j -ésima por la suma de la fila (columna) j -ésima más la fila (columna) k -ésima ($k \neq j$) multiplicada por una constante r .*
3. *Multiplicar una fila (columna) por una unidad u de A .*

DEMOSTRACIÓN: Realizar una de estas operaciones equivale a multiplicar la matriz por la izquierda o por la derecha (según sea sobre filas o columnas) por una de las siguientes matrices regulares (todas tienen unos en la diagonal y ceros fuera de ella salvo cuando se indica explícitamente lo contrario):

$$\begin{array}{c}
 \begin{array}{cc} & \begin{array}{cc} j & k \end{array} \\ j & \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 0 & & 1 \\ & & & \ddots & \\ k & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{array} , & \begin{array}{cc} & k \\ j & \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \ddots & & r \\ & & & \ddots & \\ & & & & \ddots & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{array} , \\
 \\
 & \begin{array}{cc} & j \\ j & \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & u & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \end{array} .
 \end{array}$$

■

El teorema siguiente nos resuelve el problema de determinar cuándo dos matrices son equivalentes:

Teorema 6.23 Sea E un dominio euclídeo y $A \in \text{Mat}_{m \times n}(E)$. Entonces A es equivalente a una única matriz de la forma

$$\begin{pmatrix} d_1 & & & & & \\ & \ddots & & & & \\ & & d_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

donde cada $d_i \neq 0$ y $d_i \mid d_{i+1}$. La unicidad se entiende salvo sustitución de los d_i por asociados (o sea, salvo unidades). Los elementos d_i se llaman factores invariantes de A .

DEMOSTRACIÓN: La prueba que vamos a ver nos da un algoritmo para calcular los factores invariantes de cualquier matriz dada. Llamemos ϕ a la norma euclídea del anillo E .

Si $A = 0$ ya es de la forma requerida. En otro caso, sea a_{ij} un coeficiente de A no nulo con norma mínima. Intercambiando filas y columnas podemos llevarlo a la posición $(1, 1)$, es decir, pasamos a una matriz equivalente donde $a_{11} \neq 0$ tiene norma mínima.

Para cada $k > 1$ dividimos $a_{1k} = a_{11}b_k + b_{1k}$, de modo que $b_{1k} = 0$ o bien $\phi(b_{1k}) < \phi(a_{11})$.

Restamos a la columna k -ésima la primera multiplicada por b_k , con lo que la primera fila se convierte en $(a_{11}, b_{12}, \dots, b_{1n})$.

Si algún b_{1k} es no nulo llevamos a la posición $(1, 1)$ el de menor norma y repetimos el proceso. Como cada vez la norma del coeficiente $(1, 1)$ disminuye, el proceso no puede continuar indefinidamente, por lo que al cabo de un número finito de pasos llegaremos a una primera fila de la forma $(a_{11}, 0, \dots, 0)$.

Repetiendo el proceso con la primera columna llegamos a una matriz equivalente de la forma:

$$\left(\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array} \right)$$

Si a_{11} no divide a alguno de los restantes coeficientes a_{ij} , entonces hacemos $a_{ij} = a_{ij}c + d$ con $d \neq 0$ y $\phi(d) < \phi(a_{11})$, llevamos d a la posición $(1, 1)$ y repetimos el proceso de hacer ceros. Como la norma sigue disminuyendo, tras un número finito de pasos llegaremos a una matriz como la anterior y en la que a_{11} divide a todos los coeficientes restantes.

Ahora repetimos el proceso con la matriz B , lo cual no altera los ceros de la fila y la columna primera ni el hecho de que a_{11} divide a todos los coeficientes. De este modo llegamos a una matriz como la del enunciado.

Probemos que si dos matrices como la del enunciado son equivalentes, entonces son iguales (salvo multiplicación de sus coeficientes por unidades).

Sea A una tal matriz. Sea $f : E^m \longrightarrow E^n$ el homomorfismo que en ciertas bases, digamos $\{e_1, \dots, e_m\}$ y $\{f_1, \dots, f_n\}$, tiene matriz A , es decir, $f(e_i) = d_i f_i$ para $i = 1, \dots, r$ y $f(e_i) = 0$ en otro caso.

Entonces $\text{Im } f = \langle d_1 f_1 \rangle \oplus \dots \oplus \langle d_r f_r \rangle$, y

$$\begin{aligned} E^n / \text{Im } f &= \langle f_1 \rangle \oplus \dots \oplus \langle f_n \rangle / (\langle d_1 f_1 \rangle \oplus \dots \oplus \langle d_r f_r \rangle \oplus 0 \oplus \dots \oplus 0) \\ &\cong \langle f_1 \rangle / \langle d_1 f_1 \rangle \oplus \dots \oplus \langle f_r \rangle / \langle d_r f_r \rangle \oplus \langle f_{r+1} \rangle \oplus \dots \oplus \langle f_n \rangle. \end{aligned}$$

Resulta que $E^n / \text{Im } f$ es un módulo finitamente generado sobre un dominio euclídeo con rango $n - r$ (el rango de la parte libre) y, respecto al submódulo de torsión, está generado por los elementos $f_i + \langle d_i f_i \rangle$ de orden claramente igual a d_i . Por lo tanto d_1, \dots, d_r son los factores invariantes del módulo $E^n / \text{Im } f$.

Si otra matriz B del mismo tipo es equivalente a A , entonces B es la matriz de f en otras bases, luego sus coeficientes son los factores invariantes del mismo módulo $E^n / \text{Im } f$, luego son los mismos salvo unidades. ■

Así pues, dos matrices son equivalentes si y sólo si tienen los mismos factores invariantes. La matriz equivalente a una matriz dada A que tiene la forma indicada en el teorema anterior se llama *forma canónica* de A , de modo que dos matrices son equivalentes si y sólo si tienen la misma forma canónica.

En el caso en que el anillo es un cuerpo la situación es mucho más simple. Como todos los elementos no nulos son unidades, todos los factores invariantes de una matriz A pueden tomarse iguales a 1, luego lo único que puede variar es su número r . Este número se llama *rango* de A . Tenemos, pues, que dos matrices sobre un cuerpo son equivalentes si y sólo si tienen el mismo rango.

El rango de A tiene una interpretación muy sencilla: sea f una aplicación lineal de matriz A . Las filas de A son las coordenadas de las imágenes de los vectores de la primera base respecto a la segunda base. Estas imágenes generan el subespacio $\text{Im } f$, luego contienen exactamente $\dim \text{Im } f$ vectores independientes. Como la aplicación que a cada vector le asigna sus coordenadas es un isomorfismo, resulta que A tiene $\dim \text{Im } f$ filas independientes. La forma canónica de A es también una matriz de f y tiene r filas independientes, luego $\dim \text{Im } f = r$ y el rango de una matriz es el número de filas independientes que contiene.

Por otra parte es obvio que si dos matrices son equivalentes sus traspuestas también lo son, y la traspuesta de una forma canónica es ella misma, luego el rango de una matriz es el mismo que el de su traspuesta. Por lo tanto:

El rango de una matriz A con coeficientes en un cuerpo es el número de filas y el número de columnas linealmente independientes.

Con esto tenemos clasificadas las matrices de los homomorfismos entre módulos libres, es decir, sabemos reconocer si dos matrices corresponden al mismo homomorfismo en bases distintas y a cada homomorfismo le sabemos encontrar una matriz especialmente simple (en forma canónica). Nuestro objetivo es obtener una teoría análoga para endomorfismos de un espacio vectorial, donde ahora imponemos la condición de que no queremos considerar dos bases para

el mismo espacio, sino una sola. Si $f : M \rightarrow M$ es un endomorfismo de un módulo libre M y S, T son las matrices de f respecto a ciertas bases B y C , el mismo razonamiento que cuando teníamos dos módulos nos da ahora que $T = P^{-1}SP$, para cierta matriz regular P (P es la matriz del cambio de base de B a C y P^{-1} es la matriz del cambio de base de C a B). Por ello definimos:

Definición 6.24 Sea A un anillo conmutativo y unitario. Diremos que dos matrices $S, T \in \text{Mat}_n(A)$ son *semejantes* si existe una matriz $P \in \text{GL}(n, A)$ tal que $T = P^{-1}SP$.

Es evidente que la semejanza de matrices es una relación de equivalencia en el conjunto $\text{Mat}_n(A)$. Dos matrices de un mismo endomorfismo de A -módulos libres son semejantes, y si S es la matriz de un endomorfismo f y T es semejante a S , entonces T es la matriz de f en otra base.

Dos matrices semejantes son equivalentes, pero el recíproco no es cierto. Por ejemplo, si $T = P^{-1}SP$, tomando determinantes, $|T| = |P|^{-1}|S||P| = |S|$, y es fácil encontrar matrices equivalentes con determinantes distintos, luego no semejantes.

La idea fundamental en el estudio de un endomorfismo de un espacio V es encontrarle subespacios invariantes, es decir, encontrar subespacios W tales que $h[W] \subset W$. Por ejemplo, si h es un giro en \mathbb{R}^3 (cuyo eje pasa por 0), podemos encontrar dos subespacios invariantes: el plano de giro, donde h es un giro en \mathbb{R}^2 , y el eje de giro, donde h es la identidad. La matriz de h será la más simple si tomamos una base que tenga dos vectores en el plano de giro y el tercero en el eje.

Para encontrar subespacios invariantes podemos valernos de la teoría de módulos gracias al planteamiento siguiente: Sea V un espacio vectorial de dimensión finita sobre un cuerpo K . Entonces el conjunto $\text{End}_K(V)$ de todos los endomorfismos de V tiene estructura de anillo (no conmutativo) con la suma definida punto a punto: $(f + g)(v) = f(v) + g(v)$ y tomando como producto la composición de aplicaciones. Si $h \in \text{End}_K(V)$ definimos el homomorfismo $K[x] \rightarrow \text{End}_K(V)$ que a cada polinomio $p(x)$ le asigna $p(h)$. Notar que la unidad de $\text{End}_K(V)$ es el endomorfismo identidad, por lo que la imagen del polinomio 1 es dicha identidad.

Ahora definimos una operación $K[x] \times V \rightarrow V$ dada por $p(x)v = p(h)(v)$. Así, por ejemplo, $xv = h(v)$, $(x^2 + 2)v = h(h(v)) + 2v$, etc.

Es fácil ver que V , con su suma de espacio vectorial y esta operación, es un $K[x]$ -módulo.

En resumen, hemos dotado a V de estructura de $K[x]$ -módulo de modo que la multiplicación por elementos de K es la que ya teníamos en V como espacio vectorial y la multiplicación por x consiste en aplicar h . Esto hace que los subespacios invariantes que estamos buscando sean precisamente los submódulos de V . El teorema siguiente recoge este hecho junto con los resultados que garantizan que podemos aplicar los teoremas de estructura de módulos (notemos que $K[x]$ es un dominio euclídeo).

Teorema 6.25 Sea V un espacio vectorial de dimensión finita n sobre un cuerpo K , sea h un endomorfismo de V . Entonces

1. V es un $K[x]$ -módulo finitamente generado.
2. Sus submódulos son los subespacios vectoriales W tales que $h[W] \subset W$.
3. El núcleo $N(h)$ es un submódulo de V .
4. V es un módulo de torsión.

DEMOSTRACIÓN: 1) Notar que el producto de un polinomio constante por un elemento de V es el producto dado de V como espacio vectorial. Por ello una combinación lineal con coeficientes en K también puede considerarse con coeficientes en $K[x]$, luego una base de V como espacio vectorial es un generador de V como módulo.

2) Si W es un submódulo de V , entonces la suma de elementos de W está en W y el producto de un escalar por un elemento de W está en W , luego W es un subespacio vectorial. Más aún, si $v \in W$, $xv = h(v) \in W$, luego $h[W] \subset W$.

Recíprocamente, si W cumple estas condiciones entonces W es estable para la suma y para el producto por escalares y por x , de donde fácilmente se sigue que W es estable para el producto por cualquier polinomio.

3) es consecuencia de 2).

4) Si $v \in V$, entonces los vectores $v, xv, \dots, x^n v$ han de estar repetidos o ser linealmente dependientes, luego existen escalares no todos nulos tales que $t_0 v + t_1 xv + \dots + t_n x^n v = 0$, o sea, $(t^n x^n + \dots + t_1 x + t_0)v = 0$, luego v es un elemento de torsión. ■

Los teoremas de estructura de módulos nos garantizan ahora que V se descompone en la forma

$$V = \langle v_1 \rangle_{K[x]} \oplus \dots \oplus \langle v_r \rangle_{K[x]},$$

donde la notación $\langle v \rangle_{K[x]}$ representa al submódulo generado por v , (mientras que $\langle v \rangle$ representará al subespacio vectorial generado por v).

Nuestra intención es describir h en términos de su acción sobre estos submódulos, que son subespacios invariantes donde la acción de h es la más simple posible. Vamos a probar que la acción de h sobre cada $\langle v_i \rangle_{K[x]}$ está determinada por un polinomio.

Definición 6.26 Sea V un K -espacio vectorial de dimensión finita y h un endomorfismo de V . Para cada $v \in V$ llamaremos *polinomio mínimo* de v (pol mín v) a $o(v)$, es decir, al polinomio que divide a todos los polinomios que anulan a v . En principio pol mín v está determinado salvo unidades, pero es único si lo tomamos mónico.

Veamos en primer lugar que pol mín v determina la dimensión como espacio vectorial del submódulo generado por v .

Teorema 6.27 Sea K un cuerpo y V un K -espacio vectorial de dimensión finita n . Sea h un endomorfismo de V . Entonces $\dim \langle v \rangle_{K[x]} = \text{grad pol m} \, v$.

DEMOSTRACIÓN: Sea $p(x) = \text{pol m} \, v$ y sea m su grado. Un elemento cualquiera de $\langle v \rangle_{K[x]}$ es de la forma $q(x)v$ con $q(x) \in K[x]$.

Dividamos $q(x) = c(x)p(x) + r(x)$, con $\text{grad } r(x) < m$. Entonces

$$q(x)v = c(x)(p(x)v) + r(x)v = c(x)0 + r(x)v = r(x)v, \quad (6.2)$$

que es combinación lineal de $v, xv, \dots, x^{m-1}v$.

Estos vectores han de ser distintos y linealmente independientes, pues lo contrario significa que hay un polinomio $q(x) \neq 0$ y de grado a lo sumo $m-1$ tal que $q(x)v = 0$, pero entonces $p(x) \mid q(x)$, lo cual es imposible según los grados. Por lo tanto $v, xv, \dots, x^{m-1}v$ es una base de $\langle v \rangle_{K[x]}$ como espacio vectorial. ■

Ahora ya estamos en condiciones de caracterizar la acción de h sobre un submódulo monógeno:

Teorema 6.28 Sea K un cuerpo, V un K -espacio vectorial de dimensión finita y h un endomorfismo de V . Son equivalentes:

1. Existe un $v \in V$ tal que $V = \langle v \rangle_{K[x]}$ y $\text{pol m} \, v = \sum_{i=0}^n a_i x^i$ (con $a_n = 1$).
2. $\dim_K V = n$ y existe una base de V en la cual la matriz de h es

$$\begin{pmatrix} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{n-2} & -a_{n-1} \end{pmatrix}$$

DEMOSTRACIÓN: 2) equivale a que exista una base $\{v_0, \dots, v_{n-1}\}$ de V de manera que $h(v_i) = v_{i+1}$ para $i = 0, \dots, n-2$ y $h(v_{n-1}) = -\sum_{i=0}^{n-1} a_i v_i$.

Si se cumple 1), entonces $\{v, xv, \dots, x^{n-1}v\}$ es una base que cumple 2).

Si $\{v_0, \dots, v_{n-1}\}$ cumple 2), entonces con $v = v_0$ se cumple que $v_i = x^i v_0$, así como que $x^n v = -\sum_{i=0}^{n-1} a_i x^i v$, es decir, $p(x) = \sum_{i=0}^n a_i x^i$ cumple $p(x)v = 0$, luego es un múltiplo del polinomio mínimo de v , pero por otro lado es obvio que $V = \langle v \rangle_{K[x]}$, luego el grado del polinomio mínimo de v ha de ser $n = \text{grad } p(x)$. Así pues, $p(x) = \text{pol m} \, v$. ■

En el caso general en que V no es necesariamente monógeno, la descomposición (6.2) nos permite obtener una descripción de h a partir de su acción sobre cada subespacio. Antes de entrar en ello debemos tener en cuenta que

los teoremas de estructura de módulos nos garantizan que los generadores v_i se pueden escoger de modo que sus polinomios mínimos p_i sean potencias de primo o bien de modo que cada uno divida al siguiente. Lo más importante es que, si los escogemos así, dichos polinomios mínimos están completamente determinados por V (luego por h) y se llaman divisores elementales en el primer caso y factores invariantes en el segundo.

Definición 6.29 Sea V un K -espacio vectorial de dimensión finita y h un endomorfismo de V . Entonces tenemos que V es un $K[x]$ -módulo de torsión finitamente generado y $K[x]$ es un dominio euclídeo. Llamaremos *divisores elementales* y *factores invariantes* de h a los de V como módulo.

Si $p(x)$ es el último factor invariante, entonces $p(x)$ es múltiplo de los polinomios mínimos de todos los generadores, luego $p(x)$ los anula a todos, y con ellos a todos los elementos de V . Es decir, $p(x)v = p(h)(v) = 0$ para todo vector v , o sea, $p(h) = 0$.

A este último factor invariante se le llama *polinomio mínimo* de h (pol mín h).

Notar que si $p(x)$ cumple $p(h) = 0$, en particular $p(x)$ anula al generador del submódulo asociado al último factor invariante, luego $\text{pol mín } h \mid p(x)$. Así pues, el polinomio mínimo de h es el menor polinomio que anula a h .

Teorema 6.30 Sea K un cuerpo y V un K -espacio vectorial de dimensión finita n . Sea h un endomorfismo de V . Las afirmaciones siguientes son equivalentes:

1. Los factores invariantes (o divisores elementales) de h son p_1, \dots, p_r .
2. Los polinomios p_1, \dots, p_r cumplen $p_i \mid p_{i+1}$ (o que cada p_i es potencia de primo) y, en una cierta base, la matriz de h es de la forma

$$\begin{pmatrix} M_1 & & & \\ & M_2 & & \\ & & \ddots & \\ & & & M_r \end{pmatrix}$$

donde cada M_i es la matriz asociada a p_i según el teorema 6.28.

DEMOSTRACIÓN: Consideramos una descomposición de V de tipo (6.2) de modo que los polinomios mínimos de los generadores sean los factores invariantes o los divisores elementales. Como los submódulos son subespacios invariantes, la restricción de h a cada uno de ellos es un endomorfismo y podemos aplicar el teorema 6.28 para obtener una base de cada uno de ellos de modo que la matriz de la restricción de h sea la indicada, o sea, la M_i del enunciado. La unión de las bases de los sumandos directos forma una base de V y es claro que la matriz de h en esta base es la indicada.

Recíprocamente, si la matriz de h en una base es de la forma indicada, podemos expresar V como suma directa de los subespacios generados por los

vectores de la base correspondientes a cada una de las cajas M_i de la matriz. Es claro que estos subespacios son invariantes y que M_i es la matriz de la restricción de h al subespacio i -ésimo. Por el teorema 6.28, cada uno de estos subespacios está generado por un v_i cuyo polinomio mínimo es p_i . Por la unicidad concluimos que los p_i son los factores invariantes o los divisores elementales de h . ■

Las matrices de la forma descrita en el teorema anterior son formas canónicas para la relación de semejanza. En efecto, si $A \in \text{Mat}_n(K)$ y V es cualquier K -espacio vectorial de dimensión n , entonces A es la matriz de un endomorfismo h de V en una base cualquiera de V , pero en otra base h tiene una matriz de la forma del teorema anterior (con factores invariantes), luego A es semejante a una matriz de este tipo.

Por otra parte, si A y B son matrices del tipo del teorema anterior (con factores invariantes) y son semejantes, entonces son matrices de un mismo endomorfismo h de V , y por la unicidad de los factores invariantes se ha de cumplir que $A = B$.

Para divisores elementales se cumple lo mismo salvo por el hecho de que los factores invariantes están ordenados por la divisibilidad, mientras que los divisores elementales no están ordenados de ningún modo, luego las matrices que resultan de cambiar el orden de las cajas (con divisores elementales) son semejantes entre sí sin dejar de ser del tipo del teorema 6.30.

Definición 6.31 Si K es un cuerpo y $A \in \text{Mat}_n(K)$, llamaremos 1ª *forma canónica* de A (respectivamente, 2ª *forma canónica*) a la única matriz del tipo indicado en el teorema 6.30 para factores invariantes (respectivamente para divisores elementales, con la consiguiente pérdida parcial de unicidad) que es semejante a la matriz A .

De este modo, dos matrices son semejantes si y sólo si sus formas canónicas son iguales. Los teoremas siguientes nos permiten calcular la forma canónica de cualquier matriz.

Teorema 6.32 Sea K un cuerpo y $A \in \text{Mat}_n(K)$ una matriz del tipo del teorema 6.28 para el polinomio mónico $p(x)$. Entonces los factores invariantes (los construidos en el teorema 6.23) de la matriz $xI_n - A \in \text{Mat}_n(K[x])$ son todos iguales a 1 excepto el último, que es $p(x)$.

DEMOSTRACIÓN: Sea $p(x) = \sum_{i=0}^n a_i x^i$, con $a_n = 1$. La matriz $xI_n - A$ es

$$\begin{pmatrix} x & -1 & & & & \\ & x & -1 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & x & -1 \\ a_0 & a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} + x \end{pmatrix}$$

Sumamos a la primera columna la segunda multiplicada por x , la tercera multiplicada por x^2 , etc. Luego a la segunda columna le sumamos la tercera multiplicada por x , la cuarta multiplicada por x^2 , etc. y así sucesivamente. El resultado es:

$$\begin{pmatrix} 0 & -1 & & \\ & \ddots & \ddots & \\ & & & -1 \\ \sum_{i=0}^n a_i x^i & \sum_{i=1}^n a_i x^i & \cdots & a_{n-1} + x \end{pmatrix}$$

Sumando a la fila n -sima la i -ésima multiplicada por el coeficiente adecuado queda:

$$\begin{pmatrix} 0 & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ \sum_{i=0}^n a_i x^i & 0 & \cdots & 0 \end{pmatrix}$$

Finalmente multiplicamos todas las filas salvo la última por -1 y reordenamos las columnas, con lo que llegamos a la forma canónica de la matriz:

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \sum_{i=0}^n a_i x^i \end{pmatrix}$$

■

Teorema 6.33 Sea K un cuerpo, V un K -espacio vectorial de dimensión finita n y h un endomorfismo de V . Sea A la matriz de h en 1ª forma canónica. Entonces los factores invariantes de h son los factores invariantes no unitarios de la matriz $xI_n - A$.

DEMOSTRACIÓN: Por comodidad llamaremos $[A_1, \dots, A_r]$ a la matriz formada por las cajas A_1, \dots, A_r situadas sobre su diagonal. Sean p_1, \dots, p_r los factores invariantes de h . Entonces $xI_n - A = [M_1, \dots, M_r]$, donde cada M_i tiene forma indicada en el teorema 6.28 para el polinomio p_i .

Por el teorema anterior M_i es equivalente a una matriz diagonal de la forma $N_i = [1, \dots, 1, p_i]$. Por lo tanto existen matrices $P_i, Q_i \in \text{GL}_n(K[x])$ tales que $P_i M_i Q_i = N_i$.

Sean $P = [P_1, \dots, P_r]$ y $Q = [Q_1, \dots, Q_r]$. Es fácil ver que P y Q son regulares, así como que

$$\begin{aligned} P(xI_n - A)Q &= [P_1 M_1 Q_1, \dots, P_r M_r Q_r] = [N_1, \dots, N_r] \\ &= [1, \dots, 1, p_1, \dots, 1, \dots, 1, p_r], \end{aligned}$$

luego la forma canónica de $xI_n - A$ es $[1, \dots, 1, p_1, \dots, p_r]$. ■

El resultado definitivo es el siguiente:

Teorema 6.34 Sea K un cuerpo, V un K -espacio vectorial de dimensión finita n y h un endomorfismo de V . Sea A la matriz de h en cualquier base. Entonces los factores invariantes de h son los factores invariantes no unitarios de la matriz $xI_n - A$.

DEMOSTRACIÓN: Sea B la matriz de h en 1ª forma canónica. Como A y B son matrices de h , son semejantes, es decir, existe una matriz $P \in \text{GL}_n(K)$ tal que $B = P^{-1}AP$. Entonces

$$P^{-1}(xI_n - A)P = xP^{-1}I_nP - P^{-1}AP = xI_n - B.$$

Por lo tanto, las matrices $xI_n - A$ y $xI_n - B$ son equivalentes, luego tienen los mismos factores invariantes y, por el teorema anterior, los no unitarios son los factores invariantes de h . ■

Definición 6.35 Sea K un cuerpo y $A \in \text{Mat}_n(K)$. Llamaremos *factores invariantes* de A a los factores invariantes de $xI_n - A$ (en el sentido de 6.23).

Hemos demostrado que los factores invariantes de un endomorfismo son los factores invariantes no unitarios de cualquiera de sus matrices.

También es obvio ahora que dos matrices son semejantes si y sólo si tienen los mismos factores invariantes (si tienen los mismos factores invariantes tienen la misma forma canónica).

Observar que en principio tenemos dos definiciones de factores invariantes de una matriz de $\text{Mat}_n(K)$, la dada en el teorema 6.23 y la que acabamos de dar, pero sucede que la definición según el teorema 6.23 no tiene interés para matrices sobre un cuerpo, ya que todos los factores invariantes en este sentido son iguales a 1. Lo único que interesa es su número, o sea, el rango.

Los factores invariantes de una matriz A se calculan sin dificultad mediante el algoritmo dado en el teorema 6.23.

Los divisores elementales se calculan descomponiendo los factores invariantes en potencias de primos.

Llamaremos *polinomio mínimo* de una matriz A a su último factor invariante. Así, si A es la matriz de un endomorfismo h , se cumple que $\text{pol mín } A = \text{pol mín } h$. Dado el isomorfismo entre endomorfismos y matrices, es claro que si $p(x) = \text{pol mín } A$, entonces $p(A) = 0$, y si $q(x) \in K[x]$, entonces $q(A) = 0$ si y sólo si $p(x) \mid q(x)$.

Ejemplo Vamos a calcular los factores invariantes de la matriz

$$A = \begin{pmatrix} 8 & 2 & 10 & -6 \\ 12 & 6 & 20 & -12 \\ 2 & 3 & 7 & -4 \\ 13 & 9 & 25 & -15 \end{pmatrix}$$

Para ello aplicamos el algoritmo del teorema 6.23 a $xI_4 - A$:

$$\begin{pmatrix} x-8 & -2 & -10 & 6 \\ -12 & x-6 & -20 & 12 \\ -2 & -3 & x-7 & 4 \\ -13 & -9 & -25 & x+15 \end{pmatrix}$$

En primer lugar intercambiamos las dos primeras columnas para situar en la posición (1,1) un polinomio de norma (o sea, de grado) mínima (el -2). Dividiendo la fila 1 entre -2 tenemos un 1 en el lugar (1,1) y restando a cada columna la primera multiplicada por su primer coeficiente llegamos hasta la matriz

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ x-6 & \frac{x^2}{2} - 7x + 12 & -5x + 10 & 3x - 6 \\ -3 & -\frac{3}{2}x + 10 & x + 8 & -5 \\ -9 & -\frac{9}{2}x + 23 & 20 & x + 12 \end{pmatrix}$$

Ahora restamos a cada fila la primera multiplicada por su primer coeficiente y hacemos ceros debajo del primer 1 (el resto de la matriz no se modifica). Como el 1 divide a todos los coeficientes restantes podemos continuar con la submatriz 3×3 .

Llevamos el -5 al lugar (2,2) (otra opción sería llevar el 20). Dividimos la segunda fila entre -5 y obtenemos otro 1, con el que hacemos ceros a su derecha y bajo él. El resultado es:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3x^2 - 7x + 2 & -2x^2 + 4x \\ 0 & 0 & 2x^2 - 8x + 8 & -3x^2 - 10 \end{pmatrix}$$

El cociente de $-2x^2 + 4x$ entre $3x^2 - 7x + 2$ es $-2/3$ y el resto $-(2/3)(x-2)$. Restamos a la cuarta columna la tercera multiplicada por $-2/3$ con lo que en la posición (3,4) queda $-(2/3)(x-2)$. Lo pasamos a la posición (3,3) y multiplicamos por $-3/2$:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & x-2 & -\frac{9}{2}x^2 + \frac{21}{2}x - 3 \\ 0 & 0 & -\frac{5}{3}x^2 + \frac{17}{3}x - \frac{11}{3} & 2x^2 - 8x + 8 \end{pmatrix}$$

Al dividir el polinomio de la posición (3,4) entre $x-2$ la división nos da exacta, a saber: $-(3/2)(3x-1)(x-2)$. Restamos a la cuarta fila la primera multiplicada por $-(3/2)(3x-1)$ y nos queda (factorizando los polinomios):

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & x-2 & 0 \\ 0 & 0 & -\frac{1}{3}(5x-7)(x-2) & -\frac{15}{2}(x-2)(x-1)^2 \end{pmatrix}$$

Restamos a la cuarta fila la tercera multiplicada por $-(1/3)(5x-7)$ y queda un 0 en el lugar $(4,3)$. Multiplicando la última fila por $-2/15$ queda:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & x-2 & 0 \\ 0 & 0 & 0 & (x-2)(x-1)^2 \end{pmatrix}$$

Por lo tanto los factores invariantes de la matriz A de partida son

$$x-2 \quad y \quad (x-2)(x-1)^2.$$

Los divisores elementales son $x-2$, $x-2$ y $(x-1)^2$. El polinomio mínimo de A es $\text{pol m} A = (x-2)(x-1)^2 = x^3 - 4x^2 + 5x - 2$. Las formas canónicas de A son:

$$\left(\begin{array}{c|ccc} 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 2 & -5 & 4 \end{array} \right) \quad \left(\begin{array}{c|ccc} 2 & 0 & 0 & 0 \\ \hline 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 2 \end{array} \right)$$

■

Hemos visto que si A y B son matrices semejantes, o sea, si existe $P \in \text{GL}_n(K)$ tal que $B = P^{-1}AP$, entonces también $xI_n - B = P^{-1}(xI_n - A)P$, luego tomando determinantes queda $|xI_n - B| = |P|^{-1}|xI_n - A||P| = |xI_n - A|$.

Definición 6.36 Llamaremos *polinomio característico* de $A \in \text{Mat}_n(K)$ al polinomio $\text{pol car } A = |xI_n - A|$. Es fácil ver que es mónico de grado n .

Acabamos de probar que las matrices semejantes tienen el mismo polinomio característico, por lo que podemos definir el polinomio característico de un endomorfismo como el de cualquiera de sus matrices.

El polinomio característico es mucho más fácil de calcular que los factores invariantes o los divisores elementales y en ocasiones proporciona suficiente información sobre una matriz dada.

Por ejemplo, el polinomio característico de la matriz A anterior resulta ser $\text{pol car } A = (x-2)^2(x-1)^2$.

Este ejemplo muestra que el polinomio característico está muy relacionado con los factores invariantes y los divisores elementales. La relación exacta nos la da el teorema siguiente:

Teorema 6.37 Sea K un cuerpo y $A \in \text{Mat}_n(K)$. Entonces el polinomio característico de A es el producto de sus factores invariantes (luego también el de los divisores elementales).

DEMOSTRACIÓN: Supongamos primero que A es de la forma indicada en el teorema 6.28, es decir, que tiene un único factor invariante no unitario igual a

$$\sum_{i=0}^n a_i x^i.$$

Es fácil ver que en este caso $|xI_n - A| = \sum_{i=0}^n a_i x^i$ (basta desarrollar el determinante por la primera columna).

Ahora, si una matriz A está formada por dos cajas ($A = [M, N]$ con la notación del teorema 6.33) es fácil ver que $[M, N] = [M, I][I, N]$, donde I representa en cada caso a la matriz identidad de las dimensiones adecuadas. Por lo tanto $|A| = |[M, I]| |[I, N]| = |M| |N|$. De aquí se sigue en general que $|[M_1, \dots, M_r]| = |M_1| \cdots |M_r|$.

Si A es una 1ª forma canónica, entonces $A = [M_1, \dots, M_r]$, donde las matrices M_i son del tipo del teorema 6.28 para cada factor invariante de A .

Entonces $xI_n - A = [xI - M_1, \dots, xI - M_r]$, luego

$$\text{pol car } A = |xI - M_1| \cdots |xI - M_r|.$$

Por el caso particular ya probado estos factores son los factores invariantes de A , luego el teorema es cierto para formas canónicas.

Finalmente, si A es cualquier matriz, su polinomio característico es el mismo que el de su 1ª forma canónica, que es el producto de sus factores invariantes, que son también los factores invariantes de A . ■

Una consecuencia es el llamado *teorema de Cayley*, según el cual toda matriz es raíz de su polinomio característico (porque éste es un múltiplo de su polinomio mínimo).

Las matrices semejantes tienen el mismo polinomio característico, aunque el recíproco no es cierto, pues el polinomio característico no determina los factores invariantes o los divisores elementales.

Por ejemplo, sabiendo que $\text{pol car } A = (x-2)^2(x-1)^2$, los invariantes de A podrían ser:

Factores invariantes	Divisores elementales
$(x-2)^2(x-1)^2$	$(x-2)^2, (x-1)^2$
$x-2, (x-2)(x-1)^2$	$x-2, x-2(x-1)^2$
$x-1, (x-2)^2(x-1)$	$x-1, x-1, (x-2)^2$
$(x-2)(x-1), (x-2)(x-1)$	$x-2, x-2, x-1, x-1$

Los coeficientes del polinomio característico $p(x)$ de una matriz A son invariantes por semejanza. El primero y el último de estos coeficientes (aparte del director, que es igual a 1) son especialmente simples.

El término independiente es $p(0) = |0I_n - A| = |-A| = (-1)^n |A|$. De aquí se sigue algo que ya sabíamos: que las matrices semejantes tienen el mismo determinante.

Para calcular el coeficiente de x^{n-1} observamos que, según la definición de determinante, uno de los sumandos de $p(x)$ es

$$(x - a_{11}) \cdots (x - a_{nn}) = x^n - (a_{11} + \cdots + a_{nn})x^{n-1} + \cdots$$

y ningún otro sumando tiene monomios de grado $n-1$.

Así pues, el coeficiente de grado $n-1$ de $p(x)$ es $-(a_{11} + \cdots + a_{nn})$.

Definición 6.38 Llamaremos *traza* de una matriz $A \in \text{Mat}_n(K)$ a

$$\text{Tr}(A) = a_{11} + \cdots + a_{nn}.$$

Hemos probado que las matrices semejantes tienen el mismo determinante y la misma traza. Por ello podemos definir el determinante y la traza de un endomorfismo como el de cualquiera de sus matrices.

Las raíces del polinomio característico de un endomorfismo contienen información muy importante sobre el mismo. Vamos a comprobarlo. Sea h un endomorfismo de un espacio vectorial V sobre un cuerpo K y sea A su matriz en una base. Sea $\alpha \in K$ una raíz de $\text{pol car } A$. Entonces $|\alpha I_n - A| = 0$, lo que significa que la matriz $\alpha I_n - A$ no es regular, luego el endomorfismo que determina, es decir, $\alpha I - h$, no es inyectivo, luego su núcleo no es trivial, luego existe un vector $v \in V$ no nulo tal que $h(v) = \alpha v$. Recíprocamente, si existe un vector v no nulo tal que $h(v) = \alpha v$, invirtiendo el razonamiento llegamos a que α es raíz de $\text{pol car } A$.

Definición 6.39 Sea h un endomorfismo de un espacio vectorial V sobre un cuerpo K . Si un vector $v \in V$ no nulo cumple $h(v) = \alpha v$ para cierto $\alpha \in K$, diremos que v es un *vector propio* de h y que α es un *valor propio* de h .

Acabamos de probar que los valores propios de un endomorfismo h son exactamente las raíces de su polinomio característico.

Cada vector propio tiene asociado un único valor propio. Recíprocamente, si $\alpha \in K$ es un valor propio de h , entonces α tiene asociado el espacio

$$S(h, \alpha) = \{v \in V \mid h(v) = \alpha v\},$$

que es un subespacio vectorial no trivial de V al que llamaremos *espacio fundamental* de α .

Un caso especialmente interesante es el subespacio $S(h, 1)$, que está formado por los puntos fijos de h .

Ejemplo Sea h un endomorfismo de \mathbb{R}^4 cuya matriz en la base canónica sea la matriz A del ejemplo anterior. Hemos calculado

$$\text{pol car } A = (x - 2)^2(x - 1)^2,$$

luego sus valores propios son 1 y 2.

El sistema de ecuaciones lineales $(u, v, w, x)A = (u, v, w, x)$ tiene solución $(-4v, v, -5v, 2v)$ para cualquier v , luego

$$S(h, 1) = \langle (-4, 1, -5, 2) \rangle.$$

Igualmente se calcula

$$S(h, 2) = \langle (-2, 1, 0, 0), (-3, 0, -6, 2) \rangle.$$

■

Como el polinomio característico es el producto de los divisores elementales, los valores propios son las raíces de los divisores elementales. Como éstos son potencias de irreducibles, es obvio que α es un valor propio de h si y sólo si uno de los divisores elementales de h tiene la forma $(x - \alpha)^e$. El teorema siguiente muestra exactamente la relación entre los valores propios y los divisores elementales.

Teorema 6.40 *Sea K un cuerpo, V un K -espacio vectorial de dimensión finita y h un endomorfismo de V . Sea $V = \langle v_1 \rangle_{k[x]} \oplus \cdots \oplus \langle v_r \rangle_{k[x]}$ la descomposición de V en submódulos asociada a los divisores elementales, es decir, tal que $\text{pol m} \, v_i = p_i(x)^{e_i}$ con $p_i(x)$ irreducible. Sea $\alpha \in K$ un valor propio de h y supongamos que $p_i = x - \alpha$ para $i = 1, \dots, s$. Entonces:*

1. $S(h, \alpha) = \langle (x - \alpha)^{e_1-1}v_1, \dots, (x - \alpha)^{e_s-1}v_s \rangle_k$.
2. $\dim S(h, \alpha) = s$.

Si $\alpha_1, \dots, \alpha_t$ son todos los valores propios de h , entonces se tiene la suma directa

$$S = S(h, \alpha_1) \oplus \cdots \oplus S(h, \alpha_t).$$

Al espacio S se le llama espacio fundamental de h .

DEMOSTRACIÓN: Un vector $v \neq 0$ es propio con valor propio α si y sólo si $(x - \alpha)v = 0$, luego es inmediato que los vectores $(x - \alpha)^{e_1-1}v_1, \dots, (x - \alpha)^{e_s-1}v_s$ están en $S(h, \alpha)$. Por lo tanto tenemos la inclusión

$$\langle (x - \alpha)^{e_1-1}v_1, \dots, (x - \alpha)^{e_s-1}v_s \rangle_k \leq S(h, \alpha).$$

Veamos ahora la otra inclusión. Si $v \in S(h, \alpha)$, por la descomposición de V tenemos en principio que $v = f_1(x)v_1 + \cdots + f_r(x)v_r$ para ciertos polinomios $f_i(x)$. Según el teorema 6.27 podemos exigir que $\text{grad } f_i(x) < e_i$.

Como v es un vector propio, se cumple que $(x - \alpha)v = 0$, o sea,

$$(x - \alpha)f_1(x)v_1 + \cdots + (x - \alpha)f_r(x)v_r = 0.$$

Como la suma es directa cada sumando es nulo y, por definición de polinomio mínimo, resulta que $p_i(x)^{e_i} \mid (x - \alpha)f_i(x)$ para $i = 1, \dots, r$.

Si $i > s$ tenemos que $p_i(x) \neq x - \alpha$, luego $p_i(x)^{e_i} \nmid f_i(x)$ y así $f_i(x)v_i = 0$, con lo que la expresión de v se reduce a

$$v = f_1(x)v_1 + \cdots + f_s(x)v_s.$$

Para $i \leq s$ tenemos que $(x - \alpha)^{e_i} \mid (x - \alpha)f_i(x)$, de donde $(x - \alpha)^{e_i-1} \mid f_i(x)$, es decir, $f_i(x) = q_i(x)(x - \alpha)^{e_i-1}$, pero tenemos también que $\text{grad } f_i(x) < e_i$, luego la única posibilidad para $q_i(x)$ es que sea una constante $q_i \in K$. Con esto la expresión de v queda en

$$v = q_1(x - \alpha)^{e_1-1}v_1 + \cdots + q_s(x - \alpha)^{e_s-1}v_s,$$

lo que prueba la igualdad.

Estos generadores son independientes, pues cada uno está contenido en un sumando directo distinto. Por la misma razón los distintos espacios fundamentales tienen suma directa, pues están contenidos en sumandos directos distintos (los correspondientes a los divisores elementales con raíz el valor propio correspondiente). ■

La matriz A de los ejemplos anteriores tiene divisores elementales $x - 2$, $x - 2$ y $(x - 1)^2$, por lo que el teorema anterior nos da que $\dim S(h, 2) = 2$ y $\dim S(h, 1) = 1$ (como ya habíamos comprobado).

Consideremos por último el caso en que V es un espacio vectorial euclídeo. En estos espacios conviene trabajar únicamente con bases ortonormales. Esto nos obliga a restringir como sigue la relación de semejanza:

Definición 6.41 Diremos que dos matrices $S, T \in \text{Mat}_n(\mathbb{R})$ son *congruentes* si existe una matriz $P \in O(n)$ tal que $T = P^{-1}SP$.

La interpretación de esta relación es sencilla. Supongamos que S y T son dos matrices de un mismo endomorfismo h en dos bases ortonormales B y C . Entonces sabemos que $T = P^{-1}SP$, donde P es la matriz de cambio de base entre B y C o, en otros términos, la matriz de la identidad en \mathbb{R}^n respecto de las bases B y C . Ahora bien, como la identidad es una isometría y las bases son ortonormales, sabemos que P ha de ser una matriz ortogonal, luego S y T son congruentes. Recíprocamente, si S es la matriz de h en una base ortonormal y T es una matriz congruente con S , entonces existe una base ortonormal de V respecto a la cual la matriz de h es T .

La teoría anterior no es aplicable directamente a este contexto, pues las formas canónicas de una matriz dada no son necesariamente congruentes con ella, luego no nos sirven como formas canónicas para la congruencia. En lugar de obtener resultados genéricos sobre endomorfismos vamos a restringirnos a isometrías, pues la situación se simplifica considerablemente y, al fin y al cabo, es lo único que vamos a necesitar. La clave la proporciona el teorema siguiente:

Teorema 6.42 Sea A una matriz ortogonal. Entonces las raíces complejas del polinomio mínimo de A son simples y de módulo 1.

DEMOSTRACIÓN: Vamos a considerar $V = \mathbb{C}^n$ como \mathbb{C} -espacio vectorial y $h : V \rightarrow V$ el isomorfismo definido por A en la base canónica. Así, $h(v) = vA$.

Veamos ante todo que el polinomio mínimo de A es el mismo tanto si la consideramos como matriz real o como matriz compleja. Sea $p(x)$ el polinomio mínimo real y $q(x)$ el polinomio mínimo complejo. Como $p(A) = 0$ es claro que $q(x) \mid p(x)$ (el recíproco no vale porque $p(x)$ divide a todos los polinomios reales que anulan a A). Descomponemos $q(x) = q_1(x) + q_2(x)i$, con $q_1(x), q_2(x) \in \mathbb{R}[x]$.

Como $q(A) = q_1(A) + q_2(A)i = 0$, ha de ser $q_1(A) = q_2(A) = 0$, luego $p(x)$ divide a ambos, luego $p(x) \mid q(x)$.

Sea α una raíz del polinomio mínimo de A . Entonces es un valor propio, luego existe un $v \in V$ no nulo tal que $vA = \alpha v$. Trasponiendo y conjugando esta ecuación obtenemos $A^t v^t = \alpha v^t$ y $\bar{v}A = \bar{\alpha}\bar{v}$, donde \bar{v} es el vector que resulta de

conjugar cada coordenada de v . La matriz A no necesita ser conjugada porque sus coeficientes son reales. La linealidad de la conjugación hace que la igualdad siga siendo cierta.

Ahora multiplicamos las ecuaciones: $\bar{v}AA^t v^t = \alpha \bar{\alpha} \bar{v} v^t$, y como A es ortogonal la expresión se reduce a $\bar{v} v^t = |\alpha|^2 \bar{v} v^t$.

Pero $\bar{v} v^t$ es la suma de los cuadrados de los módulos de las coordenadas de v . Como v no es nulo, $\bar{v} v^t \neq 0$, luego $|\alpha|^2 = 1$ y así, $|\alpha| = 1$.

Supongamos que el polinomio mínimo tiene una raíz múltiple α . Entonces un divisor elemental de A es de la forma $(x - \alpha)^r$ con $r \geq 2$. Sea $v \in V$ tal que su polinomio mínimo sea $(x - \alpha)^r$ y sea $u = (x - \alpha)^{r-2}$. Entonces $(x - \alpha)u \neq 0$, pero $(x - \alpha)^2 u = 0$.

En otros términos, si $w = (x - \alpha)u$, tenemos que $w \neq 0$ pero $(x - \alpha)w = 0$, o también:

$$w = uA - \alpha u \neq 0 \quad \text{y} \quad wA = \alpha w.$$

Despejamos en la primera ecuación y trasponemos: $A^t u^t = w^t + \alpha u^t$, conjugamos la segunda: $\bar{w}A = \bar{\alpha} \bar{w}$, las multiplicamos: $\bar{w}AA^t u^t = \bar{\alpha} \bar{w} w^t + \alpha \bar{\alpha} \bar{w} u^t$, aplicamos que A es ortogonal y que α tiene módulo 1, con lo que llegamos a $\bar{w} u^t = \bar{\alpha} \bar{w} w^t + \bar{w} u^t$.

O sea, $\bar{\alpha} \bar{w} w^t = 0$, de donde $\bar{w} w^t = 0$, pero w^t es la suma de los cuadrados de los módulos de las coordenadas de $w \neq 0$, luego no puede ser cero. ■

Así pues, los factores irreducibles (en $\mathbb{R}[x]$) del polinomio mínimo de una matriz ortogonal han de ser o bien $x - 1$, o bien $x + 1$, o bien de la forma

$$(x - (\cos \theta + i \operatorname{sen} \theta))(x - (\cos \theta - i \operatorname{sen} \theta)) = x^2 - 2 \cos \theta x + 1,$$

para un cierto $\theta \in]0, \pi[$ (aquí es esencial la clausura algebraica de \mathbb{C}). Además han de ser distintos dos a dos. Los divisores elementales dividen al polinomio mínimo, luego han de tener todos exponente 1, es decir, han de ser irreducibles.

El polinomio característico es el producto de los divisores elementales, luego éstos son precisamente los factores irreducibles del polinomio característico, luego el polinomio característico de una matriz ortogonal determina los divisores elementales y consecuentemente también los factores invariantes.

Ejemplo Consideremos la matriz

$$B = \begin{pmatrix} 1/9 & -4/9 & 8/9 \\ -4/9 & 7/9 & 4/9 \\ 8/9 & 4/9 & 1/9 \end{pmatrix}$$

Es fácil comprobar que se trata de una matriz ortogonal. Podríamos calcular sus factores invariantes con el algoritmo que conocemos, pero no es necesario. Basta calcular el polinomio característico, que resulta ser $(x + 1)(x - 1)^2$. Como sabemos que los divisores elementales son irreducibles, han de ser $x + 1$, $x - 1$ y $x - 1$, de donde los factores invariantes son $x - 1$ y $(x + 1)(x - 1)$. ■

Esto significa que dos matrices ortogonales son semejantes si y sólo si tienen el mismo polinomio característico, pero vamos a probar que de hecho en tal caso son congruentes (con lo que la semejanza equivaldrá a la congruencia).

Teorema 6.43 Sea V un espacio vectorial euclídeo y $h : V \longrightarrow V$ una isometría con divisores elementales p_1, \dots, p_r . Entonces V admite una descomposición de la forma

$$V = \langle v_1 \rangle_{k[x]} \perp \cdots \perp \langle v_r \rangle_{k[x]},$$

donde $\text{pol m}ín v_i = p_i$.

DEMOSTRACIÓN: Vamos a probar por inducción sobre la dimensión de V que existe una descomposición como la indicada de modo que los polinomios mínimos de los generadores son irreducibles (sin afirmar que sean los divisores elementales de h). Esto implica inmediatamente el teorema por la unicidad de los divisores elementales.

Si la dimensión de V es 1 es evidente.

Sea p un divisor elemental de h y sea $v \in V$ tal que $\text{pol m}ín v = p$. Sea $W = \langle v \rangle_{k[x]}$. Entonces $V = W \perp W^\perp$. Tenemos que W es invariante y como h es biyectiva de hecho $h[W] = W$. Puesto que h conserva el producto escalar, es claro que $h[W^\perp] = h[W]^\perp = W^\perp$, luego podemos restringir h a una isometría de W^\perp y aplicar la hipótesis de inducción. ■

Si tomamos una base ortonormal de cada uno de los subespacios de la descomposición del teorema anterior, su unión nos da una base ortonormal de V en la cual la matriz de h está formada por cajas $[M_1, \dots, M_r]$, donde M_i es la matriz de la restricción de h al subespacio i -ésimo.

Como los divisores elementales tienen grado 1 o 2, éstas son las dimensiones de las cajas. Si $p_i = x - 1$ podemos tomar como base el vector v_i y, puesto que $h(v_i) = v_i$, la caja correspondiente es un 1.

Si $p_i = x + 1$ tomamos igualmente la base v_i y ahora $h(v_i) = -v_i$, luego la caja es un -1 .

Si $p_i = x^2 - 2\cos\theta x + 1$, entonces la caja M_i es una matriz de $O(2)$ cuyo único divisor elemental (luego también su único factor invariante, luego su polinomio característico) es p_i , luego su determinante es el término independiente 1 y su traza es $2\cos\theta$. Estas condiciones determinan completamente la matriz. En efecto, dada una matriz cualquiera

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

la ecuación $AA^t = I$ implica $a^2 + b^2 = 1$, $c^2 + d^2 = 1$, $ac + bd = 0$. La ecuación $A^t A = I$ nos da también $a^2 + c^2 = 1$, $b^2 + d^2 = 1$, de donde se sigue inmediatamente $c = \pm b$ y $d = \pm a$. La ecuación $ac + bd = 0$ implica que los signos han de ser distintos, es decir, $d = \pm a$ y $c = \mp b$. Si exigimos que el determinante sea 1, es decir, $ad - bc = 1$, llegamos a que $d = a$ y $c = -b$. La matriz es, por lo tanto,

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Al considerar la traza vemos que $a = \cos\theta$ y la ecuación $a^2 + b^2 = 1$ implica que $b = \pm \sin\theta$. Cambiando el signo al segundo vector de la base si es preciso,

podemos exigir que $b = \sin \theta$. En definitiva obtenemos la matriz

$$M_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

que definimos en el capítulo IV y que corresponde al giro de ángulo θ . Vamos a recoger lo que hemos obtenido.

Definición 6.44 Diremos que una matriz ortogonal está en *forma canónica* si es de la forma $[M_1, \dots, M_r]$, donde cada caja M_i es ± 1 o bien M_θ , para un ángulo $\theta \in]0, \pi[$.

Hemos probado que toda isometría admite en una cierta base ortonormal una matriz en forma canónica con tantas cajas iguales a 1 como divisores elementales iguales a $x - 1$, tantas cajas iguales a -1 como divisores elementales iguales a $x + 1$ y tantas cajas de la forma M_θ como divisores elementales de la forma $x^2 - 2\cos\theta x + 1$ (para $\theta \in]0, \pi[$).

El hecho de que las cajas estén determinadas por (y de hecho determinen) los divisores elementales implica que dos formas canónicas no son congruentes (salvo que se diferencien en una permutación de las cajas).

Si a esto añadimos que los divisores elementales están determinados por el polinomio característico, concluimos:

Teorema 6.45 *Dos matrices ortogonales son congruentes si y sólo si son semejantes, si y sólo si tienen el mismo polinomio característico.*

Dedicamos la sección siguiente a interpretar geoméricamente estos hechos.

6.6 Clasificación de isometrías

Los resultados de la sección anterior contienen esencialmente la clasificación de las isometrías (afines) con un punto fijo, pues ésta se reduce a estudiar la isometría lineal asociada. El teorema siguiente nos reduce a este caso el problema general.

Teorema 6.46 *Sea E un espacio afín euclídeo y $f \in \text{Is}(E)$. Entonces f se descompone de forma única como $f = g \circ T_{\vec{v}}$, donde g es una isometría con un conjunto no vacío G de puntos fijos y $\vec{v} \in \vec{G}$. Además en tal caso $\vec{G} = S(\vec{f}, 1)$ y $f = T_{\vec{v}} \circ g$.*

DEMOSTRACIÓN: Probemos primero la unicidad. Dada una descomposición $f = g \circ T_{\vec{v}}$ en las condiciones del enunciado, es claro que $\vec{f} = \vec{g}$, luego tenemos que $\vec{G} = S(\vec{g}, 1) = S(\vec{f}, 1)$. Si P es un punto fijo de g se cumple que $f(P) = P + \vec{v}$, luego $\overrightarrow{Pf(P)} = \vec{v} \in S(\vec{f}, 1)$.

Si $f = g' \circ T_{\vec{v}'}$ es otra descomposición en las condiciones del enunciado y P' es un punto fijo de g' , tendremos igualmente que $\overrightarrow{P'f(P')} = \vec{v}' \in S(\vec{f}, 1)$. Un simple cálculo nos da

$$\vec{v} - \vec{v}' = \overrightarrow{PP'} - \vec{f}(\overrightarrow{PP'}) \in S(\vec{f}, 1) \cap \text{Im}(\vec{f} - 1).$$

Si probamos que $E = S(\vec{f}, 1) \perp \text{Im}(\vec{f} - 1)$ tendremos que $\vec{v} = \vec{v}'$, luego también $g = g'$.

Notar que $S(\vec{f}, 1)$ es el núcleo de $\vec{f} - 1$, luego basta probar que los dos espacios son ortogonales y por las dimensiones su suma será todo E . Ahora bien, dados $\vec{x} \in S(\vec{f}, 1)$ e $\vec{y} = \vec{f}(\vec{z}) - \vec{z} \in \text{Im}(\vec{f} - 1)$, se cumple

$$\vec{x}\vec{y} = \vec{x}\vec{f}(\vec{z}) - \vec{x}\vec{z} = \vec{f}(\vec{x})\vec{f}(\vec{z}) - \vec{x}\vec{z} = \vec{x}\vec{z} - \vec{x}\vec{z} = 0,$$

pues \vec{f} conserva el producto escalar.

Para probar la existencia fijamos un punto O arbitrario y descomponemos

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

con $\vec{v} \in S(\vec{f}, 1)$ y $\vec{f}(\vec{w}) - \vec{w} \in \text{Im}(\vec{f} - 1)$ (usando la descomposición de E que acabamos de obtener). Entonces, para todo punto P se cumple

$$\overrightarrow{Pf(P)} = \overrightarrow{PO} + \overrightarrow{Of(O)} + \vec{f}(\overrightarrow{OP}) = \vec{v} + (\vec{f} - 1)(\vec{w} + \overrightarrow{OP}).$$

En particular, si tomamos A de modo que $\overrightarrow{OA} = -\vec{w}$ la fórmula anterior se reduce a $\overrightarrow{Af(A)} = \vec{v} \in S(\vec{f}, 1)$. Equivalentemente, $f(A) = A + \vec{v}$.

Definimos $g = f \circ T_{-\vec{v}}$. Así g es ciertamente una isometría, se cumple $f = g \circ T_{\vec{v}}$ y además $g(A) = f(A) - \vec{v} = A$, luego el conjunto G de sus puntos fijos es no vacío. Más aún, es claro que $\vec{G} = S(\vec{g}, 1) = S(\vec{f}, 1)$ y por consiguiente $\vec{v} \in \vec{G}$.

Por último, se cumple que $f = T_{\vec{v}} \circ g$ porque para todo punto P tenemos

$$g(P + \vec{v}) = g(P) + \vec{g}(\vec{v}) = g(P) + \vec{f}(\vec{v}) = g(P) + \vec{v} = f(P).$$

■

Un par de observaciones complementarias: en el teorema anterior se cumplirá $\vec{v} = \vec{0}$ si y sólo si f tiene un punto fijo. Por el teorema 6.2, si f no tiene puntos fijos entonces $\dim G \geq 1$.

Existen tres clases básicas de isometrías: las traslaciones, los giros y las simetrías. Ya conocemos las primeras y parcialmente los giros. Vamos a analizar con más detalle los giros y las simetrías y después veremos en qué sentido se reducen a estos tres tipos las isometrías arbitrarias.

Definición 6.47 Una isometría f de un espacio afín euclídeo es un *giro* de ángulo θ si tiene un punto fijo y la forma canónica de \vec{f} es $[M_\theta, 1, \dots, 1]$, donde

$$M_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$

Observamos que M_θ es la identidad si y sólo si θ es múltiplo de 2π , en cuyo caso f es la identidad. Exceptuado este caso, si el espacio E tiene dimensión n , el espacio fundamental $S(\vec{f}, 1)$ tiene dimensión $n - 2$, luego f tiene una variedad G de puntos fijos de dimensión $n - 2$.

Si E es un plano, entonces G se reduce a un punto O y f es un giro de ángulo θ alrededor de O en el sentido que ya discutimos en el capítulo IV. El punto O se llama *centro* del giro. Recordemos que hay dos giros de ángulo θ con un mismo centro O , que sobre un mismo sistema de referencia de origen O se corresponden con las matrices $M_{\pm\theta}$ (de modo que al cambiar el signo de un vector de la base el giro de matriz M_θ pasa a tener matriz $M_{-\theta}$ y viceversa).

Si la dimensión de E es 3 entonces la variedad de puntos fijos de un giro f distinto de la identidad es una recta r llamada *eje* del giro. Sea $\vec{v}_1, \vec{v}_2, \vec{v}_3$ la base ortonormal en la que \vec{f} tiene matriz $[M_\theta, 1]$. Si O es cualquier punto del eje de giro r , entonces $r = O + \langle \vec{v}_3 \rangle$ y el plano ortogonal a r dado por $O + \langle \vec{v}_1, \vec{v}_2 \rangle$ es fijado por f . En el sistema de referencia $(O; \vec{v}_1, \vec{v}_2)$ la matriz de la aplicación lineal asociada a la restricción es M_θ . Así pues, f gira los puntos de cada plano perpendicular a r respecto al centro O en el cual r corta a tal plano.

En general, si G es una variedad de dimensión $n - 2$ en E , un giro de ángulo θ respecto a G es una isometría que deja invariante a cada punto de G y que se comporta como un giro (plano) de ángulo θ sobre cada plano perpendicular⁴ a G (con centro en el punto de intersección). Hay exactamente dos giros en estas condiciones.

Dada la expresión en coordenadas de un giro (no necesariamente la canónica) resolviendo un sistema de ecuaciones lineales se obtiene la variedad de puntos fijos. El ángulo de giro estará determinado salvo signo por el hecho de que $2 \cos \theta + n - 2$ es la traza de la matriz asociada (pues esto es cierto para la matriz en forma canónica y la traza es la misma en todas las matrices).

Introducimos ahora el concepto de simetría. La idea es que la imagen de un punto P por una simetría respecto a una variedad L se obtiene trazando la recta perpendicular a L por P y tomando en ésta el punto P' que se encuentra a la misma distancia que P del punto de corte. De este modo, el punto de corte es el punto medio del segmento $\overline{PP'}$. Vamos a formular analíticamente esta descripción de las simetrías.

En primer lugar, dado un espacio afín euclídeo E y una variedad L distinta de todo E , por cada punto P exterior a L pasa una única recta r perpendicular a L . En efecto, dado $Q \in L$ descomponemos $\overrightarrow{PQ} = \vec{v} + \vec{w}$, con $\vec{v} \in \vec{L}$, $\vec{w} \in \vec{L}^\perp$. Sea $R = P + \vec{w} = Q - \vec{v}$. Claramente $R \in L$ y $r = PR$ es perpendicular a L y pasa por P . Además es única, pues si una recta PR' cumple lo mismo, es decir, $\overrightarrow{PR'} \perp \vec{L}$, entonces $\overrightarrow{RR'}$ sería ortogonal a \overrightarrow{PR} y $\overrightarrow{PR'}$, lo cual es contradictorio salvo si $R = R'$.

Por otra parte, dados dos puntos P y P' , tenemos que $P' = P + \overrightarrow{PP'}$, luego el *punto medio* entre ambos será $M = P + \frac{1}{2}\overrightarrow{PP'}$, que es el único punto que cumple $\overrightarrow{PM} = \overrightarrow{MP'}$. Esto sirve como definición de punto medio en cualquier espacio

⁴Aquí entendemos que dos variedades L y L' son perpendiculares si tienen un punto en común, necesariamente único, y $\vec{L} \perp \vec{L}'$. Notar que en este sentido fuerte dos planos del espacio tridimensional nunca pueden ser perpendiculares.

afín,⁵ no necesariamente euclídeo. Una definición equivalente pero simétrica es

$$M = \frac{1}{2}P + \frac{1}{2}P'$$

(ver pág 87). Esta definición tiene sentido incluso si $P = P'$, en cuyo caso el punto medio es $M = P = P'$.

Notar que dados dos puntos P y M siempre existe un único punto P' tal que M sea el punto medio de P y P' . Necesariamente P' ha de ser el baricentro $P' = 2M - P$.

Con todas estas observaciones ya podemos justificar la existencia de simetrías respecto a variedades cualesquiera:

Definición 6.48 En un espacio afín euclídeo E , llamaremos *simetría* respecto a una variedad L de E a la única aplicación $s_L : E \rightarrow E$ que fija a todos los puntos de L y que a cada punto P exterior a L le asigna el único punto $s_L(P)$ tal que la perpendicular a L por P corta a L en el punto medio entre P y $s_L(P)$.

En particular la simetría respecto a E es la identidad. No es necesario distinguir la forma en que una simetría respecto a una variedad L actúa sobre los puntos de L y sobre los puntos exteriores a L . En cualquier caso, el punto $P' = s_L(P)$ está determinado por las relaciones

$$\overrightarrow{PP'} \in \vec{L}^\perp \quad \text{y} \quad \frac{1}{2}P + \frac{1}{2}P' \in L.$$

Las simetrías respecto a puntos, rectas y planos se llaman respectivamente simetrías *puntuales*, *axiales* (lat. ‘respecto a un eje’) y *especulares* (lat. ‘respecto a un espejo’). Las simetrías respecto a hiperplanos se llaman *reflexiones*.

Teorema 6.49 Las simetrías de un espacio afín euclídeo son isometrías. Además una isometría f es una simetría si y sólo si $f \circ f = 1$.

DEMOSTRACIÓN: Sea E un espacio afín euclídeo y L una variedad lineal de E , que podemos suponer distinta de todo el espacio. Tomemos una base ortonormal $\vec{v}_1, \dots, \vec{v}_r$ de \vec{L} y completémosla hasta una base ortonormal $\vec{v}_1, \dots, \vec{v}_n$ de \vec{E} . Tomemos un punto O en L y definamos $s_L(P) = O + \vec{s}_L(\overrightarrow{OP})$, donde \vec{s}_L es la isometría lineal que en la base anterior tiene matriz $\underbrace{[1, \dots, 1]}_r, \underbrace{[-1, \dots, -1]}_{n-r}$.

Dado un punto $P = O + \alpha_1 \vec{v}_1 + \dots + \alpha_n \vec{v}_n$, es claro que

$$\overrightarrow{Ps_L(P)} = -2\alpha_{r+1}\vec{v}_{r+1} - \dots - 2\alpha_n\vec{v}_n \in \vec{L}^\perp,$$

$$\frac{1}{2}P + \frac{1}{2}s_L(P) = O + \frac{1}{2}\overrightarrow{OP} + \frac{1}{2}\overrightarrow{Os_L(P)} = O + \alpha_1\vec{v}_1 + \dots + \alpha_r\vec{v}_r \in L.$$

⁵En realidad hemos de suponer que el cuerpo tiene característica distinta de 2. En caso contrario no existen puntos medios ni, como se ve fácilmente, tampoco simetrías. En lo sucesivo supondremos tácitamente $\text{car}K \neq 2$.

Por consiguiente la isometría s_L es la simetría respecto a L .

Con esto hemos visto que las simetrías son las isometrías con al menos un punto fijo y cuya aplicación lineal asociada tiene como matriz canónica una de la forma $[1, \dots, 1, -1, \dots, -1]$. Es claro que al multiplicar una matriz de este tipo por sí misma se obtiene la matriz identidad. Recíprocamente, si $f \circ f = 1$, sea $f = g \circ T_{\vec{v}}$ en las condiciones del teorema 6.46. Entonces $g \circ g \circ T_{2\vec{v}} = 1$, luego $\vec{g} \circ \vec{g} = \vec{f} \circ \vec{f} = 1$ y $2\vec{v} = 0$, luego $\vec{v} = 0$ (y por lo tanto f tiene un punto fijo) y la forma canónica de \vec{f} ha de constar sólo de cajas ± 1 , pues si contuviera cajas M_θ , su cuadrado contendría cajas $M_{2\theta}$, que son distintas de la identidad salvo que $\theta = 1, \pi$, en cuyo caso $M_\theta = [\pm 1, \pm 1]$. ■

Ahora podemos clasificar las isometrías de los espacios de dimensiones bajas. En general, cuando hablemos del determinante, la traza, etc. de una isometría nos referimos, naturalmente, a los de su aplicación lineal asociada.

Isometrías de una recta Si f es una isometría de una recta euclídea, entonces la forma canónica de \vec{f} ha de ser $[\pm 1]$. Si el signo es positivo se tratará de una traslación, mientras que si es negativo se tratará de una simetría puntual (notar que f tendrá un punto fijo por el teorema 6.2).

Isometrías del plano Si f es una isometría del plano y su determinante es positivo entonces la forma canónica de \vec{f} ha de ser de la forma M_θ (incluyendo aquí las posibilidades $[\pm 1, \pm 1]$). Si f tiene un punto fijo entonces será un giro (o la identidad), mientras que si f no tiene puntos fijos entonces \vec{f} ha de ser la identidad (por el teorema 6.2), luego f es una traslación.

Si el determinante es negativo, la forma canónica de \vec{f} ha de ser $[1, -1]$. Si f tiene un punto fijo, entonces se trata de una simetría axial, mientras que si no tiene puntos fijos, el teorema 6.46 implica que f es la composición de una simetría axial con una traslación respecto a un vector en la dirección del eje de simetría.

Isometrías del espacio Si f es una isometría del espacio (tridimensional) y su determinante es positivo, entonces la forma canónica de \vec{f} ha de ser de la forma $[M_\theta, 1]$. Si f tiene un punto fijo será un giro (quizá la identidad). Si no lo tiene aplicamos el teorema 6.46 para concluir que f es la composición de un giro con una traslación respecto a un vector en la dirección del eje de giro. Si el giro es la identidad tenemos simplemente una traslación. En caso contrario la isometría es lo que se llama un *giro helicoidal*, pues describe el movimiento de una hélice, que gira y avanza al mismo tiempo.

Si el determinante es negativo entonces la forma canónica de \vec{f} será de la forma $[M_\theta, -1]$. Si $M_\theta \neq I$, entonces \vec{f} no tiene puntos fijos, luego f ha de tener uno exactamente, luego f es la composición de un giro con una simetría especular respecto a un plano perpendicular al eje de giro (el punto fijo de f es la intersección entre el eje y el plano). Si $M_\theta = I$ entonces o bien f tiene puntos fijos y se trata de una simetría especular, o bien no los tiene, y se trata de una

simetría especular compuesta con una traslación en dirección perpendicular al plano de simetría.

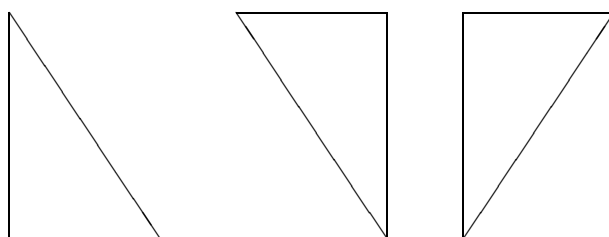
La tabla siguiente resume la clasificación:

Dim.	Det.	Con punto fijo	Sin puntos fijos
1	+1	Identidad	Traslación
	-1	Simetría puntual	—
2	+1	Giro	Traslación
	-1	Simetría axial	Simetría con traslación
3	+1	Giro	Traslación o Giro helicoidal
	-1	Simetría especular	Simetría con traslación o Giro y simetría

Ejercicio: ¿Por qué no aparecen en la tabla las simetrías puntuales en dimensión 2 y 3 ni las simetrías axiales en dimensión 3?

Ejercicio: Clasificar las isometrías en dimensión 4.

El signo del determinante de una isometría tiene una interpretación geométrica intuitivamente muy simple, pero que aún no estamos en condiciones de justificar matemáticamente. Por ejemplo, los tres triángulos de la figura son congruentes,



el primero se transforma en el segundo mediante un giro (que tiene determinante 1) y es posible realizar la transformación gradualmente sin salir del plano, es decir, podemos desplazar uno de los triángulos hasta que se superponga con el otro. En cambio, el segundo triángulo se transforma en el tercero mediante una simetría axial, y no es posible superponer uno al otro sin sacarlo del plano. Concretamente, para superponerlos es necesario aplicarle a uno un giro en el espacio tomando como eje el eje de simetría.

Lo mismo sucede con figuras tridimensionales. Por ejemplo, es posible llevar un guante derecho sobre otro guante derecho mediante un movimiento gradual que termine superponiéndolos, mientras que no es posible superponer un guante derecho con un guante izquierdo, pese a que también son congruentes (a través de una simetría especular). Para superponerlos tendríamos que girar uno de los guantes respecto al plano de simetría (el espejo) en un espacio de cuatro dimensiones.

En general, las isometrías con determinante positivo se llaman *isometrías directas* o *movimientos*, y son las isometrías que pueden realizarse mediante un movimiento continuo que desplace las figuras sin deformarlas a través del espacio. Las isometrías de determinante negativo se llaman *inversas*, y es imposible realizarlas sin “salirse” del espacio.

Llamaremos $\text{Is}^+(E)$, $\text{Is}^-(E)$ a los conjuntos de las isometrías directas e inversas, respectivamente, de un espacio afín euclídeo E y $\text{O}^+(V)$, $\text{O}^-(V)$ a los conjuntos de las isometrías lineales directas e inversas de un espacio vectorial euclídeo V . Notar que $\text{Is}^+(E)$ y $\text{O}^+(V)$ son grupos, mientras que $\text{Is}^-(E)$ y $\text{O}^-(V)$ no lo son. Los giros y las traslaciones son isometrías directas. En un espacio de dimensión n , el signo de una simetría respecto a una variedad de dimensión m es $(-1)^{n-m}$. En particular las reflexiones son isometrías inversas.

Es obvio que $\text{Is}(E)$ está generado por las traslaciones, los giros y las reflexiones, mientras que $\text{Is}^+(E)$ está generado por las traslaciones y los giros. Los teoremas siguientes nos permitirán refinar estos hechos.

Teorema 6.50 *Si L es una variedad lineal en un espacio afín euclídeo, \vec{v} un vector ortogonal a \vec{L} y $L' = \vec{v} + L$, entonces $s_L \circ s_{L'} = T_{2\vec{v}}$.*

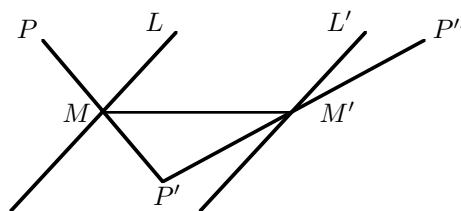
DEMOSTRACIÓN: Tomemos un punto arbitrario P . Sea $P' = s_L(P)$ y sea $P'' = s_{L'}(P')$. Por definición de simetría

$$M = \frac{1}{2}P + \frac{1}{2}P' \in L \quad \text{y} \quad \overrightarrow{PP'} \in \vec{L}^\perp.$$

Llamemos $M' = M + \vec{v} \in L'$. Entonces $\overrightarrow{P'M'} = \overrightarrow{P'M} + \overrightarrow{MM'} = \frac{1}{2}\overrightarrow{P'P} + \vec{v} \in \vec{L}^\perp$, y de nuevo por la definición de simetría $\overrightarrow{P'P''} = 2\overrightarrow{P'M'}$. Así pues,

$$\overrightarrow{PP''} = \overrightarrow{PP'} + \overrightarrow{P'P''} = 2\overrightarrow{MP'} + 2\overrightarrow{P'M'} = 2\overrightarrow{MM'} = 2\vec{v}.$$

■



El teorema anterior implica que toda traslación se puede expresar como composición de dos simetrías (respecto a dos variedades paralelas de dimensión arbitraria). Por consiguiente el grupo $\text{Is}(E)$ está generado por los giros y las reflexiones, y si $\dim E \geq 2$, el grupo $\text{Is}^+(E)$ está generado por los giros (pues cada traslación puede sustituirse por dos simetrías respecto a dos variedades de dimensión $n - 2$, y éstas son giros de amplitud π).

Teorema 6.51 Sea E un espacio afín euclídeo de dimensión $n \geq 2$. Entonces el grupo $\text{Is}(E)$ está generado por las reflexiones y el grupo $\text{Is}^+(E)$ está generado por los giros.

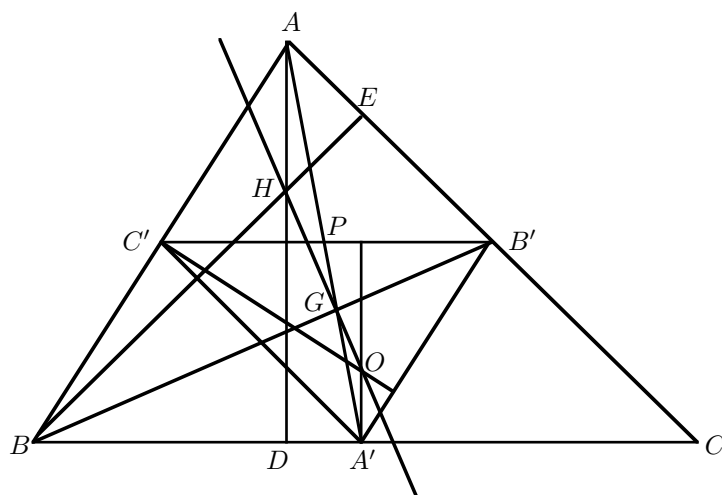
DEMOSTRACIÓN: La segunda parte es evidente. Para probar la primera afirmación es suficiente ver que todo giro se expresa como composición de dos reflexiones. Sea f un giro y $(O; \vec{v}_1, \dots, \vec{v}_n)$ un sistema de referencia ortonormal tal que O sea un punto fijo y \vec{f} tenga forma canónica $[M_\theta, 1, \dots, 1]$ respecto a la base dada. Sea $V = \langle \vec{v}_1, \vec{v}_2 \rangle$. Claramente V es un subespacio invariante para \vec{f} , su restricción a V tiene matriz M_θ y restringida a V^\perp es la identidad.

Sea g la reflexión que fija a O y tal que en la base dada \vec{g} tiene matriz $[-1, 1, \dots, 1]$. Entonces \vec{g} es también la identidad en V^\perp y su restricción a V tiene matriz $[-1, 1]$. Es claro entonces que V y V^\perp son invariantes para $\vec{f} \circ \vec{g}$, así como que ésta es la identidad en el segundo y tiene matriz $M_\theta[-1, 1]$ en V . Como esta matriz es ortogonal y tiene determinante -1 , su forma canónica es $[-1, 1]$, de donde se sigue que la forma canónica de $\vec{f} \circ \vec{g}$ es $[-1, 1, \dots, 1]$, luego $h = f \circ g$ es también una reflexión, y $f = h \circ g$ es composición de dos reflexiones. ■

Ejercicio: Probar que si $\dim E \geq 3$ el grupo $\text{Is}^+(E)$ está generado por las simetrías respecto a variedades de dimensión $n - 2$.

6.7 Aplicaciones

Concluimos el capítulo con algunas aplicaciones de los resultados que hemos visto. En primer lugar probaremos una interesante propiedad de los triángulos descubierta por Euler: El circuncentro, el ortocentro y el baricentro son siempre colineales. Llegaremos a ello estudiando el llamado *triángulo medial*, que es el triángulo que une los pies de las medianas.



En primer lugar observamos que el teorema de Tales implica que los lados del triángulo medial $\widehat{A'B'C'}$ son paralelos a los de \widehat{ABC} , así como que los del primero miden la mitad que los del segundo. El teorema de Desargues implica que ambos son homotéticos, y la razón de la homotecia es -2 . El teorema de Tales prueba también que la mediana AA' pasa por el punto medio P de $B'C'$, luego PA' es una mediana del triángulo medial. En general, las medianas del triángulo medial están contenidas en las del triángulo de partida, luego ambos triángulos tienen el mismo baricentro G .

Es fácil ver que una homotecia conserva el punto medio de un segmento, luego el punto medio P del lado $B'C'$ es el homólogo del punto medio A' del lado BC . Por lo tanto, la mediana PA' del triángulo medial se transforma en la mediana $A'A$ del triángulo dado, es decir, la homotecia deja fijas las medianas (comunes) de ambos triángulos. En particular fija al baricentro G , luego éste ha de ser el centro de la homotecia.

Por otra parte, es obvio que las alturas del triángulo medial son las bisectrices de \widehat{ABC} , luego el circuncentro O de \widehat{ABC} es el ortocentro de su triángulo medial.

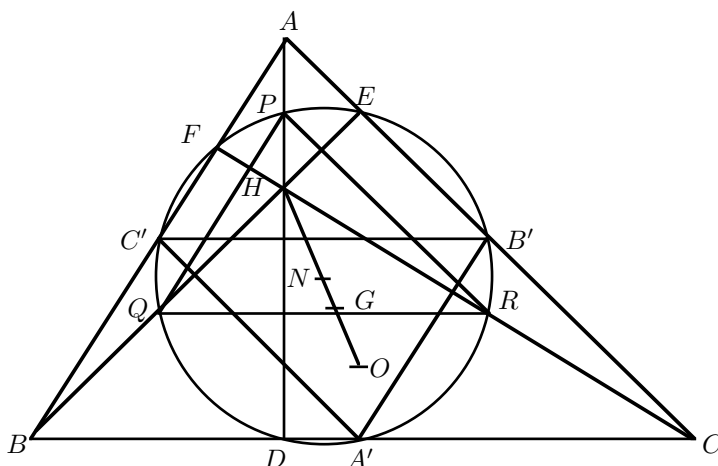
La recta homotética de la altura PA' del triángulo medial es su paralela por el punto A , es decir, la altura AD del triángulo dado. En general, cada altura se corresponde con una altura y por lo tanto el circuncentro O (ortocentro de $\widehat{A'B'C'}$) es el homólogo del ortocentro H de \widehat{ABC} .

Esto prueba que O , H y G se encuentran sobre la misma recta. Más aún, G está entre O y H y $HG = 2GO$.

Hay un caso especial, en el que $G = O$ (y necesariamente entonces $G = H$). Esto equivale a que las mediatrices coincidan con las alturas, lo que a su vez equivale a que el triángulo sea equilátero.

Definición 6.52 Se llama *recta de Euler* de un triángulo no equilátero a la recta que contiene al su circuncentro, su baricentro y su ortocentro.

Todavía podemos llevar más lejos nuestro análisis. Sea N el punto medio del segmento \overline{OH} . Es claro que su homólogo a través de la homotecia anterior es el circuncentro O de \widehat{ABC} , luego N ha de ser el circuncentro del triángulo medial.



Consideramos ahora la homotecia de centro H y razón $1/2$. Ésta transforma el circuncentro O de \widehat{ABC} en el punto N , luego transforma la circunferencia circunscrita de \widehat{ABC} en una circunferencia ω de centro N que contiene a los puntos medios de los segmentos \overline{HA} , \overline{HB} y \overline{HC} , llamémoslos P , Q , R . El triángulo $\widehat{A'B'C'}$ se transforma en \widehat{PQR} mediante la homotecia de centro G y razón -2 seguida de la homotecia de centro H y razón $1/2$ (pasando por \widehat{ABC}), pero estas transforman a N en O y O en N de nuevo, luego la composición fija a N , pero ha de ser una homotecia o una traslación y, como fija a N , ha de ser una homotecia de centro N y razón ± 1 . Es fácil ver que P no puede coincidir con A' , luego la homotecia no puede ser la identidad, luego en definitiva se trata de la homotecia de centro N y razón -1 o, si se prefiere, la simetría puntual respecto de N .

Esto implica que el círculo ω contiene también a los pies de las medianas A' , B' y C' . Más aún, los segmentos $\overline{PA'}$, $\overline{QB'}$, $\overline{RC'}$ son diámetros de ω .

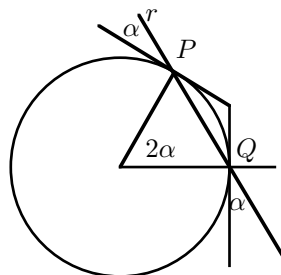
Finalmente, si el pie de la altura D es distinto de A' el triángulo $\widehat{PDA'}$ es rectángulo, luego su circuncentro es N , luego D está también en ω , y lo mismo es válido para E y F . Con esto hemos probado:

Teorema 6.53 *En todo triángulo $T = \widehat{ABC}$ con ortocentro H , los puntos medios de los segmentos \overline{HA} , \overline{HB} y \overline{HC} , los pies de las alturas y los pies de las medianas están sobre una misma circunferencia ω cuyo centro es el punto medio entre el ortocentro y el circuncentro de T .*

Poncelet llamó a la circunferencia del teorema anterior el *círculo de los nueve puntos*⁶ del triángulo T .

El círculo de los nueve puntos de un triángulo T es la circunferencia circunscrita tanto a su triángulo medial como a su triángulo órtico. Veamos otra propiedad interesante. Para enunciarla necesitamos definir el ángulo entre una circunferencia y una recta secante:

Definición 6.54 Dada una circunferencia ω y una recta r que la corte en dos puntos P y Q , llamaremos *ángulo* entre r y ω en P al ángulo (menor) determinado por r y la tangente a ω por P . Si r es tangente a ω en P convendremos en que el ángulo entre r y ω es nulo.



Es fácil ver que r determina el mismo ángulo α tanto en P como en Q , así como que éste es la mitad del arco menor determinado por P y Q en ω , y por lo tanto igual a cualquier ángulo inscrito en ω que abarque el arco \widehat{PQ} .

Teorema 6.55 *Los ángulos con que un lado BC de un triángulo corta a su círculo de los nueve puntos tiene amplitud $|\hat{B} - \hat{C}|$.*

⁶Existe una cierta tradición de llamar círculos tanto a los círculos propiamente dichos como a las circunferencias. Conservamos el nombre usual, si bien hemos de tener presente que el círculo de los nueve puntos es en realidad una circunferencia.

DEMOSTRACIÓN: Con la notación anterior, el lado corta a la circunferencia en los punto A' y D . Es fácil ver que $D = A'$ si y sólo si $B = C$ (en cuyo caso el lado es tangente a la circunferencia), por lo que podemos suponer que $D \neq A'$. Hemos de calcular cualquier ángulo inscrito en la circunferencia que abarque a $\widehat{DA'}$, por ejemplo

$$\widehat{DPA'} = \widehat{DPN} = \widehat{DAO},$$

pues las rectas PN y AO son paralelas (son homotéticas desde H).

Claramente $\widehat{DAC} = \pi/2 - \hat{C}$, mientras que

$$\widehat{OAC} = \widehat{OCA} = \frac{\pi}{2} - \frac{\widehat{AOC}}{2} = \frac{\pi}{2} - \hat{B},$$

pues \hat{B} está inscrito en la circunferencia circunscrita a \widehat{ABC} .

Esto implica que $\widehat{DAO} = |\hat{B} - \hat{C}|$. ■

Capítulo VII

La geometría afín

Nuestra forma de percibir y concebir la realidad contiene las leyes de la geometría euclídea, de modo que somos capaces de reconocer inmediatamente como verdaderos muchos de sus principios, pero sucede que la lógica de nuestra intuición no coincide con la lógica interna (algebraica) de la geometría que percibimos. Por ejemplo, el hecho de que las rectas estén ordenadas aparece como un principio básico porque nos es dado como algo evidente e inmediato, mientras que desde un punto de vista lógico es algo completamente accesorio (la geometría afín contiene el núcleo principal de la geometría y en ella es irrelevante que el cuerpo esté o no ordenado); recíprocamente, el teorema de Desargues dista mucho de ser evidente, y sin embargo veremos que es un auténtico pilar de la geometría afín.

Hasta aquí hemos presentado la geometría partiendo de unos axiomas que pretendían reflejar nuestra forma de percibirla, partiendo de lo intuitivamente más simple hacia lo intuitivamente más complejo. La única excepción ha sido que hemos pospuesto hasta el último lugar al axioma de las paralelas (Axioma E) para mostrar que una gran parte de la teoría es independiente de él, cuando en realidad es tan intuitivamente evidente como los axiomas de incidencia que han constituido nuestro punto de partida. En este capítulo mostraremos una exposición alternativa de la geometría, debida a Artin, y que no está guiada por la intuición, sino por la simplicidad lógica, de modo que nos permitirá comprender el papel real que juega cada elemento en la teoría general.

7.1 Incidencia y paralelismo

Los axiomas básicos de la geometría afín son los axiomas de incidencia juntamente con el axioma de las paralelas. Al incluir este último axioma, los primeros pueden simplificarse ligeramente. Partimos, pues, de un conjunto de puntos en el que hemos destacado dos familias de subconjuntos a los que llamaremos rectas y planos. Los conceptos de colinealidad, etc. se definen como es usual, pero conviene que modifiquemos la noción de paralelismo para aceptar que dos rectas o planos coincidentes se consideren paralelos. De este modo, dos rectas serán

paralelas si están contenidas en un mismo plano y son coincidentes o disjuntas, dos planos serán paralelos si son coincidentes o disjuntos y una recta y un plano serán paralelos si son disjuntos o bien la recta está contenida en el plano.

Definición 7.1 Un *espacio afín* (tridimensional) es un conjunto E , a cuyos elementos llamaremos *puntos*, junto con dos familias de subconjuntos no vacíos de E a cuyos elementos llamaremos *rectas* y *planos*, de modo que se satisfagan los axiomas siguientes:

Axioma A1 *Por cada par de puntos distintos P y Q pasa una única recta, que representaremos por PQ .*

Axioma A2 *Por un punto exterior a una recta pasa una única paralela.*

Axioma A3 *Por cada tres puntos no colineales P, Q, R pasa un único plano, que representaremos por PQR .*

Axioma A4 *Si una recta r tiene dos puntos en común con un plano π , entonces r está contenida en π .*

Axioma A5 *Existen cuatro puntos distintos no coplanares.*

Axioma A6 *Si dos planos tienen un punto en común, entonces tienen dos puntos en común.*

Tenemos, pues, dos axiomas que describen las rectas, dos axiomas que describen los planos, un axioma de existencia de puntos y un axioma de clausura, que impone la tridimensionalidad del espacio.

Del axioma A1 se sigue inmediatamente que si dos rectas se cortan, lo hacen en un único punto. En el capítulo I exigíamos que cada plano contuviera tres puntos no colineales. En realidad esto puede probarse con un argumento técnico que allí preferimos evitar.

Teorema 7.2 *Todo plano contiene tres puntos no colineales*

DEMOSTRACIÓN: Dado un plano π , por definición es un conjunto de puntos no vacío. Sea, pues, $P \in \pi$. Por el axioma A5 existen puntos Q y R no colineales con P . Si ambos están en π ya hemos terminado, en caso contrario el plano PQR corta a π en P y por A6 en otro punto más. No perdemos generalidad si suponemos que es Q , con lo que $R \notin \pi$. De nuevo por A6 existe un punto S no coplanar con P, Q, R . El plano PRS corta a π en P y en otro punto P' que no puede estar en PQ , pues PQ está contenida en PQR (por A4) y en tal caso los planos PRS y PQR tendrían en común los puntos P, R y P' , luego serían iguales (por A3) y P, Q, R, S serían coplanares. Por lo tanto, los puntos P, Q, P' son tres puntos no colineales en π . ■

Por equilibrar la simplicidad con la generalidad hemos dado los axiomas para la geometría tridimensional, aunque sería posible trabajar axiomáticamente con espacios de cualquier dimensión. Es importante destacar el caso bidimensional. Los axiomas para la geometría plana se obtienen de sustituir el axioma A5 por

la existencia de tres puntos no colineales y postular que todos los puntos están contenidos en un mismo plano. Alternativamente podemos suprimir todos los axiomas que nombren planos, con lo que el sistema se reduce a tan sólo tres axiomas:

Axioma A1 *Por cada par de puntos distintos P y Q pasa una única recta, que representaremos por PQ .*

Axioma A2 *Por un punto exterior a una recta pasa una única paralela.*

Axioma A3 *Existen tres puntos distintos no colineales.*

El teorema anterior prueba que todo plano de un espacio afín tridimensional (con las rectas que contiene) es un espacio afín bidimensional (o simplemente un plano afín). Todos los resultados que vamos a probar sobre objetos contenidos en un mismo plano se pueden probar directamente a partir de estos axiomas y son válidos en cualquier plano afín.

Teorema 7.3 *Toda recta contiene al menos dos puntos.*

DEMOSTRACIÓN: Dada una recta r , por definición es un conjunto no vacío, luego contiene un punto P . Sean Q y R dos puntos no colineales con P . Sea s la paralela a PQ por R . Si r no contuviera más punto que P , entonces estaría contenida en el plano PQR , al igual que s , y de hecho serían paralelas, pues P no está en s . Pero entonces r y PQ serían dos paralelas a s por el punto P , lo cual es imposible. ■

Con esto tenemos demostrados todos los axiomas de incidencia de que partimos en el capítulo I, luego tenemos a nuestra disposición todos los teoremas que a partir de ellos probamos allí. También es válida la prueba del teorema 3.5, que enunciamos aquí en una forma equivalente:

Teorema 7.4 *El paralelismo es una relación de equivalencia en el conjunto de todas las rectas del espacio*

La existencia de planos paralelos la demostramos usando el concepto de perpendicularidad. Veamos que es innecesario:

Teorema 7.5 *Por un punto exterior a un plano pasa un único plano paralelo.*

DEMOSTRACIÓN: Sea un plano $\pi = PQR$ y sea P' un punto exterior. Sea $P'Q'$ la paralela a PQ por P' y sea $P'R'$ la paralela a PR por P' . Las rectas $P'Q'$ y $P'R'$ son paralelas a π , pues, por ejemplo, el plano que contiene a PQ y $P'Q'$ corta a π en PQ , y $P'Q'$ no tiene puntos en común con PQ , luego tampoco con π .

Los puntos P' , Q' y R' no son colineales, pues en tal caso $P'Q' = P'R'$, con lo que PQ y PR deberían ser paralelas entre sí, lo cual es absurdo. Sea $\pi' = P'Q'R'$. Veamos que es paralelo a π . En caso contrario la intersección es una recta t que no puede ser paralela a $P'Q'$ y a $P'R'$ simultáneamente, pero es

imposible que $P'Q'$ corte a t , o de lo contrario $P'Q'$ cortaría a π , cuando hemos visto que son paralelos. Lo mismo le ocurre a $P'R'$.

Supongamos ahora que π'' es otro plano paralelo a π que pasa por P' . El plano $P'PQ$ corta a π' y a π'' en dos rectas paralelas a PQ que pasan por P , luego ambas son $P'Q'$ y así $Q' \in \pi''$. Del mismo modo tenemos que $R' \in \pi''$, luego $\pi'' = P'Q'R' = \pi'$. ■

Ahora es inmediato que el paralelismo de planos es una relación de equivalencia. En la prueba del teorema anterior está implícito que si una recta r está contenida en un plano π , entonces la paralela a r por un punto P está contenida en el plano paralelo a π por P .

Definición 7.6 Un *haz de rectas paralelas* es una clase de equivalencia de rectas paralelas, bien de todo el espacio bien en un plano prefijado. Un *haz de rectas concurrentes* es el conjunto de todas las rectas (del espacio o de un plano) que pasan por un mismo punto.

Teorema 7.7 *Todas las rectas tienen el mismo número de puntos.*

DEMOSTRACIÓN: Sean r y s dos rectas. Podemos suponer que tienen un punto en común P , pues dadas dos rectas arbitrarias, podemos tomar una tercera que pase un punto de cada una.

Digamos que $r = PQ$ y $s = PR$. Si $X \in r$, la recta paralela a QR que pasa por X está contenida en el plano de r y s y no puede ser paralela a s , luego la corta en un único punto $f(X)$. Esto define claramente una biyección entre los puntos de r y los de s . ■

Ejercicio: Probar que si las rectas de un espacio afín tienen un número finito p de puntos, entonces todos los planos tienen p^2 puntos y el espacio tiene p^3 puntos.

7.2 Homotecias y traslaciones

El teorema 6.8 caracteriza las homotecias y las traslaciones de un espacio afín como las biyecciones que transforman cada recta en una paralela. Lo más destacable es que no hace falta suponer que las biyecciones sean afines, por lo que estamos en condiciones de tomar esta caracterización como definición.

Definición 7.8 Dado un espacio afín E , llamaremos $HT(E)$ al conjunto de todas las aplicaciones biyectivas de E en sí mismo que transforman cada recta en una paralela. Es claro que $HT(E)$ es un grupo con la composición de aplicaciones. Un elemento de $HT(E)$ es una *homotecia* si tiene al menos un punto fijo y es una *traslación* si es la identidad o bien no tiene puntos fijos.

Probamos en primer lugar un resultado general:

Teorema 7.9 *Si dos elementos de $HT(E)$ coinciden en dos puntos distintos cualesquiera, entonces son iguales.*

DEMOSTRACIÓN: Sean P y Q dos puntos de E y $f \in \text{HT}(E)$. Veamos que la imagen de cualquier otro punto R está determinada por las de P y Q . Supongamos primeramente que R no está en la recta PQ . Entonces $f[PR]$ es una recta paralela a PR , luego es exactamente la recta paralela a PR que pasa por $f(P)$. Similarmente $f[QR]$ es la recta paralela a QR que pasa por $f(Q)$. Como PR y QR no son paralelas entre sí, las rectas $f[PR]$ y $f[QR]$ tampoco lo son. En particular no son coincidentes, pero tienen en común el punto $f(R)$. Por lo tanto $f(R)$ es necesariamente la intersección de la paralela a PR por $f(P)$ con la paralela a QR por $f(Q)$, lo cual sólo depende de $f(P)$ y $f(Q)$. Con esto hemos probado que dos elementos cualesquiera de $\text{HT}(E)$ que coincidan sobre PQ han de coincidir sobre E salvo quizá la recta PQ . Tomando una recta paralela a PQ , sobre la que ya tenemos que coinciden, el razonamiento anterior prueba que también coinciden en PQ . ■

En particular cada homotecia distinta de la identidad tiene un único punto fijo, al que llamaremos *centro* de la homotecia.

Probamos ahora una caracterización de las homotecias y traslaciones mediante una propiedad más débil que nos será útil en la sección siguiente:

Teorema 7.10 *Sea $f : E \rightarrow E$ una aplicación tal que para todo par de puntos P y Q , $f(Q)$ está sobre la recta paralela a PQ que pasa por $f(P)$. Entonces f es constante o bien $f \in \text{HT}(E)$.*

DEMOSTRACIÓN: Supongamos que existen dos puntos P y Q tales que $f(P) = f(Q) = X$. Entonces, para todo punto R exterior a la recta PQ , tenemos que $f(R)$ ha de estar en la recta paralela a PR por X y en la recta paralela a QR por X . Como estas rectas no son paralelas entre sí, se cortan sólo en X , luego $f(R) = X$ para todo punto fuera de PQ . Sustituyendo P y Q por dos puntos en una recta paralela a PQ , el razonamiento anterior nos da que $f(R) = X$ para todo punto de PQ , luego f es constante.

Dicho de otro modo, si f no es constante, entonces es inyectiva. En tal caso consideremos dos puntos cualesquiera P y Q y sea R' un punto exterior a la recta $f(P)f(Q)$. Por hipótesis las rectas PQ y $f(P)f(Q)$ son paralelas, luego están contenidas en un mismo plano π . Sea π' el plano paralelo a $f(P)f(Q)R'$ por P . Éste corta a π en una recta paralela a $f(P)f(Q)$ por P , es decir, en la recta PQ . El plano $Pf(P)R'$ corta a π' en una recta r paralela a $f(P)R'$, y el plano $Qf(Q)R'$ corta a π' en una recta s paralela a $f(Q)R'$ por Q . Como ambas están contenidas en π' y no pueden ser paralelas (porque $f(P)R'$ y $f(Q)R'$ no lo son) necesariamente se cortan en un punto R . Por hipótesis $f(R)$ está en la paralela a PR por $f(P)$ y en la paralela a QR por $f(Q)$, luego está en $r \cap s$ y por lo tanto $f(R) = R'$. Esto prueba que todos los puntos exteriores a la recta $f(P)f(Q)$ tienen antiimagen por f . Cambiando de puntos probamos lo mismo para los puntos de la recta exceptuada. Por lo tanto f es biyectiva.

Finalmente, dada una recta r y un punto $P \in r$, la imagen de cualquier punto $Q \in r$ ha de estar en la paralela a r por $f(P)$, llamémosla r' . Recíprocamente, todo punto de r' es de la forma $f(Q)$, para un cierto punto Q que ha de estar

en r , o en caso contrario $f(Q)$ estaría en la paralela a PQ por $f(P)$, que sería una recta distinta a r' . Por lo tanto $f[r] = r'$ y así $f \in \text{HT}(E)$. ■

Definición 7.11 Una *traza* de $f \in \text{HT}(E)$ es una recta r tal que $f[r] = r$.

Tres hechos obvios sobre trazas son los siguientes:

1. Las trazas de la identidad son todas las rectas.
2. Si $f \in \text{HT}(E)$, entonces cada punto P que no sea un punto fijo para f está en una única traza, a saber en la recta $Pf(P)$.
3. Si dos trazas se cortan en un punto, entonces éste es un punto fijo.

Teorema 7.12 *Las trazas de una homotecia distinta de la identidad son el haz de rectas que pasan por su centro. Las trazas de una traslación distinta de la identidad forman un haz de rectas paralelas al que llamaremos dirección de la traslación.*

DEMOSTRACIÓN: Si f es una homotecia de centro P , entonces toda recta PQ es una traza, pues $f[PQ] = Pf(Q)$ ha de ser una recta paralela a PQ , luego ha de ser PQ . Recíprocamente, si QR es una traza pero no coincide con PQ , que también es una traza, entonces $f(Q) = Q$, luego $f(Q) = Q$, luego $Q = P$. En cualquier caso la recta QR pasa por P .

Si f es una traslación, las trazas de f son claramente las rectas de la forma $Pf(P)$. Hemos de ver que dos cualesquiera de ellas son paralelas. Sea $Qf(Q)$ una recta distinta. Como f no tiene puntos fijos, las dos rectas son disjuntas. Hemos de probar que están sobre un mismo plano. Ahora bien, $f[PQ]$ es una recta paralela a PQ , luego P , Q , $f(P)$ y $f(Q)$ están sobre un mismo plano, luego $Pf(P)$ y $Qf(Q)$ también. ■

Teorema 7.13 *Si dos traslaciones coinciden sobre un punto son iguales.*

DEMOSTRACIÓN: Sea f una traslación y P un punto arbitrario. Sea Q un punto fuera de $Pf(P)$. Entonces las trazas $Pf(P)$ y $Qf(Q)$ son paralelas, luego $Qf(Q)$ es exactamente la paralela a $Pf(P)$ por Q . Igualmente $f(P)f(Q)$ es la paralela a PQ por $f(P)$. Como $Pf(P)$ no es paralela a PQ , las rectas $Qf(Q)$ y $f(P)f(Q)$ no pueden ser paralelas, luego no son coincidentes, luego $f(Q)$ es exactamente la intersección de la paralela a $Pf(P)$ por Q y la paralela a PQ por $f(P)$. Esto prueba que dos traslaciones que coincidan sobre P coinciden también sobre un segundo punto Q , luego son iguales por 7.9. ■

Teorema 7.14 *Las traslaciones forman un subgrupo normal de $\text{HT}(E)$, al que representaremos por $T(E)$. Las traslaciones con una misma dirección forman también un subgrupo normal.*

DEMOSTRACIÓN: Es obvio que la inversa de una traslación es una traslación (si f no tiene puntos fijos, f^{-1} tampoco los tiene).

Si f y g son traslaciones, entonces $f \circ g$ es una traslación, pues si tiene un punto fijo $(f \circ g)(P) = P$, tenemos que $f(P) = g^{-1}(P)$, luego $f = g^{-1}$, ya que ambas son traslaciones y coinciden en un punto. Por consiguiente $f \circ g = 1$, que también es una traslación.

Esto prueba que $T(E)$ es un subgrupo de $HT(E)$. Veamos que es normal. Hemos de ver que si $f \in T(E)$ y $g \in HT(E)$, entonces $g^{-1}fg \in T(E)$. En efecto, si $(g^{-1}fg)(P) = P$, entonces $f(g^{-1}(P)) = g^{-1}(P)$, luego f tiene un punto fijo y es, por tanto, la identidad. Así $g^{-1}fg = 1 \in T(E)$.

Es fácil ver que la composición de dos traslaciones con la misma dirección es la identidad o una traslación con la misma dirección, así como que la inversa de una traslación no trivial tiene la misma dirección.

Si $f \neq 1$ se cumple que $g^{-1}fg$ es una traslación con la misma dirección que f . En efecto, dado $P \in E$, la recta $g^{-1}(P), f(g^{-1}(P))$ es una traza de f , luego tiene la dirección de f . Su imagen por g es $P(g^{-1}fg)(P)$, que es paralela a la anterior y es una traza de $(g^{-1}fg)$, luego ambas traslaciones tienen la misma dirección. ■

Teorema 7.15 *Si existen traslaciones en dos direcciones distintas entonces $T(E)$ es abeliano.*

DEMOSTRACIÓN: Sean f y g dos traslaciones cualesquiera. Hemos de probar que $fg = gf$ o, equivalentemente, que $f^{-1}g^{-1}fg = 1$. Podemos suponer que $f \neq 1 \neq g$. Supongamos que f y g tienen direcciones distintas. Si $f^{-1}g^{-1}fg \neq 1$, entonces $g^{-1}fg$ tiene la misma dirección que f , al igual que f^{-1} y también la composición $f^{-1}g^{-1}fg$ (supuesta no trivial), pero del mismo modo se razona que su dirección es la de g , con lo que tenemos una contradicción.

Si f y g tienen la misma dirección, por hipótesis existe otra traslación h con dirección distinta. Por la parte ya probada $fh = hf$. Por otra parte, gh no puede tener la misma dirección que f , o de lo contrario $g^{-1}(gh) = h$ tendría la misma dirección. Así pues, $fgh = ghf$. Uniendo todo esto tenemos que $fgh = ghf = gfh$, luego $fg = gf$. ■

7.3 Vectores y escalares

La insólita hipótesis del teorema anterior nos hace reparar en que ninguno de los teoremas anteriores nos garantiza que el grupo $HT(E)$ sea no trivial. Garantizar la existencia de homotecias y traslaciones es un problema técnico del que nos ocuparemos en la sección siguiente. Para sortear el obstáculo momentáneamente introduciremos dos axiomas técnicos de carácter provisional. El primero se refiere a traslaciones:

Axioma B1 *Dados dos puntos P y Q , existe una traslación T_{PQ} que transforma P en Q .*

Puesto que las traslaciones están determinadas por la imagen de un único punto, la traslación T_{PQ} es única. Ahora es obvio que existen traslaciones con direcciones distintas, por lo que el grupo de traslaciones es abeliano.

Con esto podemos recuperar el conjunto de los vectores libres. Llamaremos $\vec{E} = T(E)$, usaremos notación aditiva, de modo que $\vec{0}$ será la identidad. La traslación T_{PQ} la representaremos por \vec{PQ} , de modo que todo vector de \vec{E} se puede expresar de forma única como \vec{OP} para cualquier punto prefijado O .

La imagen de un punto R por la traslación \vec{PQ} la representaremos por $R + \vec{PQ}$. Las afirmaciones siguientes son reformulaciones de hechos ya probados:

1. $\vec{PP} = \vec{0}$,
2. $\vec{PQ} + \vec{QR} = \vec{PR}$,
3. $(P + \vec{v}) + \vec{w} = P + (\vec{v} + \vec{w})$.

Ahora buscamos un cuerpo con el que convertir a \vec{E} en un espacio vectorial. En el capítulo anterior probamos que en un espacio afín (en el sentido algebraico) K^* es isomorfo al grupo de homotecias lineales $H(\vec{E})$. Más aún, si añadimos la aplicación nula a $H(\vec{E})$, el conjunto que obtenemos puede ser dotado de forma natural de estructura de cuerpo isomorfo a K . Basta definir la suma de dos homotecias f y g como la aplicación dada por $(f + g)(\vec{v}) = f(\vec{v}) + g(\vec{v})$. Fijando una base, esta suma y la composición se corresponden con la suma y el producto de matrices, pero las matrices de las homotecias lineales (más la aplicación nula) son las de la forma αI , donde $\alpha \in K$, luego efectivamente, con estas operaciones tenemos un cuerpo isomorfo a K .

Nuestro problema es esencialmente definir las homotecias lineales en nuestro contexto, pero una homotecia lineal se caracteriza porque envía cada vector a un múltiplo suyo y, vistos como traslaciones, esto significa que ambos tienen la misma dirección.

Definición 7.16 Una *homotecia lineal* en un espacio afín E es un homomorfismo de grupos $\alpha : \vec{E} \rightarrow \vec{E}$ tal que para todo $\vec{v} \in \vec{E}$, se cumple $\alpha\vec{v} = \vec{0}$ o bien $\alpha\vec{v}$ tiene la misma dirección que \vec{v} . Llamaremos K al conjunto de todas las homotecias lineales de E . A los elementos de K los llamaremos también *escalares*.

Observar que el hecho de que una homotecia lineal α sea un homomorfismo de grupos equivale a que $\alpha(\vec{v} + \vec{w}) = \alpha\vec{v} + \alpha\vec{w}$.

El homomorfismo nulo, determinado por $0\vec{v} = \vec{0}$ para todo $\vec{v} \in \vec{E}$, es trivialmente una homotecia lineal. Lo mismo le sucede a la identidad, determinada por $1\vec{v} = \vec{v}$. Llamaremos -1 a la aplicación dada por $(-1)\vec{v} = -\vec{v}$. Es claro que también es una homotecia lineal.

Definición 7.17 Sean $\alpha, \beta \in K$, definimos $\alpha + \beta$ como la aplicación $\vec{E} \rightarrow \vec{E}$ dada por $(\alpha + \beta)\vec{v} = \alpha\vec{v} + \beta\vec{v}$. Definimos $\alpha\beta$ como la aplicación dada por $(\alpha\beta)\vec{v} = \beta(\alpha\vec{v})$.

Notar que, vistos como homotecias lineales, el producto de dos escalares es simplemente su composición.

Teorema 7.18 *El conjunto de escalares K es un anillo unitario con las operaciones de la definición anterior.*

DEMOSTRACIÓN: Claramente,

$$(\alpha + \beta)(\vec{v} + \vec{w}) = \alpha(\vec{v} + \vec{w}) + \beta(\vec{v} + \vec{w}) = \alpha\vec{v} + \alpha\vec{w} + \beta\vec{v} + \beta\vec{w} = (\alpha + \beta)\vec{v} + (\alpha + \beta)\vec{w},$$

luego $\alpha + \beta$ es un homomorfismo. Si $\vec{v} = \vec{0}$ entonces $(\alpha + \beta)\vec{v} = \vec{0}$. Si $\vec{v} \neq \vec{0}$ entonces $\alpha\vec{v}$ y $\beta\vec{v}$ tienen la misma traza que \vec{v} , luego su suma también, es decir, $(\alpha + \beta)\vec{v}$ tiene la misma dirección que \vec{v} , luego $\alpha + \beta \in K$.

Del mismo modo se prueba que $\alpha\beta \in K$. La asociatividad y la conmutatividad de la suma de siguen de las de la suma en \vec{E} . El neutro es 0 y el simétrico de α es $(-1)\alpha$, pues

$$(\alpha + (-1)\alpha)\vec{v} = \alpha\vec{v} + -\alpha\vec{v} = \vec{0} = 0\vec{v},$$

luego $\alpha + (-1)\alpha = 0$.

Tampoco presenta ningún problema probar la propiedad distributiva (por la izquierda y la derecha), la asociatividad del producto y el hecho de que 1 es el elemento neutro del producto (por la izquierda y la derecha). ■

Para justificar la existencia de inversos mostraremos primero la relación entre las homotecias y las homotecias lineales:

Teorema 7.19 *Sea $\alpha \in K$ no nulo y P un punto dado. Existe una única homotecia f_α de centro P tal que $T_{\alpha\vec{v}} = f_\alpha^{-1}T_{\vec{v}}f_\alpha$, para todo $\vec{v} \in \vec{E}$.*

DEMOSTRACIÓN: Observar que usamos la doble notación $\vec{v} = T_{\vec{v}}$ según si vemos la traslación como elemento de \vec{E} (con notación aditiva) o como elemento de $\text{HT}(E)$ (con notación multiplicativa).

Veamos en primer lugar la unicidad. Sea Q un punto arbitrario y $\vec{v} = T_{PQ}$. Aplicando la fórmula del enunciado al punto P tenemos

$$P + \alpha\overrightarrow{PQ} = T_{\alpha\vec{v}}(P) = f_\alpha(T_{\vec{v}}(f_\alpha^{-1}(P))) = f_\alpha(T_{\vec{v}}(P)) = f_\alpha(Q).$$

Esto prueba que $f_\alpha(Q)$ está completamente determinado por α y P . Además la fórmula anterior nos indica cómo hemos de construir f_α .

Definimos f_α mediante

$$f_\alpha(Q) = P + \alpha\overrightarrow{PQ}. \quad (7.1)$$

Dados dos puntos distintos Q y R , tenemos que $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$, luego $P + \alpha\overrightarrow{PQ} + \alpha\overrightarrow{QR} = P + \alpha\overrightarrow{PR}$. Por (7.1) esto implica

$$f_\alpha(Q) + \alpha\overrightarrow{QR} = f_\alpha(R). \quad (7.2)$$

Si $\alpha \overrightarrow{QR} \neq \vec{0}$ entonces esta ecuación significa que la recta $f_\alpha(Q)f_\alpha(R)$ es una traza de $\alpha \overrightarrow{QR}$, luego también de \overrightarrow{QR} , luego es paralela a la recta QR , que es otra traza de la misma traslación. Si $\alpha \overrightarrow{QR} = \vec{0}$ tenemos que $f_\alpha(Q) = f_\alpha(R)$. En ambos casos $f_\alpha(R)$ está sobre la recta paralela a QR por $f_\alpha(Q)$. Podemos aplicar el teorema 7.10 para concluir que f_α es constante o bien una homotecia (no puede ser una traslación no trivial porque tiene a P como punto fijo). Ahora bien, si fuera constante, la fórmula (7.2) implicaría que $\alpha \overrightarrow{QR} = \vec{0}$ para todo vector \overrightarrow{QR} , luego sería $\alpha = 0$, en contra de la hipótesis. Así pues, f_α es una homotecia de centro P .

Veamos que f_α cumple lo pedido. Puesto que $f_\alpha(P) = P$, podemos escribir (7.1) como

$$f_\alpha(Q) = f_\alpha(P) + \alpha \overrightarrow{PQ} = T_{\alpha \vec{v}}(f_\alpha(P)),$$

donde $\vec{v} = \overrightarrow{PQ}$ es un vector arbitrario, luego

$$Q = f_\alpha^{-1}(T_{\alpha \vec{v}}(f_\alpha(P))).$$

Esto significa que la traslación $f_\alpha T_{\alpha \vec{v}} f_\alpha^{-1}$ envía P a Q , luego

$$f_\alpha T_{\alpha \vec{v}} f_\alpha^{-1} = T_{PQ} = T_{\vec{v}}.$$

Despejando obtenemos la fórmula del enunciado: $T_{\alpha \vec{v}} = f_\alpha^{-1} T_{\vec{v}} f_\alpha$. ■

Teorema 7.20 *Cada homotecia se expresa de forma única como f_α , para un cierto escalar α no nulo.*

DEMOSTRACIÓN: Sea f una homotecia de centro P . Sea α la homotecia lineal dada por $T^\alpha = f^{-1} T f$ (usamos notación exponencial para los escalares cuando usamos notación multiplicativa para las traslaciones). Es fácil ver que α es un homomorfismo: $(TT')^\alpha = f^{-1} T T' f = f^{-1} T f f^{-1} T' f = T^\alpha T'^\alpha$. Además α conserva las direcciones por el teorema 7.14. Por lo tanto α es un escalar, claramente no nulo. Cambiando la notación en la definición de α tenemos que $T_{\alpha \vec{v}} = f^{-1} T_{\vec{v}} f$ para todo vector \vec{v} , luego por la unicidad ha de ser $f = f_\alpha$.

La unicidad se sigue de (7.1), pues si $f_\alpha = f_\beta$, entonces $\alpha \overrightarrow{PQ} = \beta \overrightarrow{PQ}$ para todo vector \overrightarrow{PQ} , luego $\alpha = \beta$. ■

Con esto ya estamos en condiciones de probar:

Teorema 7.21 *El conjunto de escalares K es un anillo de división.*

DEMOSTRACIÓN: Recordemos que un anillo de división es un cuerpo no necesariamente conmutativo. Sólo falta probar que todo escalar no nulo tiene un inverso. Más en general, observamos que la biyección $\alpha \mapsto f_\alpha$ entre K^* y el grupo de las homotecias de centro un punto fijo P conserva el producto, es decir, $f_{\alpha\beta} = f_\alpha f_\beta$. En efecto, se cumple

$$(f_\alpha f_\beta)^{-1} T_{\vec{v}} f_\alpha f_\beta = f_\beta^{-1} T_{\alpha \vec{v}} f_\beta = T_{\beta(\alpha \vec{v})} = T_{(\alpha\beta) \vec{v}}.$$

Por la unicidad $f_\alpha f_\beta = f_{\alpha\beta}$. Esto implica que K^* es un grupo isomorfo al grupo de homotecias de centro P . En particular todo escalar no nulo tiene un inverso. ■

Con esto tenemos probado que \vec{E} es un K espacio vectorial por la derecha, es decir, con la asociatividad mixta en la forma $(\alpha\beta)\vec{v} = \beta(\alpha\vec{v})$. Enseguida veremos que la dimensión de \vec{E} es 3 (o 2 si E es un plano afín). Con ello tendremos probado que E tiene estructura de espacio afín en el sentido algebraico del capítulo IV, salvo por el hecho de que K no es necesariamente conmutativo. No podemos aspirar a demostrar la conmutatividad de K a partir de nuestros axiomas, pues es fácil ver que si K es cualquier anillo de división entonces K^3 satisface los axiomas de espacio afín tomando como rectas y planos las clases de congruencia módulo los subespacios vectoriales de dimensión 1 y 2, y el anillo de división que se obtiene según la construcción anterior es isomorfo a K , luego no es necesariamente conmutativo. La geometría afín básica es válida en el caso no conmutativo, pero hay algunas diferencias, como por ejemplo que las homotecias no son necesariamente afinidades.

Para calcular la dimensión de \vec{E} necesitamos un axioma adicional B2 que introducimos también con carácter provisional. En la sección siguiente demostraremos los axiomas B1 y B2 a partir del grupo A.

Axioma B2 *Si \vec{v} y \vec{w} son dos vectores distintos y no nulos con la misma dirección, entonces existe un $\alpha \in K$ tal que $\vec{w} = \alpha\vec{v}$.*

Los casos exceptuados se cumplen trivialmente: si $\vec{v} = \vec{w}$ sirve $\alpha = 1$ y si $\vec{w} = \vec{0}$ sirve $\alpha = 0$. Por lo tanto sólo hay que exigir que \vec{v} sea no nulo. La estructura vectorial de \vec{E} implica que α es único.

En primer lugar daremos una forma equivalente del axioma B2. Si el axioma B1 garantizaba la existencia de traslaciones, el axioma B2 postula el hecho análogo sobre homotecias.

Axioma B2P (para un punto P) *Para todo par de puntos Q, R tales que P, Q, R estén alineados y sean distintos, existe una homotecia de centro P que envía Q a R .*

Teorema 7.22 *El axioma B2 para un punto P equivale al axioma B2 para todo punto P y también al axioma B2.*

DEMOSTRACIÓN: Supongamos el axioma B2 para un punto P y sean \vec{v}, \vec{w} en las condiciones del axioma B2 (global). Sean $Q = P + \vec{v}$, $R = P + \vec{w}$. Como los dos vectores tienen la misma dirección, los tres puntos son colineales. Como son distintos y no nulos, los tres puntos P, Q, R son distintos. Por B2P existe una homotecia de centro P , que será de la forma f_α , tal que $f_\alpha(Q) = R$, es decir, $P + \alpha\vec{PQ} = R$, luego $\alpha\vec{PQ} = \vec{PR}$, o también, $\alpha\vec{v} = \vec{w}$.

Supongamos ahora B2 y probemos B2P para un punto arbitrario. Sean P, Q, R tres puntos colineales distintos. Entonces \vec{PQ} y \vec{PR} son vectores no

nulos distintos con la misma dirección. por B2 existe un escalar no nulo tal que $\overrightarrow{PR} = \alpha \overrightarrow{PQ}$. Equivalentemente, $P + \alpha \overrightarrow{PQ} = R$, luego $f_\alpha(Q) = R$. ■

Teorema 7.23 *Si P, Q, R, S son cuatro puntos no coplanares, entonces los vectores $\overrightarrow{PQ}, \overrightarrow{PR}, \overrightarrow{PS}$ son una base de \vec{E} .*

DEMOSTRACIÓN: Sea \overrightarrow{PX} un vector arbitrario. Los planos PQX y PRS son distintos, luego se cortan en una recta r que contiene a P . La paralela a PQ por X está contenida en el primero y ha de cortar a r en un punto $X' \in PRS$. El vector $\overrightarrow{X'X}$ es nulo o tiene la misma dirección que \overrightarrow{PQ} , luego por B2 existe un escalar α tal que $\overrightarrow{X'X} = \alpha \overrightarrow{PQ}$. Por lo tanto

$$\overrightarrow{PX} = \overrightarrow{PX'} + \overrightarrow{X'X} = \alpha \overrightarrow{PQ} + \overrightarrow{PX'}.$$

La paralela a PR por X' está contenida en PRS y no es paralela a PS , luego se cortan en un punto X'' . El vector $\overrightarrow{X''X'}$ es nulo o tiene la misma dirección que \overrightarrow{PR} , luego $\overrightarrow{X''X'} = \beta \overrightarrow{PR}$, para cierto $\beta \in K$. Así

$$\overrightarrow{PX'} = \overrightarrow{PX''} + \overrightarrow{X''X'} = \beta \overrightarrow{PR} + \overrightarrow{PX''}.$$

Así mismo, el vector $\overrightarrow{PX''}$ es nulo o tiene la dirección de \overrightarrow{PS} , luego existe un $\gamma \in K$ tal que $\overrightarrow{PX''} = \gamma \overrightarrow{PS}$. En total obtenemos

$$\overrightarrow{PX} = \alpha \overrightarrow{PQ} + \overrightarrow{PX'} = \alpha \overrightarrow{PQ} + \beta \overrightarrow{PR} + \overrightarrow{PX''} = \alpha \overrightarrow{PQ} + \beta \overrightarrow{PR} + \gamma \overrightarrow{PS}.$$

Si $\alpha \overrightarrow{PQ} + \beta \overrightarrow{PR} + \gamma \overrightarrow{PS} = \vec{0}$, entonces

$$P - \gamma \overrightarrow{PS} = P + \alpha \overrightarrow{PQ} + \beta \overrightarrow{PR}.$$

El punto $X = P + \alpha \overrightarrow{PQ}$ está en la recta PQ , contenida en PQR , el punto $P + \alpha \overrightarrow{PQ} + \beta \overrightarrow{PR}$ está en la paralela a PR por X , luego está también en PQR , mientras que el punto $P - \gamma \overrightarrow{PS}$ está en la recta PS , cuyo único punto en común con PQR es P . Por lo tanto, $P - \gamma \overrightarrow{PS} = P$, lo que implica que $\gamma = 0$. Del mismo modo se concluye que $\alpha = \beta = 0$. ■

7.4 Los teoremas de Desargues y Papos-Pascal

En esta sección veremos que los axiomas del grupo B se pueden demostrar a partir de los del grupo A. Más concretamente, resultan ser equivalentes al teorema de Desargues. Recordemos su enunciado:

Teorema de Desargues Sean r_1, r_2 y r_3 tres rectas distintas paralelas o concurrentes en un punto P . Sean R, R' puntos en r_1 , sean S, S' puntos en r_2 y T, T' puntos en r_3 distintos de P si las rectas se cortan. Si RS es paralela a $R'S'$ y RT es paralela a $R'T'$ entonces ST es paralela a $S'T'$.

Este enunciado es aparentemente más fuerte que el que dimos en el capítulo VI, pero observemos que si los puntos R, S, T están alineados, lo mismo sucede con R', S', T' , luego el teorema es trivialmente cierto. Así mismo, si $Q = Q'$ entonces $R = R'$ y $T = T'$.

Vamos a probar este teorema a partir de los axiomas del grupo A, pero antes probaremos su equivalencia con los axiomas B1 y B2. Llamaremos DA al enunciado anterior cuando las rectas son paralelas y DP al enunciado anterior cuando las rectas se cortan en el punto P . Trabajamos en un espacio afín E sin suponer más que los axiomas del grupo A.

Teorema 7.24 *DA equivale al axioma B1 y DP equivale a B2P.*

DEMOSTRACIÓN: Suponiendo B1 (o B2P) tomamos como f la traslación (o la homotecia de centro P) determinada por $f(Q) = Q'$. Entonces las rectas r_1, r_2, r_3 son trazas de f . La recta $f(Q)f(R)$ es paralela a QR y pasa por $f(Q) = Q'$, luego es $Q'R'$, y como $f(R)$ ha de estar en esta recta y en r_2 , ha de ser $f(R) = R'$, e igualmente $f(S) = S'$. Por consiguiente $R'S' = f[RS]$ es paralela a RS .

Veamos ahora que DA implica B1. Tomemos dos puntos distintos Q y Q' . Vamos a construir una traslación que envíe uno sobre el otro. En primer lugar definimos una aplicación $T^{QQ'}$ definida sólo sobre los puntos exteriores a la recta QQ' . Dado un punto R en estas condiciones, tomamos la recta r paralela a QQ' por R (con lo que $r \neq QQ'$). Las rectas QQ', QR y r están en un mismo plano, luego éste contiene también a la recta paralela a QR por Q' , que cortará a r en un punto R' . Definimos $T^{QQ'}(R) = R'$.

El punto R' está caracterizado por el hecho de que RR' es paralela a QQ' y QR es paralela a $Q'R'$. Es claro entonces que $T^{RR'}(Q) = Q'$. Si probamos que $T^{QQ'}$ y $T^{PP'}$ coinciden en su dominio común, entonces definirán una aplicación sobre todo E , a la que llamaremos $T_{QQ'}$ y probaremos que es la traslación buscada.

Sea, pues, S un punto exterior a QQ' y RR' y sea $S' = T^{QQ'}(S)$. Entonces tenemos tres rectas paralelas distintas, QQ', RR' y SS' . Además QS es paralela a $Q'S'$ y QR es paralela a $Q'R'$. Por DA podemos concluir que RS es paralela a $R'S'$, lo que implica que $S' = T^{RR'}(S)$.

Para probar que $T_{QQ'}$ es una traslación basta probar que cumple las condiciones del teorema 7.10, pues ciertamente $T_{QQ'}$ no es constante y tiene dos trazas paralelas, a saber, QQ' y RR' , luego será una traslación. Tomamos dos puntos U, V y hemos de probar que $V' = T_{QQ'}(V)$ está en la recta paralela a UU' por $U' = T_{QQ'}(U)$. Si U está en QQ' y V está en RR' , podemos tomar un punto S exterior al plano que contiene a estas dos. La aplicación $T^{SS'}$ coincidirá con $T_{QQ'}$ en su dominio común. En cualquier caso, tomando bien, QQ' , bien RR' o bien SS' , encontramos una recta que no contiene a ninguno de los puntos U y V . Podemos suponer que es QQ' .

Si UV es paralela a QQ' , entonces UV contiene también a U' y V' , con lo que ciertamente V' está en la paralela a UV por U' . Si UV no es paralela a

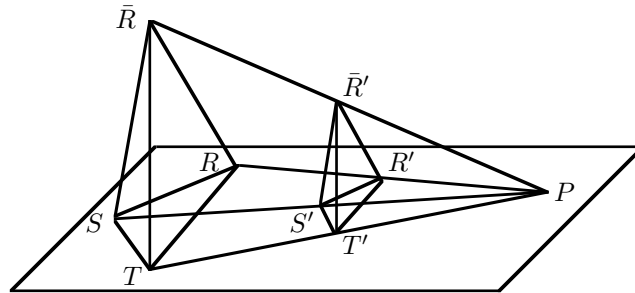
QQ' , entonces Q , U y V están dispuestos como antes Q , R y S , y el mismo razonamiento prueba que $U'V'$ es paralelo a UV .

La prueba de que DP implica el axioma $B2P$ es completamente análoga, cambiando traslaciones por homotecias y líneas paralelas por líneas que pasan por P . ■

Nota Todos los resultados que hemos probado hasta ahora se adaptan sin dificultad (de hecho se simplifican) para el caso de un plano afín. La adaptación del teorema anterior requiere una consideración adicional. En un punto de la prueba hemos usado que siempre es posible tomar un punto exterior a dos rectas paralelas (o concurrentes en el caso $DP \Rightarrow B2P$). Para ello hemos tomado un punto exterior al plano que las contiene, lo cual no es viable si el espacio es plano. Si toda recta contiene al menos tres puntos no hay ningún problema, pues basta tomar una recta que corte a las dos dadas y tomar un tercer punto en ella. El único caso peculiar se da si las rectas tienen exactamente dos puntos, pero entonces el plano consta únicamente de cuatro puntos (y seis rectas). En este caso el axioma $B2P$ se cumple trivialmente y la existencia de traslaciones se prueba sin dificultad.

DEMOSTRACIÓN: (del teorema de Desargues) Supongamos primero que los planos RST y $R'S'T'$ son distintos. Entonces son paralelos, pues las rectas $R'S'$ y $R'T'$ están contenidas en el plano paralelo a RST por R' , luego éste es $R'S'T'$. El plano que contiene a r_2 y r_3 corta a RST en ST y a $R'S'T'$ en $S'T'$, luego ambas rectas están contenidas en un mismo plano y son paralelas, pues están contenidas en planos paralelos.

Supongamos ahora que $RST = R'S'T'$. Tomemos un punto \bar{R} fuera de este plano. Sea \bar{r} la recta paralela a las r_i por \bar{R} si éstas son paralelas o bien $\bar{r} = \bar{R}P$ si concurren en P .



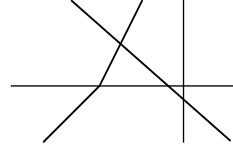
La recta paralela a $R\bar{R}$ por R' está contenida en el plano que contiene a r_1 y a \bar{r} , luego ha de cortar a \bar{r} en un punto \bar{R}' . Los planos $\bar{R}RS$ y $\bar{R}'R'S'$ son distintos y podemos aplicar el caso ya probado del teorema de Desargues para concluir que $\bar{R}S$ es paralela a $\bar{R}'S'$. Del mismo modo concluimos que $\bar{R}T$ es paralela a $\bar{R}'T'$. A su vez esto nos permite aplicar la parte ya probada a los puntos $\bar{R}ST$ y $\bar{R}'S'T'$, lo que nos da que ST es paralela a $S'T'$, como queríamos probar. ■

La demostración anterior es esencialmente tridimensional. El ejemplo siguiente muestra que no es posible demostrar el teorema de Desargues a partir de los axiomas de plano afín.

Ejemplo Vamos a construir un plano afín que no satisface el teorema de Desargues. Tomamos $\Pi = \mathbb{R}^2$ como conjunto de puntos, aunque las rectas de Π no serán las usuales. Una recta usual en \mathbb{R}^2 está formada por los puntos (x, y) que cumplen una ecuación del tipo $ax + by = c$. Si $b = 0$ tenemos una recta vertical de la forma $x = n$, mientras que si $b \neq 0$ la ecuación es equivalente a una de la forma $y = mx + n$, donde m y n están unívocamente determinados por la recta. Tomamos como rectas de Π las rectas usuales de la forma $x = n$ (verticales), las de la forma $y = mx + n$ con $m \leq 0$ (lo cual incluye a las horizontales) y, para cada $m, n \in \mathbb{R}$, $m > 0$, los conjuntos de pares (x, y) que cumplen

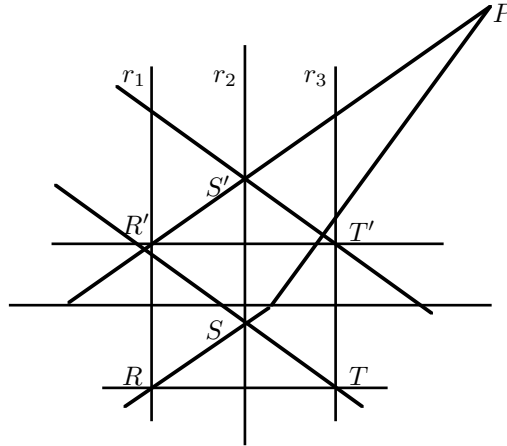
$$y = \begin{cases} 2mx + n & \text{si } y \geq 0 \\ mx + n & \text{si } y < 0 \end{cases}$$

cada uno de estos conjuntos es una línea quebrada formada por la unión de dos semirrectas con origen común en el punto $(n, 0)$, una contenida en el semiplano superior $y \geq 0$ y otra en el inferior. La figura muestra algunas rectas de Π .



La comprobación de los tres axiomas del plano afín no presenta ninguna dificultad. El menos evidente es la existencia de una única recta que pasa por dos puntos (a, b) y (c, d) cuando $b < 0$, $d > 0$ y $a < b$. Tal recta ha de ser necesariamente de las quebradas, y basta ver que las ecuaciones $ma + n = b$, $2mc + n = d$ tienen solución única (m, n) con $m > 0$.

La figura siguiente muestra un contraejemplo al teorema de Desargues: vemos dos triángulos en posición de Desargues con dos pares de lados paralelos, mientras que el tercero se corta en P .



Dejamos que el lector complete los detalles. ■

Ejercicio: Probar que en Π no hay traslaciones cuyas trazas sean las rectas verticales. Mostrar un contraejemplo del teorema de Desargues en el que las rectas r_i sean concurrentes.

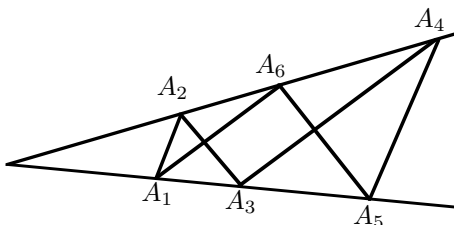
Definición 7.25 Un plano afín es *arguesiano* si en él se cumple el teorema de Desargues.

Si un plano afín Π es sumergible en un espacio afín E (en el sentido de que $\Pi \subset E$ y las rectas de E contenidas en Π son exactamente las rectas de Π) entonces Π es arguesiano, pues E cumple el teorema de Desargues. Recíprocamente, un plano afín arguesiano es isomorfo a un plano afín de la forma K^2 , para un cierto anillo de división K , que puede sumergirse en K^3 , que es un espacio afín. Así pues, los planos arguesianos son exactamente los planos sumergibles en espacios.

Ahora mostraremos un enunciado puramente geométrico equivalente a la conmutatividad del anillo de división asociado a un espacio afín (o un plano arguesiano). Ya hemos observado que el grupo K^* es isomorfo al grupo de las homotecias con un mismo centro P . Así pues, K será un cuerpo si y sólo si este grupo es abeliano.

Definición 7.26 Un *hexágono* es un conjunto de seis puntos ordenados $A_1, A_2, A_3, A_4, A_5, A_6$. Llamaremos *lados* del hexágono a las seis rectas $A_1A_2, A_2A_3, A_3A_4, A_4A_5, A_5A_6, A_6A_1$. Dos lados de un hexágono son *contiguos* si tienen un vértice en común, y en caso contrario son *opuestos*.

Teorema 7.27 (Teorema de Pappos–Pascal) Sean r y s dos rectas secantes y consideremos un hexágono tal que los vértices A_1, A_3, A_5 están sobre r y A_2, A_4, A_6 están sobre s (todos ellos distintos del punto de intersección). Si dos pares de lados opuestos son paralelos, el tercer par también lo es.



Vamos a probar que este teorema se cumple exactamente en los espacios afines cuyo anillo de división asociado es un cuerpo. Puesto que todos sus elementos están contenidos en un plano, de hecho se cumple en los espacios afines de dimensión arbitraria.

DEMOSTRACIÓN: Supongamos en primer lugar el teorema de Pappos y veamos que K es conmutativo. Sean r y s dos rectas que se corten en un punto P . Sea A_1 un punto sobre r distinto de P . Sea f una homotecia arbitraria de centro P (no trivial). Entonces $f(A_1) = A_3 \neq A_1$ está en r y eligiendo la homotecia apropiada podemos hacer que sea cualquier punto de r prefijado.

Tomemos otra homotecia no trivial g de centro P y un punto A_2 en s distinto de P . Sea $A_6 = g(A_2) \neq A_2$, que también estará en s y es un punto arbitrario si elegimos g .

Sean $A_5 = g(f(A_1)) = g(A_3)$ y $A_4 = f(g(A_2)) = f(A_6)$. Claramente A_1A_6 es paralela a $f(A_1)f(A_6) = A_3A_4$ y A_2A_3 es paralela a $g(A_2)g(A_3) = A_6A_5$. Por el teorema de Papos también A_1A_2 es paralela a A_4A_5 .

La conmutatividad $gf = fg$ equivale a que $f(g(A_1)) = g(f(A_1))$, es decir, $f(g(A_1)) = A_5$, pero $f(g(A_1))$ es el punto donde la paralela a A_1A_2 por $f(g(A_2)) = A_4$ corta a r , es decir, se trata efectivamente de A_5 .

Recíprocamente, dado un hexágono en las condiciones del teorema de Papos, digamos con A_1A_6 paralela a A_3A_4 y con A_2A_3 paralela a A_6A_5 , consideramos la homotecia f de centro P que cumple $f(A_1) = A_3$ y la homotecia g de centro P que cumple $g(A_2) = A_6$. Invertiendo el razonamiento y usando que $fg = gf$ concluimos que A_1A_2 es paralela a A_4A_5 . ■

En la sección siguiente veremos que una forma de garantizar geoméricamente la conmutatividad de K es mediante axiomas de ordenación. El teorema siguiente prueba que todos los espacios afines finitos (de dimensión mayor que 2 junto todos los planos arguesianos finitos) son conmutativos.

Teorema 7.28 (Teorema de Wedderburn) *Todo anillo de división finito es un cuerpo.*

DEMOSTRACIÓN: Sea K un anillo de división finito. Sea Z el conjunto de todos los elementos de K que conmutan con todos los demás elementos de K . Es fácil ver que se trata de un cuerpo, por lo que K es un espacio vectorial sobre Z . Si éste tiene q elementos, entonces el número de elementos de K es q^n , donde n es la dimensión de K sobre Z .

Para cada $\alpha \in K$, el conjunto C_α formado por todos los elementos de K que conmutan con α es también un anillo de división que contiene a Z , luego su número de elementos es q^d , donde $d \mid n$ (por la transitividad de grados). Indicamos con un asterisco el conjunto que resulta de eliminar el 0. Entonces Z^* es el centro del grupo K^* y C_α^* es el centralizador de $\alpha \neq 0$. La ecuación de clases que probamos en el apéndice del capítulo V es en este caso:

$$q^n - 1 = q - 1 + \sum_d \frac{q^n - 1}{q^d - 1},$$

donde d recorre ciertos divisores de n , posiblemente repetidos, distintos de n .

Sea $c_n(X)$ el polinomio ciclotómico de orden n . Es claro que

$$x^n - 1 = c_n(x)f(x), \quad \frac{x^n - 1}{x^d - 1} = c_n(x)g(x),$$

para ciertos polinomios $f(x), g(x) \in \mathbb{Z}[x]$.

Evaluando en q vemos que $c_n(q)$ divide a todos los términos de la ecuación de clases. En particular $c_n(q) \mid q - 1$.

Por otra parte, $c_n(q)$ es el producto de factores de la forma $q - \zeta$, donde ζ recorre las raíces n -simas primitivas de la unidad. Claramente

$$|q - \zeta| \geq |q - |\zeta|| = |q - 1| \geq 1,$$

y como $q \geq 2$ la igualdad sólo se da si $\zeta = 1$, y entonces $n = 1$. En otro caso concluimos que $|c_n(q)| > q - 1$, lo que nos da una contradicción. Así pues, $n = 1$ y $K = \mathbb{Z}$ es un cuerpo. ■

7.5 Axiomas de ordenación

Una vez fundamentada la geometría afín en los meros axiomas de incidencia y paralelismo vamos a ver el papel que juegan los axiomas de ordenación. Conviene estudiar antes algunas propiedades de los cuerpos ordenados. Trabajamos con anillos de división porque así los axiomas de ordenación nos servirán para obligar al anillo de división asociado a un espacio afín a ser isomorfo a \mathbb{R} o a un subcuerpo de \mathbb{R} , garantizando así su conmutatividad.

Definición 7.29 Una *ordenación* de anillo de división K es un subconjunto P de K (a cuyos elementos llamaremos *positivos*) que cumpla las propiedades siguientes:

1. $K = -P \cup \{0\} \cup P$ y la unión es disjunta.
2. $P + P \subset P$
3. $PP \subset P$

Un anillo de división ordenado es un anillo en el que se ha seleccionado una ordenación.

Si K es un anillo de división ordenado podemos definir $a < b$ como $b - a \in P$, y es fácil ver que esta relación resulta ser un orden total en K de modo que P es el conjunto de los elementos mayores que 0.

Es fácil probar las propiedades usuales de compatibilidad con el orden. Por ejemplo, si $a < 0$ entonces $-a > 0$, luego $a^2 = (-a)^2 > 0$, y trivialmente $a^2 > 0$ si $a > 0$. Por lo tanto los cuadrados de los elementos no nulos son positivos. En particular $1 = 1^2 > 0$, luego $-1 < 0$. De aquí que $1 + 1, 1 + 1 + 1$, etc. sean todos positivos, luego la característica de K ha de ser 0.

En geometría las rectas aparecen ordenadas de tal modo que es imposible distinguir una ordenación de su inversa. La definición siguiente se acerca más a lo que nos vamos a encontrar y enseguida veremos que, salvo un caso trivial, es equivalente a la anterior.

Definición 7.30 Un anillo de división K está *débilmente ordenado* si sobre él hay definida una relación de orden total de modo que

1. para cada $a \in K$, la aplicación $x \mapsto x + a$ conserva o invierte el orden.
2. para cada $a \in K, a \neq 0$, la aplicación $x \mapsto ax$ conserva o invierte el orden.

Cuando en este contexto digamos que los elementos a, b, c, \dots están ordenados deberá entenderse que $a < b < c, \dots$ o bien $\dots < c < b < a$.

Es fácil ver que todo anillo de división ordenado está débilmente ordenado. Vamos a ver que el recíproco es cierto salvo en un caso: si K es el cuerpo de dos elementos entonces es obvio que está débilmente ordenado, pues sólo admite dos ordenaciones mutuamente inversas y por lo tanto cualquier biyección conserva o invierte el orden. Sin embargo no puede ser ordenado.

Teorema 7.31 *Si K es un anillo de división con más de dos elementos y está débilmente ordenado, entonces está ordenado.*

DEMOSTRACIÓN: En primer lugar, la característica de K no puede ser 2. En efecto, sean $0, 1, a$ tres elementos distintos en K . Si están ordenados en la forma $0, 1, a$, entonces sumamos 1 y obtenemos la ordenación $1, 0, a+1$, mientras que si sumamos a obtenemos la ordenación $a, a+1, 0$. Estas dos ordenaciones implican que $1, 0, a+1, a$, lo que contradice la ordenación inicial.

Si la ordenación es $a, 0, 1$ entonces al sumar 1 y a obtenemos $a+1, 1, 0$ y $0, a, a+1$. Las dos primeras implican $a, 0, 1, a+1$, que contradice a la tercera.

Dado $a \in K^*$ consideramos las posibles ordenaciones de $0, a/2$ y $-a/2$. Si se da $-a/2, 0, a/2$ entonces sumando $\pm a/2$ obtenemos $-a, -a/2, 0$ y $0, a/2, a$, lo que implica $-a, -a/2, 0, a/2, a$. En particular $-a, 0, a$.

Si por el contrario $0, -a/2, a/2$ obtenemos $-a/2, -a, 0$ y $a/2, 0, a$, de donde se sigue igualmente $-a, 0, a$.

Cambiando la ordenación de K por su inversa si es necesario, K sigue siendo un cuerpo débilmente ordenado con $0 < 1$. Acabamos de probar que si P es el conjunto de los elementos mayores que 0, entonces $-P$ es el conjunto de los elementos menores que 0, luego se cumple el primer axioma de la definición de cuerpo ordenado.

Para probar el segundo axioma tomamos $0, a, b$ y sumamos a , con lo que obtenemos $a, 2a, a+b$. Sumando a en $-a, 0, a$ obtenemos también $0, a, 2a$, luego $0, a, 2a, a+b$. En particular es claro que $P + P \subset P$.

Finalmente, si $a, b \in P$, entonces tenemos la ordenación $0, 1, b$ o bien $0, b, 1$, luego también $0, a, ab$ o bien $0, ab, a$. En ambos casos $ab \in P$. ■

Para introducir la ordenación de un espacio afín hemos de definir la proyección paralela entre dos rectas: si r y s son dos rectas contenidas en un plano afín y P, Q son un punto de cada una de ellas, entonces para cada punto X en r , la recta paralela a PQ por X ha de cortar a s en un punto Y , de modo que la correspondencia $X \mapsto Y$ es claramente biyectiva. A esta aplicación la llamaremos *proyección* de r en s *paralelamente* a PQ .

Definición 7.32 Diremos que un espacio afín está *ordenado* si cada recta tiene asociado un orden de modo que las proyecciones paralelas conservan o invierten los órdenes respectivos.

Teorema 7.33 *Sea E un espacio afín sobre un anillo de división K cuyas rectas tengan al menos tres puntos. Entonces E admite una ordenación si y sólo si la admite K .*

DEMOSTRACIÓN: Supongamos que E está ordenado. Podemos biyectar K con los puntos de una recta mediante la correspondencia $\alpha \mapsto P + \alpha\vec{v}$, donde P y \vec{v} son un punto y un vector prefijados. Probaremos que este orden es un orden débil en K y por el teorema anterior él o su inverso (el orden que hace $1 > 0$) es una ordenación de K . Antes veremos que el orden no depende de la elección de P o \vec{v} .

En primer lugar cambiamos \vec{v} por otro vector $\vec{w} \neq \vec{0}$. Consideremos la homotecia dada por $f_\alpha(Q) = P + \alpha\vec{PQ}$. En particular

$$P + \alpha\vec{v} = f_\alpha(P + \vec{v}), \quad P + \alpha\vec{w} = f_\alpha(P + \vec{w}).$$

Esto implica que las rectas que pasan por $P + \alpha\vec{v}$ y $P + \alpha\vec{w}$ forman un haz de rectas paralelas o, equivalentemente, cada $P + \alpha\vec{w}$ se obtiene de $P + \alpha\vec{v}$ por proyección paralela. Por hipótesis el orden se conserva (o invierte) y en ambos casos el orden en K (que hace $1 > 0$) es el mismo.

Si reemplazamos el punto P por otro punto Q , tomamos un vector \vec{v} de dirección distinta a PQ . Entonces las rectas $P + \alpha\vec{v}$ y $Q + \alpha\vec{v}$ son paralelas, al igual que las rectas que pasan por cada par de puntos de esta forma. Por lo tanto $Q + \alpha\vec{v}$ se obtiene de $P + \alpha\vec{v}$ por proyección paralela y concluimos del mismo modo que ambas rectas inducen el mismo orden en K .

Para probar que la ordenación de K es un orden débil observamos que, fijados P y \vec{v} , la aplicación $\alpha \mapsto \alpha + \delta$ en K se corresponde con la aplicación $P + \alpha\vec{v} \mapsto P + (\alpha + \delta)\vec{v} = (P + \delta\vec{v}) + \alpha\vec{v}$, luego la ordenación de K se convierte en la inducida fijando $P + \alpha\vec{v}$ y \vec{v} que, según hemos visto, es la misma o la opuesta.

Similarmente, la aplicación $\alpha \mapsto \delta\alpha$ en K se corresponde con la aplicación $P + \alpha\vec{v} \mapsto P + \alpha(\delta\vec{v})$, luego el orden de K se transforma en el inducido por P y $\delta\vec{v}$, que es el mismo o el inverso.

Si K está (débilmente) ordenado, usamos las mismas biyecciones entre K y los puntos de cada recta para definir ordenaciones en éstas. Se comprueba de modo similar que las ordenaciones no dependen de la elección de P y \vec{v} , así como que se conservan o invierten por proyecciones paralelas. ■

Definición 7.34 Un anillo de división ordenado K tiene la *propiedad de Arquímedes* si para todo $a > b > 0$ existe un número natural n tal que $a < nb$. Un espacio afín sobre un anillo de división K tiene la *propiedad de Arquímedes* si la tiene K .

Es fácil expresar la propiedad de Arquímedes en términos geométricos: un espacio afín tiene la propiedad de Arquímedes si y sólo si fijado un punto P en una recta r y una traslación \vec{v} de dirección r , aplicando a P la traslación \vec{v} un número suficiente de veces podemos superar cualquier punto Q de r .

Teorema 7.35 Un anillo de división ordenado K tiene la propiedad de Arquímedes si y sólo si es isomorfo a un subcuerpo de \mathbb{R} con el orden natural.

DEMOSTRACIÓN: Como K está ordenado tiene característica 0, luego contiene a \mathbb{Q} . Dados $a < b$ en K , por la propiedad de Arquímedes existe un número

natural n tal que $1/(b-a) < n$, luego $1/n < b-a$. Existe un natural m tal que $a < m/n$ y es fácil ver entonces que $a < m/n < b$. Así pues, entre dos elementos cualesquiera de K hay un número racional. Es fácil ver que la aplicación que a cada $a \in K$ le asigna el supremo en \mathbb{R} del conjunto $\{r \in \mathbb{Q} \mid r < a\}$ es un monomorfismo de cuerpos.

Recíprocamente, es inmediato que cualquier subcuerpo de \mathbb{R} es arquimediano con el orden inducido desde \mathbb{R} . Más aún, la propiedad de Arquímedes en \mathbb{R}^n visto como espacio euclídeo equivale a la que introducimos en el capítulo II. ■

En particular todo espacio afín Arquimediano es conmutativo. Es fácil ver que el axioma D que usamos en el capítulo II implica que el cuerpo asociado es exactamente \mathbb{R} .

Capítulo VIII

La geometría proyectiva

Según hemos visto en el capítulo anterior, una buena parte de las propiedades geométricas del espacio que percibimos, las que no dependen de ordenación ni de distancias, son, desde un punto de vista matemático, las propiedades algebraicas de los espacios afines. De hecho las propiedades de ordenación también se pueden interpretar en términos algebraicos como una ordenación arquimediana del anillo de división asociado, lo cual trae como consecuencia que éste es de hecho un cuerpo y se verifica el teorema de Pappus–Pascal. Esta visión algebraica de la geometría nos da una perspectiva más objetiva de la misma, independiente de la forma en que tenemos de percibirla por nuestra estructura psicológica. Por ejemplo, desde un punto de vista algebraico, \mathbb{R}^3 y \mathbb{R}^4 no son, ciertamente, una misma cosa, pero tienen mucho en común. Desde un punto de vista intuitivo, en cambio, la diferencia es que \mathbb{R}^3 es “todo” y \mathbb{R}^4 es “nada”. Quien entienda la geometría en términos algebraicos se desenvolverá en \mathbb{R}^4 casi con la misma facilidad que en \mathbb{R}^3 , mientras que en términos intuitivos es imposible concebir, digamos, dos planos con un único punto en común.

Sin embargo, en este capítulo veremos que la estructura afín no es la forma más natural de describir la geometría desde un punto de vista algebraico, sino que aún está sesgada por la forma de nuestra intuición. Existe otra estructura algebraica todavía más simple, en el sentido de más simétrica y cuyas propiedades se siguen de principios formalmente más sencillos, que describe más eficientemente el espacio intuitivo, junto con una amplia familia de espacios algebraicos de interés.

Estamos hablando de la geometría proyectiva, cuyo principio fundamental es que dos rectas paralelas se cortan “en el infinito”, de modo que un espacio proyectivo consiste en un espacio afín al que hemos añadido un conjunto de puntos ideales (puntos infinitos) de modo que cada par de rectas paralelas se corten en uno de estos puntos. La geometría proyectiva en sentido moderno surgió a principios del siglo XIX, con los trabajos de Poncelet, Chasles, Cayley, entre otros, pero sus raíces se encuentran en los estudios iniciados en el renacimiento sobre la representación en perspectiva.

El problema de la perspectiva es determinar cómo debe ‘pintarse’ un espacio

En general, los puntos situados detrás del cuadro se biyectan con los puntos bajo la línea del horizonte. Los puntos situados delante del cuadro pero posteriores al pie del observador se biyectan con los puntos del cuadro situados bajo el suelo. Si sustituyéramos el suelo por un cristal que nos permitiera ver la parte del cuadro situada bajo tierra, el efecto óptico sería que el cuadro vertical llega hasta nuestros pies. Por otro lado, la línea paralela al cuadro y que pasa por nuestros pies no tiene cabida en el cuadro, pues las líneas que unen sus puntos con el ojo son paralelas al cuadro.

Aunque esto ya no tiene interés en pintura, desde un punto de vista matemático la correspondencia entre puntos de ambos planos puede extenderse de forma natural asignando a los puntos situados detrás del ojo los puntos situados sobre la línea del horizonte, siguiendo el mismo criterio. De este modo tenemos

una biyección entre ambos planos salvo por dos excepciones: en el cuadro ha aparecido la línea del horizonte y ha desaparecido la línea situada al pie del observador.

La interpretación de estos fenómenos desde el punto de vista de la geometría proyectiva es la siguiente: todos los planos (proyectivos) tienen una recta adicional, invisible en el sentido ordinario, donde concurren las rectas paralelas. Las proyecciones perspectivas biyectan los planos proyectivos y pueden “hacer finitas” las rectas infinitas, como ocurre en el ejemplo anterior. La recta del horizonte es la proyección de la recta infinita del plano horizontal, y el cuadro muestra efectivamente cómo concurren en ella las rectas paralelas. La recta situada al pie del observador ha desaparecido porque la proyección la ha enviado al infinito. En efecto, dos líneas que concurrieran en un punto de esta recta aparecerían como paralelas en el cuadro.

8.1 Espacios proyectivos

En esta sección veremos que las ideas anteriores se corresponden con una teoría matemática rigurosa. Comenzamos describiendo axiomáticamente lo que, de acuerdo con ellas, debe ser un plano proyectivo.

Definición 8.1 Un *plano proyectivo* es un conjunto E , a cuyos elementos llamaremos *puntos*, junto con una familia de subconjuntos no vacíos de E a cuyos elementos llamaremos *rectas*, de modo que se satisfagan los axiomas siguientes:

Axioma P1 *Por cada par de puntos distintos P y Q pasa una única recta, que representaremos por PQ .*

Axioma P2 *Todo par de rectas tienen un punto en común.*

Axioma P3 *Existen tres puntos distintos no colineales.*

Axioma P4 *Toda recta contiene al menos tres puntos.*

Los axiomas P1 y P3 coinciden con los correspondientes para el plano afín, el axioma P4 sirve para eliminar casos triviales que no se comportan adecuadamente, pero el axioma más notable es P2, que niega la existencia de rectas paralelas. Aunque aparentemente la geometría proyectiva contradice a la geometría afín, lo cierto es que la extiende, como veremos enseguida:

Definición 8.2 Sea E un plano afín. Para cada recta r de E , llamaremos P_r al haz de rectas paralelas a r . Sea r_∞ el conjunto de todos los haces P_r . Llamaremos *compleción proyectiva* de E al conjunto $E^\infty = E \cup r_\infty$. A los elementos de E los llamaremos *puntos finitos* de E^∞ , mientras que los de r_∞ serán los *puntos infinitos*. Llamaremos *rectas* de E^∞ a los conjuntos de la forma $\bar{r} = r \cup \{P_r\}$, donde r es una recta de E , más el conjunto r_∞ , al que llamaremos *recta infinita* de E^∞ . Por contraposición, a las rectas \bar{r} las llamaremos *rectas finitas* de E^∞ .

En definitiva, la compleción proyectiva se obtiene añadiendo un punto infinito a cada recta, de modo que dos rectas son paralelas si y sólo si tienen el mismo punto infinito.

Teorema 8.3 *La compleción de un plano afín es un plano proyectivo.*

DEMOSTRACIÓN: Es claro que por dos puntos finitos pasa una única recta, concretamente la extensión de la recta afín que pasa por ellos. Si P es un punto finito y P_r un punto infinito, entonces la recta s paralela a r por P se extiende a una recta finita \bar{s} que pasa por P y por $P_s = P_r$. Recíprocamente, una recta que pase por estos dos puntos ha de ser finita, luego de la forma \bar{t} , de modo que t pasa por P y es paralela a r , luego es la misma \bar{s} . Finalmente, si P_r y P_s son dos puntos infinitos, la única recta que pasa por ellos es la recta infinita r_∞ . Esto prueba P1

Veamos que dos rectas cualesquiera se cortan en un punto. Si una es la recta infinita es claro, pues por construcción todas las rectas contienen un punto infinito. Si las dos son finitas, digamos \bar{r} y \bar{s} , entonces o bien r y s tienen un punto en común, en cuyo caso \bar{r} y \bar{s} también, o por el contrario r y s son paralelas, en cuyo caso \bar{r} y \bar{s} tienen en común el punto infinito $P_r = P_s$. Esto prueba P2.

Es claro que tres puntos no colineales en el plano afín de partida siguen siendo no colineales en la compleción, luego se cumple P3

Toda recta afín r tiene al menos dos puntos, y su compleción \bar{r} tiene además un tercer punto infinito, luego todas las rectas finitas tienen al menos tres puntos. Para probar P4 falta ver que la recta infinita tiene al menos tres puntos, pero tomando tres puntos finitos no colineales P, Q, R obtenemos tres rectas no paralelas dos a dos, a saber, PQ, PR, QR , y cada una de ellas se extiende con un punto infinito distinto. ■

Recíprocamente, es fácil ver que si E es un plano proyectivo y r_∞ es cualquiera de sus rectas entonces $E \setminus r_\infty$ se convierte en un plano afín tomando como rectas los conjuntos que resultan de quitar a cada recta de E distinta de r_∞ su punto de intersección con ésta. Hay que destacar el hecho de que cualquier recta puede tomarse como recta infinita, pues significa que desde el punto de vista de la geometría proyectiva no hay diferencia entre puntos finitos e infinitos, sino que la distinción la introducimos nosotros al seleccionar una recta.

En la práctica, si en un plano proyectivo E marcamos una recta r_∞ como recta infinita, podemos decir que dos rectas son paralelas si se cortan en un punto de r_∞ . De este modo podemos usar el lenguaje afín sin necesidad de eliminar la recta infinita, pero hemos de tener presente que ahora la noción de paralelismo es relativa. Cualquier par de rectas pueden ser paralelas, sin más que tomar como recta infinita una que pase por su intersección.

En lugar de describir axiomáticamente los espacios proyectivos de dimensión arbitraria haremos como en el caso afín: los definiremos algebraicamente y después caracterizaremos axiomáticamente el caso tridimensional. Aunque en principio podríamos completar cualquier espacio afín por el mismo procedimiento directo que hemos empleado con los planos, lo cierto es que un enfoque

más algebraico nos dará modelos más manejables. Nuestra referencia es que deseamos añadir un punto a cada recta, de modo que dos rectas se completen con el mismo punto si y sólo si son paralelas. En general, cada variedad deberá ser completada con los puntos infinitos de las rectas que contiene.

Sea E un espacio afín sobre un cuerpo K . Para completar E comenzaremos identificándolo con un hiperplano de un espacio vectorial \hat{E} que no contenga al vector nulo. Siempre es posible hacer esto, la aplicación que a cada punto de E le asigna sus coordenadas en un sistema de referencia prefijado es una biyección afín entre E y K^n , y a su vez la aplicación que completa cada n -tupla de K^n con un 1 final es una biyección afín entre K^n y el hiperplano $x_{n+1} = 1$ de K^{n+1} . En lo sucesivo, \hat{E} será cualquier espacio vectorial que contenga a E como hiperplano que no pase por 0. Pronto veremos que no importa la elección.

Estamos considerando a \hat{E} como espacio afín de la forma usual, es decir, tomando como espacio vectorial asociado el propio \hat{E} , con lo que $\overrightarrow{PQ} = Q - P$. En particular tenemos que \vec{E} es un subespacio de \hat{E} . Más concretamente, si $P \in E$, tenemos $E = P + \vec{E}$, y necesariamente $P \notin \vec{E}$, pues en caso contrario $O = P - P \in E$. Como E es un hiperplano, ha de ser $\hat{E} = \langle P \rangle \oplus \vec{E}$. Notar también que $E \cap \vec{E} = \emptyset$, pues E es una clase módulo \vec{E} que no contiene a 0.

Teorema 8.4 *Si E y F son dos espacios afines sobre un mismo cuerpo K , toda afinidad $f : E \rightarrow F$ se extiende a una única aplicación lineal $\hat{f} : \hat{E} \rightarrow \hat{F}$. Además su restricción a \vec{E} es \vec{f} .*

DEMOSTRACIÓN: Tenemos $\hat{E} = \langle P \rangle \oplus \vec{E}$ y $\hat{F} = \langle f(P) \rangle \oplus \vec{F}$. La condición $P \mapsto f(P)$ define una aplicación lineal entre los sumandos izquierdos. La aplicación \hat{f} es la suma directa de esta aplicación y \vec{f} . Efectivamente, dado $P + \vec{v} \in E$, tenemos

$$\hat{f}(P + \vec{v}) = \hat{f}(P) + \hat{f}(\vec{v}) = f(P) + \vec{f}(\vec{v}) = f(P + \vec{v}),$$

luego \hat{f} extiende a f . Toda aplicación lineal g que cumpla las condiciones del teorema ha de cumplir $g(P) = f(P)$, luego coincide con f sobre el primer sumando. Por otra parte, si $\vec{v} \in \vec{E}$, entonces

$$g(P + \vec{v}) = g(P) + g(\vec{v}) = f(P) + \vec{f}(\vec{v}),$$

luego $g(\vec{v}) = \vec{f}(\vec{v})$. Esto prueba que $g = \hat{f}$. ■

Es claro que si f es biyectiva también lo es su extensión. Por lo tanto, si tenemos un mismo espacio afín E sumergido como hiperplano en dos espacios vectoriales (sin contener al 0) entonces la identidad en E se extiende a un isomorfismo entre ambos espacios vectoriales, con lo que ambas estructuras son algebraicamente indistinguibles. De todos modos, el ejercicio siguiente muestra una forma de construir \hat{E} de forma canónica, es decir, sin elegir sistemas de referencia ni aplicaciones arbitrarias.

Ejercicio: Sea E un espacio afín sobre un cuerpo K , sea V el conjunto de las aplicaciones afines de E en K . Probar que V es un espacio vectorial con las operaciones

definidas puntualmente. Sea $\hat{E} = V^*$ el espacio dual de V . Sean $j : E \longrightarrow \hat{E}$ e $\bar{j} : \vec{E} \longrightarrow \hat{E}$ las aplicaciones dadas por $j(P)(f) = f(P)$, $\bar{j}(\vec{v})(f) = f(\vec{v})$. Probar que j es una afinidad cuya aplicación lineal asociada es \bar{j} . Probar que $j[E]$ es un hiperplano de \hat{E} que no pasa por 0.

Gran parte de la simplicidad de la geometría proyectiva se debe a la “linealización” que estamos empezando a encontrar: tenemos que las afinidades se extienden a aplicaciones lineales, ahora traduciremos a conceptos lineales otros conceptos afines.

Teorema 8.5 Sea E un espacio afín

1. Los baricentros

$$\frac{1}{\lambda}(\lambda_1 P_1 + \cdots + \lambda_n P_n)$$

son las combinaciones lineales en \hat{E} .

2. Los puntos P_0, \dots, P_n son afínmente independientes en E si y sólo si son linealmente independientes en \hat{E} . Las coordenadas baricéntricas de un punto de E en esta base son sus coordenadas en el sentido vectorial.
3. Si $(O, \vec{v}_1, \dots, \vec{v}_n)$ es un sistema de referencia afín de E , entonces es una base de \hat{E} , y si $(\alpha_1, \dots, \alpha_n)$ son las coordenadas cartesianas de un punto de E en el sistema, entonces $(1, \alpha_1, \dots, \alpha_n)$ son sus coordenadas en la base.

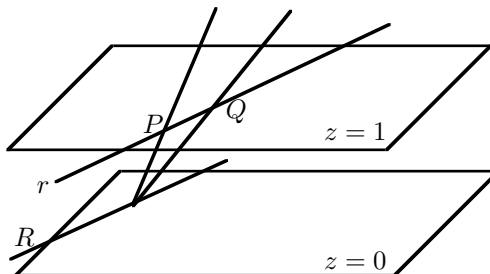
DEMOSTRACIÓN: 1) Recordemos que, fijado cualquier origen O , el baricentro es

$$O + \frac{1}{\lambda} \sum_{i=1}^n \lambda_i \overrightarrow{OP_i} = O + \frac{1}{\lambda} \sum_{i=1}^n (\lambda_i P_i - \lambda_i O) = O + \frac{1}{\lambda} \sum_{i=1}^n \lambda_i P_i - O = \frac{1}{\lambda} \sum_{i=1}^n \lambda_i P_i.$$

2) Es consecuencia inmediata de 1) por el teorema 4.13.

3) es evidente. ■

Finalmente estamos en condiciones de construir los espacios proyectivos. Comenzaremos con un ejemplo. Identifiquemos cada punto (x, y) de \mathbb{R}^2 con el punto $(x, y, 1)$ del hiperplano $z = 1$ de \mathbb{R}^3 . Podemos biyectar cada punto $(x, y, 1)$ de \mathbb{R}^3 con la recta que pasa por él y por el origen $O = (0, 0, 0)$. Recíprocamente, toda recta que pasa por 0 y que no esté contenida en el plano $z = 0$ se corresponde con un único punto de \mathbb{R}^2 . La figura muestra dos puntos P y Q y sus rectas asociadas.



A cada recta r en \mathbb{R}^2 le podemos asociar el plano π_r que la contiene y pasa por el origen O . Los puntos de r se corresponden con todas las rectas contenidas en π excepto una de ellas, la recta R donde π_r corta al plano $z = 0$. Dos rectas r y s son paralelas si y sólo si los planos π_r y π_s cortan al plano $z = 0$ en una misma recta R (pues las intersecciones han de ser paralelas con un punto en común).

Llamemos $P^2(\mathbb{R})$ al conjunto de todos los subespacios vectoriales de \mathbb{R}^3 de dimensión 1, es decir, todas las rectas que pasan por el origen. Entonces hemos visto que los puntos de \mathbb{R}^2 se pueden biyectar con los elementos de $P^2(\mathbb{R})$ que no están contenidos en el plano $z = 0$, y esta biyección transforma cada recta r de \mathbb{R}^2 en el subconjunto de $P^2(\mathbb{R})$ formado por los subespacios contenidos en un plano π_r distintos del único subespacio R contenido en el plano $z = 0$.

Pues bien, para cada recta r , llamamos \bar{r} al conjunto de todos los elementos de $P^2(\mathbb{R})$ contenidos en π_r . De este modo, \bar{r} contiene las rectas asociadas a los puntos de r más un punto infinito R , con la propiedad de que a todas las rectas paralelas a r les añadimos el mismo punto R . Así pues, $P^2(\mathbb{R})$ es la compleción proyectiva de \mathbb{R}^2 si identificamos los puntos de \mathbb{R}^2 (puntos finitos) con los elementos de $P^2(\mathbb{R})$ que no están contenidos en el plano $z = 0$, y los puntos infinitos con los que sí lo están.

Más en general, para subespacio vectorial π de \mathbb{R}^3 de dimensión 2 (cada plano que pasa por el origen), definimos la recta $\bar{\pi}$ como el conjunto de elementos de $P^2(\mathbb{R})$ contenidos en π . Si el plano π es distinto de $z = 0$, entonces $\bar{\pi}$ es la extensión de una recta afín, concretamente de la recta donde π corta a $z = 1$. Si π es el plano $z = 0$, entonces $\bar{\pi}$ está formado por todos los puntos infinitos, es decir, $\bar{\pi}$ es la recta infinita.

Definición 8.6 Sea V un espacio vectorial de dimensión $n + 1 \geq 2$ sobre un cuerpo K . Llamaremos *espacio proyectivo* asociado a V al conjunto $P(V)$ formado por todos los subespacios de V de dimensión 1. Para cada subespacio W de V observamos que $P(W) \subset P(V)$. A los subconjuntos de $P(V)$ de la forma $P(W)$ los llamaremos *variedades lineales proyectivas*.

Notar que $P(W)$ determina a W , pues

$$W = \bigcup_{P \in P(W)} P.$$

A W lo llamaremos *espacio soporte* de $P(W)$ y a $P(W)$ la llamaremos *variedad inducida* por W . Definimos $\dim P(W) = \dim W - 1$.

Si $W = 0$ entonces tenemos $P(W) = \emptyset$, pues ningún elemento de $P(V)$ (ningún subespacio de V de dimensión 1) está contenido en 0 . Por lo tanto el conjunto vacío es la (única) variedad lineal de dimensión -1 . Si $\dim W = 1$ entonces $P(W) = \{W\}$ es un punto de $P(V)$, luego las variedades lineales de dimensión 0 son los puntos.¹ Llamaremos rectas, planos e hiperplanos de $P(V)$

¹En sentido estricto son los conjuntos con un solo punto. En cuanto a la interpretación esto es irrelevante, pero técnicamente será útil llamar "puntos" no a los elementos de $P(V)$ sino a sus subconjuntos con un solo punto.

a las variedades de dimensión 1, 2 y $n - 1$. La única variedad de dimensión n se obtiene tomando $W = V$, y es $P(W) = P(V)$.

Es fácil ver que la intersección de variedades lineales es una variedad lineal, concretamente la que tiene por espacio soporte la intersección de los espacios soporte. A su vez esto implica que todo subconjunto $X \subset P(V)$ genera una mínima variedad lineal que lo contiene, y que representaremos por $\langle X \rangle$. La menor variedad que contiene a dos variedades dadas $P(W_1)$ y $P(W_2)$ la representaremos por $P(W_1) + P(W_2) = P(W_1 + W_2)$. Puesto que $P(W_1) \cap P(W_2) = P(W_1 \cap W_2)$, la conocida relación entre las dimensiones nos da el teorema siguiente:

Teorema 8.7 *Si E es un espacio proyectivo y L_1, L_2 son dos variedades lineales, entonces*

$$\dim(L_1 + L_2) = \dim L_1 + \dim L_2 - \dim(L_1 \cap L_2).$$

Por ejemplo, si L_1 y L_2 son dos rectas y están contenidas en un plano, entonces la fórmula nos da $2 = 1 + 1 - \dim(L_1 \cap L_2)$, luego $L_1 \cap L_2$ tiene dimensión 0 y es un punto (las rectas son secantes). Si no están contenidas en un plano el miembro izquierdo es ≥ 3 , luego la fórmula nos da que es exactamente 3 y que $\dim(L_1 \cap L_2) = -1$, luego las rectas son disjuntas (se cruzan). Esto prueba que todo plano proyectivo en el sentido algebraico satisface el axioma P2.

Definición 8.8 Dados $m + 1$ puntos P_0, \dots, P_m en un espacio proyectivo, diremos que son *proyectivamente independientes* si la menor variedad proyectiva que los contiene es de dimensión m .

Si $P_i = \langle v_i \rangle$, entonces la menor variedad que los contiene es la inducida por $\langle v_0, \dots, v_m \rangle$, luego la independencia de los P_i equivale a que este subespacio tenga dimensión $m + 1$ o, equivalentemente, a que los vectores v_i sean linealmente independientes.

Es claro que dos puntos distintos son siempre proyectivamente independientes, tres puntos son independientes si y sólo si no son colineales, cuatro puntos son independientes si y sólo si no son coplanares, etc. La variedad generada por $m + 1$ puntos independientes en la única variedad de dimensión m que las contiene, pues está contenida en cualquier otra y una variedad no puede estar contenida en otra distinta de la misma dimensión (pues lo mismo debería suceder con los espacios soporte).

Definición 8.9 Sea E un espacio afín. Llamaremos *compleción proyectiva* de E al espacio proyectivo $P(E) = P(\hat{E})$.

Teorema 8.10 *Sea E un espacio afín de dimensión n . Sea $i : E \rightarrow P(E)$ la aplicación dada por $i(P) = \langle P \rangle$. Entonces, para cada variedad lineal afín L de E , existe una única variedad proyectiva $P(L)$ de $P(E)$ de la misma dimensión tal que $i[L] \subset P(L)$, y además $P(L) \setminus i[L]$ es una variedad proyectiva de dimensión una unidad menor.*

DEMOSTRACIÓN: Es claro que la aplicación i es inyectiva. Dada la variedad L de dimensión m , sea $Q \in L$. Entonces $L = Q + \vec{L}$. Tenemos que $Q \notin \vec{L}$, luego $W = \langle Q, \vec{L} \rangle$ es un subespacio vectorial de \hat{E} de dimensión $m+1$ y contiene a L como hiperplano que no pasa por el 0. Tomamos $P(L) = P(W)$, que no es sino la compleción proyectiva de L , luego es una variedad proyectiva de dimensión m . Si $P \in L$, entonces $\langle P \rangle \subset W$, luego $i(P) \in P(L)$. Así pues, $i[L] \subset P(L)$.

Por otra parte, L contiene $m+1$ puntos afínmente independientes, que son linealmente independientes en \hat{E} , luego sus imágenes por i son puntos proyectivamente independientes de $P(E)$ contenidos en $i[L]$. Por consiguiente $P(L)$ es la variedad generada por $i[L]$, luego es la única variedad de dimensión m que contiene a $i[L]$.

Un punto arbitrario de $P(L)$ es de la forma $P = \langle v \rangle$ con $v \in W$. Por lo tanto $v = \alpha Q + l$, donde $l \in \vec{L}$. Si $\alpha \neq 0$ entonces $\alpha^{-1}v \in Q + \vec{L}$, luego $P = \langle \alpha^{-1}v \rangle \in i[L]$. Por el contrario, si $\alpha = 0$ entonces $P \subset \vec{L}$ y \vec{L} es disjunto con L , luego $P \notin i[L]$. Así pues, $P(L) \setminus i[L] = P(\vec{L})$, que es una variedad proyectiva de dimensión $m-1$. ■

En definitiva, al completar un espacio afín estamos añadiendo un punto infinito a cada recta, una recta infinita a cada plano (formada por todos los puntos infinitos de las rectas que contiene), un plano infinito a cada espacio tridimensional, etc. En total los puntos infinitos forman un hiperplano de la compleción.

No obstante, desde el punto de vista proyectivo no hay diferencia alguna entre el hiperplano infinito y cualquier otro hiperplano. En efecto, si $X = P(V)$ es un espacio proyectivo arbitrario, un hiperplano arbitrario es de la forma $\Pi = P(\vec{E})$, donde \vec{E} es un hiperplano vectorial de V , es decir un subespacio de dimensión una unidad menor. Si v es cualquier punto de $V \setminus \vec{E}$, entonces $E = v + \vec{E}$ es un hiperplano afín de V y es fácil ver que $X = P(E)$. Por lo tanto $X \setminus \Pi$ tiene estructura de espacio afín, cuyas variedades afines son los conjuntos $L \setminus \Pi$, donde L es cualquier variedad proyectiva de X no contenida en Π .

La elección del punto v a la hora de determinar la estructura afín de $X \setminus \Pi$ es irrelevante por el motivo siguiente. Dados X y Π , elegimos $n+1$ puntos proyectivamente independientes en $X \setminus \Pi$, digamos, O, P_1, \dots, P_n . Tomamos representantes $O = \langle v_0 \rangle$, $P_i = \langle v_i \rangle$ de modo que $v_i \in E$ (sólo hay una posibilidad). Entonces los vectores $\vec{v}_i = v_i - v_0$ están en \vec{E} y determinan un sistema de referencia afín $(v_0; \vec{v}_1, \dots, \vec{v}_n)$ en $X \setminus \Pi$. Para calcular las coordenadas de un punto $P = \langle w \rangle$ en dicho sistema tomamos α tal que $\alpha w \in E$ y calculamos las coordenadas de αw en la base $(v_0, \vec{v}_1, \dots, \vec{v}_n)$, que serán de la forma $(1, \alpha_1, \dots, \alpha_n)$. Al eliminar el 1 inicial tenemos las coordenadas que buscamos.

Si en lugar de tomar el hiperplano $v + \vec{E}$ hubiéramos tomado otro cualquiera, que será de la forma $\beta v + \vec{E}$, tendríamos que sustituir cada v_i por βv_i y cada \vec{v}_i por $\beta \vec{v}_i$. Entonces el representante adecuado de P será $\beta \alpha w$ y es claro que llegamos a las mismas coordenadas. En resumen hemos probado el teorema siguiente:

Teorema 8.11 Sea X un espacio proyectivo de dimensión n y sea $\Pi = P(\vec{E})$ un hiperplano. Entonces las coordenadas cartesianas de un punto finito respecto al sistema de referencia determinado por $n+1$ puntos independientes de $X \setminus \Pi$ no depende de la elección del hiperplano $E = v + \vec{E}$ con que se determine la estructura afín.

Notar, no obstante, que el vector concreto $\overrightarrow{PQ} \in \vec{E}$ que corresponda a un par de puntos finitos P y Q sí depende de la elección del hiperplano. Observar también que un conjunto de puntos de $X \setminus \Pi$ es afínmente independiente si y sólo si es proyectivamente independiente en X .

Como en el caso bidimensional, podemos decir que dos variedades proyectivas son paralelas respecto a un hiperplano Π si se cortan en una variedad contenida en Π , de modo que esta noción de paralelismo se corresponde con la noción usual en la geometría afín de $X \setminus \Pi$.

8.2 Homografías y coordenadas homogéneas

Introducimos ahora los conceptos proyectivos análogos a las afinidades y las coordenadas cartesianas en los espacios afines.

Definición 8.12 Sea $u : V \longrightarrow W$ un isomorfismo entre espacios vectoriales. Llamaremos *homografía* inducida por u a la aplicación $H(u) : P(V) \longrightarrow P(W)$ dada por $H(u)(\langle v \rangle) = \langle u(v) \rangle$.

Obviamente $H(u)$ es una biyección entre los espacios proyectivos que hace corresponder las variedades lineales. La restricción de $H(u)$ a una variedad lineal es la homografía inducida por la restricción de u al espacio soporte de la variedad. También es claro que la composición de homografías y la inversa de una homografía son de nuevo homografías.

Llamaremos *grupo proyectivo* de un espacio proyectivo $X = P(V)$ al grupo de todas las homografías de X en sí mismo. Lo representaremos² por $GP(X)$ o $GP(V)$.

Es claro que $H : GL(V) \longrightarrow GP(V)$ es un epimorfismo de grupos. Su núcleo está formado por todas las aplicaciones lineales que dejan invariante a cada recta de V que pasa por 0. Es claro que tales aplicaciones son exactamente las homotecias lineales de V . Así pues, $GP(V) \cong GL(V)/H(V)$.

Si $f : E \longrightarrow F$ es una afinidad biyectiva entre dos espacios afines, hemos probado que f se extiende a una única aplicación lineal $\tilde{f} : \hat{E} \longrightarrow \hat{F}$. Llamaremos $H(f) = H(\tilde{f}) : P(E) \longrightarrow P(F)$. Es fácil ver que si $i : E \longrightarrow P(E)$ y $j : F \longrightarrow P(F)$ son las inclusiones canónicas, entonces $i \circ H(f) = f \circ j$. Si identificamos a E y F con sus imágenes en sus compleciones, esto se interpreta como que $H(f)$ es una extensión de f . Consecuentemente $H(f)$ fija al hiperplano infinito de $P(E)$.

²La notación usual es $PGL(V)$, por *projective general linear group*.

Recíprocamente, si $H(u)$ es una homografía de $P(E)$ que fija al hiperplano infinito $P(\vec{E})$, entonces su restricción a E es una biyección afín. En efecto, es claro que $u[\vec{E}] \subset \vec{E}$, luego de hecho $u[\vec{E}] = \vec{E}$. Llamemos \vec{f} a la restricción de u a \vec{E} y f a la restricción de $H(u)$ a E .

Sea $E = P + \vec{E}$, de modo que $\hat{E} = \langle P \rangle \oplus \vec{E}$. Entonces $u(P) = \alpha P + \vec{v}$, donde $\vec{v} \in \vec{E}$. Sustituyendo u por $\alpha^{-1}u$ tenemos otro isomorfismo que induce la misma homografía, luego podemos suponer que $u(P) = P + \vec{v}$.

Entonces, si $Q = P + \vec{w}$ es cualquier punto de E , tenemos que

$$u(Q) = u(P) + \vec{f}(\vec{w}) = P + \vec{v} + \vec{f}(\vec{w}) \in E.$$

Si identificamos Q con su imagen en $P(V)$, es decir, con $\langle Q \rangle$, entonces la igualdad anterior es

$$f(Q) = H(u)(Q) = P + \vec{v} + \vec{f}(\overrightarrow{PQ}),$$

luego f es una afinidad. Con esto hemos probado el teorema siguiente:

Teorema 8.13 *Sea X un espacio proyectivo y Π un hiperplano en X . Sea $E = X \setminus \Pi$. Entonces la restricción induce un isomorfismo entre el grupo $GP_{\Pi}(X)$ de las homografías de X que dejan invariante a Π y el grupo $GA(E)$ de todas las biyecciones afines de E .*

(La restricción es inyectiva porque cada afinidad se extiende a una única homografía. Una prueba directa consiste en notar que si una homografía es la identidad en E , entonces fija a todas las rectas de E , luego fija al punto infinito de cada recta, pero todo punto de Π es el punto infinito de una recta de E , luego la homografía es la identidad en X .)

Sabemos que una afinidad en un espacio de dimensión n está determinada por su imagen sobre $n + 1$ puntos independientes. Esto no es cierto para las homografías. Hemos de tener en cuenta $n + 2$ puntos.

Definición 8.14 Un *sistema de referencia* en un espacio proyectivo X de dimensión n es un conjunto de $n + 2$ puntos tales que $n + 1$ cualesquiera de ellos son proyectivamente independientes.

Si $X = P(V)$ y v_1, \dots, v_{n+1} es una base de V , entonces los puntos

$$\langle v_1 \rangle, \dots, \langle v_{n+1} \rangle, \langle v_1 + \dots + v_{n+1} \rangle \quad (8.1)$$

son un sistema de referencia en X . Recíprocamente, si $\langle v_1 \rangle, \dots, \langle v_{n+1} \rangle, \langle v_{n+2} \rangle$ es un sistema de referencia en X , entonces v_1, \dots, v_{n+1} es una base de V , luego

$$v_{n+2} = \alpha_1 v_1 + \dots + \alpha_{n+1} v_{n+1},$$

para ciertos escalares, todos no nulos, pues si uno de ellos fuera nulo los vectores restantes serían linealmente dependientes, en contra de la definición de sistema de referencia. Así pues, sustituyendo v_i por $\alpha_i v_i$ tenemos el mismo conjunto de puntos, pero ahora

$$v_{n+2} = v_1 + \dots + v_{n+1},$$

luego todo sistema de referencia es de la forma (8.1).

Teorema 8.15 *Toda biyección entre dos sistemas de referencia en dos espacios proyectivos se extiende a una única homografía entre ellos.*

DEMOSTRACIÓN: Sean

$$\langle v_1 \rangle, \dots, \langle v_{n+1} \rangle, \langle v_1 + \dots + v_{n+1} \rangle \quad \text{y} \quad \langle v'_1 \rangle, \dots, \langle v'_{n+1} \rangle, \langle v'_1 + \dots + v'_{n+1} \rangle$$

dos sistemas de referencia en dos espacios proyectivos $P(V)$ y $P(V')$. Consideremos el isomorfismo $u : V \rightarrow V'$ determinado por $u(v_i) = v'_i$. Es claro que la homografía que induce transforma un sistema en otro.

Recíprocamente, si $H(u')$ es una homografía que transforma un sistema en otro, entonces $u'(v_i) = \alpha_i v'_i$, e igualando las imágenes del último componente,

$$\alpha_1 v'_1 + \dots + \alpha_{n+1} v'_{n+1} = \alpha v'_1 + \dots + \alpha v'_{n+1},$$

luego $\alpha_i = \alpha$ para todo i , es decir, $u' = \alpha u$, luego $H(u') = H(u)$. \blacksquare

Ejercicio: Encontrar una homografía de un plano proyectivo en sí mismo que deje fijos a tres puntos no colineales sin ser la identidad.

Puesto que todo espacio vectorial de dimensión $n+1$ sobre un cuerpo K es isomorfo a K^{n+1} , todo espacio proyectivo de dimensión n sobre un cuerpo K es homográfico a $P^n(K) = P(K^{n+1})$. Si fijamos un sistema de referencia P_1, \dots, P_{n+2} en un espacio proyectivo X , entonces existe una única homografía f que cumple

$$f(P_1) = \langle (1, \dots, 0) \rangle, \dots, f(P_{n+1}) = \langle (0, \dots, 1) \rangle, f(P_{n+2}) = \langle (1, \dots, 1) \rangle.$$

A la imagen de cada punto P por esta homografía la llamaremos sistema de *coordenadas homogéneas* de P respecto al sistema de referencia dado. Conviene pensar que a cada punto P le estamos asignando una $n+1$ -tupla de coordenadas que, a diferencia de las coordenadas cartesianas en un espacio afín, no están unívocamente determinadas por el punto, sino que dos $n+1$ -tuplas corresponden al mismo punto si y sólo si son proporcionales.

Es fácil ver que si $P_i = \langle v_i \rangle$ para $i = 1, \dots, n+1$ y $P_{n+2} = \langle v_1 + \dots + v_{n+1} \rangle$ entonces las coordenadas homogéneas de un punto $P = \langle v \rangle$ son simplemente las coordenadas de v en la base v_1, \dots, v_{n+1} . Si Π es un hiperplano de X , entonces su soporte está formado por los vectores de V cuyas coordenadas en la base v_1, \dots, v_{n+1} satisfacen una ecuación de la forma $\alpha_1 x_1 + \dots + \alpha_{n+1} x_{n+1} = 0$, donde no todos los coeficientes son nulos. Por consiguiente, Π está formado por los puntos de X cuyas coordenadas homogéneas satisfacen esta misma ecuación. Así pues, si fijamos un sistema de referencia, cada hiperplano se corresponde con una ecuación lineal homogénea no nula, con la única salvedad de que dos ecuaciones corresponden al mismo hiperplano si y sólo si son una un múltiplo de la otra.

Veamos la relación entre las coordenadas cartesianas de un espacio afín y las coordenadas homogéneas de su compleción proyectiva. Sea X un espacio

proyectivo de dimensión n en el que hemos seleccionado un hiperplano infinito $\Pi = P(\vec{E})$. Hemos visto que un sistema de referencia afín de $X \setminus \Pi$ está formado por $n + 1$ puntos independientes O, P_1, \dots, P_n . Un hiperplano $E = v + \vec{E}$ determina representantes $O = \langle v_0 \rangle$, $P_i = \langle v_i \rangle$ con la condición de que $v_i \in E$. Sea $\vec{v}_i = v_i - v_0$. El sistema de referencia proyectivo

$$\langle v_0 \rangle, \langle \vec{v}_1 \rangle, \dots, \langle \vec{v}_n \rangle, \langle v_0 + \vec{v}_1 + \dots + \vec{v}_n \rangle$$

es independiente de la elección del hiperplano E , pues si tomamos $\alpha v + \vec{E}$ todos los vectores se multiplican por α y el sistema queda inalterado.

Cada sistema de referencia afín en $X \setminus \Pi$ determina de este modo un sistema de referencia proyectivo en X , y es fácil ver que si las coordenadas cartesianas de un punto P son (x_1, \dots, x_n) , entonces sus coordenadas homogéneas son $(1, x_1, \dots, x_n)$. La ecuación de Π en coordenadas homogéneas es $x_0 = 0$. Recíprocamente, fijado un sistema de referencia proyectivo, podemos tomar como hiperplano infinito el de ecuación $x_0 = 0$. Si un punto P tiene coordenadas homogéneas (x_0, \dots, x_n) con $x_0 \neq 0$ entonces $(x_1/x_0, \dots, x_n/x_0)$ son las coordenadas cartesianas de P respecto al sistema de referencia afín determinado por los puntos de coordenadas homogéneas

$$(1, 0, 0, \dots, 0), (1, 1, 0, \dots, 0), \dots, (1, 0, \dots, 1).$$

Fijemos un sistema de referencia (8.1) en un espacio X , sea $H(u)$ una homografía en X y sea A la matriz de u en la base v_1, \dots, v_{n+1} . Es claro que si x es un sistema de coordenadas homogéneas de un punto P de X , entonces $y = xA$ es un sistema de coordenadas homogéneas de $H(u)(P)$. A la matriz A la llamaremos *matriz de coordenadas* de la homografía $H(u)$ en el sistema de referencia dado. La matriz A está determinada por la homografía y por el sistema de referencia salvo productos por una constante.

Consideremos un espacio afín E y sea $(O; \vec{v}_1, \dots, \vec{v}_n)$ un sistema de referencia y sea

$$\langle O \rangle, \langle \vec{v}_1 \rangle, \dots, \langle \vec{v}_n \rangle, \langle O + \vec{v}_1 + \dots + \vec{v}_n \rangle$$

su sistema de referencia asociado en $P(E)$.

Si f es una biyección afín dada por $f(P) = O + \vec{v} + \vec{f}(\overrightarrow{OP})$, entonces la matriz de la extensión de f a \hat{E} es

$$\left(\begin{array}{c|ccc} 1 & b_1 & \cdots & b_n \\ \hline 0 & & & \\ \vdots & & A & \\ 0 & & & \end{array} \right),$$

donde A es la matriz de \vec{f} en la base $\vec{v}_1, \dots, \vec{v}_n$ y (b_1, \dots, b_n) son las coordenadas de \vec{v} en dicha base.

Ejemplo Veamos la expresión coordenada de las homografías de una recta proyectiva X sobre un cuerpo K . Fijamos un punto infinito y consideramos un sistema de referencia afín en el resto de la recta. Éste determina a su vez un sistema de referencia proyectivo, de modo que la coordenada cartesiana x se corresponde con las coordenadas homogéneas $(1, x)$. El punto infinito tiene coordenadas homogéneas $(0, 1)$.

La expresión en coordenadas homogéneas de una homografía H es

$$H(x) = (1, x) \begin{pmatrix} d & b \\ c & a \end{pmatrix} = (cx + d, ax + b) = \left(1, \frac{ax + b}{cx + d}\right)$$

luego su restricción al espacio afín (en coordenadas cartesianas) es

$$H(x) = \frac{ax + b}{cx + d}, \quad \text{con } ad - bc \neq 0.$$

Si $d \neq 0$ esta fórmula no tiene sentido para $x = -c/d$, porque en la expresión homogénea se ve que la imagen de $(1, -c/d)$ es $(0, a - bc/d) = (0, 1)$, es decir, el punto infinito. A su vez, la imagen del punto infinito es $H(0, 1) = (d, b) = (1, b/d)$, o sea, el punto b/d , salvo si $d = 0$, en cuyo caso el punto infinito queda fijo y la homografía es una afinidad. ■

Terminamos la sección con una caracterización de las homografías de los espacios proyectivos reales análoga a la caracterización de las afinidades que vimos en el capítulo VI.

Teorema 8.16 (Teorema fundamental de la geometría proyectiva) *Sea f una biyección entre dos espacios proyectivos reales de la misma dimensión $n \geq 2$ tal que la imagen de tres puntos alineados cualesquiera sean tres puntos alineados. Entonces f es una homografía.*

DEMOSTRACIÓN: Sea $f : X \longrightarrow X'$. Como en el caso afín, dividimos la prueba en varios pasos:

1. Si $A = \{P_0, \dots, P_m\}$ de un conjunto de $m + 1$ puntos proyectivamente independientes en X , entonces $f[\langle A \rangle] \subset \langle f[A] \rangle$.

Por inducción sobre m . Para $m = 1$ se trata de la hipótesis del teorema. Supongámoslo cierto para conjuntos de m puntos. Si $P \in \langle A \rangle$, entonces la recta PP_m corta a $\langle P_0, \dots, P_{m-1} \rangle$, que es un hiperplano del espacio proyectivo $\langle A \rangle$. Sea Q el punto de corte. Por hipótesis de inducción

$$f(Q) \in f[\langle P_0, \dots, P_{m-1} \rangle] \subset \langle f[P_0], \dots, f[P_{m-1}] \rangle \subset \langle f[A] \rangle.$$

A su vez, $f(P)$ está en la recta $f(Q)f(P_m)$, que está en $\langle f[A] \rangle$.

2. La imagen de un conjunto proyectivamente independiente es un conjunto proyectivamente independiente.

En efecto, el conjunto A puede completarse hasta un conjunto A' de $n + 1$ puntos proyectivamente independientes. Por el apartado anterior tenemos

que $f[X] \subset \langle f[A'] \rangle$, luego teniendo en cuenta las dimensiones $\langle f[A'] \rangle = X'$, luego $f[A']$ es proyectivamente independiente y $f[A]$ también. De aquí se sigue obviamente:

3. *La inclusión en 1) es de hecho una igualdad. En particular la imagen de un hiperplano es un hiperplano.*

Si fijamos un hiperplano Π en X y $\Pi' = f[\Pi]$, entonces f induce una biyección $X \setminus \Pi \rightarrow X' \setminus \Pi'$ en las hipótesis del teorema fundamental de la geometría afín, luego es una afinidad, que según sabemos se extiende a una homografía de X en X' . La extensión es necesariamente f pues, fijado un punto $O \in X \setminus \Pi$, para todo $P \in \Pi$, tomamos un tercer punto Q en OP y así, tanto $f(P)$ como la imagen de P por la homografía están en Π' y en la imagen de la recta OQ , que es una recta, luego ambos son el único punto de intersección entre la recta y el hiperplano. ■

8.3 Perspectividades

Definición 8.17 Sea X un espacio proyectivo, sean Π y Π' dos hiperplanos distintos en X y sea O un punto exterior a ambos. Llamaremos *proyección perspectiva* de centro O entre ambos hiperplanos a la aplicación $\pi_O : \Pi \rightarrow \Pi'$ que a cada punto $P \in \Pi$ le hace corresponder el punto donde la recta OP corta a Π' . La intersección entre los dos hiperplanos se llama *eje* de la proyección.

La aplicación que transforma un plano modelo en su imagen en un cuadro es una proyección perspectiva. La geometría proyectiva es el marco idóneo para estudiar este tipo de aplicaciones, pues si quisiéramos hacerlo desde un punto de vista afín tendríamos que distinguir constantemente casos particulares a causa del paralelismo. Por ejemplo, no todo punto tendría imagen por una proyección perspectiva.

Teorema 8.18 *Toda proyección perspectiva es una homografía.*

DEMOSTRACIÓN: Consideremos una proyección $\pi_O : \Pi \rightarrow \Pi'$ en un espacio X de dimensión n . El eje de la proyección es una variedad de dimensión $n - 2$, que con el punto O genera un hiperplano Π_∞ . Sea $E = X \setminus \Pi_\infty$. Las trazas de Π y Π' en E son dos hiperplanos paralelos y las rectas que pasan por O forman un haz de rectas paralelas. Es fácil ver entonces que la restricción a E de π_O es simplemente la restricción a Π de una traslación, luego se extiende a una afinidad de E (la traslación) que a su vez se extiende a una homografía H en X que fija a Π_∞ . Por otra parte, π_O fija a los puntos de $\Pi \cap \Pi'$, luego H extiende a π_O , que es, por consiguiente, una homografía. ■

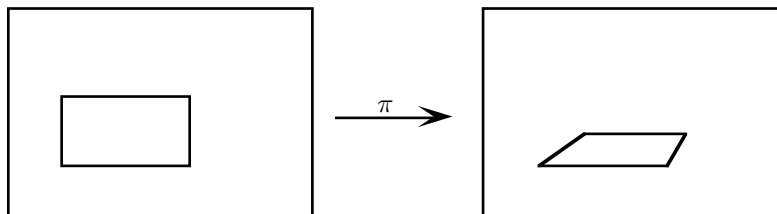
En la práctica es más cómodo sustituir una proyección perspectiva por una homografía que la extienda. Ello nos lleva al concepto de perspectividad:

Definición 8.19 Una *perspectividad* en un espacio proyectivo X es una homografía distinta de la identidad que fije a todos los puntos de un hiperplano Π , llamado *eje* de la perspectividad, y a todos los hiperplanos que pasen por un punto O , llamado *centro* de la perspectividad. Si O está en Π la perspectividad se llama *elación*, mientras que si O no está en Π se llama *homología*.

Al restringir una perspectividad de eje Π a $X \setminus \Pi$ obtenemos una afinidad que fija a cada punto infinito, luego envía cada recta a una paralela. Por consiguiente se trata de una traslación o una homotecia (según que la perspectividad sea una elación o una homología). En cualquier caso es imposible que fije a los puntos de otro hiperplano, luego Π está unívocamente determinado. Así mismo, O es la intersección de las trazas de la homotecia-traslación, luego también está unívocamente determinado. También es claro que existe una única perspectividad con un centro O y un eje Π dados que envíe un punto P exterior a Π a cualquier otro punto Q exterior a Π , distinto de P y colineal con O y P .

En la prueba del teorema 8.18 hemos visto que toda proyección perspectiva entre hiperplanos se extiende a una elación. Es fácil ver de modo similar que también puede extenderse a una homología. Recíprocamente, es fácil ver que una perspectividad H en un espacio X fija a todas las rectas que pasan por su centro (expresándolas como intersecciones de hiperplanos). De aquí que si tomamos un hiperplano Π distinto del eje y que no contenga al centro, la restricción de H es una proyección perspectiva de Π en $H[\Pi]$.

Si las perspectividades tridimensionales se interpretan como extensiones de proyecciones perspectivas, las perspectividades planas también tienen una interpretación importante. Consideremos de nuevo un plano vertical Π en el que queremos representar en perspectiva un plano horizontal Π' . Sea r la recta donde ambos se cortan. Consideremos la aplicación $\pi : \Pi \rightarrow \Pi$ que a cada punto de Π lo abate sobre Π' (mediante un giro alrededor de r) y después lo proyecta sobre Π . En otras palabras, la aplicación π pone en perspectiva cada figura de Π :



Pues bien, la aplicación π es la composición de (la extensión de) un giro con una perspectividad, luego es una homografía. Más aún, fija a todos los puntos de r , luego restringida a $\Pi \setminus r$ transforma cada recta en una paralela. Por lo tanto es una homotecia o una traslación, y en cualquier caso fija a un punto O , finito o infinito, con lo que π es una perspectividad plana.

Una perspectividad plana está determinada por su eje r , su centro O y la imagen de un punto. En la práctica (a efectos de dibujar en perspectiva)

el punto cuya imagen se fija de antemano es el punto infinito de las rectas verticales. Su imagen F se llama *punto de fuga*, y es el punto del horizonte donde concurren las rectas verticales (las rectas que en el modelo horizontal se alejan frontalmente del observador). El eje de perspectiva representa el punto de contacto entre el cuadro y el suelo (que en la práctica puede quedar fuera del cuadro, y entonces conviene evitar su uso explícito). El centro O depende de la posición del observador respecto al cuadro. Lo habitual es que se sitúe sobre la vertical de F por debajo del eje (si se sitúa fuera de esta vertical el cuadro deberá ser observado desde un ángulo y no frontalmente).

Con esta base se puede desarrollar matemáticamente toda la teoría de la perspectiva, incluyendo el problema de representar objetos tridimensionales, pero no vamos a entrar en ello. En su lugar obtendremos algunos resultados teóricos de interés sobre perspectividades y proyecciones perspectivas.

No toda homografía entre dos hiperplanos Π y Π' es una proyección perspectiva, pues si fijamos un punto O exterior a ambos y dos puntos P y Q en Π , es claro que π_O es la única perspectividad que transforma P y Q en $\pi_O(P)$ y $\pi_O(Q)$, mientras que hay infinitas homografías con esta propiedad. Por otro lado, la composición de proyecciones perspectivas no es necesariamente una proyección perspectiva. Vamos a probar que toda homografía entre dos hiperplanos (posiblemente iguales) es composición de proyecciones perspectivas. De este modo, los invariantes proyectivos son exactamente los invariantes por tales proyecciones.

La prueba se basa en el siguiente teorema geométrico y algunos resultados algebraicos que veremos tras él.

Teorema 8.20 *Sea Π un hiperplano en un espacio proyectivo X . Toda homología de Π es composición de dos proyecciones perspectivas en X .*

DEMOSTRACIÓN: Sean P y r el centro y el eje de una homología f en Π y $Q \in \Pi \setminus r$ un punto distinto de P . Entonces $f(Q)$ es un punto de $\Pi \setminus r$ distinto de Q . Sea Π' un hiperplano de X distinto de Π que pase por r . Sea O un punto exterior a ambos.

Consideremos la perspectividad $\pi_O : \Pi \rightarrow \Pi'$. Sean P' y Q' las imágenes de P y Q . Como los puntos P , Q y $f(Q)$ son colineales, los tres son coplanares con O , P' y Q' . La recta $f(Q)Q'$ corta a OP en un punto O' distinto de O . Consideremos la perspectividad $\pi'_O : \Pi' \rightarrow \Pi$ y veamos que $f = \pi_O \circ \pi'_O$.

Si llamamos g a la composición, tenemos que g fija a cada punto de r , lo que significa que g transforma cada recta de $\Pi \setminus r$ en una paralela. Por lo tanto se trata de una homotecia o una traslación, pero $g(P) = P$, luego es una homotecia de centro P . Como $g(Q) = f(Q)$, necesariamente $f = g$. ■

Ahora probaremos que toda homografía de un espacio Π se puede expresar como composición de homologías.

Fijemos un espacio vectorial V sobre un cuerpo K y sea H un hiperplano vectorial de V . Sea $G = \text{GL}(V)$ y llamemos G_H al subgrupo de G formado por los automorfismos que fijan a cada punto de H . Vamos a describir sus elementos. Sea, pues, $u \in G_H$. Fijemos $v \in V \setminus H$, de modo que $V = H \oplus \langle v \rangle$.

Sea $u(v) = h + \alpha v$, con $h \in H$ y $\alpha \in K$. Notar que $\alpha = \det u$. Basta tomar una base de H y completarla con v . La matriz de u en dicha base es triangular, y en la diagonal tiene una entrada igual a α y las restantes iguales a 1.

Si $\alpha \neq 1$ tomamos $w = h + (\alpha - 1)v$, y es fácil ver que $u(w) = \alpha w$.

Definición 8.21 Sea V un espacio vectorial y H un hiperplano vectorial de V . Una *dilatación* de V de hiperplano H , dirección $w \notin H$ y razón $\alpha \neq 0$ es un automorfismo $u \in \text{GL}(V)$ que fija a cada punto de H y que cumple $u(w) = \alpha w$.

Equivalentemente, u es una dilatación si en cierta base su matriz es

$$\begin{pmatrix} \alpha & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

Continuando nuestro argumento, si $\alpha = 1$ entonces $u(v) = h + v$. Si $u \neq 1$ ha de ser $h \neq 0$. Para cada $x \in V$ existe un $f(x) \in K$ tal que $x = h' + f(x)v$, para cierto $h' \in H$, con lo que $u(x) = h' + f(x)(h + v) = x + f(x)h$. Es fácil ver que f es una forma lineal de V de núcleo H .

Definición 8.22 Sea V un espacio vectorial y H un hiperplano vectorial de V . Una *transvección* de V de hiperplano H es un automorfismo $u \in \text{GL}(V)$ tal que existe un $h \in H$ no nulo y una forma lineal f de núcleo H de modo que $u(x) = x + f(x)h$ para todo $x \in V$.

Si completamos h hasta una base de H y ésta a una base de V , la matriz de u tiene la forma

$$\begin{pmatrix} 1 & & & \\ \alpha & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

Recíprocamente, todo automorfismo que tenga una matriz de este tipo en una base es una transvección. Hemos demostrado el teorema siguiente:

Teorema 8.23 Sea V un espacio vectorial, H un hiperplano de V y $u \in \text{GL}(V)$ que fija a cada punto de H . Entonces u es una dilatación o una transvección, según si $\det u$ es distinto o igual a 1.

Ahora probamos:

Teorema 8.24 Sea V un espacio vectorial. El grupo $\text{GL}(V)$ está generado por las dilataciones. El subgrupo de $\text{GL}(V)$ formado por los automorfismos de determinante 1 está generado por las transvecciones.

DEMOSTRACIÓN: Las matrices de la forma

$$j \begin{pmatrix} & & & k \\ & & & \\ & & & \alpha \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & 1 \end{pmatrix}$$

corresponden a transvecciones. Multiplicar una matriz por la derecha por una de estas matrices equivale a sumarle a una columna otra multiplicada por α . Mediante estas operaciones podemos convertir cualquier matriz regular en una de la forma

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \alpha \end{pmatrix} \quad (8.2)$$

En efecto, dada una matriz regular (a_{ij}) , si $a_{12} = 0$ sumamos a la segunda columna otra adecuada para que a_{12} sea no nulo. Multiplicamos la segunda columna por $(1 - a_{11})/a_{12}$ y la sumamos a la primera, con lo que obtenemos $a_{11} = 1$. Ahora es fácil convertir la primera fila en $(1, 0, \dots, 0)$. Del mismo modo podemos hacer la segunda fila igual a $(a_{21}, 1, 0, \dots, 0)$ sin deshacer la primera, y de aquí pasamos a $(0, 1, 0, \dots, 0)$. Continuando así llegamos a la forma indicada.

Por lo tanto, todo $u \in \text{GL}(V)$ se descompone en producto de transvecciones y una aplicación de matriz (8.2). Claramente $\alpha = \det u$, luego si u tiene determinante 1 entonces es producto de transvecciones, y en el caso general es producto de transvecciones y una dilatación. Basta probar que toda transvección es producto de dos dilataciones. Hemos de suponer que $K \neq \{0, 1\}$, o de lo contrario no hay dilataciones.

Si u es una transvección, tomamos una dilatación v con el mismo hiperplano que u y de determinante distinto de 1. Entonces $w = uv^{-1}$ fija a cada punto del hiperplano de u , pero su determinante es distinto de 1, luego es una dilatación, y en consecuencia $u = wv$ es producto de dilataciones. ■

Si u es una dilatación en un espacio vectorial V de hiperplano H y tal que $u(w) = \alpha w$, entonces la homografía inducida por u fija a cada punto del hiperplano $\Pi = P(H)$ y al punto $O = \langle w \rangle$, luego se trata de una homología. Así pues:

Teorema 8.25 *Toda homografía de un espacio proyectivo es producto de homologías. En particular de perspectivas.*

Uniendo esto a los resultados anteriores:

Teorema 8.26 *Toda homografía entre dos hiperplanos Π y Π' de un espacio proyectivo X es producto de proyecciones perspectivas.*

DEMOSTRACIÓN: Lo tenemos probado si $\Pi = \Pi'$. En otro caso, dada H consideramos una proyección perspectiva cualquiera $\pi : \Pi \longrightarrow \Pi'$ y tenemos que $H \circ \pi^{-1}$ es producto de proyecciones perspectivas, luego H también. ■

8.4 Caracterización axiomática

Veamos ahora la caracterización axiomática de la geometría proyectiva tridimensional, análoga a la que en el capítulo anterior obtuvimos para la geometría afín.

Definición 8.27 Un *espacio proyectivo* (tridimensional) es un conjunto E , a cuyos elementos llamaremos *puntos*, junto con dos familias de subconjuntos no vacíos de E a cuyos elementos llamaremos *rectas* y *planos*, de modo que se satisfagan los axiomas siguientes:

Axioma P1 *Por cada par de puntos distintos P y Q pasa una única recta, que representaremos por PQ .*

Axioma P2 *Todo par de rectas contenidas en un mismo plano tienen un punto en común.*

Axioma P3 *Por cada tres puntos no colineales P, Q, R pasa un único plano, que representaremos por PQR .*

Axioma P4 *Si una recta r tiene dos puntos en común con un plano π , entonces r está contenida en π .*

Axioma P5 *Existen cuatro puntos distintos no coplanares.*

Axioma P6 *Si dos planos tienen un punto en común, entonces tienen dos puntos en común.*

Axioma P7 *Toda recta contiene al menos tres puntos.*

Como en el caso bidimensional, los axiomas son los de la geometría afín cambiando el axioma de existencia de paralelas por su negación. Añadimos también la condición de que toda recta tenga tres puntos, que es necesaria porque toda recta ha de resultar de adjuntar un punto infinito a una recta afín, que tendrá al menos dos puntos. Veamos algunas consecuencias sencillas de los axiomas.

Teorema 8.28 *Dos planos distintos cualesquiera tienen una recta en común.*

DEMOSTRACIÓN: Sean π_1 y π_2 dos planos distintos. Basta probar que tienen un punto en común, pues entonces por P6 y P4 tendrán una recta en común, y por P3 no pueden tener más puntos en común.

Por definición son conjuntos no vacíos. Tomemos $P \in \pi_1$ y $Q \in \pi_2$. Si $P = Q$ ya tenemos lo que queríamos. Si $P \neq Q$ tomamos un punto R no colineal con P y Q (existe por P5) entonces el plano PQR corta a π_1 en P , luego en una recta r , y corta a π_2 en Q , luego en una recta s . Las dos rectas están contenidas en

un mismo plano, luego por P2 tienen un punto en común T , que está en π_1 y en π_2 . ■

Teorema 8.29 *Si r es una recta y π es un plano, entonces r está contenida en π o bien r corta a π en un único punto.*

DEMOSTRACIÓN: Por P7 tenemos que r contiene dos puntos distintos, digamos $r = PQ$. Sea R un punto de π . Si $PQR = \pi$ entonces $r \subset \pi$. En caso contrario PQR corta a π en una recta s (pues tienen un punto en común R). Las rectas r y s están contenidas en el plano PQR , luego tienen un punto en común T . Como son distintas, T es su único punto en común. ■

Teorema 8.30 *Todo plano contiene tres puntos no colineales.*

DEMOSTRACIÓN: Dado un plano π , por P5 existe un punto P exterior a π y por este mismo axioma y P3 existe un plano π' que pasa por P . La intersección de ambos planos ha de ser una recta r . Basta probar que π contiene un punto exterior a r . Por P5 existen dos puntos no contenidos en r , digamos A y B , luego la recta AB corta a π en un punto exterior a r . ■

Este teorema prueba que los planos contenidos en un espacio proyectivo son planos proyectivos en el sentido axiomático que hemos visto al comienzo del capítulo. Finalmente probamos que todo espacio proyectivo es la completión de un espacio afín.

Teorema 8.31 *Sea E un espacio proyectivo y Π un plano de E . Tomemos $E^* = E \setminus \Pi$. Entonces E^* es un espacio afín si tomamos como rectas y planos de E^* las intersecciones con E^* de las rectas y planos de E (cuando éstas son no vacías).*

DEMOSTRACIÓN: Cada recta r no contenida en Π corta a este plano en un único punto P_r . Llamaremos $\hat{r} = r \setminus \{P_r\}$. Cada plano $\pi \neq \Pi$ corta a Π en una única recta R_π . Llamaremos $\hat{\pi} = \pi \setminus R_\pi$.

Dados dos puntos P y Q en E^* , es claro que \widehat{PQ} es la única recta que pasa por ellos. Dada una recta \hat{r} y un punto P exterior a ella, r y P están contenidos en un único plano π , y $\hat{\pi}$ es el único plano que contiene a \hat{r} y a P . Toda recta de E^* contenida en $\hat{\pi}$ es de la forma \hat{s} , donde s está contenida en π , y sólo hay una paralela a \hat{r} , a saber, la que pasa por P y P_r .

Los axiomas A3 y A4 son evidentes. Sean P', Q', R' tres puntos no colineales en Π . Sea S un punto exterior a Π . Entonces las rectas SP', SQ', SR' contienen un tercer punto cada una, digamos P, Q, R . Es claro que los cuatro puntos P, Q, R, S son no coplanares y están en E^* .

Si dos planos $\hat{\pi}$ y $\hat{\pi}'$ tienen un punto en común, entonces π y π' tienen una recta en común, que contiene tres puntos, de los cuales al menos dos están en $\hat{\pi} \cap \hat{\pi}'$. ■

Los resultados del capítulo anterior nos permiten asignar a cada punto de E^* una terna de coordenadas (x, y, z) , fijado un sistema de referencia. Vamos a

probar que E es equivalente a $P^3(K)$, en el sentido de que existirá una biyección entre ambos que conservará las rectas y los planos.

Definimos las *coordenadas homogéneas* de un punto P de E^* respecto al sistema de referencia escogido a cualquier cuádrupla $(w, x, y, z) \in K^4$ tal que $w \neq 0$ y la terna $(xw^{-1}, yw^{-1}, zw^{-1})$ sea la terna de coordenadas cartesianas de P en el sistema de referencia dado.

Claramente, dos cuádruplas (w, x, y, z) y (w', x', y', z') con $w \neq 0 \neq w'$ son coordenadas homogéneas de un mismo punto P si y sólo si existe un $t \in K$ no nulo tal que $(w', x', y', z') = (w, x, y, z)t$. En definitiva, la aplicación que a cada punto de E^* le hace corresponder su espacio de coordenadas homogéneas es una aplicación inyectiva $i : E^* \rightarrow P^3(K)$. Con más precisión, si Π' es el plano de $P^3(K)$ inducido por el hiperplano $w = 0$, entonces la imagen de E^* es $E' = P^3(K) \setminus \Pi'$. Ahora bien, se comprueba sin dificultad que $P^3(K)$ verifica los axiomas de espacio proyectivo, luego hemos probado que E' es un espacio afín cuyas variedades son las intersecciones con E' de las variedades de E .

Un plano arbitrario π de E^* está formado por los puntos cuyas coordenadas cartesianas satisfacen una ecuación de la forma $ax + by + cz + d = 0$. Equivalentemente, un punto P está en π si y sólo si sus coordenadas homogéneas satisfacen la ecuación $dw + ax + by + cz = 0$. Toda ecuación de este tipo (con la terna (a, b, c) no nula) es la ecuación de un plano en coordenadas homogéneas. Por otra parte, esta ecuación determina un hiperplano W en K^4 , de modo que los puntos de π se corresponden con los puntos de $P(W)$ (que no están en Π').

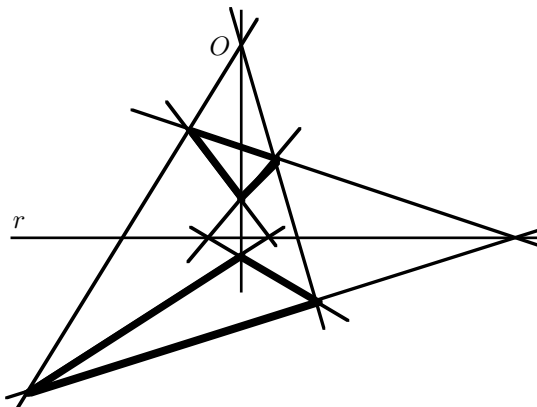
Por consiguiente, i biyecta los planos de E^* con los de E' . Puesto que toda recta afín se expresa como intersección de dos planos disjuntos, tenemos que i también biyecta las rectas con las rectas. Más aún, es obvio que i hace corresponder rectas paralelas con rectas paralelas.

Dado un punto P de Π , podemos tomar una recta que lo contenga y que no esté contenida en Π . La imagen de la restricción a E^* de esta recta será una recta de $P^3(K)$ menos un punto $Q \in \Pi'$. Podemos definir $i(P) = Q$. Esto es consistente, pues dos rectas distintas que pasen por P son paralelas en E^* , luego se transforman en rectas paralelas en E' que tienen en común el mismo punto de Π' . Es claro que así tenemos una biyección $i : E \rightarrow P^3(K)$ que transforma rectas en rectas. Es fácil ver que también transforma planos en planos, con lo que los espacios E y $P^3(K)$ son geoméricamente indistinguibles.

La introducción de coordenadas se basa en el teorema de Desargues, que probamos en el capítulo anterior a partir de los axiomas de espacio afín. Es muy ilustrativo dar una prueba directa a partir de los axiomas de espacio proyectivo, pues de hecho podemos probar un resultado formalmente más simple que el teorema de Desargues y mucho más general que éste.

Definición 8.32 Un *triángulo* en un espacio proyectivo E es un conjunto de tres puntos no colineales, a los que llamaremos *vértices*. Los *lados* del triángulo serán las rectas que pasan por cada par de vértices. Diremos que dos triángulos ABC y $A'B'C'$, con sus vértices en este orden, tienen a un punto O como

centro de perspectiva si los vértices A y A' , B y B' , C y C' están contenidos en tres rectas concurrentes en O . Diremos que tienen a una recta r como *eje de perspectiva* si cada par de lados AB y $A'B'$, AC y $A'C'$, BC y $B'C'$ se cortan en un punto de r .



Así, si dos triángulos tienen un centro de perspectiva finito, sus vértices están sobre rectas concurrentes, mientras que si tienen un centro de perspectiva infinito, sus vértices están sobre rectas paralelas. Si tienen un eje de perspectiva infinito entonces sus lados son paralelos dos a dos.

Teorema 8.33 (Teorema de Desargues) *En un espacio proyectivo, dos triángulos tienen un centro de perspectiva si y sólo si tienen un eje de perspectiva.*

DEMOSTRACIÓN: Sean dos triángulos ABC y $A'B'C'$. Si dos vértices coinciden, por ejemplo, $A = A'$, el resultado es obvio. Por ejemplo, las rectas BB' y CC' (o dos rectas que pasen por estos puntos, si coinciden) se cortan en un punto O , al igual que la recta OA , luego O es un centro de perspectiva para los triángulos. Del mismo modo se prueba que tienen un eje de perspectiva. Supongamos, pues, $A \neq A'$, $B \neq B'$, $C \neq C'$. Un razonamiento similar nos permite suponer que los lados son diferentes, es decir, $AB \neq A'B'$, $AC \neq A'C'$, $BC \neq B'C'$.

Sea π el plano que pasa por A, B, C y π' el plano que pasa por A', B', C' . Supongamos primeramente que $\pi \neq \pi'$. Entonces se cortan en una recta s . Supongamos que los triángulos tienen un eje de perspectiva r .

Sean π_1, π_2 y π_3 los planos que contienen respectivamente a los lados AB y $A'B'$, AC y $A'C'$, BC y $B'C'$. Como $\pi \neq \pi'$, el plano π_1 corta a π exactamente en la recta AB , mientras que los otros dos lo hacen en AC y BC . En particular los tres planos son distintos. π_1 corta a π_2 en una recta, que ha de ser AA' . Similarmente $\pi_1 \cap \pi_3 = BB'$ y $\pi_2 \cap \pi_3 = CC'$.

La recta AA' no puede estar contenida en π_3 , pues entonces cortaría a π en un punto de BC , luego A sería colineal con BC . Así pues, $AA' \cap \pi_3$ es un único punto O . Claramente $O = \pi_1 \cap \pi_2 \cap \pi_3$, luego las tres rectas AA' , BB' y CC' han de concurrir en O , que es, por lo tanto, un centro de perspectiva de los triángulos.

Supongamos ahora que tienen un centro de perspectiva O , Necesariamente O ha de ser distinto de los seis vértices. El plano OAB corta a π' en una recta, que necesariamente ha de ser $A'B'$. Similarmente, el plano OAC corta a π' en $A'C'$ y el plano OBC corta a π' en $B'C'$. Así pues, AB y $A'B'$ están en un mismo plano, luego se cortan en un punto, que necesariamente está en s . Lo mismo sucede con los otros pares de lados, luego s es un eje de perspectiva.

Ahora supongamos que los dos triángulos están en un mismo plano π y supongamos que tienen un eje de perspectiva s . Sea π' otro plano que corte a π en s . Sea O un punto fuera de π y π' (una recta que pase por un punto de π y otro de π' tiene un tercer punto fuera de ambos planos). Sean A'' , B'' y C'' los puntos donde las rectas OA' , OB' y OC' cortan a π' . Es fácil ver que estos tres puntos no son colineales. Los triángulos $A''B''C''$ y $A'B'C'$ tienen a O como centro de perspectiva. Por la parte ya probada tienen un eje de perspectiva, y concretamente es la recta s . Entonces $A''B''$ y $A'B'$ se cortan en un punto de s , que ha de ser también el punto donde se cortan AB y $A'B'$. Por lo tanto $A''B''$ y AB se cortan en un punto de s y lo mismo vale para los demás lados. Por consiguiente s es un eje de perspectiva para ABC y $A''B''C''$. Por la parte ya probada ambos triángulos tienen un centro de perspectiva O' . No puede ser $O' = O$, pues entonces $A = A'$, $B = B'$, $C = C'$.

Sea O'' el punto donde OO' corta a π . Veamos que O'' es un centro de perspectiva para los triángulos originales. Tenemos que los puntos O' , A y A'' son colineales. El plano que contiene a estos tres y a O corta a π en una recta que contiene a O'' , a A y a A' , luego los tres están alineados. Similarmente sucede con los otros vértices.

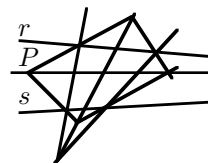
Supongamos que los triángulos tienen un centro de perspectiva O . Sea O' un punto exterior a π . Sea \bar{A} un tercer punto en la recta $O'A$. Sea \bar{A}' la intersección de $O\bar{A}$ con $O'\bar{A}'$ (ambas rectas están en el plano OAA'). El punto O es el centro de perspectiva de los triángulos $\bar{A}BC$ y $\bar{A}'B'C'$, y éstos no están sobre el mismo plano, luego por la parte ya probada ambos triángulos tienen un eje de perspectiva r , que concretamente es la intersección de los planos de los dos triángulos. O' no está en r , pues la recta $O\bar{A}$ corta a $\bar{A}BC$ en \bar{A} (si estuviera contenida entonces A estaría en BC). Sea r' la intersección con π del plano que contiene a O y a r . Veamos que r' es un eje de perspectiva para los triángulos originales. Las rectas $\bar{A}B$ y $\bar{A}'B'$ se cortan en un punto P de r . Sea P' el punto donde $O'P$ corta a π . Se trata de un punto de r' . El plano $O'PB$ contiene también a P' , a \bar{A} y a A , luego su intersección con π es la recta AB , que también contiene a P' . Del mismo modo se prueba que $A'B'$ pasa por P' , luego los lados AB y $A'B'$ se cortan en r' . Lo mismo vale para AC y $A'C'$. Respecto a BC y $B'C'$, se cortan en un punto de r que ha de estar también en π , luego en r' ■

De este resultado general se deducen inmediatamente las versiones afines que vimos en los capítulos anteriores. Por ejemplo, si las líneas que unen los vértices de dos triángulos son concurrentes o paralelas, eso significa que los triángulos tienen un centro de perspectiva (finito o infinito), luego también tienen un eje de perspectiva. Si dos pares de lados son paralelos, su eje de perspectiva pasa

por dos puntos infinitos, luego es una recta infinita, luego el tercer par de lados se corta también en el infinito, luego son paralelos. Recíprocamente, si dos triángulos tienen los lados paralelos dos a dos, entonces se cortan en tres puntos infinitos, contenidos en la intersección de los planos de ambos triángulos, es decir, en una recta infinita que es, por lo tanto, un eje de perspectiva. Por consiguiente los triángulos tienen un centro de perspectiva, lo que significa que las rectas que unen sus vértices son paralelas o concurrentes.

Como en el caso afín, un plano proyectivo no cumple necesariamente el teorema de Desargues, lo que nos lleva a distinguir entre planos proyectivos *arguesianos* y no *arguesianos*. Es claro que un plano proyectivo es arguesiano si y sólo si se puede sumergir como variedad en un espacio proyectivo. También es claro que un plano afín es arguesiano si y sólo si lo es su compleción proyectiva (construida directamente, como hemos visto al comienzo del capítulo). Puesto que hemos visto un ejemplo de plano afín no arguesiano, también tenemos probada la existencia de planos proyectivos no arguesianos.

Ejercicio: En un papel tenemos dos rectas no paralelas, pero cuya intersección queda fuera de los límites de la hoja. Dado un punto P , trazar la recta que pasa por P y por la intersección de las rectas sin usar puntos inaccesibles.



Ejercicio: Probar la versión proyectiva del teorema de Pappus–Pascal (para todo plano proyectivo sobre un cuerpo): Sean r y r' dos rectas distintas, sean A, B, C tres puntos distintos en r y A', B', C' tres puntos distintos en r' . Entonces los puntos $CA' \cap AC'$, $AB' \cap BA'$ y $BC' \cap CB'$ están alineados.

8.5 Dualidad

Nos ocupamos ahora de una de las propiedades que contribuyen a que el enfoque proyectivo de un hecho geométrico ofrezca a menudo una visión más profunda del mismo. Comenzaremos ilustrándola con un ejemplo sencillo. Consideremos por ejemplo el axioma P1 de los planos proyectivos:

Dados dos puntos distintos, hay una única recta que los contiene.

Ahora transformamos la afirmación anterior con la siguiente manipulación formal: intercambiamos las palabras ‘punto’ y ‘recta’ a la vez que invertimos las inclusiones. El resultado es:

Dadas dos rectas distintas, hay un único punto contenido en ellas.

Vemos que el resultado es otra propiedad de los planos proyectivos, concretamente el axioma P2 más la unicidad que se sigue de P1. Llamaremos *afirmación dual* de una afirmación dada sobre un plano proyectivo a la afirmación que resulta de intercambiar las palabras ‘punto’ y ‘recta’ e invertir las inclusiones. El *principio de dualidad* afirma que la afirmación dual de cualquier teorema sobre

planos proyectivos es también un teorema. La razón es sencilla: el dual de cada axioma es un teorema (ya lo hemos comprobado para P1 y para los restantes es igual de simple), luego la demostración de la afirmación dual de un teorema se obtiene sin más que reemplazar cada afirmación intermedia de la prueba por su afirmación dual correspondiente. Este principio sigue siendo válido si incorporamos como axioma el teorema de Desargues. En efecto, al dualizar un triángulo obtenemos un triángulo (un triángulo son tres puntos no colineales y las tres rectas que pasan por cada par de ellos, luego su dual son tres rectas no concurrentes y los tres puntos donde se cortan, o sea, otro triángulo) y el dual de dos triángulos con un centro de perspectiva son dos triángulos con un eje de perspectiva. Esto implica que el dual del teorema de Desargues es él mismo. Más concretamente, las dos implicaciones recíprocas son duales entre sí.

Los espacios tridimensionales también satisfacen un principio de dualidad, pero ahora las afirmaciones duales se obtienen intercambiando las palabras ‘punto’ y ‘plano’ e invirtiendo las inclusiones (las rectas son autoduales). Por ejemplo, dos afirmaciones duales son

Dados tres puntos no colineales existe un único plano que los contiene

Tres planos que no contengan una misma recta se cortan en un único punto.

Vamos a interpretar algebraicamente este principio de dualidad, lo que nos permitirá generalizarlo a espacios de dimensión arbitraria. Nos basamos en las propiedades de los espacios duales.

Definición 8.34 Sea V un espacio vectorial de dimensión $n+1$ sobre un cuerpo K . Sea V^* su espacio vectorial dual. Si W es un subespacio de V definimos

$$W^\circ = \{f \in V^* \mid f(v) = 0 \text{ para todo } v \in W\},$$

que es claramente un subespacio de V^* .

Sea w_1, \dots, w_m una base de W y completémosla hasta una base w_1, \dots, w_{n+1} de V . Sea w_1^*, \dots, w_{n+1}^* la base dual. Es fácil ver entonces que

$$W^\circ = \langle w_{m+1}^*, \dots, w_{n+1}^* \rangle.$$

Por consiguiente $\dim W + \dim W^\circ = n + 1$.

Similarmente, si W es un subespacio de V^* definimos

$$W^\circ = \{f \in V \mid f(v) = 0 \text{ para todo } f \in W\},$$

que es un subespacio de V . Similarmente se prueba la relación entre las dimensiones. Ahora es claro que si W es un subespacio de V o de V^* se cumple $W^{\circ\circ} = W$ (una inclusión es obvia y las dimensiones dan la igualdad). De aquí se sigue que las correspondencias $W \mapsto W^\circ$ entre los subespacios de V y V^* son

biyectivas y mutuamente inversas. Otro hecho obvio es que $W_1 \subset W_2$ si y sólo si $W_2^\circ \subset W_1^\circ$.

Consideremos ahora $X = P(V)$ y sea $X^* = P(V^*)$. Sea $u : V \rightarrow V^*$ un isomorfismo. Llamemos $\mathcal{P}(X)$ al conjunto de todas las variedades proyectivas de X . Definimos $p : \mathcal{P}(X) \rightarrow \mathcal{P}(X)$ como la aplicación dada por

$$p(P(W)) = P(u[W]^\circ).$$

Es claro que p es biyectiva, así como que invierte el orden: $L_1 \subset L_2$ si y sólo si $p(L_2) \subset p(L_1)$. Observar que p no se altera si sustituimos u por un múltiplo suyo, luego p depende sólo de la homografía H inducida por u .

Definición 8.35 Una *correlación* en un espacio proyectivo X es una biyección $p : \mathcal{P}(X) \rightarrow \mathcal{P}(X)$ entre las variedades lineales de X que invierte el orden.

Hemos probado que cada homografía $H : X \rightarrow X^*$ induce una correlación entre las variedades lineales de X . A las correlaciones inducidas por homografías las llamaremos *correlaciones proyectivas*. Veremos enseguida que la existencia de correlaciones implica el principio de dualidad. Antes probamos las propiedades básicas de las correlaciones:

Teorema 8.36 Sea X un espacio proyectivo de dimensión n y consideremos en él una correlación $p : \mathcal{P}(X) \rightarrow \mathcal{P}(X)$. Entonces

1. $L_1 \subset L_2$ si y sólo si $p(L_2) \subset p(L_1)$.
2. $\dim L + \dim p(L) = n$.
3. $p(L_1 + L_2) = p(L_1) \cap p(L_2)$, $p(L_1 \cap L_2) = p(L_1) + p(L_2)$.

DEMOSTRACIÓN: 1) es parte de la definición de correlación.

2) Dada una variedad L podemos formar una sucesión de variedades

$$L_{-1} \subset L_0 \subset L_1 \subset \cdots \subset L_n = X,$$

donde $\dim L_i = i$ para todo i y $L = L_i$ para algún i . Al aplicar p obtenemos una sucesión decreciente de variedades, luego necesariamente $\dim p(L_i) = n - i$.

3) es consecuencia inmediata de que p es biyectiva e invierte el orden, pues necesariamente ha de enviar la menor variedad que contiene a dos dadas en la mayor variedad contenida en dos dadas, y viceversa. ■

Ahora es claro que un espacio proyectivo X de dimensión n satisface el *principio de dualidad* siguiente:

Dada una afirmación sobre X , la afirmación que resulta de intercambiar las variedades de dimensión m por las variedades de dimensión $n - m$ e invertir las inclusiones es equivalente a la dada.

Por ejemplo, partamos de que toda recta contiene tres puntos y vamos a probar que toda variedad de dimensión $n - 2$ está contenida en tres hiperplanos. Fijamos una correlación p . Dada una variedad L de dimensión $n - 2$, entonces $p(L)$ es una recta, luego contiene tres puntos, que podemos expresar en la forma $p(L_1)$, $p(L_2)$, $p(L_3)$, donde L_1 , L_2 y L_3 son tres hiperplanos distintos. Por consiguiente $L \subset L_1 \cap L_2 \cap L_3$. ■

Hay una forma equivalente de expresar las correlaciones proyectivas que permite introducir una propiedad de simetría. Un isomorfismo $u : V \rightarrow V^*$ induce una forma bilineal regular $F : V \times V \rightarrow K$ mediante $F(v, w) = u(v)(w)$. Recíprocamente, cada forma bilineal regular F induce un isomorfismo u por la relación anterior. De este modo tenemos una biyección entre isomorfismos y formas regulares, y cada correlación proyectiva está inducida por una forma regular.

Con más detalle, si F es la forma bilineal inducida por u y W es un subespacio de V , entonces

$$u[W]^\circ = \{v \in V \mid F(w, v) = 0 \text{ para todo } w \in W\}.$$

Representaremos este conjunto por W^F . De este modo, la correlación inducida por una forma bilineal F viene dada por $p(P(W)) = P(W^F)$. Hemos probado que

$$\dim W + \dim W^F = n + 1$$

Diremos que F es *ortosimétrica* si $F(v, w) = 0$ equivale a $F(w, v) = 0$, para todo par de vectores v y w . Una forma regular en un espacio V es ortosimétrica si y sólo si $W^{FF} = W$ para todo subespacio W de V , pues si F es ortosimétrica entonces $W \subset W^{FF}$, y como las dimensiones son iguales se da la igualdad, recíprocamente, si $F(v, w) = 0$ entonces $w \in \langle v \rangle^F$ y $v \in \langle w \rangle^{FF} \subset \langle w \rangle^F$, luego $F(w, v) = 0$. A su vez esto equivale a que la correlación inducida cumpla $p(p(L)) = L$, para toda variedad L , es decir, $p \circ p$ es la identidad.

Definición 8.37 Una correlación p es *simétrica* si $p \circ p$ es la identidad.

Tenemos que una correlación proyectiva es simétrica si y sólo si la forma bilineal que la induce es ortosimétrica. En vista de esto es interesante caracterizar las formas bilineales regulares ortosimétricas:

Teorema 8.38 Sea F una forma bilineal ortosimétrica regular en un espacio vectorial V . Entonces F está en uno de los dos casos siguientes:

1. F es simétrica, es decir, $F(v, w) = F(w, v)$ para todo $v, w \in V$.
2. F es antisimétrica, es decir, $F(v, w) = -F(w, v)$ para todo $v, w \in V$. En tal caso $\dim V$ ha de ser par.

DEMOSTRACIÓN: Para cada $y \in V$ llamemos $u(y)$ y $v(y)$ a las formas de V^* dadas por $u(y)(x) = F(x, y)$ y $v(y)(x) = F(y, x)$. El hecho de que F sea regular implica que u y v son dos isomorfismos entre V y V^* . La hipótesis de

ortosimetría implica que $u(y)$ y $v(y)$ tienen el mismo núcleo, luego existe un escalar t_y tal que $v(y) = t_y u(y)$.

Si y_1 e y_2 son vectores linealmente independientes, entonces

$$v(y_1 + y_2) = t_{y_1+y_2} u(y_1 + y_2) = t_{y_1} u(y_1) + t_{y_2} u(y_2),$$

de donde se sigue que $t_{y_1} = t_{y_1+y_2} = t_{y_2}$. Es fácil ver entonces que el escalar t_y no depende de y (si V tiene dimensión 1 se llega inmediatamente a la misma conclusión). Por consiguiente $v(y) = t u(y)$, para todo y .

Distinguimos dos casos. Si existe un vector $a \in V$ tal que $F(a, a) \neq 0$, entonces $u(a) = v(a) \neq 0$, de donde $t = 1$, luego

$$F(x, y) = u(y)(x) = v(y)(x) = F(y, x),$$

es decir, F es simétrica. Si $F(a, a) = 0$ para todo vector a , entonces

$$0 = F(v + w, v + w) = F(v, w) + F(w, v),$$

luego $F(v, w) = -F(w, v)$ para todo para de vectores, luego F es antisimétrica. Falta ver que en este caso la dimensión del espacio es par.

Sea e_1, \dots, e_n una base de V y e_1^*, \dots, e_n^* su base dual. La matriz A de u en estas bases tiene en el lugar (i, j) a $u(e_i)(e_j) = F(e_j, e_i)$. La antisimetría nos da entonces que $A^t = -A$. Tomando determinantes $|A| = (-1)^n |A|$, luego n ha de ser par, pues $|A| \neq 0$. ■

Esta clasificación es satisfactoria porque ahora basta conocer la matriz de F en una base para saber si es o no ortosimétrica (según si la matriz es simétrica o antisimétrica).

En particular resulta que las correlaciones proyectivas simétricas de los planos proyectivos son las inducidas por formas bilineales simétricas (no pueden ser antisimétricas porque están definidas sobre un espacio tridimensional).

Dado un espacio proyectivo X , llamaremos $\mathcal{H}(X)$ al conjunto de todos los hiperplanos de X . Cada correlación se restringe a una biyección $p : X \rightarrow \mathcal{H}(X)$. A tales biyecciones se las llama *dualidades*, y si la correlación es simétrica se llaman *polaridades*. Por el teorema 8.36, cada dualidad determina la correlación que la define. Una polaridad p empareja cada punto P de X con un hiperplano Π . Diremos que P es el *polo* de Π y que Π es el hiperplano *polar* de P .

Si U es la $n+1$ -tupla de coordenadas homogéneas de un punto Q respecto a un sistema de referencia y A es la matriz de una dualidad proyectiva p en dicho sistema (es decir, la matriz de una forma bilineal que la induzca), entonces $p(Q)$ es el hiperplano de ecuación $UAX^t = 0$.

El ejemplo más simple lo tenemos cuando A es la matriz identidad. Fijado un sistema coordenada, el hiperplano polar del punto (a_1, \dots, a_{n+1}) es el determinado por la ecuación $a_1 x_1 + \dots + a_{n+1} x_{n+1} = 0$, y viceversa.

Notemos que si H es una homografía en X , entonces H se extiende a una biyección $H : \mathcal{P}(X) \rightarrow \mathcal{P}(X)$ que conserva las inclusiones ($H(L) = H[L]$). El

comportamiento de las homografías así extendidas es muy similar al de las correlaciones, salvo que no invierten las inclusiones. El teorema siguiente muestra la compatibilidad entre ambas nociones:

Teorema 8.39 *Sea H una homografía en un espacio proyectivo X y sean p y q dos correlaciones proyectivas en X . Entonces $H \circ p$ y $p \circ H$ son correlaciones proyectivas $p \circ q$ es una homografía.*

DEMOSTRACIÓN: Los tres hechos se prueban de forma similar. Veamos tan sólo que $p \circ q$ es una homografía. Fijemos un sistema de referencia proyectivo y sean A y B las matrices de p y q en dicho sistema. Sea U la $n+1$ -tupla de coordenadas homogéneas de un punto Q . Entonces $p(Q)$ es el hiperplano de ecuación $UAX^t = 0$. La imagen por q de este hiperplano es UAB^{-1} , luego $p \circ q$ es la homografía inducida por el isomorfismo de matriz AB^{-1} . ■

Si p y q son correlaciones no necesariamente proyectivas, lo único que podemos decir de $p \circ q$ es que es una biyección que conserva las colineaciones. Si estamos en un espacio proyectivo real podemos concluir que se trata de una homografía por el teorema fundamental. En particular, si p es una correlación arbitraria y q es proyectiva, tenemos que $p \circ q = H$ es una homografía, luego $p = q \circ H^{-1}$ es también proyectiva. En definitiva:

Teorema 8.40 *Toda correlación en un espacio proyectivo real es proyectiva.*

El principio de dualidad se aplica a enunciados que contienen homografías. Para verlo, dada una homografía $H \in \text{GP}(X)$ y una dualidad p , definimos $H^p : X \rightarrow X$ como la homografía dada por $H^p(Q) = p(H[p^{-1}(Q)])$.

Veamos un ejemplo. Diremos que una homografía H tiene un *centro de perspectiva* O si fija a los hiperplanos que pasan por el punto O , y diremos que H tiene un *eje de perspectiva* Π fija a todos los puntos del hiperplano Π . Recordemos que hemos definido una perspectividad como una homografía distinta de la identidad que tiene un centro y un eje de perspectiva. Ahora bien, es claro que se cumple la afirmación siguiente:

Toda homografía con un eje de perspectiva tiene un centro de perspectiva.

La razón es que al restringir la homografía al complementario de su eje tiene que ser una traslación o una homotecia. Vamos a probar el teorema dual:

Toda homografía con un centro de perspectiva tiene un eje de perspectiva.

Fijamos una correlación proyectiva p . Basta observar que si H tiene a O por centro de perspectiva, entonces H^p tiene a $p(O)$ como eje de perspectiva: en efecto, si $Q \subset p(O)$, entonces $O \subset p^{-1}(Q)$, luego $H[p^{-1}(Q)] = p^{-1}(Q)$, luego $Q = H^p(Q)$. Por consiguiente H^p tiene un centro de perspectiva E , y un razonamiento similar muestra que $p^{-1}(E)$ es un eje de perspectiva de H . ■

En general podemos dualizar cualquier afirmación que haga referencia a homografías considerando a éstas autoduales.

8.6 Razones dobles y separación armónica

Introducimos ahora otro de los conceptos fundamentales de la geometría proyectiva. Antes conviene establecer un convenio útil para trabajar con perspectivas y homografías. La notación

$$A_1, \dots, A_n \underset{\wedge}{\sim} B_1, \dots, B_n$$

indicará que los puntos colineales distintos A_1, \dots, A_n se transforman respectivamente en B_1, \dots, B_n por una homografía. Cuando la homografía sea una proyección perspectiva de centro O y queramos indicarlo explícitamente escribiremos

$$A_1, \dots, A_n \underset{\wedge}{\overset{O}{\sim}} B_1, \dots, B_n.$$

Notemos que un sistema de referencia proyectivo en una recta está formado simplemente por tres puntos distintos. Por lo tanto, una homografía H entre rectas queda completamente determinada en cuanto conocemos la imagen de tres puntos distintos P, Q, R , es decir, que $H : PQR \underset{\wedge}{\sim} P'Q'R'$ sirve como definición de una homografía H (siempre que las dos ternas de puntos sean colineales).

Si consideramos un cuarto punto S , entonces existe un único punto S' tal que $PQRS \underset{\wedge}{\sim} P'Q'R'S'$. Para determinarlo basta observar que S' ha de tener las mismas coordenadas respecto a P', Q', R' que S respecto a P, Q, R . Esto nos lleva a la definición de razón doble:

Definición 8.41 Dados cuatro puntos distintos P, Q, R, S en una recta proyectiva X sobre un cuerpo K , llamaremos *razón doble* entre ellos a

$$\mathcal{R}(P, Q, R, S) = \frac{\alpha}{\beta} \in K \cup \{\infty\},$$

donde (α, β) son las coordenadas homogéneas de S respecto al sistema de referencia P, Q, R .

Estamos adoptando el convenio de que $\alpha/\beta = \infty$ si y sólo si $\beta = 0$. Veamos una caracterización sencilla pero práctica. Tomemos como punto infinito de $P^2(K)$ a $\infty = (1, 0)$. Los demás puntos (α, β) pueden identificarse con los escalares α/β . Dados P, Q, R, S , existe un único $S' \in P^2(K)$ tal que $PQRS \underset{\wedge}{\sim} \infty 0 1 S'$. Entonces $\mathcal{R}(P, Q, R, S) = S'$, pues claramente la homografía $PQR \underset{\wedge}{\sim} \infty 0 1$ no es más que el sistema de coordenadas determinado por los puntos P, Q, R .

Visto así es obvio que la razón doble se conserva por homografías, es decir, que si $PQRS \underset{\wedge}{\sim} P'Q'R'S'$, entonces

$$\mathcal{R}(P, Q, R, S) = \mathcal{R}(P', Q', R', S').$$

El teorema siguiente explica el nombre de “razón doble”.

Teorema 8.42 Sean P_1, P_2, P_3, P_4 cuatro puntos distintos de una recta proyectiva cuyas coordenadas cartesianas en un sistema de referencia dado sean x_1, x_2, x_3, x_4 respectivamente (pueden ser finitas o infinitas). Entonces

$$\mathcal{R}(P_1, P_2, P_3, P_4) = \frac{x_3 - x_1}{x_4 - x_1} : \frac{x_3 - x_2}{x_4 - x_2}.$$

DEMOSTRACIÓN: Podemos suponer que x_1, x_2 y x_3 son finitas. Consideremos la aplicación $H : \mathbb{P}^2(K) \rightarrow \mathbb{P}^2(K)$ dada por

$$H(x) = \frac{x_3 - x_1}{x - x_1} : \frac{x_3 - x_2}{x - x_2}.$$

Se trata de una homografía, pues puede expresarse como

$$H(x) = \frac{(x_3 - x_1)x - x_2(x_3 - x_1)}{(x_3 - x_2)x - x_1(x_3 - x_2)},$$

y su determinante es $(x_3 - x_1)(x_3 - x_2)(x_2 - x_1)$.

Además se cumple claramente $H(x_1) = \infty$, $H(x_2) = 0$, $H(x_3) = 1$. La composición del sistema de coordenadas dado con la aplicación H es una homografía que hace corresponder $P_1 P_2 P_3 P_4 \xrightarrow{H} \infty 0 1 H(x_4)$, luego la razón doble es $H(x_4)$, lo que nos da la fórmula indicada. ■

Es obvio que existe una homografía que hace corresponder cuatro puntos colineales con otros cuatro puntos colineales si y sólo si sus razones dobles coinciden. Veamos ahora la interpretación geométrica de la razón doble.

Sea X un espacio proyectivo en el que hemos seleccionado un hiperplano infinito Π . Entonces $X \setminus \Pi$ tiene estructura de espacio afín. Sea L una recta en X no contenida en Π . Sean P_1, P_2, P_3, P_4 cuatro puntos distintos en L . Supongamos que todos ellos son finitos. Consideremos un sistema de referencia afín $(O; \vec{v}_1, \dots, \vec{v}_n)$ tal que O y $O + \vec{v}_1$ estén en L . Entonces las coordenadas de cada P_i son de la forma $(x_i, 0, \dots, 0)$. Es claro que x_i es la coordenada de P_i en el sistema de referencia $(O; \vec{v}_1)$, luego podemos calcular la razón doble mediante la fórmula del teorema anterior. Es claro que $\overrightarrow{P_1 P_3} = (x_3 - x_1)\vec{v}_1$, etc. luego la razón doble es

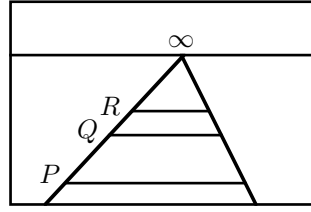
$$\mathcal{R}(P_1, P_2, P_3, P_4) = \frac{\overrightarrow{P_1 P_3}}{\overrightarrow{P_1 P_4}} : \frac{\overrightarrow{P_2 P_3}}{\overrightarrow{P_2 P_4}}.$$

Supongamos ahora que $P_2 = \infty = L \cap \Pi$. Entonces $(x_3 - x_2)/(x_4 - x_2) = 1$ y la expresión se reduce a

$$\mathcal{R}(P_1, \infty, P_3, P_4) = \frac{\overrightarrow{P_1 P_3}}{\overrightarrow{P_1 P_4}}.$$

En particular, si $\mathcal{R}(P_1, \infty, P_3, P_4) = -1$ entonces P_1 es el punto medio del segmento $\overrightarrow{P_3 P_4}$ (suponiendo $\text{car } K \neq 2$).

Las tres rectas que atraviesan el camino de la figura están equiespaciadas, por lo que los puntos originales cumplen $\mathcal{R}(Q, \infty, P, R) = -1$. Sin embargo, la perspectiva transforma el punto infinito de la recta PR en un punto finito, luego la razón doble, aunque sigue siendo -1 , ya no se interpreta como la proporción entre los vectores y vemos que, en efecto, Q ya no es el punto medio del segmento \overline{PR} . Vemos, pues, que las homografías no conservan los puntos medios de los segmentos, ni en general la proporción entre vectores. Sin embargo, las afinidades envían puntos infinitos en puntos infinitos, de donde se sigue que, como ya sabemos, sí conservan las proporciones.



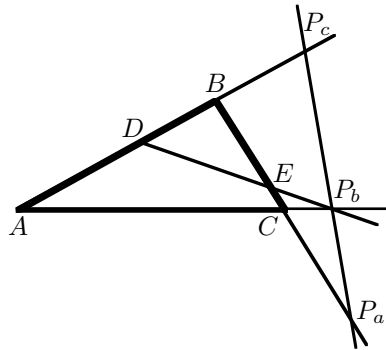
La vista humana es capaz de estimar razones dobles, pues podemos estimar (incluso con un solo ojo) la proporción entre dos segmentos vistos oblicuamente.

El hecho de que las homografías conservan las razones dobles contiene implícitamente el teorema de Tales. En efecto, consideremos un triángulo \widehat{ABC} y una recta r en su plano que no pase por sus vértices. Digamos que r corta a los lados AB , AC y BC en los puntos P_c , P_b y P_a . Entonces $CP_aB \stackrel{P_b}{\wedge} AP_cB$. Si tomamos un punto D en BA y un punto E en BC es evidente que la recta DE pasará por el punto P_b si y sólo si $CP_aBE \stackrel{P_b}{\wedge} AP_cBD$, lo que a su vez equivale a que $\mathcal{R}(C, P_a, B, E) = \mathcal{R}(A, P_c, B, D)$.

En particular, si tomamos como r la recta infinita, lo que hemos probado equivale a que las rectas AC y DE son paralelas si y sólo si

$$\frac{\overrightarrow{CB}}{\overrightarrow{CE}} = \frac{\overrightarrow{AB}}{\overrightarrow{BD}},$$

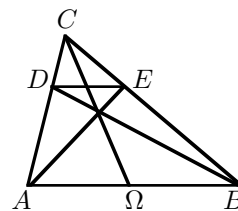
que es justo lo que afirma el teorema de Tales.



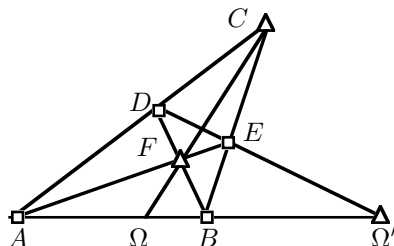
Vamos a seguir esta línea de argumentos para probar que las medianas de un triángulo se cortan en un punto. Más en general, probaremos el teorema siguiente:

Teorema Si la recta DE de la figura es paralela a AB , entonces Ω es el punto medio de AB .

Para probarlo lo transformaremos primero en un teorema proyectivo, es decir, eliminaremos el concepto de paralelismo. Tracemos una figura similar pero con el punto infinito de las rectas paralelas proyectado en un punto finito Ω' .



Esta configuración es muy importante en geometría proyectiva, por lo que hemos de introducir algunas definiciones.



Definición 8.43 Un *cuadrilátero completo* en un plano proyectivo es una figura formada por cuatro puntos no colineales tres a tres, llamados *vértices* junto con las seis rectas que unen cada par de ellos, llamadas *lados*. Dos lados se dicen *contiguos* u *opuestos* según si tienen o no un vértice en común. Las tres intersecciones de lados opuestos se llaman *puntos diagonales* del cuadrilátero.

La figura anterior muestra un cuadrilátero completo, cuyos vértices y puntos diagonales están marcados con cuadrados y triángulos respectivamente. Es fácil ver que los tres puntos diagonales son siempre distintos entre sí.

Dados cuatro puntos colineales A, B, Ω, Ω' , diremos que el par (A, B) *separa harmónicamente* al par (Ω, Ω') si A y B son dos vértices de un cuadrilátero completo, Ω' es un punto diagonal y Ω es la intersección con AB de la recta que une los otros dos puntos diagonales. Lo representaremos $\mathcal{H}(A, B; \Omega, \Omega')$.

El teorema que queremos probar es una consecuencia inmediata de este importante resultado:

Teorema 8.44 Si A, B, Ω, Ω' son cuatro puntos colineales distintos, entonces $\mathcal{H}(A, B; \Omega, \Omega')$ si y sólo si $\mathcal{R}(A, B, \Omega, \Omega') = -1$.

DEMOSTRACIÓN: Supongamos que $\mathcal{H}(A, B; \Omega, \Omega')$. Sea entonces un cuadrilátero completo de vértices A, B, D y E de modo que Ω' sea un punto diagonal y Ω sea la intersección con AB de los otros dos, digamos C y F (ver la figura). Tomemos como sistema de referencia proyectivo los puntos

$$A(1, 0, 0), \quad B(0, 1, 0), \quad C(0, 0, 1), \quad F(1, 1, 1).$$

Respecto a este sistema, la recta AB tiene ecuación $z = 0$, mientras que CF tiene ecuación $x = y$. Por consiguiente $\Omega(1, 1, 0)$.

Del mismo modo vemos que AC es $y = 0$ y FB es $x = z$. Por consiguiente $D(1, 0, 1)$ y similarmente se calcula $E(0, 1, 1)$.

De aquí obtenemos que DE tiene ecuación $z = x + y$. Se trata de la ecuación del plano que contiene a $(0, 0, 0)$ y a las dos ternas de coordenadas. Por ejemplo se obtiene de

$$\begin{vmatrix} x & y & z \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} = 0.$$

Con esto podemos concluir que $\Omega'(1, -1, 0)$. Es claro que si eliminamos la última coordenada (nula) de los puntos de la recta AB obtenemos sus coordenadas homogéneas respecto a un sistema de referencia de la recta. Concretamente queda

$$A(1, 0) = \infty, \quad B(0, 1) = 0, \quad \Omega(1, 1) = 1, \quad \Omega'(1, -1) = -1.$$

Ahora es claro por definición que $\mathcal{R}(A, B, \Omega, \Omega') = -1$.

Recíprocamente, dados A, B, Ω , es fácil construir un punto P tal que $\mathcal{H}(A, B; \Omega, P)$. Por la parte ya probada $\mathcal{R}(A, B, \Omega, P) = -1 = \mathcal{R}(A, B, \Omega, \Omega')$, luego $P = \Omega'$. ■

Ejercicio: Sea X un plano proyectivo sobre un cuerpo K . Probar que $\text{car } K = 2$ equivale a que los puntos diagonales de un cuadrilátero completo estén alineados, y también a que los puntos diagonales de todo cuadrilátero completo estén alineados. La negación de este hecho se conoce como *axioma de Fano*.

Usando el teorema 8.42 es fácil ver que

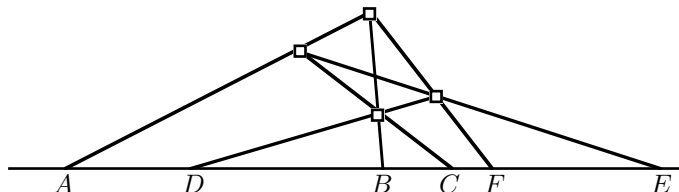
$$\mathcal{H}(A, B; \Omega, \Omega') \Leftrightarrow \mathcal{H}(B, A; \Omega, \Omega') \Leftrightarrow \mathcal{H}(A, B; \Omega', \Omega) \Leftrightarrow \mathcal{H}(\Omega, \Omega'; A, B),$$

En particular, si DE es paralela a AB , entonces el punto Ω' es infinito y $\mathcal{H}(\Omega, \Omega'; A, B)$ luego $\mathcal{R}(\Omega, \Omega', A, B) = -1$, que equivale a que Ω sea el punto medio de A y B , como queríamos probar.

Las equivalencias anteriores expresan que la separación armónica de dos pares de puntos depende únicamente de los pares, A, B y Ω, Ω' , sin importar el orden de ambos ni el orden de sus componentes. Es claro que dados tres puntos colineales distintos A, B y Ω , existe un único punto Ω' distinto de los anteriores tal que $\mathcal{H}(A, B; \Omega, \Omega')$. A dicho punto se le llama el *conjugado armónico* de Ω respecto a A, B . Así, dos puntos cualesquiera A y B dividen la recta que los contiene en pares de conjugados.

Hay una configuración más general en torno a los cuadriláteros completos que conviene introducir.

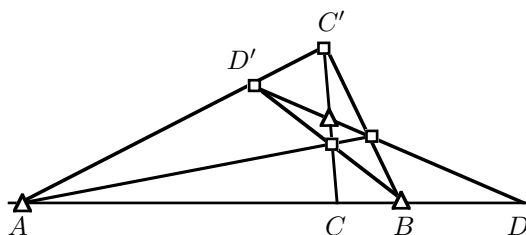
Definición 8.45 Diremos que seis puntos (no necesariamente distintos) en una recta r forman un *conjunto cuadrangular*, y lo representaremos $\mathcal{Q}(ABC; DEF)$, si existe un cuadrilátero completo con vértices exteriores a r y de modo que A, B, C sean las intersecciones con r de los lados que pasan por tres de los vértices y D, E, F sean las intersecciones con r de los lados opuestos respectivos.



Observar que si $Q(ABC; DEF)$ las únicas coincidencias posibles son $A = D$, $B = E$ o $C = F$. Veamos en primer lugar que esta configuración está relacionada con la separación armónica:

Teorema 8.46 *Cuatro puntos distintos cumplen $\mathcal{H}(A, B; C, D)$ si y sólo si $Q(ABC; ABD)$.*

DEMOSTRACIÓN: Si $Q(ABC; ABD)$ entonces A y B son dos puntos diagonales de un cuadrilátero completo. Si llamamos O al tercer punto diagonal y C', D' a los vértices del cuadrilátero de uno de los lados que pasan por A , según indica la figura, entonces $ABCD \stackrel{Q}{=} AB'C'D'$, y es claro que $\mathcal{H}(A, B'; C'D')$, luego lo mismo es válido para $AB\bar{C}D$.

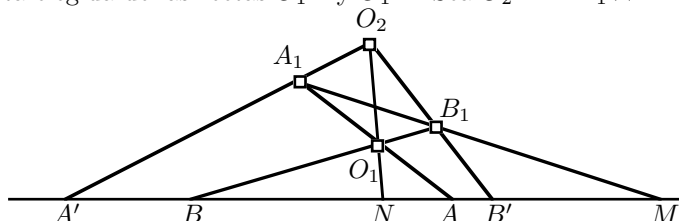


Recíprocamente, si $\mathcal{H}(A, B; C, D)$ tomamos un punto arbitrario C' fuera de la recta, elegimos D' en AC' , trazamos CC' , BD' y la recta que pasa por A y la intersección de estas dos. Con ello determinamos una figura como la anterior salvo que el punto D no es necesariamente el dado, sino otro punto D^* . Por la parte ya probada resulta que $\mathcal{H}(A, B; C, D^*)$, luego $D = D^*$ y por consiguiente $Q(ABC; ABD)$. ■

El resultado principal que vamos a probar sobre conjuntos cuadrangulares es el siguiente:

Teorema 8.47 *Si $M \neq N$ entonces la relación $MNAB \stackrel{\sim}{\sim} MNA'B'$ es equivalente a $Q(MAB; NB'A')$.*

DEMOSTRACIÓN: Supongamos $MNAB \stackrel{\sim}{\sim} MNA'B'$. Tomemos una recta por M y un punto O_1 fuera de las dos rectas. Sean A_1 y B_1 las intersecciones con la recta elegida de las rectas O_1A y O_1B . Sea $O_2 = A'A_1 \cap B'B_1$.



Entonces $MAB \xrightarrow{O_1} MA_1B_1 \xrightarrow{O_2} MA'B'$. Puesto que una homografía está determinada por la imagen de tres puntos, esta sucesión de perspectivas ha de fijar a N , lo cual sólo es posible si N está alineado con O_1 y O_2 , lo que implica $\mathcal{Q}(MAB; NB'A')$. El razonamiento se invierte sin dificultad. ■

Si aplicamos el razonamiento del teorema anterior con $M = N$, obtenemos una homografía con M como único punto fijo, y viceversa. Usaremos la notación $MMAB \frown MMA'B'$ para indicar que la homografía tiene a M como único punto fijo. Observar que $MMA \frown MMA'$ determina una única homografía, pues, considerando como infinito el punto M , una homografía que cumpla esto ha de ser una traslación, luego está determinada por la imagen de un punto A . Una leve modificación del razonamiento anterior nos da el teorema siguiente:

Teorema 8.48 *La relación $MMAB \frown MMA'B'$ equivale a $\mathcal{Q}(MAB; MB'A')$.*

Ejercicio: Probar que cinco puntos de un conjunto cuadrangular determinan el sexto.

Ejercicio: Probar que una relación $\mathcal{Q}(ABC, ABC)$ equivale a que la característica del cuerpo sea 2.

Es inmediato que en una relación $\mathcal{Q}(ABC; A'B'C')$ podemos permutar las tres componentes de uno de los grupos siempre y cuando permutemos igualmente las otras tres. Un hecho no inmediato pero que ya estamos en condiciones de probar es el siguiente:

Teorema 8.49 *La relación $\mathcal{Q}(ABC; A'B'C')$ equivale a $\mathcal{Q}(A'B'C'; ABC)$.*

DEMOSTRACIÓN: Si $A = A'$, $B = B'$, $C = C'$ no hay nada que probar. Por lo tanto podemos suponer que se cumplen a lo sumo dos igualdades. Si de hecho se dan dos igualdades aplicamos 8.46. Si se cumple a lo sumo $A = A'$ podemos aplicar uno de los dos últimos teoremas para concluir que $AA'BC \frown AA'C'B'$. Tomando la homografía inversa tenemos que $A'AB'C' \frown A'ACB$, lo que a su vez equivale a $\mathcal{Q}(A'B'C'; ABC)$. ■

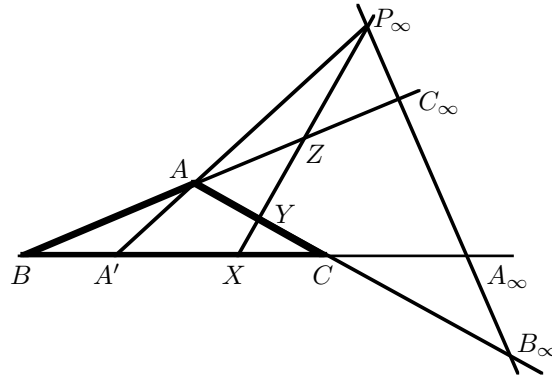
Veamos ahora un par de aplicaciones más de las razones dobles y la separación armónica.

Teorema 8.50 (Teorema de Menelao) *Tres puntos X, Y, Z situados en los lados BC, AC y AB de un triángulo son colineales si y sólo si*

$$\frac{\overrightarrow{BX}}{\overrightarrow{XC}} \frac{\overrightarrow{CY}}{\overrightarrow{YA}} \frac{\overrightarrow{AZ}}{\overrightarrow{ZB}} = -1.$$

DEMOSTRACIÓN: Sean $A_\infty, B_\infty, C_\infty$ y P_∞ los puntos infinitos de las rectas BC, AC, AB y XY . Sea $A' = AP_\infty \cap BC$. Tenemos $YB_\infty CA \xrightarrow{P_\infty} XA_\infty CA'$, luego

$$\frac{\overrightarrow{YC}}{\overrightarrow{YA}} = \mathcal{R}(Y, B_\infty, C, A) = \mathcal{R}(X, A_\infty, C, A') = \frac{\overrightarrow{XC}}{\overrightarrow{XA'}}.$$



La colinealidad de X, Y, Z equivale a $ZC_\infty AB \stackrel{P_\infty}{\wedge} XA_\infty A'B$, o también a

$$\frac{\overrightarrow{ZA}}{\overrightarrow{ZB}} = \mathcal{R}(Z, C_\infty, A, B) = \mathcal{R}(X, A_\infty, A', B) = \frac{\overrightarrow{XA'}}{\overrightarrow{XB}}.$$

Por lo tanto X, Y, Z son colineales si y sólo si

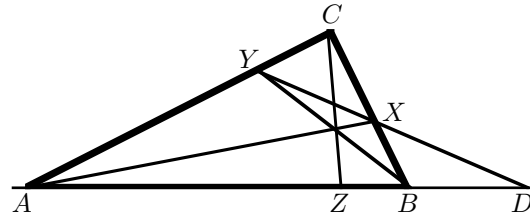
$$\frac{\overrightarrow{BX}}{\overrightarrow{XC}} \frac{\overrightarrow{CY}}{\overrightarrow{YA}} \frac{\overrightarrow{AZ}}{\overrightarrow{ZB}} = \frac{\overrightarrow{BX}}{\overrightarrow{XC}} \frac{\overrightarrow{CX}}{\overrightarrow{XA'}} \frac{\overrightarrow{A'X}}{\overrightarrow{XB}} = -1.$$

■

Teorema 8.51 (Ceva) *En un triángulo, tres cevianas AX, BY, CZ son concurrentes si y sólo si*

$$\frac{\overrightarrow{BX}}{\overrightarrow{XC}} \frac{\overrightarrow{CY}}{\overrightarrow{YA}} \frac{\overrightarrow{AZ}}{\overrightarrow{ZB}} = 1.$$

DEMOSTRACIÓN: Sea $D = XY \cap AB$. Supongamos primero que D es finito.



Entonces, por el teorema de Menelao,

$$\frac{\overrightarrow{BX}}{\overrightarrow{XC}} \frac{\overrightarrow{CY}}{\overrightarrow{YA}} \frac{\overrightarrow{AD}}{\overrightarrow{DB}} = -1$$

Es claro que las cevianas son concurrentes si y sólo si Z y D son conjugados harmónicos respecto a A y B , lo cual a su vez equivale a que

$$\frac{\overrightarrow{AZ}}{\overrightarrow{AD}} : \frac{\overrightarrow{BZ}}{\overrightarrow{BD}} = -1.$$

Multiplicando las dos igualdades obtenemos la condición de Ceva.

Si D es el punto infinito entonces $X \infty CB \stackrel{D}{\underset{\wedge}{=}} Y \infty CA$, luego

$$\mathcal{R}(X, \infty, C, B) = \mathcal{R}(Y, \infty, C, A),$$

donde ∞ representa un punto distinto en cada caso.

Por consiguiente

$$\frac{\overrightarrow{XC}}{\overrightarrow{XB}} = \frac{\overrightarrow{YC}}{\overrightarrow{YA}}.$$

Por otra parte, el hecho de que Z y D sean conjugados armónicos equivale a que Z sea el punto medio de A y B , es decir, a que

$$\frac{\overrightarrow{ZA}}{\overrightarrow{ZB}} = -1.$$

La conclusión es obvia. ■

Para terminar la sección aplicaremos los resultados sobre separación armónica al estudio de las involuciones, que definimos a continuación. Su interés se debe en gran parte a que las involuciones son las polaridades proyectivas de las rectas.

Definición 8.52 Una *involución* en una recta proyectiva X es una homografía $I : X \rightarrow X$ de orden 2, es decir, tal que $I \neq 1$ pero $I \circ I = 1$. Diremos que dos puntos P y Q de X son *conjugados* respecto a I si $I(P) = Q$.

Es claro que cada punto tiene un único conjugado respecto a una involución dada (que puede ser él mismo). Si P es el conjugado de Q , entonces Q es el conjugado de P .

Teorema 8.53 Las involuciones de una recta coinciden con sus polaridades simétricas.

DEMOSTRACIÓN: Sea A la matriz de una involución f en un sistema de referencia cualquiera. Entonces A^2 es un múltiplo de la matriz identidad. Plantando las ecuaciones a las que esto conduce llegamos fácilmente a que

$$A = \begin{pmatrix} u & v \\ w & -u \end{pmatrix}, \quad \text{donde } u^2 + vw \neq 0.$$

Consideremos la polaridad simétrica p que en el mismo sistema de referencia tiene matriz

$$B = \begin{pmatrix} v & -u \\ -u & -w \end{pmatrix}.$$

El punto polar de un punto de coordenadas (a, b) es el que tiene ecuación

$$(a, b) \begin{pmatrix} v & -u \\ -u & -w \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0,$$

es decir, $(av - bu)x + (-au - bw)y = 0$. El punto es, pues, $(au + bw, av - bu)$, que coincide con la imagen de (a, b) por f . Por consiguiente $f = p$. Recíprocamente, toda polaridad simétrica tiene matriz B en un sistema de referencia arbitrario, y la homografía de matriz A es una involución que coincide con la polaridad dada. ■

Ejercicio: Probar que la única polaridad antisimétrica de una recta proyectiva es la identidad.

Teorema 8.54 Sea X una recta proyectiva.

1. Si $H \in \text{GP}(X)$ cumple $H(P) = Q$ y $H(Q) = P$ para dos puntos distintos P y Q , entonces H es una involución.
2. Dados dos pares (A, A') , (B, B') (sin puntos comunes) existe una única involución I respecto a la cual son pares conjugados.
3. Tres pares (A, A') , (B, B') y (C, C') son conjugados respecto a una involución si y sólo si $\mathcal{Q}(A, B, C; A', B', C')$.
4. Si A y B son puntos fijos de una involución I , entonces cada par de puntos conjugados respecto de I son conjugados armónicos respecto de A y B .
5. Toda homografía en X es producto de dos involuciones.

DEMOSTRACIÓN: 1) Sea $P = \langle v \rangle$ y $Q = \langle w \rangle$. Si $H = H(u)$, entonces $u(v) = \alpha w$ y $u(w) = \beta v$. La matriz de u en la base v, w es

$$A = \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix},$$

y es claro que $A^2 = \alpha\beta I$, luego $H^2 = 1$.

2) Si uno de los pares tiene componentes distintas, digamos $A \neq A'$, entonces la homografía $AA'B \overline{\wedge} A'AB'$ es una involución por 1), y claramente es la única para la cual los pares dados son conjugados. El caso en que $A = A'$ y $B = B'$ lo deduciremos de 4)

3) Podemos suponer que al menos $A \neq A'$. Si existe la involución entonces $AA'BC \overline{\wedge} A'AB'C'$. Por otra parte $A'AB'C' \overline{\wedge} AA'C'B'$ por el apartado anterior. Componiendo ambas homografías tenemos $AA'BC \overline{\wedge} AA'C'B'$, que por el teorema 8.47 equivale a $\mathcal{Q}(ABC; A'B'C')$. Si se cumple esta condición llegamos igualmente a $AA'BC \overline{\wedge} A'AB'C'$, lo que implica la existencia de la involución.

4) Basta hacer $A = A'$, $B = B'$ en 3) y usar el teorema 8.46.

Ahora podemos completar la prueba de 2): Dados dos puntos distintos A y B , existe una única involución con A y B como puntos fijos pues, dado otro punto C , la imagen de C ha de ser necesariamente su conjugado armónico respecto de A y B , digamos C' , que cumple $ABCC' \overline{\wedge} ABC'C$, y entonces ésta es la única involución posible.

5) Sea f una homografía. Podemos suponer que es distinta de la identidad. Sea A un punto no fijado por f . Sea $f(A) = A' \neq A$ y $f(A') = A''$. Sea I la involución determinada por $AA'A'' \frown A''A'A$ (notar que A puede ser igual o distinto a A''). Entonces $J = (f \circ I)$ intercambia A y A' , luego es una involución. Por consiguiente $f = J \circ I$. ■

Conviene observar que, aunque una homografía en una recta puede tener en general 0, 1 o 2 puntos fijos, lo cierto es que una involución no puede tener un único punto fijo, pues tomándolo como punto infinito obtenemos una afinidad de orden 2, de ecuación $f(x) = a + bx$, no puede ser $b = 1$ pues entonces no tendría orden 2 (aquí hay que suponer $\text{car } K \neq 2$), y entonces la ecuación $x = a + bx$ tiene como solución un segundo punto fijo.

Ejercicio: Dos involuciones se llaman *armónicas* si su producto es una involución. Probar que dos involuciones son armónicas si y sólo si conmutan.

8.7 Espacios sobre cuerpos ordenados

Sabemos que una ordenación sobre un cuerpo se traduce en que en los espacios afines asociados cada recta tiene asociadas dos ordenaciones naturales, mutuamente inversas. En esta sección exploraremos las consecuencias que tiene en geometría proyectiva el trabajar con un cuerpo ordenado.

Para interpretar intuitivamente lo que sigue hemos de pensar que los extremos de una recta afín se unen en el punto infinito, con lo que una recta proyectiva es en realidad una circunferencia. Así, del mismo modo que dos puntos en una circunferencia no determinan un segmento, sino que la dividen en dos arcos, igualmente dos puntos en una recta proyectiva no determinan un segmento, sino dos. Si los puntos son finitos, un segmento es el usual, y el otro está formado por los puntos exteriores al segmento más el punto infinito. Si tenemos un punto finito y otro infinito, los dos segmentos que aparecen son las dos semirrectas con origen en el punto finito. Vemos, pues, que los conceptos afines de “segmento” y “semirrecta” corresponden a un mismo concepto proyectivo: una semirrecta no es más que un segmento con un extremo infinito.

A lo largo de toda la sección supondremos que los espacios proyectivos que consideramos están definidos sobre un cuerpo ordenado K (en particular de característica 0). Dada una recta proyectiva $X = P(V)$, dos puntos de X son de la forma $A = \langle v \rangle$ y $C = \langle w \rangle$, donde v y w son linealmente independientes. Definimos los conjuntos

$$S_+(v, w) = \{ \langle u \rangle \mid u = \alpha v + \beta w \text{ con } \alpha\beta \geq 0 \},$$

$$S_-(v, w) = \{ \langle u \rangle \mid u = \alpha v + \beta w \text{ con } \alpha\beta \leq 0 \}.$$

Es claro que todo punto de X distinto de A y C está en uno y sólo uno de los dos conjuntos anteriores. Ambos contienen infinitos puntos. Únicamente A y C están en ambos a la vez. Los llamaremos *segmentos* de extremos A y C .

Si cambiamos los generadores de P y Q por $v' = av$ y $w' = bv$, entonces según los signos de a y b sucederá

$$S_+(v, w) = S_+(v', w') \quad \text{y} \quad S_-(v, w) = S_-(v', w')$$

o bien

$$S_+(v, w) = S_-(v', w') \quad \text{y} \quad S_-(v, w) = S_+(v', w').$$

De este modo, cada par de puntos A y C divide la recta en dos segmentos, si bien es imposible distinguir uno de otro (la distinción depende de la elección de los generadores). Pese a esto, es posible distinguirlos en función de un sistema de referencia, pues entonces podemos elegir como generador de cada punto finito el vector que corresponde a las coordenadas homogéneas de la forma $(1, \alpha)$. Así, dados dos puntos $P(1, \alpha)$ y $Q(1, \beta)$, podemos definir

$$\begin{aligned} S_+(P, Q) &= \{a(1, \alpha) + b(1, \beta) \mid ab \geq 0\} = \left\{ \left(1, \frac{a\alpha}{a+b} + \frac{b\beta}{a+b}\right) \mid ab \geq 0 \right\} \\ &= \{(1, \lambda\alpha + (1-\lambda)\beta) \mid 0 \leq \lambda \leq 1\}, \end{aligned}$$

que es el segmento de extremos α y β en el sentido de la geometría afín, y $S_-(P, Q)$, que es el complementario del anterior.

Si consideramos un punto $P(1, \alpha)$ y el punto infinito $Q(0, 1)$, entonces

$$\begin{aligned} S_+(P, Q) &= \{a(1, \alpha) + b(0, 1) \mid ab \geq 0\} = \left\{ \left(1, \alpha + \frac{b}{a}\right) \mid ab > 0 \right\} \cup \{\infty\} \\ &= \{(1, \lambda) \mid \alpha \leq \lambda\} \cup \{\infty\}, \end{aligned}$$

que es una de las semirrectas de origen en α . Similarmente se comprueba que $S_-(P, Q)$ es la semirrecta complementaria.

Con esto hemos probado lo que habíamos afirmado: al quitarle el punto infinito a un segmento proyectivo obtenemos un segmento afín, o bien el complementario de un segmento afín, o bien una semirrecta.

Si no fijamos un sistema de referencia, para distinguir los dos segmentos con un mismo par de extremos hemos de usar un tercer punto:

Dados tres puntos distintos A, B, C en una recta, definimos \overline{ABC} como el segmento de extremos A y C que contiene a B . Veamos algunas propiedades:

Teorema 8.55 Sean A, B, C puntos de una recta proyectiva X .

1. $\overline{ABC} = \overline{CBA}$.
2. Si $D \in \overline{ABC}$, $A \neq D \neq C$, entonces $\overline{ABC} = \overline{ADC}$.
3. Si $D \notin \overline{ABC}$ entonces $\overline{ABC} \cup \overline{ADC} = X$ y $\overline{ABC} \cap \overline{ADC} = \{A, C\}$.
4. Dado \overline{ABC} , existen puntos P y Q tales que $\overline{ABC} = \overline{APB} \cup \overline{BQC}$ y $\overline{APB} \cap \overline{BQC} = \{B\}$.

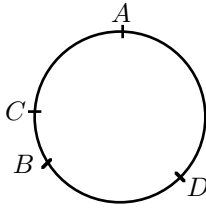
5. Si $D \in \overline{ABC}$, $D \neq A, B, C$ entonces existe P tal que $\overline{BPD} \subset \overline{ABC}$.
6. Una homografía $ABC \xrightarrow{\quad} A'B'C'$ transforma \overline{ABC} en $\overline{A'B'C'}$.

DEMOSTRACIÓN: Las tres primeras propiedades son consecuencias inmediatas de la definición de segmento. La última es inmediata y las propiedades 4) y 5) se reducen a las propiedades de los segmentos afines. Por ejemplo, para probar 4) tomamos como punto infinito un punto del segmento complementario a \overline{ABC} y lo que hay que probar es que si $B \in \overline{AC}$ entonces $\overline{AC} = \overline{AB} \cup \overline{BC}$, $\overline{AB} \cap \overline{BC} = \{B\}$, lo cual es obvio. ■

En general, cualquier propiedad básica sobre segmentos se prueba sin dificultad reduciéndola a una recta afín mediante una elección oportuna del punto infinito.

Aunque, como ya hemos dicho, no podemos ordenar las rectas proyectivas, sí podemos introducir un concepto importante a la hora de “localizar” figuras:

Definición 8.56 Sean A, B, C, D cuatro puntos distintos en una recta proyectiva. Diremos que el par AB *separa* al par CD , y lo representaremos por $AB \parallel CD$, si $D \notin \overline{ACB}$.



Esto significa que los dos segmentos de extremos A y B son \overline{ACB} y \overline{ADB} . En particular cualquiera de los segmentos de extremos A y B corta a cualquiera de los segmentos de extremos C y D . Por el contrario, si $AB \not\parallel CD$, entonces tenemos $C, D \in \overline{ACB}$, luego existe un P tal que $\overline{CPD} \subset \overline{ACB}$. Si tomamos un punto infinito en el segmento complementario de \overline{ACB} tenemos una inclusión de segmentos usuales con extremos distintos, luego es claro que \overline{CPD} es disjunto del segmento complementario de \overline{ACB} . Por lo tanto:

Teorema 8.57 Cuatro puntos colineales distintos cumplen $AB \parallel CD$ si y sólo si los dos segmentos de extremos A y B cortan a cada uno de los segmentos de extremos C y D .

Similarmente se prueba:

Teorema 8.58 La relación $AB \parallel CD$ equivale a

$$AB \parallel DC, \quad BA \parallel CD, \quad CD \parallel AB \quad \text{y} \quad AC \not\parallel BD.$$

La separación de segmentos se caracteriza de forma muy simple mediante las razones dobles:

Teorema 8.59 Cuatro puntos colineales distintos cumplen $AB \parallel CD$ si y sólo si $\mathcal{R}(A, B, C, D) < 0$.

DEMOSTRACIÓN: Consideremos un sistema de referencia en la recta de modo que las coordenadas de A y B sean respectivamente O e ∞ . Entonces

$$\mathcal{R}(A, B, C, D) = \frac{\overrightarrow{AC}}{\overrightarrow{AD}}$$

y es claro que $AB \parallel CD$ si y sólo si los vectores \overrightarrow{AC} y \overrightarrow{AD} tienen sentidos opuestos, lo que equivale a que la razón anterior sea negativa. ■

Como consecuencia vemos que la separación armónica es un caso particular de separación, pues $\mathcal{H}(A, B; C, D)$ equivale a $\mathcal{R}(A, B, C, D) = -1$.

Teorema 8.60 Si $\mathcal{H}(A, B; C, D)$ entonces $AB \parallel CD$.

Terminamos la sección con unas propiedades de las involuciones. Conviene introducir la siguiente nomenclatura habitual:

Definición 8.61 Diremos que una homografía de una recta proyectiva es *elíptica*, *parabólica* o *hiperbólica* según si tiene 0, 1 o 2 puntos fijos.

Más adelante explicaremos el motivo de estos nombres. En la sección anterior hemos visto que no existen involuciones parabólicas. Si suponemos que en el cuerpo K todos los elementos positivos tienen raíz cuadrada entonces existen sólo dos formas canónicas de matrices de involuciones. En efecto:

Teorema 8.62 Toda involución admite en un sistema de referencia adecuado una matriz de la forma

$$M = \begin{pmatrix} 0 & 1 \\ \pm 1 & 0 \end{pmatrix}. \quad (8.3)$$

Si el signo es positivo la involución es *elíptica* y si es negativo es *hiperbólica*. Más en general, una involución es *elíptica* o *hiperbólica* según si el determinante de cualquiera de sus matrices es positivo o negativo.

DEMOSTRACIÓN: Una involución en una recta $X = P(V)$ está inducida por un automorfismo f . Sea $v_1 \in V$ tal que $\langle v_1 \rangle$ no sea un punto fijo. Sea $v_2 = f(v_1)$. Entonces $f(v_2) = av_1$ para un cierto $a \in K$. Sea $v'_2 = v_2/\sqrt{|a|}$. Entonces es claro que $f(v_1) = \sqrt{|a|}v'_2$ y $f(v'_2) = \pm\sqrt{|a|}v_1$. Si cambiamos f por $f/\sqrt{|a|}$ la involución inducida es la misma, pero ahora la matriz en la base v_1, v'_2 es de la forma indicada en el enunciado.

Los puntos fijos de la involución se corresponden con los valores propios de la matriz M , que son las raíces de $\text{pol car } M = x^2 + \Delta$, donde $\Delta = \pm 1$ es el determinante de M . Así pues, M tiene dos valores propios distintos o ninguno según si $\Delta < 0$ o $\Delta > 0$. Por último observamos que, aunque el determinante no es un invariante de las matrices de una misma homografía, matrices de una misma homografía en una misma base se diferencian en un factor constante $a \neq 0$, luego sus determinantes se diferencian en $a^2 > 0$, luego el signo del determinante sí es invariante, y podemos calcularlo a partir de cualquier matriz. ■

Teorema 8.63 *Cada par de puntos conjugados por una involución elíptica separa a cualquier otro par. Si la involución es hiperbólica ningún par separa a ningún otro.*

DEMOSTRACIÓN: Supongamos primer que la involución tiene dos puntos fijos O y P . Entonces podemos tomar a P como punto infinito y sabemos que cada par de puntos conjugados A, A' ha de cumplir $\mathcal{H}(O, \infty; A, A')$, lo que significa que O es el punto medio de A y A' , es claro entonces que si B y B' son también conjugados los segmentos $\overline{AOA'}$ y $\overline{BOB'}$ son concéntricos, luego uno está contenido en el otro, y los extremos no se separan.

Supongamos ahora que existen dos pares de conjugados (A, A') y (B, B') que no se separan y veamos que la involución es hiperbólica, o sea, que su determinante es negativo. La matriz de la involución respecto a dos generadores de A y A' será de la forma (8.3). Sean (x, y) unas coordenadas homogéneas de B en dicha base. Entonces unas coordenadas de B' serán $(\pm y, x)$. Por hipótesis B y B' están en el mismo segmento de extremos A y A' , lo que se traduce en que el signo de xy es el mismo que el de $\pm xy$, luego el signo es positivo y el discriminante de M es negativo. ■

Capítulo IX

Secciones cónicas

Las figuras fundamentales de la geometría euclídea plana son, sin duda, las rectas y las circunferencias. En el capítulo anterior hemos estudiado las rectas desde el punto de vista de la geometría proyectiva, pero no hemos dicho nada acerca de circunferencias. En realidad el concepto de circunferencia es ajeno a la geometría proyectiva, pues por lo pronto depende del concepto de distancia, y la distancia euclídea no puede extenderse al espacio proyectivo. Más precisamente, lo que sucede es que, al contrario que las rectas, las circunferencias no son invariantes proyectivos o, dicho de otro modo, para representar una circunferencia en perspectiva, la figura que hemos de dibujar no es generalmente una circunferencia, sino una cónica.

En general la perspectiva achata las circunferencias convirtiéndolas en “elipses”, pero si el observador se halla sobre la circunferencia, el punto situado a sus pies será enviado al infinito, con lo que la representación será sustancialmente distinta (una parábola), y si el observador se encuentra dentro de la circunferencia la proyección tiene dos puntos en el infinito y es una hipérbola.

En este capítulo estudiaremos las cónicas que hemos introducido en la sección 4.8 desde el punto de vista de la geometría proyectiva. Para ello vamos a necesitar algunos resultados adicionales de álgebra lineal a los que dedicamos la primera sección.

9.1 Clasificación de formas bilineales simétricas

En esta sección consideramos matrices y espacios vectoriales definidos sobre un cuerpo arbitrario k de característica distinta de 2.

Sea V un espacio vectorial de dimensión n sobre k y sea $F : V \times V \longrightarrow k$ una forma bilineal simétrica en V . Sabemos que, fijada una base de V , la aplicación F está determinada por una matriz A con la propiedad de que si dos vectores v y w tienen coordenadas X e Y respecto a la base dada, entonces $F(v, w) = XAY^t$.

Definición 9.1 Diremos que dos matrices simétricas A y B son *congruentes* si existe una matriz regular M tal que $A = MBM^t$.

Es claro que todas las matrices correspondientes a una misma forma bilineal simétrica son congruentes entre sí, y que si B es la matriz de una forma F en una cierta base, cualquier matriz A congruente con B es la matriz de F en una base adecuada. En efecto, dadas dos bases de V , si X e Y son las coordenadas de v y w respecto a una de ellas, sus coordenadas respecto de la otra serán de la forma XM e YM , donde M es la matriz de cambio de base, luego $F(v, w) = XMBM^tY^t$, donde B es la matriz de F en esta segunda base. Por la unicidad de la matriz de una forma bilineal, la matriz de F en la primera base ha de ser $A = MBM^t$. Esto prueba que las matrices de F respecto de dos bases de V son congruentes. Igualmente se razona que si A es la matriz de F en una base, entonces B es la matriz de F en la base determinada por M a partir de la base dada.

Ahora encontraremos invariantes y formas canónicas para las matrices de las formas bilineales simétricas, de modo análogo a como hicimos con las matrices de los endomorfismos de un espacio vectorial. La guía del procedimiento que vamos a seguir consiste en tratar a las formas bilineales sobre un espacio vectorial como si fueran el producto escalar en \mathbb{R}^n . Así, diremos que dos vectores de un espacio vectorial V son *ortogonales* respecto a una forma bilineal simétrica $F: V \times V \rightarrow K$ si cumplen $F(v, w) = 0$. Lo representaremos $v \perp w$.

Hemos de tener presente que la semejanza sólo llega hasta cierto punto. Así, para una forma arbitraria pueden ocurrir casos como que un vector no nulo sea ortogonal a sí mismo, o incluso a todos los demás vectores. Pese a ello vamos a probar que V tiene siempre una base ortogonal, es decir, una base v_1, \dots, v_n tal que $v_i \perp v_j = 0$ cuando $i \neq j$.

Teorema 9.2 *Sea V un espacio vectorial de dimensión finita sobre un cuerpo de característica distinta de 2 y sea F una forma bilineal simétrica en V . Entonces V admite una base ortogonal respecto a F .*

DEMOSTRACIÓN: Lo probaremos por inducción sobre la dimensión de V . En dimensión 1 cualquier base es ortogonal. Supuesto cierto para espacios de dimensión $n-1$, consideremos un espacio de dimensión n . Si F es idénticamente nula cualquier base es ortogonal. En caso contrario existen vectores v, w tales que $F(v, w) \neq 0$. Entonces

$$F(v+w, v+w) = F(v, v) + 2F(v, w) + F(w, w).$$

Necesariamente, uno de los tres vectores $v, w, v+w$ no es ortogonal a sí mismo. Llamemos v_1 a este vector. Sea

$$W = \{v \in V \mid v \perp v_1\}.$$

Claramente W es un subespacio vectorial de V (el subespacio ortogonal a v_1). En general los subespacios ortogonales no tienen el buen comportamiento del caso euclídeo, pero en este caso podemos probar que $V = \langle v_1 \rangle \oplus W$. En efecto, si $w \in W \cap \langle v_1 \rangle$ entonces $w = \alpha v_1 \perp v_1$, es decir, $\alpha F(v_1, v_1) = 0$, lo que implica $\alpha = 0$. Por consiguiente $\langle v_1 \rangle \cap W = 0$.

Por otra parte, W es el núcleo de la aplicación lineal $V \rightarrow K$ dada por $v \mapsto F(v_1, v)$, cuya imagen es todo K (porque no es nula), luego $\dim W = n - 1$, y esto implica que $V = \langle v_1 \rangle \oplus W$.

Por hipótesis de inducción W tiene una base ortogonal respecto a la restricción de F , digamos v_2, \dots, v_n . Es claro que v_1, \dots, v_n es una base ortogonal de V . ■

Es obvio que la matriz de una forma bilineal en una base ortogonal es una matriz diagonal. Por lo tanto hemos probado lo siguiente:

Teorema 9.3 *Toda matriz simétrica sobre un cuerpo de característica distinta de 2 es congruente con una matriz diagonal.*

(Observar que la simetría de la matriz es necesaria. Éste es el motivo por el cual nos hemos restringido a formas simétricas.)

Las matrices diagonales no son formas canónicas, pues nada impide que dos matrices diagonales sean congruentes. El problema de encontrar invariantes y formas canónicas se vuelve extremadamente delicado a partir de este punto, y depende profundamente del cuerpo considerado. Sólo hay un sencillo resultado que podemos añadir en general: Una matriz diagonal de la forma $A = [a_1^2 b_1, \dots, a_n^2 b_n]$, con cada $a_i \neq 0$ es congruente con $B = [b_1, \dots, b_n]$, pues $A = MBM^t$, donde $M = [a_1, \dots, a_n]$, es decir, podemos eliminar cuadrados de la diagonal. Esto nos resuelve el problema en cuerpos donde todo elemento sea un cuadrado, como es el caso de \mathbb{C} y, en general, de todo cuerpo algebraicamente cerrado. En efecto:

Teorema 9.4 *Sea K un cuerpo de característica distinta de 2 en el que todo elemento sea un cuadrado. Entonces toda matriz simétrica A sobre K es congruente con una única matriz de la forma $[1, \dots, 1, 0, \dots, 0]$, donde r es el rango de A .*

En efecto, toda matriz A es congruente con una matriz diagonal, que será de la forma $[a_1^2, \dots, a_r^2, 0, \dots, 0]$, y podemos eliminar los cuadrados para convertirlos en unos. Es obvio que dos matrices congruentes son equivalentes, luego tienen el mismo rango, luego el número de ceros y unos que aparecen es fijo.

En particular dos matrices simétricas sobre \mathbb{C} (de las mismas dimensiones) son congruentes si y sólo si tienen el mismo rango, y esto es comprobable en la práctica. Nos interesa llegar a un resultado similar para matrices sobre \mathbb{R} . En este caso ya no podemos contar con que todo número real sea un cuadrado. Lo que tenemos es que todo número real es de la forma $\pm a^2$. Por lo tanto toda matriz simétrica real es congruente con una de la forma $[\pm a_1^2, \dots, \pm a_r^2, 0, \dots, 0]$, y ésta a su vez es congruente con una de la forma $[1, \dots, 1, -1, \dots, -1, 0, \dots, 0]$. Vamos a probar que estas matrices son canónicas, es decir, que dos distintas nunca son congruentes.

Teorema 9.5 (Ley de inercia de Sylvester) *Todas las matrices diagonales congruentes con una matriz real simétrica A tienen el mismo número s de coeficientes positivos, al que llamaremos *signatura* de A . El número de coeficientes negativos será necesariamente $r - s$, donde r es el rango de A .*

DEMOSTRACIÓN: Consideremos una forma bilineal F en un espacio vectorial V definida por la matriz A en una cierta base. Sean v_1, \dots, v_n y w_1, \dots, w_n dos bases de V en las que la matriz de F sea diagonal. Podemos suponer que estas matrices son $[a_1, \dots, a_r, 0, \dots, 0]$ y $[b_1, \dots, b_r, 0, \dots, 0]$, donde r es el rango de las matrices,

$$a_i > 0 \quad \text{para } i = 1, \dots, s, \quad a_i < 0 \quad \text{para } i = s+1, \dots, r,$$

$$b_i > 0 \quad \text{para } i = 1, \dots, s', \quad b_i < 0 \quad \text{para } i = s'+1, \dots, r.$$

Hemos de probar que $s = s'$.

Sean $S = \langle v_1, \dots, v_s \rangle$, $T = \langle v_{s'+1}, \dots, v_n \rangle$. Veamos que $S \cap T = 0$. En efecto, todo $v \in S$ es de la forma $v = \sum_{i=1}^s \alpha_i v_i$, luego

$$F(v, v) = \sum_{i,j=1}^s \alpha_i \alpha_j F(v_i, v_j) = \sum_{i=1}^s \alpha_i^2 F(v_i, v_i) = \sum_{i=1}^s \alpha_i^2 a_i \geq 0.$$

Además $F(v, v) = 0$ si y sólo si $\alpha_i = 0$ para todo i , si y sólo si $v = 0$.

Similarmente se prueba que si $v \in T$ entonces $F(v, v) \leq 0$ y $F(v, v) = 0$ si y sólo si $v = 0$. Ahora es claro que $S \cap T = 0$.

Tomando dimensiones vemos que $s + n - s' \leq n$, o sea, $s \leq s'$. Del mismo modo se prueba la desigualdad contraria. ■

El teorema anterior nos permite definir la *signatura* de una forma bilineal simétrica real como la de cualquiera de sus matrices. Ahora es claro que una matriz simétrica real de rango r y signatura s es congruente con la matriz diagonal

$$[\underbrace{1, \dots, 1}_s, \underbrace{-1, \dots, -1}_{r-s}, 0, \dots, 0],$$

y que dos matrices de esta forma no pueden ser congruentes, pues tienen distinto rango o signatura. Por consiguiente:

Teorema 9.6 *Dos matrices simétricas reales son congruentes si y sólo si tienen el mismo rango y la misma signatura.*

En realidad el teorema anterior es válido en cualquier cuerpo ordenado en el que los elementos positivos sean cuadrados. Esto incluye, por ejemplo, al cuerpo de los números reales constructibles con regla y compás. La obtención de formas canónicas en cuerpos como \mathbb{Q} requiere resultados importantes de teoría de números.

Con el teorema anterior tenemos resuelto el problema teórico que nos habíamos planteado, pero conviene que nos detengamos a dar un método para calcular en la práctica la signatura de una matriz simétrica real, pues la forma de hacerlo es muy simple. Vamos a probar que todo se reduce a calcular los valores propios de la matriz. Nos basamos en el teorema siguiente:

Teorema 9.7 Sea $h : V \longrightarrow V$ un endomorfismo en un espacio vectorial euclídeo. Si la matriz de h en una base ortonormal es simétrica, entonces V tiene una base ortonormal formada por vectores propios de h .

DEMOSTRACIÓN: Sea A la matriz de h a la que alude el enunciado. En primer lugar probaremos que su polinomio característico tiene todas sus raíces reales. Sea $\lambda \in \mathbb{C}$ una de sus raíces. Entonces $zA = \lambda z$ para un cierto $z \in \mathbb{C}^n$ no nulo. Conjugando, $\bar{z}A = \bar{\lambda}\bar{z}$, luego $A\bar{z}^t = \bar{\lambda}\bar{z}^t$. Multiplicando por z queda $\bar{\lambda}z\bar{z}^t = zA\bar{z}^t = \lambda z\bar{z}^t$. Es claro que $z\bar{z}^t \neq 0$, luego $\lambda = \bar{\lambda}$.

Sean $\lambda_1, \dots, \lambda_r$ las raíces distintas del polinomio característico de A , es decir, los valores propios de A . Para cada i , sea V_i el espacio de vectores propios asociado. Veamos que si $i \neq j$ entonces $V_i \perp V_j$.

Sean x e y las coordenadas de dos vectores en V_i y V_j respectivamente. Entonces $xA = \lambda_i x$, $yA = \lambda_j y$, luego $Ay^t = \lambda_j y^t$. Multiplicando obtenemos $\lambda_i xy^t = xAy^t = \lambda_j xy^t$. Puesto que $\lambda_i \neq \lambda_j$ ha de ser $xy^t = 0$, lo que prueba que los vectores correspondientes son ortogonales.

Sea $W = V_1 \perp \dots \perp V_r$. Veamos que $h[W^\perp] \subset W^\perp$. En efecto, sea $v \in W^\perp$. Sean x las coordenadas de v e y las coordenadas de un vector de V_i . Entonces $xy^t = 0$ y $yA = \lambda_i y$, luego $xAy^t = \lambda_i xy^t = 0$. Como xA son las coordenadas de $h(v)$, concluimos que $h(v)$ es ortogonal a cada V_i , luego a W .

Esto prueba que h puede restringirse a un endomorfismo de W^\perp . Es fácil ver que la matriz de h en cualquier base ortonormal es simétrica (basta probar que toda matriz de la forma MAM^t lo es). Si $W^\perp \neq 0$, la matriz de $h|_{W^\perp}$ en una base ortonormal será simétrica (basta completarla hasta una base ortonormal de V y usar que la matriz de h en esta base lo es). Aplicando lo ya probado llegamos a que h tiene un vector propio en W^\perp , pero esto es absurdo, pues todos los vectores propios de h están en W . Así pues, $W^\perp = 0$ y $V = W \perp W^\perp = W$.

Ahora basta tomar una base ortonormal en cada V_i y unir las todas. ■

Equivalentemente, hemos probado que toda matriz simétrica es semejante a una matriz diagonal y, más aún, que la matriz de semejanza puede tomarse ortogonal (pues transforma una base ortonormal en otra base ortonormal).

En términos de los invariantes de una matriz, lo que hemos probado es que las matrices simétricas reales tienen todos sus divisores elementales de grado 1, por lo que la dimensión del espacio fundamental de un valor propio es su multiplicidad en el polinomio característico. Ahora probamos el resultado que nos interesa:

Teorema 9.8 Sea F una forma bilineal simétrica en un espacio vectorial euclídeo. Entonces existe una base ortonormal de V que es ortogonal para F .

DEMOSTRACIÓN: Sea A la matriz de F en una base ortonormal de V , que será simétrica. Sea h el endomorfismo de V que en dicha base tiene matriz A . Por el teorema anterior V tiene una base e_1, \dots, e_n formada por vectores propios de h , es decir, si llamamos x_1, \dots, x_n a sus coordenadas en la base

original, $x_i A = \lambda_i x_i$, luego si $i \neq j$ se cumple $F(e_i, e_j) = x_i A x_j^t = \lambda_i x_i x_j^t = 0$, luego la base es ortogonal para F . ■

Así pues, dada una matriz simétrica real A , podemos considerar un espacio euclídeo y en él una forma bilineal que tenga matriz A en una base ortonormal. Si B es la matriz de F en la base dada por el teorema anterior, tenemos que B es diagonal y $B = M A M^t$, donde M es la matriz del cambio de base entre dos bases ortonormales, luego es ortogonal, luego $M^t = M^{-1}$. Esto significa que las matrices A y B son semejantes, luego los coeficientes de la diagonal de B son los valores propios de A . En resumen:

Teorema 9.9 *Dada una matriz simétrica real A , la matriz diagonal B formada con los valores propios de A (repetidos según su multiplicidad en el polinomio característico) es congruente con A (y la matriz de congruencia puede tomarse ortogonal). Por lo tanto la signatura de A es el número de valores propios positivos.*

9.2 Cónicas proyectivas y afines

Aunque en la sección 4.8 hemos definido las cónicas en el contexto de un plano afín euclídeo, es decir, mediante propiedades que involucran distancias, en realidad podemos llegar a las mismas curvas mediante una definición válida en cualquier plano afín. Una primera aproximación sería la definición siguiente, basada en las formas canónicas que hemos obtenido para las cónicas en el teorema 4.36:

Definición 9.10 Una *cónica afín* (real) es un conjunto de puntos del plano afín (real) que en un sistema de referencia arbitrario satisfacen una ecuación de la forma

$$a x^2 + b y^2 + d x y + e x + f y + c = 0 \quad (9.1)$$

con alguno de los coeficientes a, b, d no nulo.

Observemos que esta definición no depende del sistema de referencia considerado, es decir, que un conjunto de puntos verifica una ecuación de tipo (9.1) en un sistema de referencia si y sólo si verifica otra (no la misma) respecto de otro sistema de referencia cualquiera. Esto se debe a que las coordenadas de un punto respecto de dos sistemas de referencia dados están relacionados por ecuaciones de la forma

$$x = p + r x' + s y', \quad y = q + t x' + u y', \quad ru - st \neq 0,$$

y es claro que al sustituir estas ecuaciones en (9.1) obtenemos otra ecuación del mismo tipo. Igualmente concluimos que la imagen de una cónica por una biyección afín es de nuevo una cónica.

También es obvio que las cónicas en el sentido de la definición 4.35 son cónicas en este sentido general, incluidas las circunferencias, que admiten ecuaciones

de la forma $x^2 + y^2 = r^2$. Sin embargo, no es cierto que toda cónica en el sentido de 9.10 lo sea también en el sentido de 4.35. Por ejemplo, es obvio que, según 9.10, la ecuación $xy = 0$ define una cónica que está formada por dos rectas, mientras que $x^2 + y^2 = 0$ define una cónica formada por un solo punto. No obstante, enseguida veremos cómo se pueden excluir estos casos triviales y determinar las ecuaciones de tipo (9.1) que definen realmente elipses, parábolas o hipérbolas.

Antes de ocuparnos de esta cuestión, vamos a introducir el concepto de cónica en un plano proyectivo, que no es sino una ligera modificación de la definición anterior:

Definición 9.11 Sea X un plano proyectivo real. Una *cónica* en X es un conjunto de puntos caracterizado por que sus coordenadas en un cierto sistema de referencia satisfacen una ecuación de la forma

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz = 0 \quad (9.2)$$

con algún coeficiente no nulo.

Notemos que si un punto tiene coordenadas homogéneas (x, y, z) respecto de un sistema de referencia dado, cualquier terna $(\lambda x, \lambda y, \lambda z)$ con $\lambda \neq 0$ es también una terna de coordenadas homogéneas para el mismo punto, pero es claro que la primera satisface la ecuación (9.2) si y sólo si lo hace la segunda, es decir, que el hecho de que las coordenadas homogéneas de un punto cumplan o no la ecuación es independiente de la elección concreta de éstas.

Notemos también que toda ecuación de la forma (9.2) se puede expresar matricialmente como $XAX^t = 0$, para una cierta matriz simétrica A . Basta tomar

$$A = \begin{pmatrix} a & d/2 & e/2 \\ d/2 & b & f/2 \\ e/2 & f/2 & c \end{pmatrix}.$$

Recíprocamente, toda matriz simétrica A no nula define una cónica de ecuación $XAX^t = 0$.

Dados dos sistemas de referencia proyectivos en X , las coordenadas de un mismo punto entre ambos sistemas están relacionadas por una ecuación de la forma $X = X'M$, para una cierta matriz regular M , luego los puntos de una cónica que en un sistema de referencia está asociada a una matriz A , en el otro son los puntos de la cónica dada por $X'MAM^tX'^t = 0$. Esto prueba que la definición anterior no depende del sistema de referencia, en el sentido de que si una curva satisface la ecuación (9.2) respecto de un sistema de referencia con matriz A , entonces en cualquier otro sistema de referencia satisfará una ecuación análoga con matriz MAM^t . Análogamente se razona que la imagen de una cónica por una homografía es una cónica. Además, dos matrices corresponden a la misma cónica en dos sistemas de referencia proyectivos (o a una cónica y a su imagen por una homografía) si y sólo si son congruentes.

Al igual que sucede con la definición de cónica afín, la de cónica proyectiva también incorpora algunos casos triviales que conviene excluir. Para identificarlos empezamos observando que, dada una cónica proyectiva con matriz A , el teorema 9.5 nos da que A es congruente con una única matriz diagonal cuya diagonal está formada por ceros, unos y menos unos, y esto equivale a que, cambiando de sistema de referencia (o, equivalentemente, aplicando una homografía), podemos transformar una ecuación general de tipo (9.2) en una de las siguientes:

$$x^2 + y^2 + z^2 = 0, \quad x^2 + y^2 - z^2 = 0, \quad x^2 + y^2 = 0, \quad x^2 - y^2 = 0, \quad x^2 = 0.$$

(Notemos que no hace falta considerar, por ejemplo, $-x^2 - y^2 = 0$ porque es equivalente a la tercera ecuación de la lista anterior.)

La primera ecuación no la cumple ningún punto, la tercera la cumple sólo el punto $(0, 0, 1)$, la cuarta la cumplen los puntos de las rectas $x + y = 0$, $x - y = 0$ y la quinta la cumplen los puntos de la recta $x = 0$. Los tres últimos casos se distinguen de los dos primeros por que la matriz A es singular. Así pues, el único caso no trivial es el segundo.

Concluimos que hay cinco tipos de cónicas proyectivas. Para distinguirlos introducimos el concepto de *índice* de una matriz simétrica real A : Si A tiene rango r y signatura s , el índice de A es el mínimo entre s y $r - s$. De este modo, dos matrices A y B tienen el mismo rango y el mismo índice si y sólo si A es congruente con $\pm B$. Esto nos da la siguiente clasificación de las cónicas:

Rango	Índice	Ecuación	Tipo
3	0	$x^2 + y^2 + z^2 = 0$	Imaginaria
3	1	$x^2 + y^2 - z^2 = 0$	Real
2	0	$x^2 + y^2 = 0$	Punto
2	1	$x^2 - y^2 = 0$	Dos rectas
1	0	$x^2 = 0$	Recta

Las cónicas propiamente dichas son las que hemos llamado “reales”. Las cónicas imaginarias son vacías. Dos cónicas del mismo tipo están inducidas por matrices congruentes, luego existe una homografía que transforma una en otra. Enseguida veremos que las cónicas reales tienen más de un punto pero no contienen rectas, de donde se sigue que una cónica no puede ser de dos tipos a la vez, y que, de hecho, dos cónicas son del mismo tipo si y sólo si existe una homografía que transforma una en otra. El tipo de una cónica queda, pues, determinado por el rango y el índice de una cualquiera de sus matrices.

Las cónicas de rango menor que tres se llaman *degeneradas*. Mientras no indiquemos lo contrario, la palabra “cónica” la reservaremos para las cónicas reales. Así, desde un punto de vista proyectivo, todas las cónicas son iguales.

Teorema 9.12 *Por cinco puntos no alineados tres a tres pasa una única cónica.*

DEMOSTRACIÓN: Podemos tomar a cuatro de los puntos como sistema de referencia, de modo que sus coordenadas serán $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ y $(1, 1, 1)$. Sea A la matriz en este sistema de una cónica que pase por los cinco puntos. Para que pase por los tres primeros ha de ser de la forma

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

Para que pase por $(1, 1, 1)$ ha de cumplirse la ecuación $a + b + c = 0$. Para que A sea regular ha de ser $abc \neq 0$. Puesto que podemos sustituir la matriz A por un múltiplo sin que ello varíe la cónica que define, podemos suponer $b = 1$. En definitiva ha de ser

$$A = \begin{pmatrix} 0 & a & 1 \\ a & 0 & -1-a \\ 1 & -1-a & 0 \end{pmatrix}, \quad \text{con } a \neq 0, -1.$$

El quinto punto tendrá coordenadas (x, y, z) . La cónica pasará por él si y sólo si

$$a = \frac{yz - xz}{xy - yz}.$$

No puede ocurrir que $xy - yz = 0$, pues esto implica que el quinto punto está en la recta $y = 0$ o en la recta $x - z = 0$, luego estaría alineado con $(0, 0, 1)$ y $(1, 0, 1)$ o bien con $(0, 1, 0)$ y $(1, 1, 1)$. Por el mismo motivo no puede ocurrir que el valor de a así determinado sea 0 o -1 . ■

El teorema anterior prueba además que dos matrices corresponden a la misma cónica real si y sólo si se diferencian en un factor constante. Una variante de la prueba anterior nos da el hecho siguiente, que ya habíamos comentado:

Teorema 9.13 *Una cónica no contiene tres puntos colineales.*

DEMOSTRACIÓN: Puesto que todas las cónicas son proyectivamente equivalentes (transformables mediante homografías) el teorema anterior prueba que toda cónica contiene cinco puntos no colineales tres a tres. Si contuvieran tres puntos colineales A, B, C , podríamos tomar dos puntos más D y E de modo que B, C, D, E fueran no colineales tres a tres. Tomamos un sistema de referencia respecto al cual los cuatro últimos puntos tengan coordenadas $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ y $(1, 1, 1)$. Si el primer tiene coordenadas (x, y, z) , los cálculos del teorema anterior muestran que la matriz de la cónica en este sistema de referencia tiene determinante nulo. ■

Con esto termina la prueba de que los cinco tipos de cónicas que hemos distinguido son disjuntos dos a dos.

Nuestra definición de cónica vale igual para planos proyectivos complejos. Los tipos posibles de cónicas dependen sólo del rango de la matriz asociada. Las formas canónicas son:

$$x^2 + y^2 + z^2 = 0, \quad x^2 + y^2 = 0, \quad x^2 = 0.$$

El primer tipo es una cónica propiamente dicha (no degenerada), el segundo es el par de rectas $x - iy = 0$ y $x + iy = 0$, el tercero es la recta $x = 0$. Luego las únicas cónicas no triviales son nuevamente las no degeneradas.

Veamos ahora la relación entre las cónicas proyectivas y las cónicas afines definidas al principio de esta sección.

Si fijamos una recta infinita r en el plano proyectivo X y consideramos el espacio afín $E = X \setminus r$, sabemos que cada sistema de referencia afín en E determina un sistema de referencia en X respecto al cual las coordenadas cartesianas (x, y) se corresponden con las coordenadas homogéneas $(x, y, 1)$ y la recta infinita r tiene ecuación $z = 0$. Entonces, un punto finito $(x, y, 1)$ cumple la ecuación (9.2) si y sólo si sus coordenadas afines (x, y) cumplen la ecuación (9.1). Recíprocamente, todo plano afín E puede completarse hasta un plano proyectivo X , por lo que cada cónica afín dada por una ecuación (9.1) puede extenderse a una cónica proyectiva que cumpla la ecuación correspondiente (9.2).

Notemos que en la definición 9.10 hemos exigido que la ecuación (9.1) tuviera no nulo alguno de los coeficientes a, b, d , pero es que si los tres son nulos entonces la correspondiente cónica proyectiva es degenerada, luego con ello sólo estábamos eliminando por adelantado algunos de los casos triviales. Observemos también que cualquier cónica proyectiva degenerada se restringe a una cónica afín trivial, pues si, por ejemplo, está formada por dos rectas, su restricción también estará formada por dos rectas (o por una, si es que la otra era la recta infinita, aunque entonces no cumplirá el requisito de que $abd \neq 0$). Por lo tanto, si consideramos únicamente cónicas reales, la restricción a E de una cónica proyectiva es una cónica afín y, recíprocamente, toda cónica afín no trivial se extiende a una cónica proyectiva.

Ahora vamos a probar que el concepto de cónica real es equivalente al concepto de cónica definido en la sección 4.8, es decir, que no introduce ninguna curva que no sea una elipse, una hipérbola o una parábola. Para comparar ambas definiciones hemos de partir de un plano afín euclídeo E , que completamos hasta un plano proyectivo X . Consideremos una cónica afín en E cuya extensión a X no sea degenerada (aunque admitimos que pueda ser imaginaria, y así llegaremos a un criterio sencillo para determinar cuándo una cónica no degenerada es real y cuándo imaginaria).

Fijemos un sistema de referencia ortonormal en E respecto al cual la ecuación de la cónica sea (9.1). Vamos a reducirla a una forma canónica aplicando isometrías (o, equivalentemente, pasando a otro sistema de referencia ortonormal). La matriz asociada a la compleción proyectiva es

$$A = \left(\begin{array}{cc|c} a & d/2 & e/2 \\ d/2 & b & f/2 \\ \hline e/2 & f/2 & c \end{array} \right).$$

Observemos que la submatriz A_0 formada por las dos primeras filas y columnas es no nula, pues de lo contrario la cónica sería degenerada. En virtud del teorema 9.9, existe una matriz ortogonal M tal que MA_0M^t es diagonal. El

cambio de coordenadas

$$(x, y) = (x', y')M$$

corresponde a una isometría en E que transforma

$$ax^2 + by^2 + dxy = (x, y)A_0 \begin{pmatrix} x \\ y \end{pmatrix}$$

en

$$ux'^2 + vy'^2,$$

donde u y v son los valores propios de A_0 . Por lo tanto, transforma (9.1) en una ecuación de la forma

$$ax^2 + by^2 + ex + fy + c = 0,$$

donde los coeficientes a, b, e, f, c son distintos de los de la ecuación original. Si $a \neq 0$, observamos que $ax^2 + ex = a(x + e/2a)^2 - e^2/4a$, luego la traslación $x' = x + e/2a$ transforma la ecuación en otra de la forma

$$ax^2 + by^2 + fy + c = 0.$$

Si $b \neq 0$ podemos hacer igualmente $f = 0$. Como a y b no pueden ser simultáneamente nulos, podemos acabar con una de estas dos ecuaciones:

$$ax^2 + by^2 + c = 0, \quad by^2 + ex + c = 0. \quad (9.3)$$

En el primer caso a, b y c han de ser no nulos, o la cónica sería degenerada. Distinguimos varios casos:

- Si a, b y c tienen el mismo signo la cónica es imaginaria. Notemos que podemos suponer que los tres coeficientes son positivos y entonces la ecuación equivale a

$$-\left(\frac{x}{\sqrt{c/a}}\right)^2 - \left(\frac{y}{\sqrt{c/b}}\right)^2 = 1.$$

- En otro caso, si a y b son ambos negativos, podemos cambiar el signo para que sean ambos positivos, y entonces c es negativo. La ecuación equivale entonces a la de una elipse (incluyendo la posibilidad de que sea una circunferencia) en forma canónica:

$$\left(\frac{x}{\sqrt{-c/a}}\right)^2 + \left(\frac{y}{\sqrt{-c/b}}\right)^2 = 1.$$

- Si a y b tienen signos distintos (intercambiando x e y si es necesario, lo cual es una isometría), podemos suponer que b y c son negativos, con lo que tenemos la hipérbola

$$\left(\frac{x}{\sqrt{-c/a}}\right)^2 - \left(\frac{y}{\sqrt{c/b}}\right)^2 = 1.$$

Por último, en la segunda ecuación ha de ser $e \neq 0$ o la cónica sería degenerada. La traslación $x = x' - c/e$ la convierte en

$$by^2 + ex = 0,$$

que claramente es equivalente a $y^2 = -(e/b)x$, correspondiente a una parábola.

Con esto hemos demostrado:

Teorema 9.14 *Las cónicas en el sentido de la definición 4.35 coinciden con las cónicas afines reales.*

En efecto, hemos probado que, pasando de un sistema de referencia ortonormal a otro la ecuación pasa a ser una de las ecuaciones canónicas que habíamos encontrado para cada tipo de cónica euclídea. Equivalentemente, hemos probado que toda cónica afín no degenerada en un plano euclídeo se puede transformar mediante una isometría en una dada por una de las cuatro ecuaciones

$$-\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \quad y^2 = 4px.$$

Si estamos dispuestos a considerar biyecciones afines en general y no sólo isometrías (o, equivalentemente, sistemas de referencia afines cualesquiera, sin tener en cuenta ninguna estructura euclídea en el plano), las cuatro ecuaciones anteriores (que en realidad son infinitas, pues hay una para cada elección de a , b y p) se transforman en

$$-x^2 - y^2 = -1, \quad x^2 + y^2 = 1, \quad x^2 - y^2 = 1, \quad y^2 = x$$

mediante la transformación $(x', y') = (x/a, x/b)$ en los tres primeros casos y $x' = 4px$ en el tercero. Por consiguiente, hay tres familias de cónicas euclídeas salvo isometría pero solo tres cónicas (reales) afines salvo biyección afín.

Es posible saber a qué tipo corresponde una cónica dada sin necesidad de encontrar el cambio oportuno de sistema de referencia. Para ello observamos que si en la ecuación (9.2) hacemos $z = 0$ obtenemos la ecuación $ax^2 + by^2 + dxy = 0$, que representa los puntos de intersección de la cónica con la recta infinita $z = 0$.

La figura formada por estos puntos es el análogo a una cónica en dimensión 1, y está asociada a la submatriz A_0 . Una afinidad de E (vista como homografía en X) se restringe a una homografía en X , que transformará la intersección en otra figura determinada por una ecuación similar asociada a una matriz 2×2 congruente con A_0 (multiplicada por un escalar no nulo). Por consiguiente, el rango y el índice de esta submatriz (o sólo el rango si el plano proyectivo es complejo) son invariantes afines, es decir, se conservan al aplicar afinidades a las cónicas. Los llamaremos *rango* e *índice reducidos* de la cónica afín dada. Por supuesto, el rango y la signatura de toda la cónica (de toda la matriz A) se conservan por homografías, luego en particular por afinidades, y son también invariantes afines.

La tabla siguiente contiene los cuatro invariantes para las cuatro ecuaciones a las que hemos reducido la ecuación general (bajo la hipótesis de que fuera no degenerada).

R	i	R_0	i_0	Ecuación	Tipo
3	0	2	0	$-x^2 - y^2 = 1$	Imaginaria
3	1	1	0	$x = y^2$	Parábola
3	1	2	0	$x^2 + y^2 = 1$	Elipse
3	1	2	1	$x^2 - y^2 = 1$	Hipérbola

Vemos que conociendo los invariantes sabemos el tipo de cónica correspondiente a una ecuación dada de rango máximo. Según anunciábamos, hemos incluido las cónicas imaginarias en la discusión porque así hemos obtenido un criterio sencillo para determinar si una ecuación dada corresponde a una cónica real o imaginaria.

Aunque hemos obtenido $x^2 - y^2 = 1$ como forma canónica para la hipérbola, aplicando la afinidad $(x', y') = (x + y, x - y)$ podemos transformarla en $xy = 1$, que a menudo resulta más manejable.

Con esto tenemos probado que el concepto de cónica proyectiva (real) se corresponde exactamente con las cónicas euclídeas propiamente dichas (elipses, parábolas e hipérbolas), sin añadir ninguna otra clase de curva, pero de tal forma que cualquier cónica puede transformarse en cualquier otra mediante una homografía.

Ahora observamos que, fijada una recta infinita, hay otra forma de distinguir los tres tipos de cónicas:

Teorema 9.15 *Sea X un plano proyectivo real y sea $E = X \setminus r$ un plano afín resultante de fijar una recta infinita r en E . Entonces una cónica afín en E es una elipse, una parábola o una hipérbola si y sólo si corta a r en 0, 1 o 2 puntos.*

DEMOSTRACIÓN: El número de puntos que una cónica tenga en la recta infinita no depende del sistema de referencia afín en que expresemos su ecuación, luego basta considerar las tres ecuaciones canónicas de la tabla anterior. Para la elipse tenemos $x^2 + y^2 = 1$, que se completa hasta la cónica proyectiva $x^2 + y^2 = z^2$, cuyos puntos infinitos son los puntos de coordenadas homogéneas $(x, y, 0)$ que cumplen $x^2 + y^2 = 0$. Esta ecuación implica $x = y = 0$, pero las tres coordenadas homogéneas no pueden ser nulas, luego la elipse no tiene puntos infinitos.

Para la parábola tenemos la ecuación proyectiva $xz = y^2$, y sus puntos infinitos cumplen $y^2 = 0$, luego el único es $(0, 1, 0)$.

Para la hipérbola tenemos la ecuación $x^2 - y^2 = z^2$, cuyos puntos infinitos cumplen $(x + y)(x - y) = 0$, luego ha de ser $x = \pm 1$, lo que sólo lo cumplen dos puntos: $(1, \pm 1, 0)$. ■

Notemos que si consideráramos una cónica imaginaria, su ecuación proyectiva es $-x^2 - y^2 = z^2$ y sus puntos infinitos tendrían que cumplir $-x^2 - y^2 = 0$, luego tampoco hay. Por ello se habla más precisamente de *elipses imaginarias*, es decir, convendremos en considerar las cónicas imaginarias como elipses.

La idea que subyace en el teorema anterior es que, desde el punto de vista proyectivo, todas las cónicas son elipses. Una hipérbola es una elipse de la que —por el sistema de referencia que estamos considerando— vemos dos puntos de menos, y por eso la vemos “partida” en dos mitades, mientras que una parábola es una elipse de la que vemos un punto de menos, y por eso la vemos “abierta”, aunque de una sola pieza.

Ejemplo Consideremos la cónica de ecuación

$$x^2 + 3y^2 - 4xy + 2x + 4y + 4 = 0.$$

En coordenadas homogéneas es $x^2 + 3y^2 - 4xy + 2xz + 4yz + 4z^2 = 0$. Calculando el determinante de la matriz asociada vemos que no es degenerada. Haciendo $z = 0$ obtenemos la ecuación de sus puntos infinitos: $x^2 - 4xy + 3y^2 = 0$. Como $y = 0$ es claramente imposible podemos hacer $y = 1$, con lo que obtenemos $x^2 - 4x + 3 = 0$, cuyas soluciones son $x = 1, 3$. Así pues, la cónica pasa por los puntos infinitos $(1, 1, 0)$ y $(3, 1, 0)$ y es, por lo tanto, una hipérbola. ■

9.3 La polaridad de una cónica

En esta sección vamos a considerar un plano proyectivo \mathcal{P} sobre un cuerpo arbitrario de característica distinta de 2. Utilizaremos la siguiente definición provisional de cónica, que más adelante sustituiremos por otra más restrictiva. Notemos que es satisfecha por las cónicas reales en un plano proyectivo real, por lo que todo cuando deduzcamos de esta definición será aplicable en particular al caso que veníamos considerando hasta ahora.

Definición 9.16 Una *cónica* en un plano proyectivo \mathcal{P} es el conjunto \mathcal{C} de los puntos de \mathcal{P} cuyas coordenadas respecto de un sistema de referencia dado satisfacen una ecuación de la forma $AXX^t = 0$, para cierta matriz regular simétrica A . Supondremos además que es no trivial en el sentido de que contenga al menos cuatro puntos no colineales.

Las observaciones siguientes a la definición 9.11 son válidas en general: cambiar de sistema de referencia equivale a cambiar la matriz A por otra matriz congruente.

La matriz A define una polaridad p en \mathcal{P} que es independiente del sistema de referencia, pues si pasamos a otro, tanto la matriz de la cónica como la de la polaridad se transforman en otra congruente a través de la misma matriz. Si un punto Q tiene coordenadas X , entonces $p(Q)$ es la recta formada por los puntos cuyas coordenadas Y satisfacen $XY^t = 0$, luego un punto Q está en \mathcal{C} si y sólo si $Q \in p(Q)$, es decir, si y sólo si pertenece a su recta polar.

Consideremos ahora dos puntos distintos P y Q con coordenadas homogéneas X e Y . Los puntos de la recta R que pasa por X e Y son los de la forma $\lambda X + Y$,

donde sólo estamos perdiendo el propio punto X , que no se corresponde con ningún valor de λ . Uno de estos puntos estará en la cónica si y sólo si cumple

$$(\lambda X + Y)A(\lambda X + Y)^t = 0,$$

es decir,

$$\lambda^2 XAX^t + 2\lambda XAY^t + YAY^t = 0.$$

Si $P \notin \mathcal{C}$, entonces la ecuación de segundo grado en λ puede tener dos soluciones, una sola o ninguna, luego cada recta R corta a la cónica a lo sumo en dos puntos. Esto nos lleva a la definición siguiente:

Definición 9.17 Una recta es *tangente* a una cónica si la corta en un solo punto, y es *secante* si la corta en dos puntos.

Continuando el razonamiento anterior, supongamos que el punto P está en la cónica, con lo que $XAX^t = 0$. Vemos entonces que la ecuación anterior nos da un único valor de λ correspondiente a otro punto de la recta R que corta a la cónica, salvo si $XAY^t = 0$, es decir, salvo si $Q \subset p(P)$. Así pues la recta polar de un punto $P \in \mathcal{C}$ es la única tangente a \mathcal{C} que pasa por P , mientras que cualquier otra recta que pasa por P es secante a \mathcal{C} .

El teorema siguiente resume lo que hemos obtenido y algún hecho adicional:

Teorema 9.18 Sea \mathcal{C} una cónica en un plano proyectivo \mathcal{P} .

1. Cada recta corta a \mathcal{C} a lo sumo en dos puntos.
2. Por cada punto $P \in \mathcal{C}$ pasa una única recta tangente a \mathcal{C} , mientras que las restantes son secantes.
3. Existe una única polaridad $p_{\mathcal{C}}$ en \mathcal{P} que a cada punto $P \in \mathcal{C}$ le hace corresponder la recta tangente a \mathcal{C} por P . La llamaremos polaridad asociada a \mathcal{C} .
4. Un punto P está en \mathcal{P} si y sólo si es autopolar, es decir, si y sólo si $P \subset p_{\mathcal{C}}(P)$.
5. El polo de una recta secante a \mathcal{C} es la intersección de las tangentes en los puntos de corte.

DEMOSTRACIÓN: La unicidad de la polaridad de una cónica se debe a que estamos suponiendo que \mathcal{C} contiene cuatro puntos no colineales, y dos polaridades p y p' que coincidan sobre un sistema de referencia proyectivo han de ser iguales, pues $p \circ p'$ es una homografía que fija a un sistema de referencia, luego es la identidad.

Sólo falta demostrar la última propiedad, que se sigue de las propiedades básicas de las polaridades: como la recta pasa por los puntos de corte, las polares de estos puntos (es decir, las rectas tangentes) han de pasar por el polo de la recta. ■

Ejemplo Vamos a calcular la recta tangente a la elipse

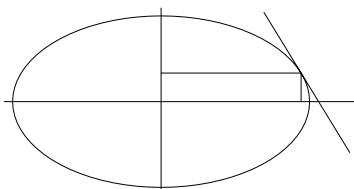
$$\frac{x^2}{3} + y^2 = 1$$

por el punto $(x, y) = (\sqrt{8/3}, 1)$. Para ello consideramos sus coordenadas homogéneas $(\sqrt{8/3}, 1, 1)$ y las multiplicamos por la matriz de la elipse:

$$(\sqrt{8/3}, 1, 1) \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = (\sqrt{8/27}, 1/3, -1).$$

La ecuación de la recta tangente es $\sqrt{8/27}x + (1/3)y - z = 0$, que en coordenadas afines es $\sqrt{8/27}x + (1/3)y - 1 = 0$ o, equivalentemente,

$$y = 3 - \sqrt{8/3}x.$$



■

En los términos que acabamos de introducir, tenemos que la recta infinita es tangente a las parábolas y es secante a las hipérbolas, mientras que no corta a las elipses.

Definición 9.19 Las tangentes a una hipérbola por sus puntos infinitos se llaman *asíntotas*. El *centro* de una cónica es el polo de la recta infinita.

Así, el centro de una cónica es un punto infinito si y sólo si pertenece a su propia recta polar, lo cual equivale a que esté en la cónica y a que la recta infinita sea tangente, es decir, a que la cónica sea una parábola. Las elipses y las hipérbolas, por el contrario tienen el centro finito. En el caso de una hipérbola, como la recta infinita es secante, su polo es el punto en el que se cortan las tangentes por los puntos de corte, es decir, el centro de una hipérbola es el punto donde se cortan sus asíntotas.

Tenemos así una definición afín del centro de una cónica, mientras que la definición que teníamos hasta ahora era euclídea (lo habíamos definido como el punto medio de los focos, y la noción de punto medio es afín, pero la de foco no). El lector puede comprobar que la nueva definición de centro coincide con la anterior sin más que calcular la polar de la recta infinita para las ecuaciones canónicas del teorema 4.36 y comprobar que obtiene $(0, 0)$ para las elipses y las hipérbolas y el punto infinito de tangencia $(1, 0, 0)$ para la parábola.

Ahora sabemos, por ejemplo, que cualquier biyección afín transforma el centro de una cónica en el centro de la imagen.

Ejemplo Consideremos de nuevo la hipérbola de ecuación

$$x^2 + 3y^2 - 4xy + 2x + 4y + 4 = 0.$$

Ya hemos visto que sus puntos infinitos son $(1, 1, 0)$ y $(3, 1, 0)$. Su centro es el punto $(x, y, 1)$ cuya recta polar tiene coordenadas $(0, 0, a)$, es decir

$$(x, y, z) \begin{pmatrix} 1 & -2 & 1 \\ -2 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix} = (0, 0, a),$$

lo que se traduce en las ecuaciones

$$\left. \begin{aligned} x - 2y + z &= 0 \\ -2x + 3y + 2z &= 0 \end{aligned} \right\}$$

La solución es $(7, 4, 1)$. En coordenadas cartesianas el centro es $(7, 4)$.

Las asíntotas son las polares de los puntos infinitos:

$$(1, 1, 0) \begin{pmatrix} 1 & -2 & 1 \\ -2 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0,$$

$$(3, 1, 0) \begin{pmatrix} 1 & -2 & 1 \\ -2 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0,$$

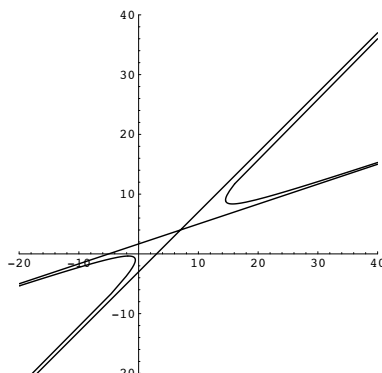
Al desarrollar obtenemos $-x + y + 3z = 0$ y $x - 3y + 5z = 0$. Las ecuaciones cartesianas son, pues,

$$y = x - 3, \quad y = \frac{1}{3}x + \frac{5}{3}.$$

Para representar gráficamente la hipérbola conviene parametrizarla del modo siguiente. Tomamos uno cualquiera de sus puntos, por ejemplo $(1, 1, 0)$ y consideramos el haz de rectas que pasa por él. Se trata del haz de rectas paralelas a la asíntota por este punto, es decir, $y = x + m - 3$, donde el parámetro m varía en \mathbb{R} . La constante -3 hace que para $m = 0$ tengamos la asíntota, mientras que para los demás valores la recta correspondiente corta a la hipérbola en un único punto finito (porque ha de ser secante y uno de los puntos de corte es infinito), que puede calcularse resolviendo el sistema formado por la ecuación de la recta y la de la cónica. El resultado es

$$\left(\frac{-3m^2 + 14m - 19}{2m}, \frac{-m^2 + 8m - 19}{2m} \right)$$

Cuando m recorre los intervalos $]-\infty, 0[$ y $]0, +\infty[$ la expresión anterior nos da cada una de las dos ramas de la hipérbola.



■

Ejercicio: Demostrar que una parábola no puede tener pares de tangentes paralelas, y que por el centro de una elipse o una hipérbola no pasan tangentes, salvo las asíntotas en el caso de una hipérbola.

Ejercicio: Demostrar a partir de las ecuaciones canónicas que la directriz de una cónica es la polar de su foco correspondiente.

9.4 El teorema de Steiner

El tratamiento algebraico de las cónicas que hemos desarrollado hasta aquí es bastante ágil en el caso de planos reales y complejos, pero es poco adecuado para otros cuerpos. Además depende fuertemente de los sistemas de coordenadas. En esta sección veremos que las cónicas se pueden caracterizar geoméricamente de un modo muy simple, lo que permite desarrollar gran parte de la teoría sobre cuerpos arbitrarios y de forma mucho más intuitiva. Necesitamos algunas propiedades sobre haces de rectas. En primer lugar sus expresiones en coordenadas:

En un plano proyectivo X , consideremos el haz de rectas de centro un punto O . Lo representaremos por \mathcal{H}_O . Éste está determinado por dos cualesquiera de sus rectas, digamos r y r' . Fijemos un sistema de referencia, de modo que las rectas tengan ecuaciones

$$ax + by + cz = 0 \quad \text{y} \quad a'x + b'y + c'z = 0.$$

Consideremos la polaridad p asociada al sistema de referencia, esto es, la que en él tiene matriz identidad. Entonces $p(r)$ y $p(r')$ son los puntos $A(a, b, c)$ y $B(a', b', c')$. La polaridad biyecta \mathcal{H}_O con los puntos de la recta $p(O)$. Las coordenadas homogéneas de un punto cualquiera de $p(O) = AB$ son las dadas por

$$(a, b, c) + \lambda(a', b', c'),$$

donde $\lambda \in K \cup \{\infty\}$, entendiendo que $\lambda = \infty$ se corresponde con el punto (a', b', c') . Volviendo a aplicar p concluimos que los miembros de \mathcal{H}_O son las

rectas de ecuación

$$(ax + by + cz) + \lambda(a'x + b'y + c'z) = 0, \quad \text{con } \lambda \in K \cup \{\infty\}.$$

Una homografía $H \in \text{GP}(X)$ induce una biyección $\bar{H} : \mathcal{H}_O \longrightarrow \mathcal{H}_{H(O)}$ dada por $r \mapsto H[r]$. A una biyección de este tipo la llamaremos una *homografía* entre los haces de rectas.

Sean O y O' dos puntos cualesquiera y r, r' dos rectas tales que O no esté en r y O' no esté en r' . Veamos que existe una correspondencia natural entre las homografías $\mathcal{H}_O \longrightarrow \mathcal{H}_{O'}$ y las homografías $r \longrightarrow r'$. Concretamente, a cada homografía H entre los haces le corresponde la homografía dada por $P \mapsto H(OP) \cap r'$. Su inversa asigna a cada homografía H entre las rectas la dada por $s \mapsto O'H(s \cap r)$.

En efecto, es obvio que ambas correspondencias son una inversa de la otra, luego sólo hay que probar que las aplicaciones así definidas son realmente homografías. Sea \bar{H} una homografía entre los haces y sea $s = H[r]$. La proyección perspectiva de s en r' de centro O' se extiende a una perspectividad f de centro O' . Como fija a las rectas que pasan por O' es claro que la homografía $H \circ f$ induce la misma homografía entre los haces de rectas, pero además transforma r en r' . Por lo tanto podemos suponer que $H[r] = r'$. En tal caso, la aplicación que hemos definido entre r y r' es simplemente la restricción de H , luego es una homografía.

Recíprocamente, si $H : r \longrightarrow r'$ es una homografía, tomamos un sistema de referencia formado por O , dos puntos A, B en r y un cuarto punto D . Por otra parte, consideramos el sistema de referencia formado por $O', H(A), H(B), D'$, donde D' es cualquier punto de la recta $O'H(D \cap r)$ no contenido en r . Sea H' la homografía del plano que transforma un sistema de referencia en otro. Entonces H' induce una homografía entre los haces de rectas y ésta a su vez induce una homografía de r en r' que coincide en tres puntos con la dada, luego es la dada. Es claro que \bar{H}' es la aplicación que le hemos asignado a H . ■

En particular esto implica que dadas tres rectas en un haz \mathcal{H}_O , existe una única homografía que las transforma en tres rectas dadas de otro haz $\mathcal{H}_{O'}$. Vamos a obtener la expresión en coordenadas de una homografía entre dos haces.

Una homografía que transforme las rectas $ax + by + cz = 0$ y $a'x + b'y + c'z = 0$ en las rectas $ux + vy + wz = 0$ y $u'x + v'y + w'z = 0$ se corresponde a través de p con una homografía que transforma los puntos (a, b, c) y (a', b', c') en los puntos (u, v, w) y (u', v', w') . Una tercera recta del haz será de la forma

$$(ax + by + cz) + \lambda_0(a'x + b'y + c'z) = 0, \quad \text{con } \lambda_0 \in K \setminus \{0\},$$

que se transformará en una recta

$$(ux + vy + wz) + \mu_0(u'x + v'y + w'z) = 0, \quad \text{con } \mu_0 \in K \setminus \{0\}.$$

La homografía asociada a través de p transformará $(a, b, c) + \lambda_0(a', b', c')$ en $(u, v, w) + \mu_0(a', b', c')$. Es claro que una (y por tanto la única) homografía que cumple estas condiciones es la dada por

$$(a, b, c) + \lambda(a', b', c') \mapsto (u, v, w) + \lambda \frac{\mu_0}{\lambda_0}(a', b', c'),$$

luego la homografía entre los haces de rectas es la dada por

$$ax + by + cx + \lambda(a'x + b'y + c'z) = 0 \mapsto ux + vy + wz + \lambda \frac{\mu_0}{\lambda_0}(u'x + v'y + w'z) = 0.$$

Notemos que podemos multiplicar por μ_0/λ_0 los coeficientes (u', v', w') sin alterar por ello la ecuación de la recta, y entonces la ecuación de la homografía entre haces se expresa simplemente como

$$ax + by + cx + \lambda(a'x + b'y + c'z) = 0 \mapsto ux + vy + wz + \lambda(u'x + v'y + w'z) = 0.$$

Es claro que toda correspondencia de este tipo es una homografía entre haces.

Diremos que una homografía $\bar{H} : \mathcal{H}_O \longrightarrow \mathcal{H}_{O'}$ entre dos haces con centros distintos $O \neq O'$ es una *proyección perspectiva* si H es una perspectividad (más exactamente, si existe una perspectividad H' , no necesariamente H , tal que $\bar{H} = \bar{H}'$).

Una homografía \bar{H} es una perspectividad si y sólo si fija la recta OO' . En efecto, si H es una perspectividad de centro P , como ha de cumplir $H(O) = O'$, los puntos P , O y O' han de ser colineales, luego H fija a OO' . Recíprocamente, si \bar{H} fija a OO' , tomamos dos rectas r y s en \mathcal{H}_O distintas de OO' y sus imágenes $r' = H[r]$ y $s' = H[s]$. Sean $Q = r \cap r'$, $Q' = s \cap s'$. Sea $P = OO' \cap QQ'$. La proyección perspectiva de centro P entre r y s' se extiende a una perspectividad de centro P que transforma O en O' y Q en Q' , luego induce una homografía entre los haces que coincide con la dada en OO' , r y s , luego es la dada.

Notar que una homografía $\bar{H} : \mathcal{H}_O \longrightarrow \mathcal{H}_{O'}$ con $O \neq O'$ no puede fijar a ninguna recta salvo a lo sumo a OO' , luego decir que no es perspectiva equivale a decir que no fija a ninguna recta.

Finalmente podemos probar:

Teorema 9.20 (Steiner) *Sean O y O' puntos distintos en un plano proyectivo real o complejo. Sea $\bar{H} : \mathcal{H}_O \longrightarrow \mathcal{H}_{O'}$ una homografía no perspectiva. Entonces cada recta $r \in \mathcal{H}_O$ corta a $\bar{H}(r)$ en un único punto P_r . La figura C formada por todos los puntos P_r es una cónica no degenerada, y toda cónica no degenerada puede obtenerse de esta forma.*

DEMOSTRACIÓN: Fijemos un sistema de referencia en el plano X . La ecuación de la homografía entre las rectas es

$$ax + by + cx + \lambda(a'x + b'y + c'z) = 0 \mapsto ux + vy + wz + \lambda(u'x + v'y + w'z) = 0.$$

Un punto de coordenadas (x, y, z) estará en C si y sólo si satisface las ecuaciones de ambos haces de rectas con el mismo valor de λ , es decir, si y sólo si

$$-\frac{ax + by + cz}{a'x + b'y + c'z} = -\frac{ux + vy + wz}{u'x + v'y + w'z},$$

o equivalentemente, si y sólo si

$$(ax + by + cz)(u'x + v'y + w'z) - (a'x + b'y + c'z)(ux + vy + wz) = 0.$$

(Hay que considerar aparte el caso $\lambda = \infty$, donde los denominadores se anulan, pero la conclusión es la misma.)

Al operar la última ecuación obtenemos ciertamente la ecuación de una cónica. Notemos que no puede ser idénticamente nula, pues no todos los puntos de X están en C . Más exactamente, cada recta que pasa por O distinta de OO' contiene un único punto adicional de C (su intersección con su imagen), teniendo en cuenta además que C es infinito, concluimos que C no puede ser una cónica degenerada (no es vacía, ni un punto, ni una recta ni un par de rectas, pues en los últimos casos habría una recta por O que contendría infinitos puntos de C).

Recíprocamente, si C' es una cónica no degenerada, tomamos en ella cinco puntos A, B, C, D, E no colineales tres a tres. Existe una única homografía de \mathcal{H}_A en \mathcal{H}_B que transforma AC, AD y AE en BC, BD y BE respectivamente. Si fuera perspectiva, es decir, de la forma \bar{H} , donde H es una perspectividad, entonces el eje de H no pasaría por O ni O' , pues $H(O) = O'$, y los puntos C, D, E tendrían que estar sobre dicho eje, pero no son colineales. Así pues, la homografía no es perspectiva, luego por la parte ya probada determina una cónica C que claramente contiene a los cinco puntos que hemos tomado, luego $C = C'$. ■

Esta caracterización de las cónicas no degeneradas es mucho más adecuada como definición de cónica en un cuerpo arbitrario, pues en cuerpos cuya estructura algebraica no sea tan simple como la de \mathbb{R} o \mathbb{C} resulta ser mucho más operativa que la que hasta ahora manejábamos. Así pues:

Definición 9.21 Sea X un plano proyectivo sobre un cuerpo arbitrario. Una *cónica* en X es la figura formada por las intersecciones de rectas homólogas de dos haces de rectas distintos a través de una homografía no perspectiva.

Para estudiar las cónicas así definidas necesitaremos suponer, como es habitual, que el cuerpo del plano proyectivo tiene característica distinta de 2. Otro caso patológico que conviene excluir es el del cuerpo de tres elementos, pues en el plano que él determina no existen cinco puntos no colineales tres a tres, y todos los teoremas sobre cónicas se vuelven triviales.

Notemos también que una implicación del teorema de Steiner es válida en general, de modo que una cónica está formada por los puntos cuyas coordenadas en un sistema de referencia dado satisfacen una ecuación de la forma $XAX^t = 0$, donde A es una cierta matriz simétrica. En general no sabemos obtener formas

canónicas para la matriz A , pero sabemos al menos que eligiendo el sistema de referencia la podemos tomar diagonal. Su determinante ha de ser no nulo, pues en caso contrario sería de la forma $[1, a, 0]$, y la ecuación correspondiente sería $x^2 + ay^2 = 0$, que corresponde al punto $(0, 0, 1)$ si $-a$ no es un cuadrado en K o a las rectas $x + by = 0$, $x - by = 0$ si $-a = b^2$, y ni un punto ni dos rectas son una cónica. Así pues:

Teorema 9.22 *Los puntos de una cónica están caracterizados por que sus coordenadas en un sistema de referencia dado satisfacen una ecuación de la forma $XAX^t = 0$, donde A es una matriz simétrica regular.*

Lo que no sabemos determinar en general es qué matrices corresponden a cónicas y cuáles no, pero al menos sabemos que las cónicas en el sentido de la definición 9.21 cumplen también la definición 9.16 que hemos usado para definir la polaridad de una cónica, luego todos los resultados sobre polaridades son válidos para la definición de cónica que finalmente hemos adoptado.¹

9.5 Propiedades de las cónicas proyectivas

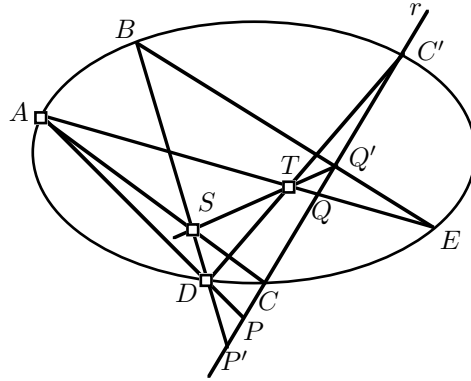
Vamos a extraer las primeras consecuencias de la definición general de cónica proyectiva que acabamos de dar, entre las cuales estarán los hechos que ya conocíamos para el caso particular de las cónicas reales y complejas. En primer lugar es importante notar lo siguiente: dados cinco puntos A, B, C, D, E no colineales tres a tres, podemos construir una cónica que los contiene a partir de la homografía entre \mathcal{H}_A y \mathcal{H}_B que hace corresponder las rectas AC, AD y AE con BC, BD y BE , respectivamente. Para abreviar, las correspondencias como ésta entre rectas de dos haces las representaremos así: $A(CDE) \xrightarrow{\wedge} B(CDE)$.

Hay que destacar que esto no prueba que por los cinco puntos dados pase una única cónica, pues para ello hemos de probar que no importa la elección de A y B como centros de los haces o, dicho de otro modo, que otra homografía como $C(ABE) \xrightarrow{\wedge} D(ABE)$ define la misma cónica. Todavía no estamos en condiciones de probarlo.

Sea \mathcal{C} la cónica determinada por la homografía $A(CDE) \xrightarrow{\wedge} B(CDE)$. Veamos cómo determinar explícitamente los demás puntos de \mathcal{C} a partir de los cinco que tenemos. Más concretamente, tomamos una recta arbitraria r que pase por C y vamos a ver que corta a \mathcal{C} en un único punto distinto de C salvo en un caso excepcional (en que r será tangente a \mathcal{C}). Podemos suponer que r no pasa por A , pues en tal caso es claro que A y C son los únicos puntos de \mathcal{C} en r . Similarmente suponemos que B no está en r .

¹El requisito de que la cónica contenga al menos cuatro puntos no colineales sólo lo hemos usado para asegurar la unicidad de la polaridad de una cónica. Sin él se demuestra igualmente que cada recta corta a una cónica a lo sumo en dos puntos, luego las cónicas no contienen ternas de puntos colineales. La existencia de cuatro puntos se cumple porque el cuerpo ha de tener al menos tres elementos, luego cada recta tiene al menos cuatro puntos, luego cada haz de rectas tiene al menos cuatro rectas, luego cada cónica tiene al menos cuatro puntos.

Según hemos visto al comienzo de la sección, la homografía entre los haces de rectas de A y B que define la cónica determina una homografía sobre r . Si P y Q son los puntos donde AD y AE cortan a r , sus imágenes por esta homografía son los puntos P' y Q' donde BD y BE cortan a r . Ciertamente $P \neq P'$ y $Q \neq Q'$, luego la homografía de r es la dada por $CPQ \bar{\wedge} CP'Q'$.



Ahora observamos que los puntos de r que están en \mathcal{C} son los puntos fijos de $CPQ \bar{\wedge} CP'Q'$. Como la homografía no es la identidad, o bien C es su único punto fijo o bien tiene uno más C' . En ambos casos existe un único punto C' (igual o no a C) tal que $C'CPQ \bar{\wedge} C'CP'Q'$ (ver la pág. 241). Según los teoremas 8.47 y 8.48, el punto C' está unívocamente determinado por la relación $\mathcal{Q}(C'PQ; CQ'P')$.

Veamos cómo construir explícitamente C' , lo que nos mostrará que la igualdad $C = C'$ se da sólo para una elección de r . Vamos a construir un cuadrilátero completo que nos dé el conjunto cuadrangular anterior. Si r pasa por D es inútil seguir, pues D es el punto buscado y es distinto de C . En caso contrario consideramos los puntos $S = BD \cap AC$ y $T = SQ' \cap AE$. Es fácil ver que no están en r . Ahora basta considerar el cuadrilátero completo de vértices A, D, S, T . El punto C' será $DT \cap r$.

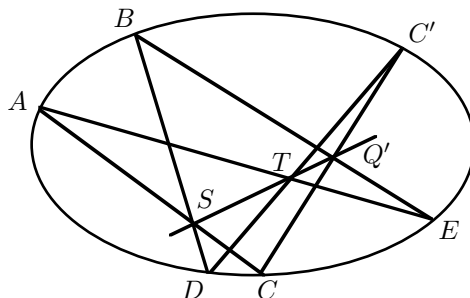
Se cumplirá $C = C'$ si y sólo si T está en DC . Si esto sucede, los puntos S y T están determinados directamente por los puntos de partida (independientemente de r), pues $S = BD \cap AC$ y $T = DC \cap AE$, luego $Q' = BE \cap ST$ también está determinado por los cinco puntos. Así, la recta r ha de ser necesariamente CQ' , donde Q' está determinado como acabamos de indicar. Recíprocamente, invirtiendo el razonamiento es claro que para esta r se cumple $C = C'$.

Si recordamos que C es cualquier punto de \mathcal{C} distinto de A y B , podemos afirmar:

Teorema 9.23 *Si \mathcal{C} es una cónica y C es cualquiera de sus puntos, entonces cada recta de \mathcal{H}_C corta a \mathcal{C} en dos puntos, excepto una, que sólo la corta en C . A dicha recta excepcional la llamaremos tangente a \mathcal{C} en C .*

Esto ya lo sabíamos, pero en realidad hemos probado mucho más. Lo que realmente hemos obtenido se puede enunciar así:

Sea \mathcal{C} una cónica determinada por una homografía entre dos haces de rectas de centros A y B . Sean C, D, E tres puntos más de \mathcal{C} . Entonces un sexto punto C' está en \mathcal{C} si y sólo si los puntos $S = AC \cap BD$, $T = AE \cap DC'$ y $Q' = BE \cap CC'$ están alineados.



Es claro que S, T y Q' son las intersecciones de los lados opuestos del hexágono $ACC'DBE$. Vamos a cambiar la notación para poner de manifiesto una importante simetría en la afirmación anterior. Hacemos $A_1 = A, B_1 = B, A_2 = D, B_2 = C, C_1 = C', C_2 = E$. El resultado es:

Sea \mathcal{C} una cónica determinada por una homografía entre dos haces de rectas de centros A_1 y B_1 . Sean A_2, B_2, C_2 tres puntos más de \mathcal{C} . Entonces un sexto punto C_1 está en \mathcal{C} si y sólo si los puntos $S = A_1B_2 \cap B_1A_2$, $T = A_1C_2 \cap A_2C_1$ y $U = B_1C_2 \cap B_2C_1$ están alineados.

Ahora el hexágono es $A_1B_2C_1A_2B_1C_2$. Lo importante es que la condición sobre S, T, U no se altera si cambiamos los subíndices 1 por los subíndices 2. Por consiguiente tenemos:

Sean $A_1, B_1, C_1, A_2, B_2, C_2$ seis puntos distintos no colineales tres a tres. Entonces C_1 está en la cónica determinada por

$$A_1(A_2B_2C_2) \frown B_1(A_2B_2C_2)$$

si y sólo si C_2 está en la cónica determinada por

$$A_2(A_1B_1C_1) \frown B_2(A_1B_1C_1).$$

Así, dados cinco puntos A_1, B_1, A_2, B_2, C_2 si llamamos \mathcal{C}_1 a la cónica determinada por $A_1(A_2B_2C_2) \frown B_1(A_2B_2C_2)$ y \mathcal{C}_2 a la cónica determinada por $A_2(A_1B_1C_2) \frown B_2(A_1B_1C_2)$, ambas resultan ser la misma cónica. En efecto, ambas contienen a los cinco puntos dados y si C_1 es cualquier otro punto de \mathcal{C}_1 , entonces la cónica dada por $A_2(A_1B_1C_1) \frown B_2(A_1B_1C_1)$ contiene también a C_2 , luego se trata de \mathcal{C}_2 , luego $C_1 \in \mathcal{C}_2$. Intercambiando los papeles obtenemos la otra inclusión.

Para probar que no importa elegir A_1, B_1 o A_2, B_2 observamos que A_1, B_1 determinan la misma cónica que C_1, C_2 y éstos la misma que A_2, B_2 .

Ahora sí podemos afirmar:

Teorema 9.24 *Por cinco puntos no colineales tres a tres pasa una única cónica.*

De hecho tenemos que cualquier cónica está definida a partir de una homografía entre los haces de rectas de centros dos cualesquiera de sus puntos. Esto nos permite expresar de forma completamente simétrica la condición que hemos obtenido para que seis puntos distintos formen parte de una cónica:

Teorema 9.25 (Teorema de Pascal) *Seis puntos no colineales tres a tres están contenidos en una misma cónica si y sólo si los lados opuestos de un hexágono cualquiera que los tenga como vértices se cortan en tres puntos colineales.*

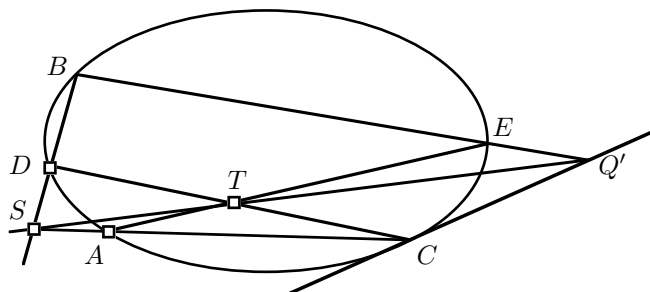
Sigamos recogiendo consecuencias: Si \mathcal{C} es una cónica y A, B son dos cualesquiera de sus puntos, entonces la cónica está determinada por una homografía entre los haces de rectas por A y B . Si $P \in \mathcal{C}$ es distinto de A y B entonces la recta AP es la homóloga de BP . Esto determina la imagen de cualquier recta por A excepto la tangente a \mathcal{C} , que no corta a la cónica en ningún otro punto P , y la recta AB . Similarmente, las rectas BP cubren todas las del haz de rectas por B excepto la tangente en B y la misma AB . Como la homografía no puede ser perspectiva y por lo tanto no fija a ninguna recta, ha de hacer corresponder AB con la tangente por B y la tangente por A con AB . Si convenimos en representar por PP la tangente a \mathcal{C} por un punto P , entonces la expresión $AP \mapsto BP$ recoge todos los casos. Expresaremos que esta correspondencia es una homografía mediante la notación $A[P] \bar{\wedge} B[P]$, donde los corchetes indican que P es una variable (que en este caso recorre los puntos de \mathcal{C}). Esto es, en un contexto general, el núcleo del teorema de Steiner:

Teorema 9.26 (Steiner) *Si \mathcal{C} es una cónica, A y B son dos cualesquiera de sus puntos y P recorre los puntos de \mathcal{C} , entonces $A[P] \bar{\wedge} B[P]$.*

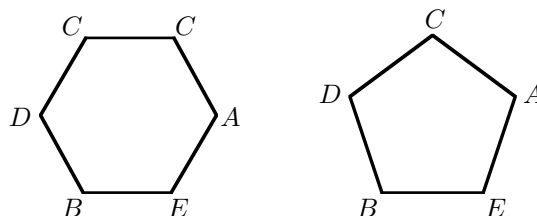
Nos ocupamos ahora del estudio de las tangentes. Hemos probado que si conocemos cinco puntos A, B, C, D, E en una cónica, podemos trazar la tangente por uno de ellos, digamos C determinando los puntos $S = BD \cap AC$, $T = DC \cap AE$ y $Q' = BE \cap ST$. La tangente es entonces CQ' .

Tenemos, pues, el teorema siguiente, que contiene la información necesaria para trazar la tangente a una cónica por un punto dado:

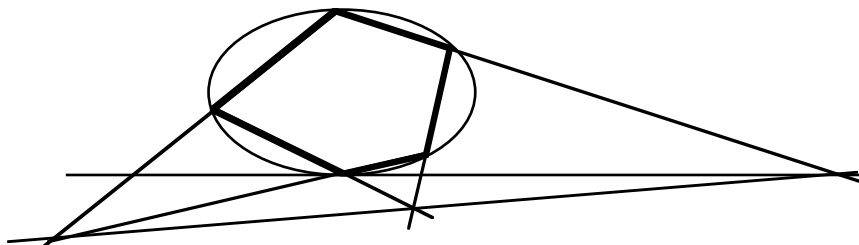
Teorema 9.27 *Sean A, B, C, D, E cinco puntos distintos en una cónica \mathcal{C} . Sea r la tangente a \mathcal{C} por C . Entonces los puntos $r \cap BE$, $AE \cap CD$ y $BD \cap CA$ son colineales.*



Este teorema puede verse como una variante del teorema de Pascal, donde hemos identificado dos puntos en uno. En efecto, imaginemos un hexágono en el que dos vértices tienen el mismo nombre:



Llamemos lados opuestos del pentágono a los mismos que en el hexágono, es decir, AC y DB ; CD y AE ; BE y CC , entendiendo que CC es la tangente a la cónica por C . Lo que afirma el teorema anterior es que si los vértices de un pentágono están contenidos en una cónica, entonces los pares de lados opuestos se cortan en puntos colineales.

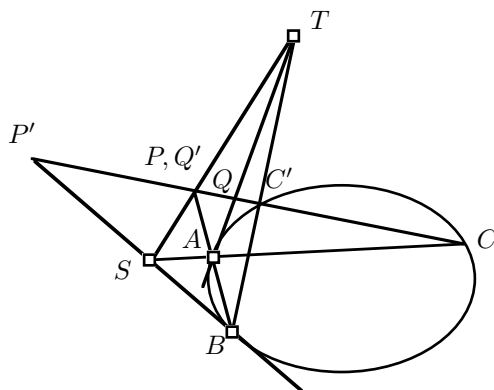


Ejercicio: Probar que existe una única cónica que pasa por cuatro puntos dados no colineales tres a tres y cuya tangente en uno de ellos es una recta dada (que no pase por los otros).

Antes de seguir conviene introducir un par de términos útiles:

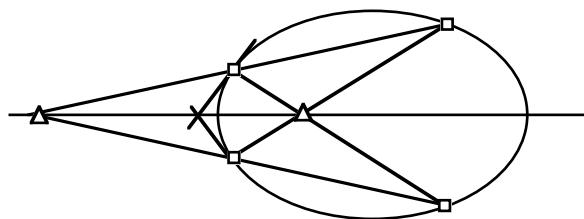
Definición 9.28 diremos que una figura formada por puntos y rectas está *inscrita* en una cónica si todos sus puntos están en la cónica. Diremos que la figura está *circunscrita* a la cónica si todas sus rectas son tangentes a la cónica.

Así por ejemplo, el teorema de Pascal implica que los pares de lados opuestos de un hexágono inscrito en una cónica se cortan en puntos colineales, y según hemos explicado lo mismo vale para pentágonos. Del mismo modo podemos identificar dos pares de lados en un hexágono, y el teorema de Pascal sigue siendo cierto. De hecho hay dos teoremas, según que los pares de lados que identifiquemos sean contiguos u opuestos. Por ejemplo, con la notación que hemos empleado para probar el teorema de Pascal, tenemos el hexágono $ACC'DBE$. Si identificamos los pares de lados contiguos $A = E$ y $B = D$ obtenemos que si un cuadrilátero $ACC'B$ está inscrito en una cónica, entonces los puntos $AB \cap CC'$, $C'B \cap AA$ y $AC \cap BB$ son colineales. La prueba es literalmente la misma del teorema de Pascal sin más que interpretar AE como la tangente por A y BD como la tangente por B . Nos limitamos a dar una figura ilustrativa:



Si identificamos dos pares de vértices opuestos obtenemos una configuración especialmente sencilla:

Teorema 9.29 *Si un cuadrilátero completo está inscrito en una cónica, las tangentes de dos vértices se cortan en un punto colineal con los dos puntos diagonales que no están sobre el lado determinado por dichos vértices.*



Ejercicio: Probar que por tres puntos dados no colineales para una única cónica cuyas tangentes en dos de ellos sean dos rectas dadas.

Ejercicio: Enunciar y demostrar la versión del teorema de Pascal para un triángulo inscrito en una cónica (identificando tres pares de vértices en un hexágono).

El razonamiento con que hemos obtenido (entre otras cosas) el teorema de Pascal tiene todavía una consecuencia más que debemos destacar.

Teorema 9.30 (Desargues) *Consideremos un cuadrilátero completo inscrito en una cónica \mathcal{C} y una recta r que no pase por ninguno de sus vértices pero que corte a \mathcal{C} en uno o dos puntos. Entonces éstos son conjugados respecto a la involución que determinan los pares de puntos donde los lados opuestos del cuadrilátero cortan a r .*

DEMOSTRACIÓN: Recordemos del teorema 8.54 que los pares de puntos donde los lados opuestos del cuadrilátero cortan a r son en efecto conjugados respecto a una (única) involución. Volviendo a la situación que describíamos al comienzo de la sección (ver la figura de la pág. 273), supongamos que r corta a \mathcal{C} en los puntos C y C' (iguales o distintos) y que los vértices del cuadrilátero

son A, B, D, E . Entonces la involución en r es la determinada por los pares (P, Q') y (P', Q) , pero por otra parte hemos visto que el cuadrilátero de vértices A, D, S, T determina la relación $\mathcal{Q}(C', P, Q; C, Q', P')$, luego sus pares de lados opuestos determinan en r la involución para la cual son conjugados (P, Q') , (P', Q) y (C, C') . Los dos primeros pares prueban que se trata de la misma involución que la del cuadrilátero dado y por lo tanto (C, C') es también un par de conjugados respecto a ella. ■

Es fácil ver que si se identifica B con D (y por lo tanto BD pasa a ser la tangente a la cónica por este punto), todo el razonamiento vale igualmente y nos lleva a este caso particular del teorema de Desargues:

Teorema 9.31 *Consideremos un triángulo inscrito en una cónica \mathcal{C} y una recta r que no pase por ninguno de ellos pero que corte a \mathcal{C} en uno o dos puntos. Entonces éstos son conjugados respecto a la involución que determinan el par de puntos donde dos lados del triángulo cortan a r y el par donde el tercer lado y la tangente al vértice opuesto cortan a r .*

La polaridad asociada a una cónica nos permite utilizar argumentos de dualidad. Para ello hemos de definir el concepto dual al de cónica:

Definición 9.32 Una *cónica de rectas* es el conjunto de todas las rectas que pasan por pares de puntos homólogos respecto a una homografía entre dos rectas que no sea una proyección perspectiva.

Es claro que una polaridad transforma una cónica en una cónica de rectas, por lo que todos los teoremas que hemos probado sobre cónicas admiten un dual sobre cónicas de rectas. En particular el concepto dual de recta tangente es lo que se llama un *punto de tangencia* de una recta r perteneciente a una cónica de puntos. Se trata, naturalmente, de un punto de r por el que no pasa ninguna otra recta de la cónica. Por dualidad tenemos que cada recta de una cónica de rectas contiene un único punto de tangencia.

Si consideramos concretamente la polaridad asociada a una cónica dada obtenemos el teorema siguiente:

Teorema 9.33 *Las tangentes a una cónica de puntos forman una cónica de rectas y los puntos de tangencia de una cónica de rectas forman una cónica de puntos.*

DEMOSTRACIÓN: Para probar la primera afirmación basta aplicar la polaridad de la cónica. Para la segunda tomamos una cónica de rectas \mathcal{C}_1 . Al aplicarle una polaridad p obtenemos una cónica de puntos \mathcal{C}_2 cuyas tangentes son las rectas polares de los puntos de tangencia de la cónica de partida. Si p' es la polaridad de \mathcal{C}_2 , al aplicarla obtenemos una nueva cónica de rectas \mathcal{C}_3 . La composición $p \circ p'$ es una homografía que transforma \mathcal{C}_1 en \mathcal{C}_3 . Si Q es un punto de tangencia de \mathcal{C}_1 , entonces $p(Q)$ es una tangente a \mathcal{C}_2 , luego $p'(p(Q))$ es un punto de \mathcal{C}_2 . Esto prueba que los puntos de tangencia de \mathcal{C}_1 forman una cónica de puntos. ■

Ejercicio: Probar el dual del teorema 9.18 aplicando una polaridad. Deducir directamente la segunda parte del teorema anterior.

Todos los teoremas que hemos visto sobre cónicas dan lugar a teoremas duales sin más que intercambiar “cónica de puntos” por “cónica de rectas” y “tangente” por “punto de tangencia”. A su vez, el teorema dual se puede aplicar a la cónica de tangentes de una cónica dada, lo que nos da un tercer enunciado. El paso del primero al tercero se obtiene conservando el concepto de “cónica” pero intercambiado “punto de la cónica” por “tangente a la cónica”.

Por ejemplo, para formular el dual del teorema de Pascal hemos de notar que el dual de “hexágono” es “hexágono” (entendiendo por hexágono la figura formada por seis puntos y las rectas que los unen cíclicamente). En efecto, una polaridad transforma el hexágono de vértices $ABCDEF$ en el hexágono de vértices $p(AB)p(BC)p(CD)p(DE)p(EF)p(FA)$. Así pues:

Teorema 9.34 *Los lados de un hexágono pertenecen a una cónica de rectas si y sólo si las tres rectas que unen vértices opuestos son concurrentes.*

Si aplicamos esto a la cónica de tangentes de una cónica dada obtenemos:

Teorema 9.35 (Brianchon) *Un hexágono está circunscrito a una cónica si y sólo si las tres rectas que unen vértices opuestos son concurrentes.*

Pascal demostró el teorema que lleva su nombre en 1640 (a los 16 años de edad). Lo probó primero para circunferencias y luego generalizó por proyección. Brianchon demostró el teorema anterior en 1806. Su prueba contenía trazas de la teoría de la dualidad.

Vamos a describir geométricamente la polaridad de una cónica. Nuestro punto de partida será la descripción que de ella nos da el teorema 9.18.

Observemos en primer lugar que tres tangentes a una cónica no pueden ser concurrentes. Esto es el dual de que tres puntos de una cónica no son colineales. Por lo tanto, dada una cónica \mathcal{C} , por cada punto Q pasan a lo sumo dos tangentes de \mathcal{C} . Supongamos que así es, es decir, supongamos que A_1 y A_2 son puntos distintos de \mathcal{C} cuyas tangentes pasan por Q . Si p es la polaridad de \mathcal{C} tenemos que $Q \subset p(A_1) \cap p(A_2)$, luego $p(Q) = A_1A_2$.

Así pues, si por un punto dado pasan dos tangentes a \mathcal{C} , la polar del punto es la recta que pasa por los puntos de tangencia.

Veamos ahora que si por Q pasa una única tangente a \mathcal{C} es necesario que Q esté en la cónica, con lo que $p(Q)$ es la tangente. En efecto, supongamos que A es un punto de \mathcal{C} cuya tangente pasa por Q . Si $A \neq Q$ entonces la polar $p(Q)$ es una recta que pasa por A distinta de la tangente $p(A)$, luego corta a \mathcal{C} en un segundo punto B . Como $B \subset p(Q)$, tenemos también $Q \subset p(B)$, es decir, Q está también en la tangente de un segundo punto de \mathcal{C} .

Resumamos en un teorema lo que hemos obtenido hasta ahora:

Teorema 9.36 *Un punto Q está en una cónica \mathcal{C} si y sólo si hay una única tangente a \mathcal{C} que pasa por Q . En caso contrario las tangentes a \mathcal{C} que pasan por Q son dos o ninguna. Si pasan dos tangentes, la recta polar de Q es la que pasa por los puntos de tangencia.*

Las cónicas de rectas admiten un tratamiento algebraico análogo al de las cónicas de puntos. Para verlo hemos de introducir el concepto de coordenadas homogéneas de una recta. Una recta r está formada por los puntos cuyas coordenadas en un sistema de referencia dado cumplen una ecuación de la forma $ax+by+cz=0$. La terna (a,b,c) determina la recta y está determinada por ella salvo un factor constante. Diremos que (a,b,c) son las *coordenadas homogéneas* de r en el sistema de referencia dado.

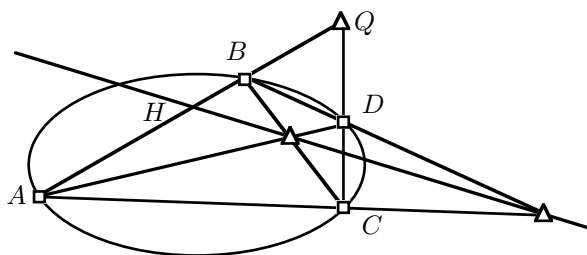
En estos términos, si A es la matriz de una polaridad en un sistema de referencia dado y X son las coordenadas homogéneas de un punto, su polar está formada por los puntos cuyas coordenadas Y cumplen $XY^t = 0$, lo que se interpreta como que XA son las coordenadas homogéneas de la recta polar. Recíprocamente, si Y son las coordenadas homogéneas de una recta, las coordenadas homogéneas de su polo han de cumplir $XA = Y$, luego el polo es YA^{-1} .

Si aplicamos esto a la polaridad de una cónica vemos que una recta de coordenadas homogéneas X será tangente a la cónica si y sólo si las coordenadas XA^{-1} de su polo cumplen la ecuación $XA^{-1}AA^{-1}X^t = 0$, o sea, $XA^{-1}X = 0$. Ésta es la ecuación de la cónica de tangentes asociada a la cónica dada. En general:

Teorema 9.37 *Dado un sistema de referencia, cada cónica de rectas tiene asociada una matriz regular simétrica A de modo que una recta pertenece a la cónica si y sólo si sus coordenadas homogéneas satisfacen la ecuación $XAX^t = 0$. La matriz de una cónica de puntos y la de su cónica de tangentes son mutuamente inversas.*

Pasemos ahora a interpretar geoméricamente la polar de cualquier punto.

Teorema 9.38 *Supongamos que un cuadrilátero completo está inscrito en una cónica \mathcal{C} . Entonces la recta polar de un punto diagonal es la recta que pasa por los otros dos puntos diagonales.*



DEMOSTRACIÓN: Aplicamos el teorema 9.29: Sea Q el punto diagonal y sean AB y CD los lados que pasan por Q . Sean D_1 y D_2 los otros dos puntos

diagonales. Entonces las tangentes a la cónica por A y B se cortan en un punto $X \subset D_1D_2$. Por lo tanto $p(X) = AB$. Como $Q \subset p(X)$, también $X \subset p(Q)$.

De igual modo, las tangentes por C y D se cortan en un punto $Y \subset D_1D_2$ que será distinto de X , o si no tendríamos cuatro tangentes concurrentes. Concluimos igualmente que $Y \subset p(Q)$, luego $p(Q) = XY = D_1D_2$. ■

Dado un punto Q que no esté en una cónica \mathcal{C} , siempre podemos encontrar un punto A en \mathcal{C} tal que la recta AQ sea secante a la cónica (a lo sumo hay que excluir dos puntos para los cuales la recta sea tangente). Si B es el otro punto donde AQ corta a \mathcal{C} , excluyendo A , B y los dos posibles puntos de tangencia, siempre podemos encontrar un tercer punto C tal que la recta QC corte a \mathcal{C} en un cuarto punto D , y así formamos un cuadrilátero completo cuyos vértices están en \mathcal{C} y que tiene a Q como punto diagonal. Esto nos permite aplicar el teorema anterior para trazar la polar de cualquier punto. Más aún, si observamos el punto H de la figura anterior concluimos inmediatamente lo siguiente:

Teorema 9.39 *Sea \mathcal{C} una cónica, Q un punto y consideremos una recta que pase por Q y corte a \mathcal{C} en dos puntos A y B distintos de Q . Entonces la recta polar de Q contiene al conjugado armónico de Q respecto de A y B .*

Ejercicio: Explicar cómo se construye el polo de una recta dada respecto a una cónica dada (dualizar la construcción de la recta polar).

Ejercicio: Un triángulo se dice *autopolar* si y sólo si cada vértice es el polo de su lado opuesto. Probar que un triángulo es autopolar respecto a una cónica si y sólo si es el triángulo diagonal de un cuadrilátero inscrito en la misma.

Los teoremas que hemos visto hasta aquí son también muy útiles en su forma dual. Dejamos a cargo del lector escribir los enunciados duales de cada uno de ellos. En lo sucesivo los usaremos cuando convenga. Definimos, no obstante, el concepto dual de un cuadrilátero completo, que aparece en muchos de estos enunciados.

Definición 9.40 Un *cuadrilátero completo dual* es una figura formada por cuatro rectas no concurrentes tres a tres, llamadas *lados*, junto con los seis puntos donde se cortan, llamados *vértices*. Dos vértices son *contiguos* si comparten un lado y *opuestos* en caso contrario. Las tres rectas que unen los pares de vértices opuestos se llaman *rectas diagonales* del cuadrilátero dual.

Veamos ahora un resultado cuya importancia se pondrá de manifiesto más adelante:

Teorema 9.41 *Sea \mathcal{C} una cónica y r una recta que no sea tangente a \mathcal{C} . La aplicación $f : r \rightarrow r$ dada por $f(Q) = p(Q) \cap r$ es una involución en r .*

DEMOSTRACIÓN: Notemos en primer lugar que f es biyectiva, pues la polaridad p biyecta los puntos de r con las rectas que pasan por su polo. Ninguna de éstas es r , pues si r pasara por su polo sería tangente a \mathcal{C} , luego cada una corta a r en un único punto distinto en cada caso.

Fijemos un sistema de referencia y sea A la matriz p . Sean Q_1 y Q_2 las coordenadas homogéneas de dos puntos A y B de r y sean R_1, R_2 las coordenadas de sus imágenes por f , de modo que $Q_1 AR_1^t = Q_2 AR_2^t = 0$.

Claramente $Q_1 + Q_2$ son las coordenadas homogéneas de un punto de r , cuya imagen por f será un punto de coordenadas $\alpha R_1 + \beta R_2$, pero como podemos cambiar cada R_i por un múltiplo sin variar el punto que representan, podemos suponer que de hecho las coordenadas de la imagen son $R_1 + R_2$, y por consiguiente $(Q_1 + Q_2)(R_1 + R_2) = 0$, de donde $Q_1 R_2 + Q_2 R_1 = 0$.

Tomemos ahora cualquier punto $C \in r$, cuyas coordenadas serán de la forma $Q = \alpha Q_1 + \beta Q_2$. Entonces $p(C)$ es la recta de ecuación $(\alpha Q_1 + \beta Q_2)AX^t = 0$.

El punto $f(C)$ será el único de r que cumple esta ecuación, pero es inmediato comprobar que esto le sucede a $\alpha R_1 + \beta R_2$.

Esto prueba que f es la homografía que transforma los puntos de coordenadas $Q_1, Q_2, Q_1 + Q_2$ en los puntos de coordenadas $R_1, R_2, R_1 + R_2$. Es inmediato comprobar que se trata de una involución. ■

Definición 9.42 Diremos que dos puntos Q y R son *conjugados* respecto de una cónica \mathcal{C} si uno está contenido en la recta polar del otro. Dos rectas son *conjugadas* si una contiene al polo de la otra.

Acabamos de probar que si r es una recta no tangente a una cónica \mathcal{C} , entonces los pares de conjugados en r respecto a \mathcal{C} son pares de conjugados respecto a una involución que llamaremos *involución inducida* por \mathcal{C} en r .

Ejercicio: Sea \mathcal{C} una cónica y O un punto no contenido en ella. Probar que la aplicación de \mathcal{H}_O en sí mismo que a cada recta le asigna su única conjugada en el haz es una involución, a la que se llama *involución inducida* por la cónica en el haz.

9.6 Homografías entre cónicas

Definición 9.43 Dadas dos cónicas \mathcal{C}_1 y \mathcal{C}_2 en dos planos proyectivos, una *homografía* entre ellas es una aplicación $H : \mathcal{C}_1 \rightarrow \mathcal{C}_2$ inducida por restricción a partir de una homografía entre los planos correspondientes.

Puesto que una cónica contiene un sistema de referencia proyectivo, es claro que cada homografía entre cónicas determina la homografía entre los planos que la induce, y a su vez ésta determina una biyección entre las tangentes de ambas cónicas. En la práctica no distinguiremos entre una homografía entre dos cónicas y la homografía (entre planos) que la induce. Vamos a probar que siempre existen homografías entre dos cónicas dadas:

Teorema 9.44 Sean \mathcal{C}_1 y \mathcal{C}_2 dos cónicas no necesariamente en el mismo plano, sean A, B, C tres puntos de \mathcal{C}_1 y A', B', C' tres puntos de \mathcal{C}_2 . Entonces existe una única homografía entre las cónicas tal que $ABC \xrightarrow{H} A'B'C'$.

DEMOSTRACIÓN: Sea P el punto donde se cortan las tangentes a \mathcal{C}_1 por A y B , sea P' el punto donde se cortan las tangentes a \mathcal{C}_2 por A' y B' . Es claro que A, B, C, P forman un sistema de referencia proyectivo del plano de \mathcal{C}_1 , y A', B', C', P' forman un sistema de referencia del plano de \mathcal{C}_2 . La homografía $ABCP \xrightarrow{\wedge} A'B'C'P'$ transforma \mathcal{C}_1 en una cónica que contiene a los puntos A', B' y C' y cuyas tangentes en A' y B' son las rectas $A'P'$ y $B'P'$, es decir, esta cónica y \mathcal{C}_2 tienen tres puntos y dos tangentes en común, luego son iguales (esto se sigue fácilmente de 9.29).

Además, cualquier homografía entre las cónicas que cumpla $ABC \xrightarrow{\wedge} A'B'C'$ ha de transformar las tangentes de la primera en las de la segunda, luego P en P' , luego es necesariamente la inducida por $ABCP \xrightarrow{\wedge} A'B'C'P'$. ■

Es claro que el conjunto de todas las homografías de una cónica \mathcal{C} en sí misma forma un grupo, al que llamaremos $\text{GP}(\mathcal{C})$ y que podemos identificar con un subgrupo de $\text{GP}(X)$, donde X es el plano de \mathcal{C} .

Si \mathcal{H}_O es el haz de rectas que pasan por un punto O , las homografías de \mathcal{H}_O forman también un grupo, al que llamaremos $\text{GP}(\mathcal{H}_O)$. Sabemos que si r es una recta que no pasa por O , cada homografía $H \in \text{GP}(\mathcal{H}_O)$ se corresponde con una homografía $\bar{H}|_r \in \text{GP}(r)$ dada por $\bar{H}|_r(P) = \bar{H}(OP) \cap r$. Además es fácil ver que esta correspondencia es un isomorfismo de grupos, luego concluimos que $\text{GP}(\mathcal{H}_O) \cong \text{GP}(r)$.

Ahora probaremos que si O es un punto de una cónica \mathcal{C} , entonces

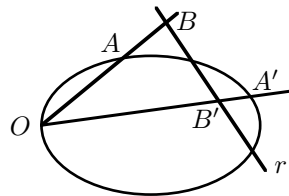
$$\text{GP}(\mathcal{C}) \cong \text{GP}(\mathcal{H}_O) \cong \text{GP}(r).$$

En efecto, para cada homografía $H \in \text{GP}(\mathcal{C})$ definimos $H^* \in \text{GP}(\mathcal{H}_O)$ mediante $H^*(OP) = OH(P)$, donde $P \in \mathcal{C}$, entendiendo, como es habitual, que OO es la tangente en O .

Veamos que H^* es realmente una homografía. Si $H(O) = O$ se trata simplemente de la homografía inducida por H (vista como homografía del plano). Si por el contrario $H(O) = O' \neq O$, entonces H^* es la composición de $\bar{H} : \mathcal{H}_O \rightarrow \mathcal{H}_{O'}$ con la homografía $\mathcal{H}_{O'} \rightarrow \mathcal{H}_O$ inducida por la cónica (teorema 9.26).

La correspondencia es obviamente inyectiva y también es suprayectiva pues, dada $\bar{H} \in \text{GP}(\mathcal{H}_O)$, consideramos tres puntos A, B, C en \mathcal{C} y los puntos A', B', C' tales que $H(OA) = OA'$, etc. La homografía $ABC \xrightarrow{\wedge} A'B'C'$ induce una homografía en $\text{GP}(\mathcal{H}_O)$ que coincide con la dada sobre las tres rectas OA, OB, OC , luego es la dada. ■

Conviene recordar la forma en que actúa la composición de los isomorfismos anteriores entre $\text{GP}(\mathcal{C})$ y $\text{GP}(r)$. Fijado un punto O en \mathcal{C} que no esté en r , si una homografía de \mathcal{C} hace corresponder los puntos A y A' , entonces su imagen en r hace corresponder los puntos B y B' .



En realidad la figura sugiere una forma de biyectar los puntos de una cónica con los de una recta. En efecto, la aplicación $f : \mathcal{C} \rightarrow r$ dada por $f(P) = OP \cap r$

es obviamente biyectiva, y el isomorfismo $\text{GP}(\mathcal{C}) \cong \text{GP}(r)$ es simplemente el dado por $H \mapsto H^f = f^{-1}Hf$.

Ahora podemos traducir a homografías sobre cónicas todos los resultados que conocemos sobre homografías sobre rectas. Por ejemplo, si una homografía en una cónica intercambia dos puntos A y B , podemos asegurar que se trata de una involución. En realidad es más interesante la dirección contraria: representar una homografía de una recta como homografía sobre una cónica hace mucho más fácil su estudio, como pronto veremos. Antes vamos a usar las biyecciones entre cónicas y rectas para introducir coordenadas homogéneas en una cónica.

Sea \mathcal{C} una cónica y fijemos en ella tres puntos P_0, P_1 y P_∞ . Tomemos una recta r que no pase por P_∞ y determinemos en ella los puntos Q_0, Q_1 y Q_∞ como las imágenes de los de \mathcal{C} por la biyección f anterior. La recta afín $r \setminus Q_\infty$ tiene estructura de cuerpo con Q_0 y Q_1 como neutros de la suma y el producto. La biyección f traslada esta estructura de cuerpo a $\mathcal{C} \setminus P_\infty$. Vamos a ver que el resultado no depende de la elección de r , sino tan sólo de los tres puntos tomados como sistema de referencia.

En efecto, dados dos puntos x e y en $\mathcal{C} \setminus P_\infty$, consideremos la homografía H dada por $P_\infty xy \frown P_\infty yx$. Se trata de una involución² que f transporta a $Q_\infty f(x)f(y) \frown Q_\infty f(y)f(x)$. Existe una única involución en r con estas propiedades, luego necesariamente es la que tiene expresión en coordenadas

$$I(\alpha) = u + v - \alpha,$$

donde u y v son las coordenadas de $f(x)$ y $f(y)$ respectivamente (pues efectivamente I permuta u y v y deja fijo a ∞). Luego $I(0) = u + v$.

Así pues, $H(P_0) = x + y$. En resumen, fijados P_0, P_1 y P_∞ , la suma de dos puntos de $\mathcal{C} \setminus P_\infty$ se calcula como $H(P_0)$, donde H es la involución $P_\infty xy \frown P_\infty yx$.

Del mismo modo se comprueba que xy se calcula como $H(P_1)$, donde H es la involución $P_\infty xy \frown P_0 yx$, pues f la transforma necesariamente en la involución cuya expresión en coordenadas es $I(\alpha) = uv/\alpha$, donde, como antes, u y v son las coordenadas de $f(x)$ y $f(y)$.

Con todo esto es inmediato el teorema siguiente:

Teorema 9.45 *Sea \mathcal{C} una cónica en un plano proyectivo X sobre un cuerpo K . Sean P_0, P_1 y P_∞ tres puntos en \mathcal{C} . Definimos en $\mathcal{C} \setminus P_\infty$ las operaciones dadas por*

$$P_\infty xy P_0 \frown P_\infty yx (x + y), \quad P_\infty xy P_1 \frown P_0 yx (xy).$$

Entonces $\mathcal{C} \setminus P_\infty$ tiene estructura de cuerpo isomorfo a K .

Aún podemos decir más. En las condiciones de la discusión previa al teorema anterior, hemos visto que el sistema de referencia P_0, P_1, P_∞ en \mathcal{C} determina el sistema de referencia Q_0, Q_1, Q_∞ en la recta r , luego podemos definir las *coordenadas homogéneas* de un punto $P \in \mathcal{C}$ como las coordenadas de $f(P)$.

²Notar que no excluimos la posibilidad $x = y$ (ver 8.54).

Esto no depende de r , pues si tomamos otra recta r' (siempre con la condición de que no pase por P_∞), entonces tenemos dos biyecciones f_r y $f_{r'}$, y es claro que $f_r \circ f_{r'}^{-1}$ es la proyección perspectiva de centro P_∞ , luego es una homografía, y transforma el sistema de referencia Q_0, Q_1, Q_∞ de r inducido por el de \mathbb{C} en el sistema de referencia en r' inducido por el de \mathbb{C} . Por consiguiente $f_r(P)$ y $f_{r'}(P)$ tienen las mismas coordenadas homogéneas en los sistemas de referencia respectivos.

Si representamos las coordenadas homogéneas de un punto P de una cónica \mathbb{C} como $(1, x)$, con $x \in K \cup \{\infty\}$, la correspondencia $P \mapsto x$ es un isomorfismo entre $\mathbb{C} \setminus P_\infty$ y K . En principio puede haber varios isomorfismos, pero sólo uno corresponde a una asignación de coordenadas.

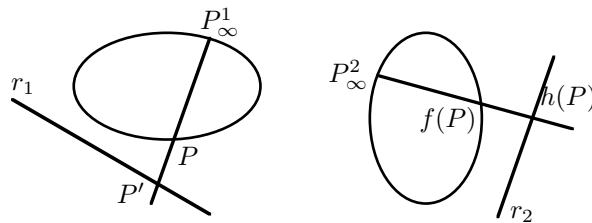
Ahora podemos caracterizar las homografías entre cónicas en términos de coordenadas homogéneas:

Teorema 9.46 *Una biyección $f : \mathbb{C}_1 \longrightarrow \mathbb{C}_2$ entre dos cónicas es una homografía si y sólo si al fijar un sistema de referencia en cada cónica la coordenada $y \in K \cup \{\infty\}$ de la imagen de un punto de coordenada x viene dada por una relación de la forma*

$$y = \frac{ax + b}{cx + d}, \quad \text{donde } ad - bc \neq 0.$$

DEMOSTRACIÓN: Si la aplicación f es una homografía consideramos la aplicación $g : \mathcal{H}_{P_\infty^1} \longrightarrow \mathcal{H}_{P_\infty^2}$ dada por $g(P_\infty^1 Q) = P_\infty^2 f(Q)$. Se trata de una homografía, pues si $f(P_\infty^1) = P_\infty^2$ la aplicación g es simplemente la homografía inducida por f , mientras que si $f(P_\infty^1) = A \neq P_\infty^2$ entonces g es la composición de $\tilde{f} : \mathcal{H}_{P_\infty^1} \longrightarrow \mathcal{H}_A$ con la homografía $\mathcal{H}_A \longrightarrow \mathcal{H}_{P_\infty^2}$ inducida por \mathbb{C}_2 .

Tomamos una recta r_1 que no pase por P_∞^1 y una recta r_2 que no pase por P_∞^2 y consideramos los sistemas de referencia inducidos desde las cónicas así como las aplicaciones f_{r_1} y f_{r_2} . Sabemos que la homografía g induce una homografía $h : r_1 \longrightarrow r_2$. La figura muestra la situación:



Si $P \in \mathbb{C}_1$, entonces $f_{r_1}(P)$ es el punto $P' \in r_1$ colineal con P y P_∞^1 , la homografía g transforma $P_\infty^1 P$ en $P_\infty^2 f(P)$ y entonces h transforma $P' = f_{r_1}(P)$ en $P_\infty^2 f(P) \cap r_2$, o sea, $h(f_{r_1}(P)) = f_{r_2}(f(P))$.

Por definición, las coordenadas homogéneas de P son las de P' y las de $f(P)$ son las de $h(P')$. Como h es una homografía la relación entre ellas es la indicada en el enunciado.

Recíprocamente, si f satisface una ecuación como la del enunciado, ésta determina una homografía h entre las rectas r_1 y r_2 , con la cual podemos construir una homografía entre las cónicas invirtiendo el proceso anterior. La homografía que obtenemos verifica la misma relación que f en coordenadas, luego ha de ser la propia f . ■

Los resultados que hemos obtenido se expresan más simplemente si definimos una *homografía* entre una recta y una cónica (o al revés) como una biyección entre ambas cuya expresión en coordenadas sea del tipo

$$y = \frac{ax + b}{cx + d}, \quad \text{donde } ad - bc \neq 0.$$

Es inmediato comprobar que si una tal biyección tiene una expresión en coordenadas de este tipo respecto a un par concreto de sistemas de referencia, entonces tiene una expresión similar respecto a cualquier otro par. También es claro que la composición de cualquier tipo de homografías (recta/cónica, cónica/recta, recta/recta, cónica/cónica) es de nuevo una homografía, así como que la inversa de una homografía es una homografía.

La aplicación que a cada punto de una cónica le asigna sus coordenadas homogéneas respecto a un sistema de referencia dado es una homografía entre la cónica y $P^1(K)$, si O es un punto de una cónica \mathcal{C} y r es una recta que no pasa por O , la aplicación $f : \mathcal{C} \rightarrow r$ dada por $f(P) = OP \cap r$ es una homografía, que podemos llamar *proyección perspectiva* de centro O .

Ejercicio: Probar que dados tres puntos en una cónica y otros tres en una recta existe una única homografía entre ambas que los hace corresponder en un cierto orden.

El teorema siguiente contiene en esencia algo que afirmamos sin prueba al definir las cónicas, y es que en un espacio euclídeo toda cónica puede obtenerse mediante una sola proyección perspectiva a partir de un círculo.

Teorema 9.47 Sean \mathcal{C}_1 y \mathcal{C}_2 dos cónicas en planos distintos π_1 y π_2 de modo que la recta $r = \pi_1 \cap \pi_2$ es tangente a ambas por un mismo punto A . Entonces existe una proyección perspectiva que transforma \mathcal{C}_1 en \mathcal{C}_2 .

DEMOSTRACIÓN: Sean B y C dos puntos más en \mathcal{C}_1 , sean P y Q los puntos donde las tangentes a \mathcal{C}_1 por B y C cortan a r . Como P y Q están sobre una tangente a \mathcal{C}_2 , existen puntos B' y C' en \mathcal{C}_2 distintos de A tales que sus tangentes pasan por P y Q . Las rectas BB' y CC' se cortan en un punto O fuera de π_1 y π_2 . Entonces $ABCPQ \xrightarrow{O} AB'C'P'Q'$, luego la proyección perspectiva desde O transforma \mathcal{C}_1 en una cónica que comparte con \mathcal{C}_2 los puntos A, B', C' y sus tangentes por todos ellos, luego ha de ser \mathcal{C}_2 . ■

El resto de la sección lo dedicamos a estudiar las homografías de una cónica. Su comportamiento es especialmente simple.

Teorema 9.48 Si $A \neq A', B, B'$ son dos pares de puntos homólogos por una homografía en una cónica, entonces $AB' \cap A'B$ está sobre una recta r independiente de la elección de los puntos A y B .

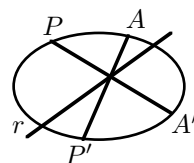
DEMOSTRACIÓN: Hemos visto que cada homografía en una cónica $[P] \xrightarrow{\sim} [P']$ induce una homografía en \mathcal{H}_A dada por $A[P] \xrightarrow{\sim} A[P']$. Si componemos ésta con la homografía del teorema de Steiner, es decir, $A'[P] \xrightarrow{\sim} A[P]$, obtenemos la homografía $A'[P] \xrightarrow{\sim} A[P']$ entre los haces $\mathcal{H}_{A'}$ y \mathcal{H}_A . Esta homografía fija a la recta AA' , luego es una proyección perspectiva, es decir, está inducida por una perspectividad H . Sea r su eje de perspectiva. Entonces cada par de rectas homólogas AB' y $B'A$ se cortan en r (para todo punto B). Falta probar que r es independiente de A . Ahora bien, si tomamos tres puntos distintos A, B, C con imágenes distintas de ellos tres A', B', C' , los puntos $AB' \cap A'B, A'C \cap C'A, BC' \cap C'B$ son colineales, por el teorema de Pascal aplicado al hexágono $AB'CA'BC'$. Si llamamos r_0 a la recta que los contiene, vemos que r_0 es la recta r determinada por A y también la determinada por B . Esto prueba que si A y B tienen imágenes distintas de ellos mismos (es decir, salvo si $A' = B$ o $B' = A$) ambos determinan la misma recta. Si $A' = B$ tomamos un tercer punto C tal que C y C' sean distintos de A, A', B, B' y concluimos que A , y B determinan ambos la misma recta que C . ■

Por dualidad tenemos también el teorema siguiente:

Teorema 9.49 Si $r \neq r', s \neq s'$ son dos pares de tangentes homólogas por una homografía en una cónica, entonces la recta que pasa por $r \cap s'$ y $r' \cap s$ es colineal con un punto O independiente de la elección de las rectas r y s .

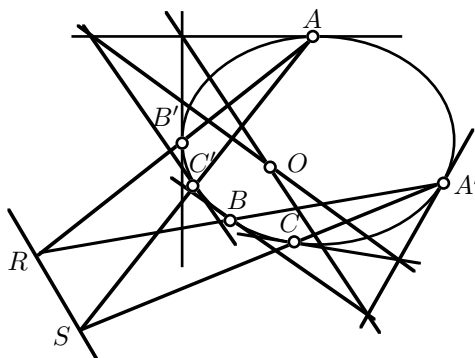
Definición 9.50 La recta r y el punto O determinados por los dos teoremas anteriores reciben el nombre de *eje* y *centro* de la homografía considerada.

Una homografía en una cónica (distinta de la identidad) queda completamente determinada en cuanto conocemos su eje y un par de puntos homólogos $A \neq A'$. En efecto, para encontrar el homólogo de otro punto P no tenemos más que unir A con $PA' \cap r$. La imagen P' será el punto de intersección de esta recta con la cónica.



La construcción dual nos permite trazar la imagen de cualquier tangente a partir de dos tangentes homólogas y del centro de la homografía. Estas construcciones muestran también que los puntos fijos de una homografía en una cónica \mathcal{C} son precisamente las intersecciones de su eje con \mathcal{C} .

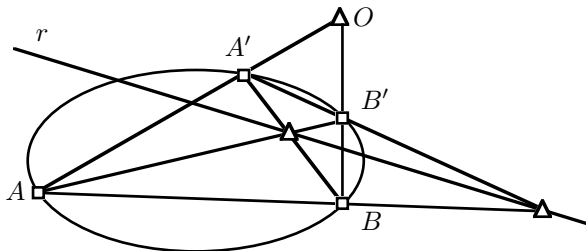
Teorema 9.51 El centro de una homografía en una cónica es el polo de su eje.



DEMOSTRACIÓN: Consideremos tres pares de puntos homólogos $A \neq A'$, $B \neq B'$, $C \neq C'$. Sean $R = A'B \cap B'A$ y $S = A'C \cap C'A$. De este modo RS es el eje de la homografía. Las tangentes en A y B' se cortan en el polo de AR , y las tangentes en A' y B se cortan en el polo de $A'R$, luego la recta polar de R es la que une ambos puntos de intersección. Similarmente se calcula la polar de S y la intersección O entre ambas es el polo de RS . Pero, por definición, el punto O así calculado es el centro de la homografía. ■

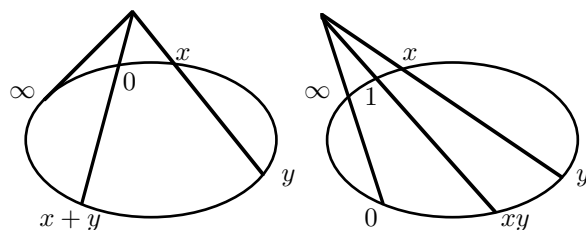
Hemos visto cómo una homografía en una cónica puede calcularse fácilmente a partir de su eje o de su centro. La situación es aún más simple en el caso de las involuciones.

Teorema 9.52 Si A y A' son conjugados por una involución en una cónica, entonces la recta AA' pasa por su centro.



DEMOSTRACIÓN: Sean $A \neq A'$ y B, B' dos pares de puntos conjugados por la involución. Entonces $AB' \cap BA'$ está en el eje, y es claro que su imagen por la homografía que genera la involución es él mismo. Todo punto del eje puede expresarse de este modo para un B adecuado, luego la homografía fija a cada punto de su eje y es, por consiguiente, una perspectividad. Basta probar que su centro de perspectiva coincide con el centro de la involución. Por el teorema anterior basta probar que el centro de perspectiva O es el polo del eje de perspectiva r . Si tomamos dos pares de conjugados $A \neq A'$, $B \neq B'$, entonces los puntos $AB' \cap A'B$ y $AB \cap A'B'$ están en r (notar que el segundo es $A''B \cap A'B'$). Ahora basta aplicar el teorema 9.38. ■

Con esto podemos construir muy fácilmente las operaciones en una cónica, pues éstas se definen en términos de involuciones. Por ejemplo, para calcular una suma $x + y$ trazamos la tangente por ∞ y la cortamos con xy , lo que nos da el centro de la involución requerida. Al unirlo con 0 obtenemos $x + y$. Similarmente con el producto:



Ejercicio: Describir las construcciones para la resta, la división y la raíz cuadrada en una cónica. Interpretar geoméricamente el hecho de que (cuando el cuerpo es \mathbb{R}) los números positivos tienen raíz cuadrada y los negativos no.

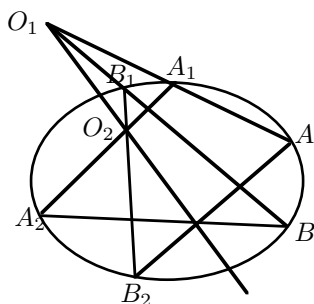
Ejercicio: Describir el grupo aditivo de una parábola cuando se toma como P_∞ su punto infinito. Describir el grupo multiplicativo de una hipérbola cuando se toma como P_0 y P_∞ sus puntos infinitos.

Es claro que todo punto O que no esté contenido en una cónica \mathcal{C} es el centro de una única involución en \mathcal{C} . Basta considerar dos rectas secantes a \mathcal{C} que pasen por O . Si éstas cortan a la cónica en A, A' y B, B' , la involución $AA'BB' \bar{\wedge} A'AB'B$ tiene claramente centro O .

En el capítulo anterior vimos que toda homografía en una recta (y por lo tanto en una cónica) es producto de dos involuciones. El teorema siguiente nos da la relación entre los centros y los ejes.

Teorema 9.53 *El eje de la composición de dos involuciones de una cónica es la recta que une sus centros y el centro de la intersección de sus ejes.*

DEMOSTRACIÓN: Tomemos dos puntos A y B de la cónica que no sean puntos fijos o conjugados de ninguna de las involuciones ni de su composición. Entonces la primera involución los transforma en otros dos puntos A_1, B_1 , y la segunda transforma éstos en otros dos puntos A_2, B_2 , de modo que los seis puntos son distintos. Si O_1 y O_2 son los centros de las involuciones, por el teorema 9.52 sabemos que $O_1 = AA_1 \cap BB_1$ y $O_2 = A_1A_2 \cap B_1B_2$. Al aplicar el teorema de Pascal al hexágono $AA_1A_2BB_1B_2$ concluimos que O_1O_2 contiene a $AB_2 \cap BA_2$, que por definición es un punto del eje de la composición. Por lo tanto dicho eje ha de ser O_1O_2 . Aplicando la polaridad de la cónica obtenemos que el centro es la intersección de los ejes. ■



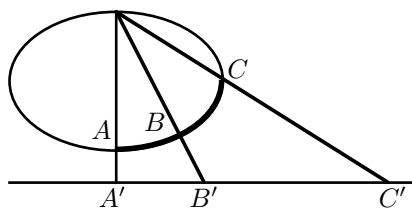
9.7 Cónicas sobre cuerpos ordenados

Nos ocupamos ahora de las propiedades adicionales que cumplen las cónicas cuando el plano que las contiene está construido sobre un cuerpo ordenado. Aunque una gran parte de la teoría es válida para un cuerpo ordenado arbitrario K , las cosas se simplifican si suponemos que todos los elementos positivos de K tienen raíz cuadrada, y así lo haremos. Esto nos permite contar con el teorema 8.63.

Todos los resultados que ya tenemos sobre ordenación de rectas se transmiten automáticamente a las cónicas a través de las homografías (y de hecho resultan más intuitivos). Por ejemplo, si seleccionamos un sistema de referencia P_0, P_1, P_∞ en una cónica \mathcal{C} , cada punto de la cónica distinto de P_∞ admite unas coordenadas homogéneas $(x, 1)$ o, equivalentemente, una coordenada cartesiana $x \in K$, de modo que el orden de K se transmite a una ordenación de $\mathcal{C} \setminus P_\infty$. Si cambiamos P_0 o P_1 la ordenación resultante es la misma o la inversa (pues las coordenadas cartesianas en dos sistemas de referencia están relacionados por una ecuación de la forma $y = ax + b$, con $a \neq 0$, que conserva o invierte el orden según el signo de a).

Si r es una recta y $H : \mathcal{C} \rightarrow r$ es una homografía, podemos definir el arco \overline{ABC} en \mathcal{C} como la antiimagen por H del segmento $\overline{H(A)H(B)H(C)}$. Esto no depende de la elección de la recta o la homografía, pues si $H' : \mathcal{C} \rightarrow r'$ es otra elección, entonces la aplicación $H'^{-1} \circ H : r' \rightarrow r$ es una homografía, luego conserva los segmentos, es decir, transforma $\overline{H'(A)H'(B)H'(C)}$ en $\overline{H(A)H(B)H(C)}$, luego la antiimagen del primero por H' coincide con la antiimagen del segundo por H .

Si tomamos como homografía una proyección perspectiva nos convenceremos de que la ordenación y los arcos en una cónica son lo que intuitivamente cabe esperar:



Todas las propiedades sobre segmentos que hemos probado para rectas son válidas para arcos en cónicas y las pruebas son todas inmediatas. La definición de separación de pares de puntos vale ahora para pares de puntos sobre cónicas, y todas las propiedades se demuestran igualmente (o se deducen de las ya probadas a través de homografías).

Introducimos ahora un concepto genuinamente bidimensional. Vamos a distinguir los puntos interiores y exteriores de una cónica. Hay varias formas equivalentes de definir estos conceptos. Tomaremos una bastante práctica y demostraremos su equivalencia con algunas otras.

Definición 9.54 Sea \mathcal{C} una cónica y sea A un punto no contenido en ella. Diremos que A es un punto *interior* (resp. *exterior*) a \mathcal{C} si por A no pasa ninguna tangente (resp. pasan dos tangentes) a \mathcal{C} .

De este modo, todo punto del plano de una cónica \mathcal{C} es interior a \mathcal{C} , está en \mathcal{C} o es exterior a \mathcal{C} según si el número de tangentes que pasan por él es 0, 1 o 2.

Hay una caracterización más abstracta pero muy útil de estos conceptos: dado un punto O que no esté contenido en una cónica \mathcal{C} , un punto fijo para la involución de centro O es un punto cuya tangente pasa por O , luego la involución

será elíptica o parabólica según que O sea un punto interior o exterior. El teorema 8.63 implica ahora:

Teorema 9.55 Sean A, B, C, D cuatro puntos distintos en una cónica. Entonces se cumple $AB \parallel CD$ si y sólo si $AB \cap CD$ es un punto exterior, luego $AB \nparallel CD$ si y sólo si $AB \cap CD$ es interior.

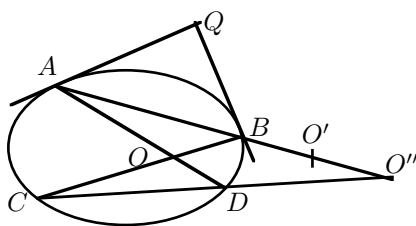
Es obvio que los puntos de una recta tangente (distintos del de tangencia) son todos exteriores, pues si pasa por un punto que no está en una cónica pasa una tangente a la misma, de hecho pasan dos. El teorema siguiente describe la distribución entre puntos interiores y exteriores en una recta secante:

Teorema 9.56 Sea \mathcal{C} una cónica y r una recta que corte a \mathcal{C} en dos puntos A y B . Entonces uno de los segmentos en r de extremos A y B está formado por puntos interiores (aparte de los extremos) y el segmento complementario está formado por puntos exteriores.

DEMOSTRACIÓN: Sea C un punto en uno de los segmentos de \mathcal{C} de extremos A y B . Consideremos la proyección perspectiva de \mathcal{C} en r de centro C . Es claro que fija A y B . Si P es cualquier punto del arco \overline{ACB} entonces tenemos por definición que $AB \nparallel CP$, luego el punto $AB \cap CP$ es exterior, pero éste es la proyección perspectiva de P desde C , es decir, la imagen del arco \overline{ACB} es uno de los segmentos de r de extremos A y B y está formado por puntos exteriores. Análogamente, si P es un punto del arco complementario de \overline{ACB} entonces $AC \parallel CP$, luego la proyección $AB \cap CP$ es interior, con lo que todos los puntos del segmento imagen en r son interiores. ■

Para completar esta descripción tendríamos que probar que si una recta no corta a una cónica entonces todos sus puntos son exteriores, pero para ello necesitamos algunas consideraciones previas. Deduciremos esto y varias cosas más a partir de una misma construcción que pasamos a describir:

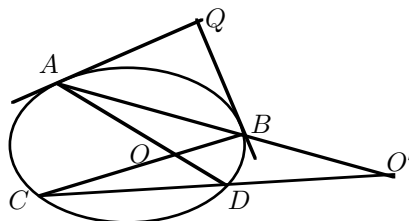
Partamos de dos puntos O y O' conjugados respecto a una cónica \mathcal{C} , es decir, tales que uno esté contenido en la polar del otro. Supongamos que O es interior. Probaremos, entre otras cosas, que O' es exterior.



En primer lugar tomamos una recta que pase por O' y que corte a \mathcal{C} en dos puntos A y B . Podemos suponer que AB no contiene al punto O (aunque no excluimos de momento que O' pueda estar en \mathcal{C} y ser, por tanto, A o B).

Como O es interior, las rectas AO y BO no pueden ser tangentes a \mathcal{C} , luego cortan a la cónica en dos puntos C y D distintos de los anteriores. La recta CD cortará a AB en un punto O'' . Vamos a probar que $O'' = O'$.

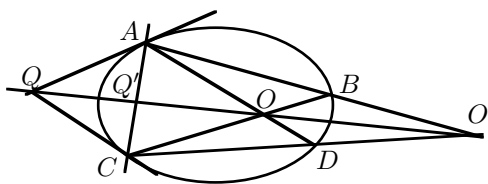
Sea Q el punto de intersección de las tangentes por A y B . Entonces tenemos que $p(Q) = O'O''$, luego Q está contenido en $p(O')$ y en $p(O'')$. Por el teorema 9.38 sabemos que $O \subset p(O'')$, y por hipótesis $O \subset p(O')$, luego en definitiva $p(O') = p(O'') = OQ$, de donde $O' = O''$ y la figura que tenemos es en realidad:



Recordemos que podemos llegar a esta situación siempre que O sea interior a \mathcal{C} y $O' \subset p(O)$. Aplicando dos veces el teorema 9.55 vemos que $AC \parallel BD$, porque O es interior, y de aquí que O' es exterior. Con esto tenemos demostrado lo siguiente:

Teorema 9.57 *La recta polar de un punto interior de una cónica contiene únicamente puntos exteriores (en particular no corta a la cónica). Equivalentemente, si una recta pasa por un punto interior de una cónica entonces su polo es un punto exterior.*

Pero en realidad tenemos más. Tomemos una recta r que pase por un punto interior O de \mathcal{C} (con lo que en particular no es tangente a \mathcal{C}). Sabemos que la aplicación que a cada punto de r le asigna su único conjugado en r (es decir, la intersección de su polar con r) es una involución. Sea O' el conjugado de O . Entonces tenemos la situación anterior. Ahora llamamos Q a la intersección de las tangentes por A y C , que se encuentra sobre la recta OO' , ya que esto equivale a que $p(Q)$ contenga a $p(OO')$ y, efectivamente, $p(Q) = AC$ y el polo de OO' es el tercer punto diagonal del cuadrilátero completo de vértices A, B, C, D . Sea $Q' = AC \cap r$. Entonces la involución que \mathcal{C} determina en r es la dada por $OO'QQ' \bar{\wedge} O'OQ'Q$.



Ahora bien, es claro que $DBAC \stackrel{A}{\bar{\wedge}} OO'QQ'$ y, como O es interior, $AD \parallel BC$, luego $DB \parallel AC$, luego $OO' \parallel QQ'$. El teorema 8.63 nos permite concluir que la involución en r es hiperbólica, lo que obviamente se traduce en que \mathcal{C} corta a r en dos puntos distintos. El recíproco es evidente, y así tenemos otra caracterización del concepto de punto interior:

Teorema 9.58 *Un punto O es interior a una cónica \mathcal{C} si y sólo si toda recta que pasa por O corta a \mathcal{C} en dos puntos distintos.*

En particular una recta que no corte ni toque a una cónica ha de estar formada completamente por puntos exteriores, tal y como queríamos probar. Más aún, por un punto exterior O siempre pasa una recta sin puntos en común con la cónica. Basta tener en cuenta que $p(O)$ es secante a la cónica (hay dos puntos en la cónica cuyas tangentes pasan por O y entonces $p(O)$ pasa por los puntos de tangencia). Por consiguiente $p(O)$ contiene un punto interior Q , luego $O \subset p(Q)$ y ya hemos visto que $p(Q)$ está formada por puntos exteriores. ■

Resumimos todo esto en el teorema siguiente:

Teorema 9.59 *Sea \mathcal{C} una cónica. Un punto P es interior a \mathcal{C} si y sólo si todas las rectas que pasan por P son secantes a \mathcal{C} , el punto P está en \mathcal{C} si y sólo si por él pasa una única tangente a \mathcal{C} , y P es exterior si y sólo si por P pasa una recta exterior (es decir, disjunta) a \mathcal{C} .*

También tenemos probado el teorema siguiente sobre polares:

Teorema 9.60 *Las polares de los puntos interiores de una cónica son las rectas exteriores, las polares de los puntos exteriores son las rectas secantes, y las polares de los puntos de la cónica son las tangentes.*

Ejercicio: Dados dos puntos A y B ambos interiores (o ambos exteriores) a una cónica, existe un segmento de extremos A y B formado completamente por puntos interiores (o exteriores).

Ejercicio: El centro de una elipse es un punto interior, mientras que el de una hipérbola es exterior.

Terminamos con unos teoremas que ilustran una técnica de trabajo con homografías sobre rectas: la de representarlas sobre cónicas.

Teorema 9.61 *En una recta real, dos involuciones distintas al menos una de las cuales sea elíptica tienen un único par común de puntos conjugados.*

DEMOSTRACIÓN: Es equivalente probar el teorema para involuciones sobre rectas que para involuciones sobre cónicas, pero en este último caso el resultado es inmediato: una involución elíptica tiene por centro un punto interior, luego la recta que une los dos centros corta a la cónica en dos puntos, que obviamente serán conjugados para ambas involuciones. La unicidad es clara. ■

Conviene destacar el siguiente caso particular:

Teorema 9.62 *Una involución elíptica en una recta real tiene un único par de puntos conjugados armónicamente separados respecto a un par de puntos prefijado.*

Si las involuciones son ambas hiperbólicas hemos de exigir un poco más:

Teorema 9.63 *Dos involuciones hiperbólicas en una recta real cuyos pares de puntos fijos no se separen tienen un único par común de puntos conjugados.*

DEMOSTRACIÓN: Podemos trabajar con una cónica en lugar de una recta. Sean A, A' y B, B' los pares de puntos fijos. Las rectas AA' y BB' se cortan en un punto exterior P , la polar de P corta a la cónica en dos puntos C y C' . La recta CC' contiene a la intersección O de las tangentes por A y A' y a la intersección O' de las tangentes por B y B' , que son los centros de las dos involuciones dadas, luego es claro que ambas conjugan a C y C' . La unicidad es obvia. ■

9.8 Complexificación

Los resultados de la sección anterior se aplican principalmente al caso de los planos proyectivos reales. Si consideramos un plano complejo perdemos todas las propiedades de ordenación, pero a cambio ganamos cierta simplicidad en algunos aspectos. Por ejemplo, toda involución en una recta compleja tiene dos puntos fijos. Para probarlo sólo hemos de adaptar de forma obvia el razonamiento del teorema 8.62. A su vez esto implica que toda recta tiene en común uno o dos puntos con una cónica. En efecto, si la recta no es tangente, entonces la cónica induce una involución sobre la recta que empareja a los puntos conjugados por su polaridad. Los puntos fijos de esta involución han de ser puntos de corte.

El estudio de los espacios proyectivos complejos tiene gran interés en geometría algebraica, aunque nosotros no entraremos en ello. Aquí nos limitaremos a usarlos como auxiliares en el estudio de los espacios reales. Para ello observamos que la inclusión $\mathbb{R} \subset \mathbb{C}$ implica evidentemente $P^n(\mathbb{R}) \subset P^n(\mathbb{C})$. Puesto que todo espacio proyectivo real es equivalente (a través de una homografía) con un espacio $P^n(\mathbb{R})$, es claro que todo espacio proyectivo real X se puede sumergir en un espacio complejo \bar{X} de la misma dimensión de modo que la relación entre X y \bar{X} sea la misma que entre $P^n(\mathbb{R})$ y $P^n(\mathbb{C})$. Es posible construir \bar{X} de forma canónica a partir de X mediante productos tensoriales, pero no será necesario entrar en ello. En su lugar nos apoyaremos en que cualquier cosa que probemos para $P^n(\mathbb{R})$ es de hecho válida para cualquier espacio proyectivo de dimensión n . Al espacio \bar{X} lo llamaremos *complexificación* de X .

Puesto que es difícil imaginar incluso un plano proyectivo complejo, quizá sea de ayuda pensar en la relación análoga que existe entre $P^2(\mathbb{Q})$ y $P^2(\mathbb{R})$ o, más simplemente aún, entre los planos afines \mathbb{Q}^2 y \mathbb{R}^2 .

Cada recta racional es de la forma $(a, b) + t(p, q)$, donde t recorre \mathbb{Q} . Es claro que está contenida en una única recta real, a saber la que resulta de permitir que t recorra todo \mathbb{R} . Sin embargo hay rectas reales irracionales, en el sentido de que no contienen ninguna recta racional. Por ejemplo, la recta $t(1, \sqrt{2})$ no contiene más punto racional que $(0, 0)$, pues siempre que la primera coordenada es racional (no nula) la segunda es irracional.

Volviendo al caso de un espacio proyectivo real X y su complexificación \bar{X} , a los puntos de X los llamaremos *puntos reales*, mientras que a los de \bar{X} los llamaremos *puntos complejos*. A los puntos complejos que no sean reales los llamaremos *imaginarios*.

Es claro que un conjunto de vectores de \mathbb{R}^{n+1} es linealmente independiente sobre \mathbb{R} si y sólo si lo es sobre \mathbb{C} visto como subconjunto de \mathbb{C}^{n+1} (si los vectores son independientes sobre \mathbb{R} y tenemos una combinación lineal nula con coeficientes en \mathbb{C} , entonces tomando partes reales y partes imaginarias obtenemos dos combinaciones lineales nulas de los mismos vectores pero con coeficientes reales, y así tanto las partes reales de los coeficientes como las imaginarias son nulas). Por consiguiente un conjunto de puntos de $P^n(\mathbb{R})$ es proyectivamente independiente visto en $P^n(\mathbb{R})$ si y sólo si lo es visto en $P^n(\mathbb{C})$.

Enunciando esto mismo en general, tenemos que si X es un espacio proyectivo real un conjunto de puntos reales es proyectivamente independiente en X si y sólo si lo es en \bar{X} . De aquí que cada variedad L en X está contenida en una única variedad \bar{L} en \bar{X} de la misma dimensión (la generada en \bar{X} por un generador independiente de L en X). Llamaremos *variedades complejas* a las variedades de \bar{X} , de entre ellas, llamaremos *variedades reales* a las que extienden a las variedades de X , mientras que a las variedades complejas que no sean reales las llamaremos *variedades imaginarias*.

Podemos identificar las variedades de X con las variedades reales de \bar{X} , pues se cumplen todas las propiedades de compatibilidad que cabe esperar, como $\overline{L_1 \cap L_2} = \bar{L}_1 \cap \bar{L}_2$, etc.

Toda homografía entre dos espacios reales se extiende de forma única a sus complexificaciones. Basta tomar un sistema de referencia real y considerar la homografía compleja que sobre dicho sistema actúa como la dada. Llamaremos *homografías reales* a las extensiones de homografías entre espacios reales. Se caracterizan por que hacen corresponder puntos reales con puntos reales (o también por que sus matrices respecto a sistemas de referencia reales son reales).

Salvo que aparezca explícitamente el adjetivo “complejo” o “imaginario”, cuando hablemos de puntos, rectas, homografías, etc. se sobrentenderá que son reales. Son los únicos que nos interesa estudiar. Los puntos imaginarios serán, como ya hemos dicho, auxiliares cuyo interés reside en que sabemos que forman un espacio proyectivo. Rara vez nos interesarán las rectas y homografías imaginarias.

Con las cónicas no sucede exactamente lo mismo. Toda cónica en un plano real se extiende a una cónica en su complexificación (la que pasa por cinco puntos de la dada). A estas extensiones las llamaremos *cónicas reales*. La diferencia es que hay ciertas cónicas imaginarias que podemos estudiar junto a las reales, y además conviene hacerlo. Se trata de las cónicas cuya ecuación en un sistema de referencia real tiene coeficientes reales. Hemos probado que cualquiera de estas cónicas, en un sistema de referencia adecuado, tiene asociada la ecuación $x^2 + y^2 + z^2 = 0$. Son las cónicas que llamábamos “imaginarias” y que entonces decíamos que correspondían al conjunto vacío. Ahora podemos decir que corresponden a ciertas cónicas de la complexificación del plano compuestas íntegramente de puntos imaginarios.

En lo sucesivo la palabra “cónica” (en un plano real) hará referencia a una cónica (compleja no degenerada) cuya ecuación en un sistema de referencia real

tenga coeficientes reales. Existen, pues, dos clases de cónicas en un espacio real, las reales y las imaginarias, en el sentido de que dos cónicas son del mismo tipo si y sólo si existe una homografía (real) que transforma una en otra.

Si \mathcal{C} es una cónica en un plano real X y A es su matriz en un sistema de referencia, entonces A es la matriz de su extensión $\bar{\mathcal{C}}$ respecto al mismo sistema de referencia. La recta polar en X de un punto Q cuyas coordenadas homogéneas sean X_0 es la recta de ecuación $X_0 A X^t = 0$, y ésta es también la ecuación de la recta polar de Q en \bar{X} , luego $p(Q)$ en \bar{X} es la extensión de $p(Q)$ en X . En este sentido podemos decir que la polaridad asociada a la extensión de \mathcal{C} extiende a la polaridad de \mathcal{C} . Más aún, es claro que un punto es real si y sólo si lo es su recta polar (si la polar es real admite una ecuación con coeficientes reales, digamos C , y entonces CA^{-1} son unas coordenadas homogéneas reales del polo).

Si aplicamos esto en particular a los puntos de la cónica, vemos que la tangente a \mathcal{C} por un punto real (vista en \bar{X}) es la extensión de su tangente en X , en otras palabras, que una tangente a una cónica real por un punto real no contiene un segundo punto imaginario en común con la cónica, o también: si una recta real corta a una cónica en dos puntos, éstos son ambos reales o ambos imaginarios.

Si seleccionamos una recta infinita r en un plano real, ésta tiene uno o dos puntos en común (reales o imaginarios) con cualquier cónica \mathcal{C} . Si \mathcal{C} tiene sólo un punto en común con r , es decir, si r es tangente a \mathcal{C} , entonces el punto de tangencia (el polo de una recta real) es un punto real, luego \mathcal{C} es una cónica real (ya que las cónicas imaginarias no contienen ningún punto real) y por consiguiente es (la extensión de) una hipérbola.

Si r corta a \mathcal{C} en dos puntos reales concluimos del mismo modo que \mathcal{C} es (la extensión de) una hipérbola.

Por último, si r corta a \mathcal{C} en dos puntos imaginarios tenemos dos posibilidades, o bien la cónica es real, y entonces se trata de una elipse, o bien la cónica es imaginaria.

En resumen, podemos redefinir los conceptos de parábola, hipérbola y elipse del modo siguiente:

Definición 9.64 En un espacio afín real, una *parábola* es una cónica con un único punto en común (necesariamente real) con la recta infinita, una *hipérbola* es una cónica que corta a la recta infinita en dos puntos reales y una *elipse* es una cónica que corta a la recta infinita en dos puntos imaginarios.

De este modo, las parábolas e hipérbolas son siempre reales y son las extensiones de las parábolas e hipérbolas en el plano real, mientras que una elipse puede ser real (y entonces es la extensión de una elipse del plano real) o bien imaginaria.

Tenemos, pues, cuatro tipos de cónicas afines: hipérbolas, parábolas, elipses reales y elipses imaginarias. Es claro que dos cónicas son del mismo tipo si y sólo si hay una afinidad (real) que transforma una en otra.

Hasta aquí hemos distinguido por claridad entre los conceptos de un plano real y sus extensiones a su complexificación. En lo sucesivo ya no volveremos a hacer tal distinción. Hemos de pensar que las rectas y cónicas contienen puntos imaginarios que no corresponden con ninguna representación intuitiva, pero que pueden usarse coherentemente en expresiones como, “la recta r corta a la elipse \mathbb{C} en dos puntos imaginarios”.

Cuando se trabaja conjuntamente con puntos reales e imaginarios, a menudo surge el problema de probar que un determinado punto es real. Ahora probaremos un resultado útil a este respecto. Éste gira en torno al concepto de puntos imaginarios conjugados.

Definición 9.65 Diremos que dos puntos son *imaginarios conjugados* si son imaginarios, están situados sobre una recta real y la involución en ella que los fija es real.

Podríamos definir igualmente “imaginarios conjugados sobre una cónica”, sin más que cambiar “recta” por “cónica”. Entonces se cumple que dos puntos imaginarios conjugados sobre una cónica son también imaginarios conjugados sobre una recta. En efecto, si P y Q son puntos imaginarios sobre una cónica real y la involución que los fija es real, entonces su eje también es una recta real que corta a la cónica en P y Q , luego la involución que la cónica induce en dicho eje por conjugación es real y tiene a P y Q por puntos fijos.

Así mismo podemos definir el concepto de rectas imaginarias conjugadas respecto a un haz de rectas sobre un punto real. Su relación con los puntos conjugados imaginarios es la siguiente:

Ejercicio: Probar que las rectas que unen un punto real con dos puntos imaginarios conjugados son imaginarias conjugadas.

El resultado que, como anticipábamos, ayuda a veces a probar que determinados puntos son reales es el siguiente:

Teorema 9.66 Si A_1, A_2 y B_1, B_2 son dos pares de puntos imaginarios conjugados sobre rectas distintas, entonces las rectas A_1B_1 y A_2B_2 se cortan en un punto real.

DEMOSTRACIÓN: Por definición de puntos imaginarios conjugados, las rectas A_1A_2 y B_1B_2 son reales, luego se cortan en un punto real C . Sea A el conjugado de C respecto a la involución que fija a A_1 y A_2 y sea B el conjugado de C respecto a la involución que fija a B_1 y B_2 . Los puntos A y B son ambos reales. Por el teorema 9.62 podemos tomar dos puntos reales P y Q separados harmónicamente por A y C y conjugados respecto a la involución (elíptica) que fija a A_1, A_2 e igualmente dos puntos reales R y S separados harmónicamente por B y C y conjugados respecto a la involución que fija a B_1 y B_2 . Como los conjuntos harmónicos son proyectivos tenemos

$$ACPQ \stackrel{\wedge}{=} BCRS \quad \text{y} \quad ACPQ \stackrel{\wedge}{=} BCSR.$$

Cada una de estas proyecciones perspectivas transforma dos pares de conjugados respecto a la involución elíptica que fija a A_1 y A_2 en dos pares de conjugados respecto a la involución elíptica que fija a B_1 y B_2 , de donde se sigue fácilmente que han de transformar todo par de conjugados por la primera en un par de conjugados por la segunda. En particular las imágenes de A_1 y A_2 han de ser B_1 y B_2 en un cierto orden. Este orden ha de ser distinto para cada una, pues de lo contrario coincidirían sobre cuatro puntos y serían la misma. Por consiguiente una de las dos transforma A_1 en B_1 y A_2 en B_2 . Si P es el centro de dicha proyección, entonces $P = A_1B_1 \cap A_2B_2$ y, como la proyección es real, su centro P también lo es. ■

Capítulo X

La geometría parabólica

Dedicamos este capítulo a estudiar la geometría euclídea desde el punto de vista de la geometría proyectiva. En realidad trabajaremos en un contexto más general, la geometría parabólica, cuyo interés principal para nosotros será que permite extender una parte de la geometría euclídea a la complexificación de un espacio real. Emplearemos técnicas sintéticas, y por simplicidad lo haremos en dimensión 2. Después esbozaremos la forma de generalizar algebraicamente los resultados a dimensiones superiores.

Para motivar la descripción proyectiva de la geometría euclídea vamos a recorrer primero el camino inverso. Partamos de un plano euclídeo E y veamos qué aporta la estructura euclídea a su compleción proyectiva $P(E)$.

Para cada punto Q de la recta infinita r podemos considerar las rectas finitas que pasan por Q , que forman un haz de rectas paralelas en E . Una recta es perpendicular a una de ellas si y sólo si lo es a las demás, y las rectas que cumplen esto forman un haz de rectas paralelas, que cortan a r en un mismo punto, al que podemos llamar $I(Q)$. Si escribimos la aplicación I en un sistema de referencia veremos fácilmente que se trata de una involución.

De este modo, todo plano euclídeo induce en su recta infinita una involución I con la propiedad de que dos rectas son perpendiculares si y sólo si sus puntos infinitos son conjugados.

Ejercicio: Probar que la involución inducida en la recta infinita es la inducida por conjugación a partir de cualquier circunferencia.

10.1 Espacios parabólicos

Definición 10.1 Un *plano parabólico* es un plano proyectivo X en el que hemos especificado una recta infinita r y en ella una involución I a la que llamaremos *involución ortogonal*. El *grupo parabólico* de X será el grupo de las afinidades f de X que conservan la involución, es decir, tales que $f(I(Q)) = I(f(Q))$. Lo representaremos por $\text{Sem}(X)$ y a sus elementos los llamaremos *semejanzas*.

de X . Diremos que dos rectas (finitas) s y t son *ortogonales* o *perpendiculares* si sus puntos infinitos son conjugados por I .

Un plano parabólico X es en particular un plano afín. Cuando hablemos de puntos y rectas de X se sobrentenderá que son finitos salvo que se indique lo contrario. Una semejanza es una afinidad que conserva la perpendicularidad de rectas. El teorema siguiente es inmediato a partir de la definición anterior:

Teorema 10.2 *En un plano parabólico X , por cada punto pasa una única perpendicular a una recta dada, dos rectas son paralelas tienen las mismas perpendiculares, dos rectas con una perpendicular común son paralelas.*

Un plano parabólico no es exactamente un plano euclídeo pues, por ejemplo, si la involución I es hiperbólica entonces existen rectas perpendiculares a sí mismas. Más detalladamente, llamaremos *puntos circulares* de X a los puntos fijos de I , si es que existen, y los representaremos por C_1 y C_2 . Las rectas finitas que pasan por los puntos circulares son perpendiculares a sí mismas, y las llamaremos *rectas isótropas*.

Según el teorema 8.62, sólo existen dos geometrías parabólicas en un plano proyectivo real, según si la involución ortogonal es elíptica o hiperbólica (es decir, según si existen o no puntos circulares). El caso elíptico corresponde, como ya hemos dicho, con la geometría euclídea usual, mientras que el caso hiperbólico tiene interés porque es el único posible en un plano complejo.

Aunque ya hemos dicho que vamos a trabajar sintéticamente con un plano parabólico arbitrario, no está de más que veamos qué es la perpendicularidad desde un punto de vista algebraico. Diremos que dos vectores \vec{u} u \vec{v} son *ortogonales* si determinan rectas perpendiculares. Lo representaremos por $\vec{u} \perp \vec{v}$, como es habitual. Convenimos que el vector nulo es ortogonal a cualquier otro.

Teorema 10.3 *Si X es un plano parabólico, existe una forma bilineal regular simétrica F sobre su espacio vectorial asociado de modo que dos vectores \vec{u} y \vec{v} son ortogonales si y sólo si $F(\vec{u}, \vec{v}) = 0$.*

DEMOSTRACIÓN: Tomemos un sistema de referencia afín arbitrario de origen O . Consideramos las coordenadas cartesianas en este sistema así como las coordenadas homogéneas asociadas, respecto a las cuales la recta infinita es $z = 0$. Los puntos $(1, 0, 0)$, $(0, 1, 0)$ y $(1, 1, 0)$ determinan un sistema de referencia en la recta infinita respecto al cual el punto $(x, y, 0)$ tiene coordenadas homogéneas (x, y) . Según el teorema 8.53, la involución ortogonal es una polaridad simétrica. Sea A su matriz. Sea F la forma bilineal que en el sistema de referencia dado tiene matriz A .

Dado un vector no nulo \vec{u} de coordenadas (a, b) , la recta que pasa por O y $O + \vec{u}$ es $-bx + ay = 0$, y su punto infinito es $(a, b, 0)$. Su conjugado ortogonal es el punto $(c, d, 0)$ determinado por

$$(a, b)A \begin{pmatrix} c \\ d \end{pmatrix} = 0.$$

Si llamamos \vec{v} al vector de coordenadas (c, d) , entonces las rectas de dirección \vec{v} son las que pasan por $(c, d, 0)$, luego tenemos que $\vec{u} \perp \vec{v}$, y por otro lado la ecuación anterior implica $F(\vec{u}, \vec{v}) = 0$.

Es claro entonces que todos los vectores \vec{v} ortogonales a \vec{u} cumplen la relación $F(\vec{u}, \vec{v}) = 0$. El recíproco se prueba invirtiendo el argumento. ■

Notemos que la involución ortogonal determina la forma bilineal F salvo un factor escalar no nulo. Recíprocamente, toda forma bilineal simétrica regular determina una involución en la recta infinita, por lo que un plano parabólico puede definirse como un plano afín \vec{E} tal que en \vec{E} se ha especificado una forma bilineal simétrica regular.

Si partimos de un espacio euclídeo E , es fácil probar que cualquier circunferencia induce en la recta infinita la involución ortogonal. Esto es tanto como decir que si dos tangentes a una circunferencia son paralelas, entonces la recta que une los puntos de tangencia es perpendicular a ambas. En el camino inverso que estamos recorriendo vamos a usar este hecho para dar una definición provisional de circunferencia.

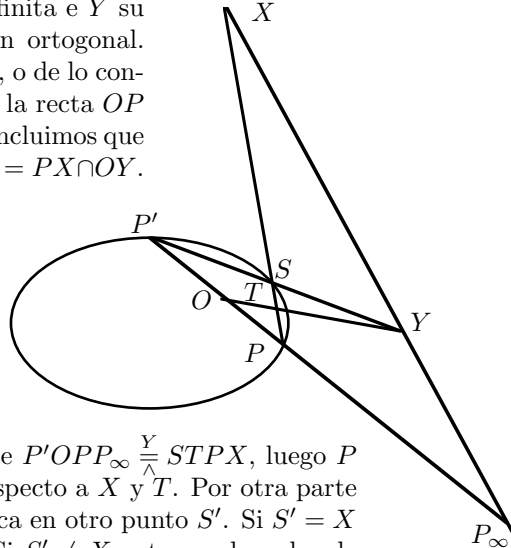
Teorema 10.4 Sean O y P dos puntos de un plano parabólico tales que la recta que determinan no sea isótropa. Entonces existe una única cónica \mathcal{C} de centro O que pasa por P y que induce en la recta infinita la involución ortogonal.

DEMOSTRACIÓN: Sea P_∞ el punto infinito de la recta OP y sea P' el conjugado harmónico de P respecto a O y P_∞ . Es claro que P' ha de pertenecer a la cónica que queremos construir (por las propiedades del centro de una cónica).

Sea X un punto de la recta infinita e Y su conjugado respecto a la involución ortogonal. Las rectas PX y $P'Y$ son distintas, o de lo contrario tendríamos $X = Y = P_\infty$, y la recta OP sería isótropa. Del mismo modo concluimos que $PX \neq OY$. Sea $S = PX \cap P'Y$ y $T = PX \cap OY$.

Si PX es tangente a la cónica que buscamos, entonces la polar de X ha de ser OP , luego $Y = P_\infty$. En este caso tenemos simplemente $S = T = P$. Supongamos ahora que PX no es tangente a la cónica que buscamos.

Entonces $Y \neq P_\infty$ y claramente $P'OPP_\infty \stackrel{Y}{\underset{\wedge}{=}} STPX$, luego P y S son conjugados harmónicos respecto a X y T . Por otra parte la recta PX ha de cortar a la cónica en otro punto S' . Si $S' = X$ entonces $S' = X = Y = S = T$. Si $S' \neq X$ entonces la polar de X , que es OY , ha de contener al conjugado harmónico de X respecto de S' y P (teorema 9.39), que necesariamente ha de ser T , luego en cualquier caso $S' = S$ y S está en la cónica.



En resumen, la cónica que buscamos está formada necesariamente por los puntos de la forma $PX \cap P'Y$, donde X recorre la recta infinita e $Y = I(X)$. Ahora bien, es claro que $P[X] \wedge P'[Y]$ y ya hemos visto que las rectas PX y $P'Y$ nunca coinciden, luego la homografía no es perspectiva, y por consiguiente los puntos de la forma indicada forman en efecto una cónica. ■

Definición 10.5 Una *circunferencia* en un plano parabólico es cualquier cónica que induzca en la recta infinita la involución ortogonal.

En estos términos hemos probado que si la recta OP no es isótropa entonces existe una única circunferencia de centro O que pasa por P . Obviamente las circunferencias son elipses si la involución ortogonal es elíptica y son hipérbolas si es hiperbólica, y en este caso todas las circunferencias pasan por los puntos circulares (de aquí su nombre).

Ahora vamos a construir las isometrías, para lo cual nos apoyaremos en que toda isometría es producto de reflexiones, que en el caso bidimensional que nos ocupa son simplemente las simetrías respecto a una recta.

Definición 10.6 Una *reflexión* es una homología de orden 2 cuyo centro C es un punto infinito y cuyo eje r corta a la recta infinita en el punto $I(O)$.

Por definición, el centro de una homología no está sobre su eje, luego no puede ser un punto circular y el eje no puede ser isótropo. Por el teorema 8.54, cada recta no isótropa r es el eje de una única reflexión. En efecto, el centro es necesariamente el conjugado ortogonal del punto infinito de r , luego la reflexión ha de fijar las rectas perpendiculares a r . Si s es cualquiera de estas rectas y P es el punto de corte con r , entonces la restricción a s de la reflexión es la única involución que fija al punto infinito y a P . Concretamente envía cada punto a su conjugado armónico respecto a estos dos. Es fácil construir explícitamente una homología con estas características.

El teorema siguiente clarifica la idea que subyace en la prueba de 10.4:

Teorema 10.7 Si \mathcal{C} es una circunferencia de centro O y f es una reflexión respecto a una recta que pasa por O , entonces f deja invariante a \mathcal{C} . Más aún, si P y S son dos puntos (finitos) de \mathcal{C} , existe una reflexión cuyo eje pasa por O y que transforma P en S .

DEMOSTRACIÓN: Sea C el centro de la reflexión. La polar de C pasa por O (porque O es el polo de la recta infinita) y por $I(C)$ (porque I es la involución inducida por \mathcal{C} en la recta infinita). Por lo tanto es el eje r . Si por C pasan tangentes a \mathcal{C} los puntos de tangencia serán las intersecciones de r con \mathcal{C} , y f los deja invariantes. Si P es cualquier otro punto de \mathcal{C} entonces la recta CP corta a \mathcal{C} en otro punto P' con la propiedad de que la polar r contiene al conjugado armónico de C respecto a P y P' , pero esto significa también que $P' = f(P)$. Así pues, $f[\mathcal{C}] = \mathcal{C}$.

Si P y S son puntos de \mathcal{C} , es claro que la recta OP no es isótropa, pues si existen puntos circulares C_1 y C_2 , entonces las rectas OC_i son tangentes a la circunferencia, luego no contienen a P . Si $S = P$ es obvio que P es su propio conjugado por la reflexión de eje OP . Si P es el punto que en la prueba de 10.4 llamábamos P' entonces S es la imagen de P por la reflexión respecto a la perpendicular a OP por O . Si S es cualquier otro punto entonces se encuentra en la situación de dicha prueba, y vemos que es el conjugado de P respecto a la reflexión de eje OY . ■

Como consecuencia inmediata obtenemos:

Teorema 10.8 *Las reflexiones son semejanzas.*

DEMOSTRACIÓN: Consideremos una reflexión f de eje r y centro C . Sea O un punto cualquiera en r y P cualquier punto fuera de r tal que la recta OP no sea isótropa. Sea \mathcal{C} la circunferencia de centro O que pasa por P . Por el teorema anterior $f[\mathcal{C}] = \mathcal{C}$.

Pensemos por un momento en una homografía $f : X \rightarrow Y$ entre dos planos proyectivos cualesquiera, de modo que \mathcal{C} es una cónica en X que induce una involución I en una recta s . Entonces es claro que $f[\mathcal{C}]$ es una cónica en Y que induce en $f[s]$ la involución $f^{-1}If$.

Si aplicamos esto al caso en que $X = Y$, f es la reflexión que tenemos, s es la recta infinita, \mathcal{C} la circunferencia e I la involución ortogonal, vemos que $f[\mathcal{C}] = \mathcal{C}$ induce en $f[s] = s$ la involución $f^{-1}If$, pero ésta tiene que ser I , luego $I = f^{-1}If$, como había que probar. ■

Definición 10.9 Definimos el *grupo de isometrías* de un plano parabólico X como el grupo generado por las reflexiones. Lo representaremos por $\text{Is}(X)$.

Por el teorema anterior sabemos que $\text{Is}(X)$ es un subgrupo de $\text{Sem}(X)$. Observemos que las homotecias y las traslaciones son obviamente semejanzas, pues fijan cada punto de la recta infinita. Veamos que las traslaciones son también isometrías:

Teorema 10.10 *El producto de dos reflexiones con ejes paralelos es una traslación. Toda traslación cuyo centro no sea un punto circular es producto de dos reflexiones de ejes paralelos. Si el centro es circular entonces la traslación es producto de cuatro reflexiones. En particular todas las traslaciones son isometrías.*

DEMOSTRACIÓN: Sean f y g dos reflexiones con ejes paralelos. Sea P_∞ el punto infinito común de sus ejes. Ambas fijan a P_∞ y a su centro $I(P_\infty) \neq P_\infty$, luego ambas inducen la misma involución en la recta infinita. Por consiguiente $f \circ g$ deja invariante cada punto de la recta infinita, así como a todas las rectas que pasan por $I(P_\infty)$, luego se trata de una traslación.

Sea T una traslación y O un punto arbitrario. Sea P el punto medio entre O y $T(O)$ y sea P_∞ el punto infinito de esta misma recta. Sean r y r' las rectas

perpendiculares a $OT(O)$ por O y P . Si P_∞ no es un punto circular podemos considerar las reflexiones f_r y $f_{r'}$ de ejes r y r' respectivamente. La composición $f_r \circ f_{r'}$ es una traslación que hace corresponder O con $T(O)$, luego $f_r \circ f_{r'} = T$.

Si T tiene como centro un punto circular es claro que podemos expresar T como suma de dos traslaciones con centros arbitrarios (no circulares), y cada una de ellas se puede descomponer en producto de dos reflexiones. ■

Entre las isometrías se encuentran también las simetrías puntuales:

Teorema 10.11 *La composición de dos reflexiones respecto a ejes perpendiculares es la simetría puntual respecto a su punto de intersección.*

DEMOSTRACIÓN: Sean f y g dos reflexiones en las condiciones indicadas. Sea O el punto de intersección de sus ejes. Entonces ambas inducen la misma involución en la recta infinita, la que fija a los puntos infinitos de sus ejes. Por lo tanto fg fija a cada punto de la recta infinita, y también a O , luego a todas las rectas que pasan por O . Por consiguiente fg es la homología de centro O y eje la recta infinita, es decir, la simetría puntual respecto a O . ■

El teorema siguiente nos permitirá estudiar la acción de las isometrías sobre las circunferencias:

Teorema 10.12 *Si una isometría fija a un punto O , entonces se expresa como producto de reflexiones cuyos ejes pasan por O .*

DEMOSTRACIÓN: Una isometría es de la forma $f_1 \cdots f_n$, donde cada f_i es una reflexión. Sea f'_i la reflexión cuyo eje es paralelo al de f_i y pasa por O (quizá $f_i = f'_i$). Entonces $f_i f'_i = T_i$ es una traslación (quizá la identidad) y $f_i = T_i f'_i$. La isometría dada es

$$f_1 \cdots f_n = T_1 f'_1 \cdots T_n f'_n.$$

El grupo de traslaciones es un subgrupo normal del grupo de las afinidades, luego si T es una traslación y f una afinidad se cumple que $f^{-1}Tf = T'$ es una traslación, y $Tf = fT'$. Usando esto repetidas veces llegamos a que

$$f_1 \cdots f_n = f'_1 \cdots f'_n T,$$

para una cierta traslación T , que cumple $T(O) = O$, luego $T = 1$ y el teorema está probado. ■

El teorema 10.7 implica ahora:

Teorema 10.13 *Si una isometría fija a un punto O , entonces fija todas las circunferencias de centro O .*

Ahora podemos probar las propiedades de rigidez de las semejanzas e isometrías:

Teorema 10.14 *Si una semejanza fija a dos puntos O y P tales que la recta OP no es isótropa, entonces es una reflexión o bien la identidad.*

DEMOSTRACIÓN: Sea f la semejanza. Como f fija también al punto infinito de la recta OP , de hecho fija a cada punto de esta recta. El conjugado de dicho punto infinito también es fijado, luego f fija a todas las rectas perpendiculares a OP .

Consideremos la circunferencia \mathcal{C} de centro O y que pasa por P . Entonces $f[\mathcal{C}]$ es una circunferencia (porque f conserva la involución ortogonal), también de centro O y que pasa por P , luego por la unicidad $f[\mathcal{C}] = \mathcal{C}$. La homografía que f induce en \mathcal{C} es la identidad o bien intercambia cada par de puntos donde una perpendicular a OP corta a la cónica, luego es una involución y f también. En este caso es obvio que se trata de la reflexión de eje OP . ■

Para isometrías el resultado es más fuerte:

Teorema 10.15 *Si una isometría fija a un punto O y a una recta no isótropa que pasa por O , entonces es la identidad, una reflexión o una simetría puntual.*

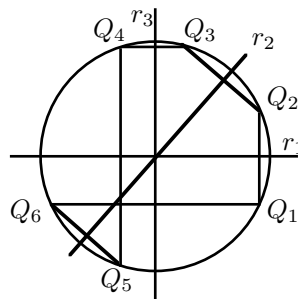
DEMOSTRACIÓN: Sea f la isometría y \mathcal{C} una circunferencia de centro O que pase por un punto P de la recta del enunciado. Dicha recta ha de cortar a \mathcal{C} en otro punto P' (no puede ser tangente porque entonces P estaría en la recta polar de O , que es la recta infinita). Sabemos que f fija a \mathcal{C} , luego fija a P y P' o bien los intercambia. Si los fija entonces el teorema anterior nos da que f es una reflexión o bien la identidad. Si los intercambia, consideremos la reflexión g respecto a la recta perpendicular a OP que pasa por O . Entonces $h = fg$ fija a O y a P , luego de nuevo por el teorema anterior se trata de la identidad o de la reflexión de eje OP , con lo que $f = gh$ es una reflexión o bien la simetría puntual respecto a O . ■

El teorema siguiente nos permite probar que la clasificación de isometrías que ya conocemos para los planos euclídeos es válida igualmente en planos parabólicos arbitrarios:

Teorema 10.16 *Sean f_1 , f_2 y f_3 tres reflexiones cuyos ejes concurren en un punto O (finito o infinito). Entonces la composición $f_1f_2f_3$ es una reflexión cuyo eje pasa por O .*

DEMOSTRACIÓN: Si O es infinito, es decir, si los ejes son paralelos, entonces f_2f_3 es una traslación. La prueba del teorema 10.10 muestra que $f_2f_3 = f_1f_4$, donde f_4 es una reflexión respecto a un eje paralelo a los otros. Por lo tanto $f_1f_2f_3 = f_1f_1f_4 = f_4$.

Supongamos ahora que O es un punto finito. Podemos suponer que los ejes de las tres reflexiones son distintos dos a dos, pues si dos coinciden el resultado es trivial. Sea P un punto cualquiera tal que la recta OP no sea isótropa. Sea \mathcal{C} la circunferencia de centro O que pasa por P . Si Q_1 es cualquier punto de \mathcal{C} , definimos $Q_2 = f_1(Q_1)$, $Q_3 = f_2(Q_2)$, $Q_4 = f_3(Q_3)$, $Q_5 = f_1(Q_4)$ y $Q_6 = f_2(Q_5)$.



Supongamos que los seis puntos así definidos son distintos dos a dos. Entonces es claro que Q_1Q_2 es paralela a Q_4Q_5 y Q_2Q_3 es paralela a Q_5Q_6 . Por el teorema de Pascal concluimos que Q_6Q_1 es paralela a Q_3Q_4 , de donde $f_3(Q_6) = Q_1$. Así pues, $(f_1f_2f_3)^2(Q_1) = Q_1$.

Si los seis puntos no son distintos hemos de estudiar qué coincidencias pueden darse. Notemos en general que $Q_i \neq Q_{i+2}$, pues el paso de Q_i a Q_{i+1} se hace mediante una reflexión de eje distinto a la que transforma Q_{i+1} en Q_{i+2} , y si se diera la igualdad ambos ejes deberían pasar por O y por el punto medio de Q_i y Q_{i+1} , luego serían el mismo. Un primer caso a considerar es si $Q_1 = Q_4$. La conclusión es entonces inmediata. Notamos también que cualquiera de las igualdades $Q_2 = Q_5$ o $Q_3 = Q_6$ implica $Q_1 = Q_4$. Podemos, pues, descartarlas a todas.

Si hacemos una tabla con las igualdades que pueden darse, veremos que las únicas son las de la forma $Q_i = Q_{i+1}$ (incluyendo $Q_6 = Q_1$). Por ejemplo, la igualdad $Q_1 = Q_5$ implica $Q_2 = Q_4$, que es imposible. Además, si dos puntos consecutivos coinciden el siguiente ha de ser distinto. Por lo tanto todas las identificaciones permiten aplicar una de las variantes del teorema de Pascal. Hay que tener en cuenta que una tangente a \mathbb{C} por un punto P es perpendicular a OP (pues la polar del punto infinito de la tangente es OP , y pasa por su conjugado ortogonal).

Así pues, tenemos que $(f_1f_2f_3)^2(Q_1) = Q_1$ para todo punto (finito) Q_1 de \mathbb{C} , luego $(f_1f_2f_3)^2 = 1$. Tenemos que $f_1f_2f_3$ es una isometría que fija a O y a la recta Q_1Q_4 (entendida como la tangente a \mathbb{C} si $Q_1 = Q_4$). También fija al punto infinito de esta recta y a la recta perpendicular que pasa por O . A su vez esto implica que $f_1f_2f_3$ es la identidad, una reflexión (cuyo eje pasa por O) o la simetría puntual de centro O . Para probar que se trata de una reflexión basta ver que su restricción a la recta infinita no es la identidad.

Sean g_1 , g_2 y g_3 las restricciones a la recta infinita de f_1 , f_2 y f_3 . Se trata de tres involuciones hiperbólicas. Si $g_1 = g_2$ la conclusión es obvia. En caso contrario, si $g_1g_2g_3 = 1$ tendríamos $g_1 = g_2g_3$. Sea P uno de los puntos fijos de g_1 . Entonces $Q = g_2(P) \neq P$, pues en caso contrario g_1 y g_2 tendrían los mismos puntos fijos (el otro es necesariamente el conjugado ortogonal de P) y serían iguales. Por lo tanto $g_3(Q) = P$. Así pues, g_2 y g_3 tienen el par (P, Q) en común, pero partiendo del segundo punto fijo de g_1 obtenemos otro par común de conjugados, luego $g_2 = g_3$ y llegamos a una contradicción. ■

Definición 10.17 Un *giro* en un plano parabólico es un producto de dos reflexiones cuyos ejes concurren en un punto O , llamado *centro* del giro. Un *movimiento* es un producto de un número par de reflexiones, una *simetría* es un producto de un número impar de reflexiones.

Teorema 10.18 Sea f una isometría en un plano parabólico y O uno de sus puntos.

1. Se cumple que $f = gT$, donde T es una traslación y g una isometría que fija a O .

2. Si f fija a O entonces f es un giro de centro O o bien una reflexión cuyo eje pasa por O .
3. Los giros de centro O forman un grupo abeliano.
4. f no puede ser a la vez un movimiento y una simetría.
5. Todo movimiento es un giro o una traslación.

DEMOSTRACIÓN: 1) Es evidente: si $f(O) = O'$ basta tomar como T la traslación que cumple $T(O) = O'$.

2) se sigue del teorema anterior y 10.12.

3) El teorema anterior implica claramente que los giros de centro O forman un grupo. Para probar que es abeliano hemos de ver que si f_1, f_2, f_3, f_4 son reflexiones cuyos ejes pasan por O , entonces $f_1 f_2 f_3 f_4 = f_3 f_4 f_1 f_2$.

Por el teorema anterior tenemos que $f_2 f_3 f_4$ tiene orden 2, luego $f_2 f_3 f_4 = f_4 f_3 f_2$. Por lo tanto $f_1 f_2 f_3 f_4 = f_1 f_4 f_3 f_2$, y del mismo modo $f_1 f_4 f_3 = f_3 f_4 f_1$, luego $f_1 f_2 f_3 f_4 = f_3 f_4 f_1 f_2$.

4) Si f se pudiera expresar como producto de un número par y de un número impar de reflexiones, componiendo con una traslación (dos o cuatro reflexiones) podemos suponer que f fija a un punto O , luego por el argumento de 2) tendríamos que f sería a la vez un giro y una reflexión, pero esto es imposible, porque una reflexión fija a una recta y un giro sólo fija a un punto.

5) Basta probar que todo movimiento f (excepto las traslaciones cuyo centro es un punto circular) se expresa como producto de dos reflexiones. Tenemos que f se descompone como $f = gT$, donde T es una traslación y g es un movimiento que fija a un punto O , luego es un giro de centro O .

Si el centro de T no es un punto circular descomponemos $T = h_1 h_2$ en producto de dos reflexiones, donde podemos exigir que el eje de h_1 pase por O . Sea $g = g_1 g_2$, donde las dos reflexiones pasan por O . Entonces $h_3 = g_1 g_2 h_1$ es una reflexión y $gT = g_1 g_2 h_1 h_2 = h_3 h_1 h_1 h_2 = h_3 h_2$.

Si el centro de T es un punto circular entonces $T = T_1 T_2$, donde los factores son traslaciones con centros no circulares. Si $g \neq 1$ entonces gT_1 es un giro y $gT_1 T_2$ también. Si $g = 1$ tenemos el caso que hemos exceptuado. ■

Para terminar con las generalidades sobre la geometría parabólica introducimos las nociones de congruencia y semejanza:

Definición 10.19 Dos figuras en un plano parabólico son *semejantes* (resp. *congruentes*) si existe una semejanza (resp. isometría) que transforma una en otra.

A menudo resulta útil esta caracterización de las isometrías:

Teorema 10.20 Una semejanza que transforma una circunferencia en otra congruente es una isometría.

DEMOSTRACIÓN: Sea f una semejanza en las condiciones del teorema. Sea g una isometría entre la circunferencia indicada y su imagen por f . Entonces fg^{-1} es una semejanza que deja fija a una circunferencia. Basta probar que fg^{-1} es una isometría o, equivalentemente, podemos suponer que f fija a la circunferencia. Sea P un punto de la circunferencia. Existe una reflexión h cuyo eje pasa por el centro O de la circunferencia y que transforma $f(P)$ en P . Por lo tanto fh fija a P (y a O). Por el teorema 10.14 concluimos que fh es una reflexión, luego f es una isometría. ■

Como no estamos suponiendo una ordenación en el cuerpo no podemos hablar de congruencia de segmentos, pero sí de congruencia de vectores:

Diremos que dos vectores \overrightarrow{AB} y \overrightarrow{CD} son *congruentes* si lo son los pares de puntos (A, B) y (C, D) . Lo representaremos por $\overrightarrow{AB} \equiv \overrightarrow{CD}$.

La definición no depende de la elección de los puntos A y B , ya que si tenemos $\overrightarrow{AB} = \overrightarrow{A'B'}$, entonces la traslación de vector $\overrightarrow{AA'}$ es una congruencia entre (A, B) y (A', B') , luego no importa qué par tomamos para compararlo con el de otro vector.

Diremos que un vector \vec{v} es *isótropo* si lo son las rectas de dirección \vec{v} . Si existen puntos circulares entonces hay dos subespacios de vectores isótropos.

Teorema 10.21 *Si \vec{v} es un vector no isótropo, entonces los únicos vectores de $\langle \vec{v} \rangle$ congruentes con \vec{v} son \vec{v} y $-\vec{v}$. Dos vectores isótropos no nulos son congruentes.*

DEMOSTRACIÓN: Sea O un punto cualquiera y $P = O + \vec{v}$. Si \vec{v} no es isótropo podemos construir la circunferencia \mathcal{C} de centro O que pasa por P . La recta OP no puede ser tangente a \mathcal{C} , luego corta a la cónica en otro punto P' , de modo que la polar de O , es decir, la recta infinita, contiene al conjugado armónico de O respecto a P y P' , es decir, O es el punto medio entre P y P' o, equivalentemente, $P' = P - \vec{v}$. Una isometría que transforme el par (O, P) en otro par (O, Q) ha de fijar a O , luego también a \mathcal{C} , luego Q ha de estar sobre \mathcal{C} , si además \overrightarrow{OQ} ha de estar en $\langle \vec{v} \rangle$, es decir, si la recta OQ ha de ser OP , necesariamente Q ha de ser P o P' , con lo que tenemos lo pedido.

Para probar la segunda parte es suficiente suponer que AB y AC son las dos rectas isótropas y probar que $\overrightarrow{AB} \equiv \overrightarrow{AC}$.

Consideremos la reflexión f cuyo centro es el punto infinito de la recta BC y cuyo eje pasa por A . Toda reflexión conjuga los puntos circulares, luego $f[AB] = AC$ y, como la recta BC ha de quedar fija, necesariamente $f(B) = C$. ■

Teorema 10.22 *La circunferencia de centro O que pasa por un punto P está formada por todos los puntos Q tales que $\overrightarrow{OP} \equiv \overrightarrow{OQ}$ (más los puntos circulares, si los hay).*

DEMOSTRACIÓN: Si $\overrightarrow{OP} \equiv \overrightarrow{OQ}$ entonces existe una isometría que fija a O y que transforma P en Q . Ésta fijará la circunferencia, luego Q estará también en ella. Por otro lado, si Q es un punto de la circunferencia según 10.7 tenemos que existe una reflexión cuyo eje pasa por O y que transforma P en Q , luego $\overrightarrow{OP} \equiv \overrightarrow{OQ}$. ■

Ejercicio: Probar que por tres puntos no colineales entre sí ni con ningún punto circular pasa una única circunferencia.

Ejercicio: Probar que si P y Q son dos puntos de una circunferencia de centro O entonces existe un único giro de centro O que transforma P en Q .

10.2 El plano euclídeo

Definimos un *plano euclídeo* como un plano parabólico real cuya involución ortogonal es elíptica. En esta sección estudiaremos los planos euclídeos así definidos y probaremos, entre otras cosas, que son los mismos que ya tenemos definidos axiomática y algebraicamente. Las primeras observaciones adicionales que podemos hacer frente a la teoría general sobre planos parabólicos son las siguientes: no hay puntos circulares, luego las circunferencias son elipses, luego el centro de una circunferencia es un punto interior, luego toda recta que pasa por el centro de una circunferencia corta a ésta en dos puntos.

Acabamos de hacer uso de la ordenación del cuerpo, con la que tampoco contábamos en la teoría general. Otra consecuencia es que podemos hablar del segmento \overline{AB} determinado por dos puntos distintos A y B . Es claro que una afinidad transforma un segmento en otro si y sólo si hace corresponder sus extremos: Según la definición de congruencia de vectores es claro entonces que

$$\overline{AB} \equiv \overline{CD} \quad \text{si y sólo si} \quad \overrightarrow{AB} \equiv \overrightarrow{CD}.$$

El teorema siguiente introduce la proporcionalidad entre segmentos.

Teorema 10.23 Si \vec{u} y \vec{v} son dos vectores no nulos existe un único número real $\alpha > 0$ tal que $\vec{v} \equiv \pm \alpha \vec{u}$.

DEMOSTRACIÓN: Podemos representar los vectores como $\vec{u} = \overrightarrow{OA}$ y $\vec{v} = \overrightarrow{OB}$. Sea \mathcal{C} la circunferencia de centro O que pasa por A . Sea A' uno de los puntos donde OB corta a \mathcal{C} . Existe una reflexión cuyo eje pasa por O y que transforma A' en A , luego transforma B en un punto B' colineal con O y A . Así pues, $\vec{v} = \overrightarrow{OB} \equiv \overrightarrow{OB'} = \alpha \overrightarrow{OA} = \alpha \vec{u}$. El teorema 10.21 prueba que podemos cambiar α por $-\alpha$ para garantizar que es positivo, a la vez que nos da la unicidad. ■

Dados dos segmentos \overline{AB} y \overline{CD} definimos la *razón* entre ambos como el único número real $\alpha > 0$ tal que $\overrightarrow{AB} \equiv \pm \alpha \overrightarrow{CD}$. La representaremos por

$$\frac{\overline{AB}}{\overline{CD}}.$$

Si los segmentos son colineales podemos conservar el signo de la razón, pero este signo se pierde irremediabilmente cuando comparamos segmentos no colineales, pues podemos obtener uno u otro según la isometría que usemos para compararlos. De la propia definición se desprende que las isometrías conservan las razones entre segmentos.

Si fijamos un segmento unidad podemos asignar una longitud a cada segmento \overline{AB} como su razón respecto a la unidad establecida. La representaremos indistintamente por \overline{AB} o por $d(A, B)$, pues también es la distancia entre A y B . De nuevo es obvio que las isometrías conservan las distancias entre puntos. Más aún, se cumple $d(A, B) = d(C, D)$ si y sólo si $\overline{AB} \equiv \overline{CD}$.

En efecto, si \overline{PQ} es el segmento unidad y $\vec{u} = \overrightarrow{PQ}$, entonces $d(A, B)$ es el único $\alpha > 0$ tal que $\overrightarrow{AB} \equiv \pm \alpha \vec{u}$, luego si $d(C, D)$ es también α , entonces $\overrightarrow{AB} \equiv \overrightarrow{CD}$.

El teorema 10.22 puede expresarse ahora como que una circunferencia de centro O está formada por todos los puntos cuya distancia a O es un valor constante, al que llamamos *radio* de la circunferencia. El radio de una circunferencia es, pues, la distancia de su centro a uno cualquiera de sus puntos. Cada punto O y cada $r > 0$ determinan una única circunferencia de centro O y radio r .

Es obvio que las isometrías transforman cada circunferencia en otra de igual radio. Esto nos lleva a la caracterización natural de las isometrías:

Teorema 10.24 *Una biyección afín en un plano euclídeo es una isometría si y sólo si conserva las distancias entre puntos.*

DEMOSTRACIÓN: Si una biyección afín conserva las distancias entonces transforma circunferencias en circunferencias (lo que implica que es una semejanza), pero como también conserva el radio, ha de ser una isometría (por 10.20, pues dos circunferencias del mismo radio son congruentes por una traslación que haga corresponder sus centros). ■

Para relacionar todo esto con los planos euclídeos en el sentido algebraico hemos de expresar vectorialmente los conceptos que estamos manejando:

Definimos la *norma* de un vector $\vec{v} = \overrightarrow{AB}$ como $\|\vec{v}\| = d(A, B)$. Puesto que las isometrías (en particular las traslaciones) conservan las distancias es claro que esta definición no depende de los puntos escogidos como origen y extremo de \vec{v} . Si leemos en sentido inverso esta definición, tenemos que $d(A, B) = \|\overrightarrow{AB}\|$.

Es claro que $\vec{v} \equiv \vec{w}$ si y sólo si $\|\vec{v}\| = \|\vec{w}\|$, por lo que en realidad no necesitaremos hablar más de equivalencia de vectores.

Diremos que dos vectores forman una *base ortonormal* si son ortogonales y tienen norma 1. Un sistema de referencia afín $(O; \vec{u}, \vec{v})$ es *ortonormal* si su base es ortonormal.

Teorema 10.25 *Sean \vec{u} y \vec{v} dos vectores cuyas coordenadas en una base ortonormal sean (a, b) y (a', b') respectivamente. Entonces $\vec{u} \perp \vec{v}$ si y sólo si $aa' + bb' = 0$, y $\|\vec{u}\| = \sqrt{a^2 + b^2}$.*

Este teorema afirma que la perpendicularidad y la norma en un plano euclídeo en el sentido que aquí le hemos dado al término coinciden con los inducidos por el producto escalar dado por $\vec{u}\vec{v} = aa' + bb'$, donde (a, b) y (a', b') son las coordenadas de \vec{u} y \vec{v} en una base ortonormal prefijada. Probaremos este teorema juntamente con el que enunciamos a continuación, que implica en particular que todo espacio euclídeo en el sentido algebraico puede obtenerse de esta forma a partir de una involución apropiada en la recta infinita:

Teorema 10.26 *Dada una base en un plano afín real X , existe una única involución elíptica sobre la recta infinita de X de modo que la base es ortonormal respecto a la estructura parabólica inducida por ella.*

DEMOSTRACIÓN: Sea \vec{u}, \vec{v} una base arbitraria. Podemos representarla en la forma $\vec{u} = \overrightarrow{OA}, \vec{v} = \overrightarrow{OB}$. Consideramos las coordenadas cartesianas respecto a esta base, así como las coordenadas homogéneas asociadas, respecto a las cuales la recta infinita es $z = 0$.

Para que $\vec{u} \perp \vec{v}$ es necesario que los puntos infinitos de las rectas OA y OB sean conjugados respecto de la involución ortogonal. Sus coordenadas son $(1, 0, 0)$ y $(0, 1, 0)$.

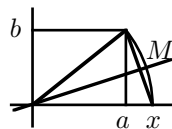
Para que $\|\vec{u}\| = \|\vec{v}\|$ es necesario que A y B estén sobre la misma circunferencia de centro O . Para ello han de ser conjugados respecto a una reflexión cuyo eje pase por O . Dicho eje ha de pasar también por el punto medio de $A(1, 0)$ y $B(0, 1)$, luego es la recta $y = x$, la cual ha de ser perpendicular a AB . Los puntos infinitos de estas rectas son $(1, 1, 0)$ y $(-1, 1, 0)$.

Por consiguiente la única involución posible es la determinada por los pares de conjugados $(1, 0, 0), (0, 1, 0)$ y $(1, 1, 0), (-1, 1, 0)$.

Los puntos $(1, 0, 0), (0, 1, 0)$ y $(1, 1, 0)$ determinan un sistema de referencia en la recta infinita respecto al cual las coordenadas homogéneas de un punto $(x, y, 0)$ son (x, y) . Una matriz de la involución ortogonal (vista como polaridad) determinada por los dos pares anteriores es la identidad, pues efectivamente conjugaba a ambos pares. De aquí se sigue que es elíptica.

Según el teorema 10.3 la ortogonalidad en X está determinada por la forma bilineal que en el sistema de referencia dado tiene matriz identidad. Esto significa que dos vectores de coordenadas (a, b) y (a', b') en la base dada son ortogonales si y sólo si $aa' + bb' = 0$. Vamos a probar que respecto a esta base y la involución que hemos fijado se cumplen las fórmulas del teorema 10.25, lo que probará que la base es efectivamente ortonormal. El teorema 10.25 estará así demostrado, pues dada una base ortonormal cualquiera en un plano euclídeo, la involución ortogonal será necesariamente la que hemos construido aquí, luego verificará las fórmulas pedidas.

Calculemos la norma de un vector de coordenadas (a, b) suponiendo que la unidad de longitud es \vec{u} . Si $b = 0$ se trata simplemente de $|a|$. Supongamos que $b \neq 0$. Consideramos la circunferencia de centro $(0, 0)$ y que pasa por (a, b) . Hemos de calcular el punto $(x, 0)$ en que ésta corta a la recta $y = 0$ y tal que $x > 0$. Los



puntos $(x, 0)$ y (a, b) son conjugados por una reflexión cuyo eje pasa por $(0, 0)$. Éste también ha de pasar por el punto medio entre ambos, que es

$$M = \left(\frac{a+x}{2}, \frac{b}{2} \right).$$

Dicho eje ha de ser perpendicular a la recta que pasa por los dos puntos, que tiene vector director $(a-x, b)$. Al desarrollar $M \perp (a-x, b)$ obtenemos

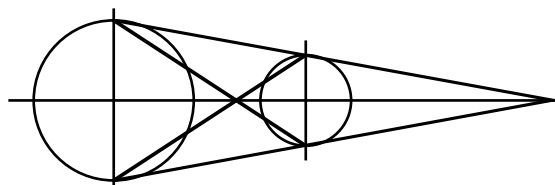
$$\frac{(a+x)(a-x)}{2} + \frac{b^2}{2} = 0,$$

luego $a^2 + b^2 = x^2$, o también, $x = \sqrt{a^2 + b^2}$. Esta fórmula vale igualmente si $b = 0$. Así pues, $\|(a, b)\| = \sqrt{a^2 + b^2}$. ■

A partir de aquí podemos contar con todos los resultados que conocemos sobre espacios euclídeos en el sentido algebraico, aunque el lector puede entretenerse deduciendo por su cuenta en este contexto las propiedades elementales de normas, distancias etc.

Sólo nos faltaría comprobar que las semejanzas tal y como las hemos definido aquí coinciden con las definidas en términos algebraicos. Lo dejamos como ejercicio.

Ejercicio: Probar que dos circunferencias cualesquiera son homotéticas. Si no son concéntricas existen dos homotecias que transforman una en la otra, salvo si tienen el mismo radio, en cuyo caso una de las homotecias tiene el centro en el infinito y es, por lo tanto, una traslación. Deducir que toda semejanza es la composición de una isometría con una homotecia.



A continuación señalaremos algunos hechos de interés sobre varios aspectos de la geometría euclídea.

La polaridad de una circunferencia Consideremos ahora una circunferencia \mathcal{C} de centro O y radio r . Vamos a estudiar su polaridad. Consideremos un punto $Q \neq O$ (pues ya sabemos que la polar de O es la recta infinita). Sea $r = p(Q)$ y P_∞ el punto infinito de r . Entonces la polar de P_∞ pasa por O (el polo de la recta infinita) y por Q (el polo de r). Así pues, se trata de OQ . El punto infinito de OQ es el conjugado de P_∞ por la involución inducida por \mathcal{C} en la recta infinita, es decir, la involución ortogonal. En otras palabras, r es perpendicular a OQ . Para determinar r basta conocer el punto donde ésta corta a OQ . Éste es el conjugado de Q respecto a la involución que \mathcal{C} induce

en OQ . Ésta está determinada por el hecho de que fija a los dos puntos en que OQ corta a \mathcal{C} . Tomemos un sistema de referencia afín (O, \vec{v}) de modo que \vec{v} tenga norma 1 y la dirección de OQ . Entonces los puntos de \mathcal{C} en OQ tienen coordenadas $\pm r$. La única involución que los fija es la que tiene ecuación

$$y = \frac{r^2}{x},$$

pues, efectivamente, ésta cumple las condiciones y es única. Vemos, por tanto, que si $p(Q) \cap OQ = R$, entonces $\|\vec{OQ}\| \|\vec{OR}\| = r^2$. Además Q y R están en la misma semirrecta respecto de O , y esto caracteriza completamente a la polar.

Definición 10.27 Diremos que dos puntos Q y R son mutuamente *inversos* respecto a la circunferencia de centro O y radio r si ambos se encuentran sobre la misma semirrecta de origen O y además $\|\vec{OQ}\| \|\vec{OR}\| = r^2$. Convenimos en que O es el inverso de todos los puntos infinitos.

En estos términos, la polar de un punto Q respecto a una circunferencia de centro O es la perpendicular a OP que pasa por el inverso de P .

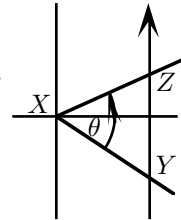
Una fórmula para la razón doble Vamos a probar una conocida fórmula para la razón doble de cuatro puntos A, B, C, D sobre una recta r . Si P es cualquier punto exterior a r entonces

$$\mathcal{R}(A, B, C, D) = \frac{\widehat{\text{sen } APC}}{\widehat{\text{sen } APD}} : \frac{\widehat{\text{sen } BPC}}{\widehat{\text{sen } BPD}}.$$

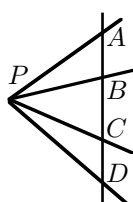
Es importante aclarar que los senos deben tomarse con el signo apropiado. Dados tres puntos no colineales X, Y, Z y fijado un sistema de referencia ortonormal de origen X existe un único ángulo $\theta \in]-\pi, \pi[$ tal que el giro de ángulo θ transforma la semirrecta \vec{XY} en \vec{XZ} . Definimos $\widehat{\text{sen } YXZ} = \text{sen } \theta$. El signo de θ depende del sistema de referencia, pero el miembro derecho de la fórmula es independiente de éste pues, si lo cambiamos, todos los signos se conservan o todos cambian, y en ambos casos la expresión queda inalterada.

Si en particular tomamos el sistema de referencia con el segundo eje paralelo a una recta r que no contenga a X , es fácil ver que podemos escoger una ordenación en r de modo que para todo par de puntos Y, Z en r el signo de $\widehat{\text{sen } YXZ}$ es positivo o negativo según si $Y < Z$ o $Z < Y$. Teniendo esto en cuenta así como la expresión

$$\mathcal{R}(A, B, C, D) = \frac{\overrightarrow{AC}}{\overrightarrow{AD}} : \frac{\overrightarrow{BC}}{\overrightarrow{BD}},$$



es claro que los signos de los dos miembros de la fórmula que queremos probar son iguales, y basta probar que los valores absolutos también lo son.



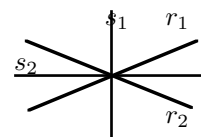
Recordemos que (ABC) es el área del triángulo \widehat{ABC} . Claramente tenemos

$$\begin{aligned}
 |\mathcal{R}(A, B, C, D)| &= \frac{\overline{AC}}{\overline{AD}} : \frac{\overline{BC}}{\overline{BD}} = \frac{(APC)(BPD)}{(APD)(BPC)} \\
 &= \left| \frac{\overline{AP} \overline{CP} \sin \widehat{APC} \cdot \overline{BP} \overline{DP} \sin \widehat{BPD}}{\overline{AP} \overline{DP} \sin \widehat{APD} \cdot \overline{BP} \overline{CP} \sin \widehat{BPC}} \right| \\
 &= \left| \frac{\sin \widehat{APC}}{\sin \widehat{APD}} : \frac{\sin \widehat{BPC}}{\sin \widehat{BPD}} \right|.
 \end{aligned}$$

■

Bisectrices Ahora caracterizaremos la bisectriz de un ángulo en términos proyectivos. Ante todo conviene definir la *razón doble* de cuatro rectas concurrentes como la de los puntos donde éstas cortan a una recta cualquiera. La elección de la recta es irrelevante, pues dos cuádruplas correspondientes a dos elecciones distintas están en perspectiva. Es claro que la razón doble es un invariante respecto a homografías entre haces. A su vez diremos que dos pares de rectas en un haz están separadas harmónicamente si su razón doble es -1 .

Dadas dos rectas cualesquiera r_1 y r_2 concurrentes en un punto P , por el teorema 9.62, que se traduce inmediatamente a haces de rectas, existen dos únicas rectas perpendiculares s_1 y s_2 que pasan por P harmónicamente conjugadas respecto a r_1 y r_2 . La reflexión de eje s_1 induce una involución en el haz que fija a s_1 y s_2 , luego ha de hacer corresponder r_1 con r_2 , y lo mismo vale para la reflexión de eje s_2 . Esto implica fácilmente que cada una de las rectas s_1 y s_2 es la bisectriz de dos de los ángulos opuestos por el vértice que determinan r_1 y r_2 .



Complexificación La involución ortogonal de un espacio euclídeo se extiende a una única involución en la recta infinita de su complexificación, que tendrá dos puntos circulares imaginarios conjugados. Es claro que las circunferencias, semejanzas e isometrías reales siguen siéndolo respecto de la geometría parabólica de la complexificación, por lo que identificaremos cada uno de estos conceptos con su extensión correspondiente.

Las coordenadas de los puntos circulares en una base ortonormal han de ser de la forma $(x, y, 0)$ de modo que $x^2 + y^2 = 0$, luego son concretamente $(1, \pm i, 0)$. De aquí se sigue que una cónica contiene a uno de ellos si y sólo si los contiene a los dos (la ecuación de la cónica tiene coeficientes reales y se satisface con las coordenadas de uno de los puntos, al aplicar la conjugación compleja obtenemos que también se satisface con el otro). Por lo tanto las circunferencias son las cónicas que pasan por al menos uno de los puntos circulares.

Ejercicio: Probar que los movimientos se diferencian de las simetrías en que los primeros fijan a los puntos circulares y las segundas los intercambian.

Veamos una importante fórmula debida a Cayley que relaciona el ángulo entre dos rectas con cierta razón doble imaginaria. Dadas dos rectas r_1 y r_2 concurrentes en un punto O , existen dos giros de centro O que transforman una en la otra. Fijado un sistema de referencia ortonormal de centro O , cada giro tiene asociado un ángulo módulo 2π (que en realidad es independiente del sistema de referencia salvo por el signo). Si uno de los giros que transforman r_1 en r_2 es el de ángulo α , el otro es el de ángulo $\alpha + \pi$. Definimos el ángulo entre r_1 y r_2 como éste valor de α , que está determinado módulo π , y lo representaremos por $\widehat{r_1 r_2}$. Si tomamos $-\pi/2 \leq \widehat{r_1 r_2} \leq \pi/2$ entonces el valor absoluto es el menor ángulo que forman las dos rectas, mientras que el signo indica si para girar r_1 hasta superponerla a r_2 —recorriendo el menor de los dos ángulos posibles—hemos de girar en sentido horario (argumentos decrecientes) o antihorario (argumentos crecientes). Este signo depende del sistema de referencia respecto al cual definimos los argumentos.

Una caracterización más manejable del ángulo entre dos rectas es la siguiente: tomamos un argumento θ_i de un punto de r_i distinto de O . Esto significa que el vector $(\cos \theta_i, \sin \theta_i)$ tiene la dirección de r_i . Es claro que θ_i está determinado módulo π . Su interpretación es que un giro de ángulo θ_i transforma el primer eje coordenado en r_i , luego un giro de ángulo $-\theta_i$ seguido de un giro de ángulo θ_2 transformará r_1 en el eje coordenado y luego en r_2 . Así pues, $\widehat{r_1 r_2} = \theta_2 - \theta_1$.

Considerando las coordenadas homogéneas asociadas al sistema de referencia elegido, es claro que los puntos infinitos de r_i son $(\cos \theta_i, \sin \theta_i, 0)$. Por otra parte, en la prueba del teorema 10.25 hemos visto que en el sistema de referencia formado por los puntos $(1, 0, 0)$, $(0, 1, 0)$, $(1, 1, 0)$, la involución ortogonal, vista como polaridad, tiene por matriz la identidad, luego las coordenadas de los puntos circulares satisfacen la ecuación $x^2 + y^2 = 0$. Por consiguiente los puntos circulares tienen coordenadas $(1, -i, 0)$ y $(1, i, 0)$. Sean I_1 e I_2 las rectas imaginarias que pasan por O y por los puntos circulares.

Calculamos la razón doble $\mathcal{R}(r_1, r_2, I_1, I_2)$ usando los puntos donde las cuatro rectas cortan a la recta infinita, que en coordenadas cartesianas son:

$$\frac{\sin \theta_1}{\cos \theta_1}, \quad \frac{\sin \theta_2}{\cos \theta_2}, \quad -i, \quad i.$$

Si llamamos $\theta = \theta_2 - \theta_1 = \widehat{r_1 r_2}$ tenemos que

$$\begin{aligned} \mathcal{R}(r_1, r_2, I_1, I_2) &= \frac{-i - \frac{\sin \theta_1}{\cos \theta_1}}{i - \frac{\sin \theta_1}{\cos \theta_1}} : \frac{-i - \frac{\sin \theta_2}{\cos \theta_2}}{i - \frac{\sin \theta_2}{\cos \theta_2}} \\ &= \frac{\cos \theta_1 \cos \theta_2 + \sin \theta_1 \sin \theta_2 + i(\cos \theta_1 \sin \theta_2 - \sin \theta_1 \cos \theta_2)}{\cos \theta_1 \cos \theta_2 + \sin \theta_1 \sin \theta_2 - i(\cos \theta_1 \sin \theta_2 - \sin \theta_1 \cos \theta_2)} \\ &= \frac{\cos(\theta_2 - \theta_1) + i \sin(\theta_2 - \theta_1)}{\cos(\theta_2 - \theta_1) - i \sin(\theta_2 - \theta_1)} = \cos^2 \theta - \sin^2 \theta + 2i \cos \theta \sin \theta \\ &= \cos 2\theta + i \sin 2\theta. \end{aligned}$$

En definitiva:

$$\widehat{r_1 r_2} = \frac{1}{2} \arg \mathcal{R}(r_1, r_2, I_1, I_2),$$

que está determinado módulo π , como ya sabíamos.¹ Es claro que el segundo miembro de esta fórmula es independiente del sistema de referencia salvo por un detalle: si cambiamos de orden I_1, I_2 la razón doble se invierte y su argumento cambia de signo. Para que la expresión de la derecha nos dé el ángulo entre las dos rectas con el signo adecuado a un sistema de referencia dado hemos de llamar I_1 al punto circular cuyas coordenadas son $(-i, 1, 0)$ y no $(i, 1, 0)$. Notar además que el segundo miembro no varía si sustituimos una recta r_i por una paralela.

Ejercicio: Probar mediante la fórmula de Cayley que $\widehat{r_1 r_3} = \widehat{r_1 r_2} + \widehat{r_2 r_3}$.

Ejercicio: Probar que $\widehat{r_1 r_2} + \widehat{r_2 r_3} + \widehat{r_3 r_1} = \pi$ y deducir que los ángulos de un triángulo suman dos rectos.

Dualización de teoremas euclídeos Terminamos con un ejemplo de una técnica muy general que permite obtener nuevos teoremas euclídeos a partir de otros conocidos. Se trata de generalizarlos a teoremas proyectivos, considerar los enunciados duales y particularizar los resultados de nuevo al plano euclídeo. Consideremos por ejemplo:

Las alturas de un triángulo se cortan en un punto.

Ejercicio: Deducir este hecho del apartado 3 del teorema 8.54.

El siguiente enunciado proyectivo es equivalente al teorema clásico que acabamos de citar:

Si T es un triángulo en un plano real, r una recta que no contiene a ninguno de sus vértices e I una involución elíptica en r , entonces las rectas que unen cada vértice de T con el conjugado de la intersección con r del lado opuesto son concurrentes.

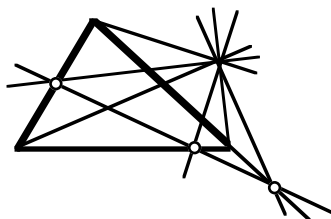
Para probarlo basta tomar a r como recta infinita y a I como involución ortogonal. Entonces las rectas descritas son las alturas del triángulo. Ahora consideramos el enunciado dual. Notemos que una polaridad transforma una involución en una recta en una involución en el haz de rectas de su polo. Así pues, queda:

Si T es un triángulo en un plano real, O un punto exterior a sus lados e I una involución elíptica en el haz de rectas de O , entonces los puntos de corte entre cada lado de T y la recta conjugada a la que une O y el vértice opuesto son colineales.

¹La expresión habitual de esta fórmula es $\widehat{r_1 r_2} = -\frac{i}{2} \log \mathcal{R}(r_1, r_2, I_1, I_2)$, donde \log representa al logaritmo complejo, que nosotros no hemos definido.

Ahora particularizamos este enunciado al caso en que I es la involución en O que hace corresponder cada recta con su perpendicular (la inducida en el haz de O desde la recta infinita). El resultado es relativamente simple:

Dado un triángulo T y un punto O no contenido en ninguno de sus lados, los puntos donde cada lado corta a la perpendicular a la recta que une O con el vértice opuesto son colineales.



Obtenemos así un enunciado “dual” cuya prueba directa no es evidente. De hecho por este método de generalización proyectiva y dualización se han obtenido teoremas euclídeos que hubiera sido difícil descubrir de otro modo.

10.3 El plano de Lorentz

Aquí estudiaremos el plano parabólico real cuando la involución ortogonal es hiperbólica. Antes de particularizar a \mathbb{R} es interesante hacer una observación general. En cada plano proyectivo existe una única geometría en estas condiciones pues, dadas dos involuciones hiperbólicas en la recta infinita, una homografía que transforme los puntos fijos de una en los puntos fijos de la otra transformará de hecho una involución en la otra, por lo que cualquier extensión al plano conservará todas las propiedades definibles a partir de dichas involuciones, y por lo tanto las geometrías respectivas tendrán exactamente las mismas características. A los planos parabólicos cuya involución ortogonal es hiperbólica se les llama en general *planos artinianos*.² El plano artiniano real recibe el nombre de *plano de Lorentz* y se trata del primer ejemplo que nos encontramos de geometría no euclídea.³ Uno de los motivos por los que este plano resulta de interés es porque contiene esencialmente la geometría de la teoría de la relatividad especial (en realidad la geometría de la relatividad especial es la del espacio de Minkowski, que es el análogo tetradimensional del plano de Lorentz, pero las características más importantes se encuentran igualmente en dos dimensiones).

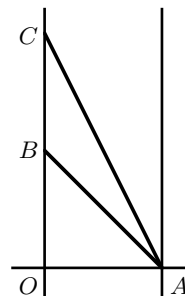
Para interpretar adecuadamente la geometría del plano de Lorentz conviene conocer algunos conceptos elementales sobre la cinemática relativista. El precio que hay que pagar por trabajar en dos dimensiones en lugar de cuatro es que hemos de suponer que el espacio físico tiene una dimensión en lugar de tres. Definimos un *suceso* como un hecho puntual tanto en el espacio como en el tiempo, es decir, un suceso es algo que sucede en un punto dado y en un momento dado. Imaginemos un observador dotado de un reloj y capaz de medir la distancia entre dos puntos cualesquiera. Éste puede identificar cada suceso con un par de números reales (x, t) , donde x representa

²Artin los llamó planos hiperbólicos, pero esta nomenclatura no encaja con la tradicional en geometría proyectiva, donde un plano hiperbólico es otra cosa.

³En cierto sentido el plano proyectivo real es también un ejemplo de geometría no euclídea, pues contradice el axioma de las paralelas, pero podemos considerarlo como una extensión del espacio euclídeo usual, cosa que no puede decirse del plano de Lorentz.

la distancia entre el observador y el punto donde ocurrió el suceso y t es el instante en qué ocurrió según su reloj. El signo de x distingue a los dos puntos del espacio (rectilíneo) situados a la misma distancia, mientras que el signo de t es positivo si el suceso ocurrió después del instante 0 y negativo en caso contrario.

De este modo, el observador puede representar la totalidad de los sucesos en un plano real. La figura muestra algunos ejemplos. El punto O es el suceso $(0, 0)$, es decir, el punto del espacio donde se encontraba el observador en el instante en que puso a 0 su reloj, juntamente con dicho instante. La recta vertical que pasa por O está formada por los sucesos de la forma $(0, t)$, es decir, son los sucesos de “la vida” del observador. A esta recta se la llama *línea de universo* del observador, pues son los sucesos del universo que el observador experimenta directamente. Cada recta horizontal representa los sucesos de la forma (x, t_0) , para un t_0 fijo, es decir, son todos los sucesos acaecidos en un mismo instante t_0 , por lo que las rectas horizontales se llaman *rectas de simultaneidad*. La recta paralela que aparece en la figura es la línea de universo de otro observador⁴ que permanece en reposo respecto al primero. Este observador medirá los sucesos con su propio criterio, por lo que las coordenadas de un mismo suceso serán distintas según el observador. Si suponemos que ambos usan las mismas unidades de longitud y de tiempo, y que ambos toman el mismo instante como origen de tiempos, el suceso (x, t) del primer observador será $(x + x_0, t)$, para el segundo, donde x_0 es la posición del primer observador medida por el segundo. Si admitimos que los observadores podrían haber adoptado criterios distintos para determinar qué semirrecta es positiva y cuál negativa, o si los observadores parten de distintos orígenes temporales o usan unidades distintas la ecuación tendrá algunas constantes adicionales, pero no vamos a entrar en ello. Lo único importante ahora es comprender que los pares (x, t) son en realidad las coordenadas de un suceso respecto a un sistema de referencia determinado por el observador.

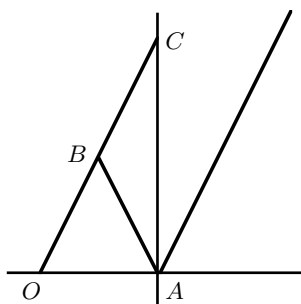


Consideremos ahora la recta AB de la figura. Se trata de la línea de universo de un tercer observador que se acerca al primero. El suceso B representa el encuentro de ambos observadores. En principio la línea de universo de un observador O' , desde el punto de vista de un observador O , no tiene por qué ser una recta. El hecho de que lo sea se traduce, por el teorema de Tales, en que O ve a O' recorrer espacios iguales en tiempos iguales, es decir, que O' se acerca o aleja de O con velocidad constante. Si la línea de universo de O' tiene ecuación $x = vt + x_0$, entonces la constante v es la velocidad con que O' se acerca o aleja de O . En el diagrama la velocidad de un observador es la inversa de la pendiente de su línea de universo. Por ejemplo, la recta AC representa a otro observador que partió del mismo punto que AB pero con la mitad de velocidad, de modo que le cuesta el doble de tiempo encontrarse con O . De este modo, para un observador O las rectas verticales son líneas de universo de observadores en reposo respecto a O , mientras que las rectas oblicuas representan observadores en movimiento uniforme, tanto más rápido cuanto más se acercuen a la horizontal. Las rectas horizontales (líneas de simultaneidad) son las únicas que no pueden ser líneas de universo de ningún observador, pues nadie puede estar simultáneamente en todos los puntos del espacio.

⁴El término “observador” no debe interpretarse literalmente. Por ejemplo, este segundo observador podría ser una silla situada a una cierta distancia de nosotros. Cuando hablamos de la forma en que el observador mide un suceso queremos decir la forma en que lo haría un observador hipotético sentado en la silla.

Ahora llegamos al punto crucial, y es determinar la relación entre las coordenadas espaciotemporales de dos observadores que se aproximan o se alejan entre sí. Por ejemplo, cabe esperar que el diagrama espaciotemporal del observador AC de la figura anterior sea el dado por la figura adjunta.

Observamos que la pendiente de AB ha aumentado, lo que significa que su velocidad ha disminuido. Esto es lógico, pues AC estaba “persiguiendo” a AB , luego la velocidad con la que el éste se aleja de aquél es en realidad la diferencia entre sus velocidades.



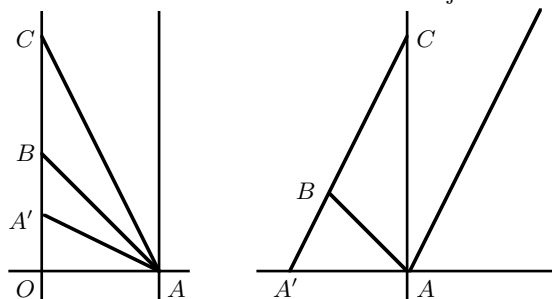
Ejercicio: Supongamos que un observador O ve la línea de universo de otro observador O' como la dada por la ecuación $x' = vx + x_0$. Determinar la relación entre las coordenadas (x, t) de un suceso para O y las coordenadas (x', t') del mismo suceso respecto de O' . Suponer que ambos observadores usan las mismas unidades y la misma orientación izquierda/derecha, pero no necesariamente el mismo origen de tiempos.

Pues bien, el último esquema es incorrecto según la teoría especial de la relatividad. Supongamos que las dos rectas verticales del primer diagrama representan las líneas de universo de dos estrellas que distan un año luz. La recta AC es la línea de universo de una nave que parte de una estrella (o de uno de sus planetas) con rumbo a la otra. Su velocidad es la mitad de la velocidad de la luz, con lo que el viaje le cuesta dos años. En el momento de su partida, la nave envía un mensaje de radio a su destino, que viaja a la velocidad de la luz y, por consiguiente, es recibido un año después. La línea de universo del mensaje es la recta AB . Si la segunda figura reflejara realmente las observaciones desde la nave, tendríamos que concluir que desde ella el mensaje de radio se ha desplazado a la mitad de la velocidad de la luz. Un mensaje similar que hubiera sido enviado en sentido contrario se habría visto moverse a 1,5 veces la velocidad de la luz. Esto tendría consecuencias muy interesantes, pues un observador podría determinar su velocidad midiendo la velocidad de la luz en distintas direcciones. Pensemos por ejemplo en la Tierra, que sigue al Sol en un movimiento a través del espacio a una velocidad desconocida en principio. Comparando la posición del Sol con la de otras estrellas es posible determinar aproximadamente la dirección y la velocidad en la que el Sol se mueve por la Galaxia, pero esto es sólo aproximado, pues las estrellas también se mueven. Si en un laboratorio midiéramos la velocidad de la luz en direcciones distintas, un cálculo sencillo nos permitiría determinar la velocidad absoluta de la tierra, así como la dirección en que se mueve.

Esto se ha intentado en la práctica. Se trata de los famosos experimentos de Michelson y Morley (1881–1887) cuyo resultado fue completamente negativo. Estas y muchas otras experiencias confirman que la velocidad de la luz es la misma en todas direcciones y respecto a todos los observadores. En realidad hay que hacer una precisión: un observador es *inercial* si no está sometido a la acción de ninguna fuerza. El *principio de relatividad* afirma que todos los observadores inerciales son equivalentes, en el sentido de que las leyes de la física se expresan con las mismas fórmulas respecto a cualquiera de ellos, y en particular la velocidad de la luz es la misma en todos ellos. El diagrama correcto será el de la figura siguiente.

Al exigir que la línea de universo del rayo de luz tenga la misma pendiente en los dos diagramas nos encontramos con que los dos observadores discrepan en muchos aspectos. Antes veamos en qué coinciden: ambos observadores estarán de acuerdo en que su encuentro se produjo en el instante C , el segundo observador está legitimado

a afirmar que su partida se produjo en A y el segundo puede afirmar que recibió el mensaje en B . Todos estos hechos son absolutos, en el sentido de que es absurdo discutirle al primer observador cuándo recibió él el mensaje.



Sin embargo el primer observador afirmará que el tiempo transcurrido entre A y B es el mismo que entre A y C (un año), por lo que indirectamente deducirá que el suceso A es simultáneo con el suceso O . Por el contrario, el segundo observador replicará que el tiempo transcurrido entre B y C es el doble que entre A y B , por lo que considerará que el instante A de su partida es simultáneo con un instante A' que, desde el punto de vista del primer observador, es medio año posterior a O . Así pues, cada observador tendrá su propia línea de simultaneidad, OA para el primero, $A'A$ para el segundo. Un análisis más detallado permite probar que ambos observadores diferirán también en la distancia recorrida y en el tiempo que ha durado el viaje, pero no entraremos ahora en ello. Insistiendo en la cuestión de la simultaneidad, vemos que una forma en que el primer observador podría probar que es él quien tiene razón sería que un tercer observador partiera del mismo punto que el segundo pero al doble de la velocidad de la luz. Si así se hiciera, encontraría al primer observador en el suceso A' , y podría atestiguar que el segundo observador ya había partido cuando él salió, luego éste no podría mantener consistentemente que partió simultáneamente a A' . Más aún, la línea de universo de este tercer observador sería AA' , luego si el segundo quisiera mantener su postura debería admitir que el tercero ha ido de A hasta A' en tiempo cero, lo cual es absurdo.

Para que la teoría de la relatividad sea consistente es necesario que este tercer observador sea físicamente imposible o, equivalentemente, que nada pueda viajar al doble de la velocidad de la luz. Cambiando las condiciones del ejemplo se ve que es necesario que nada pueda superar la velocidad de la luz. De este modo, sucesos situados a tal distancia espaciotemporal que ni siquiera una onda luminosa pueda ir de uno a otro (como es el caso de A y A' o de A y O) no están conectados causalmente, por lo que no existe ningún criterio objetivo en virtud del cual alguien pueda decir que uno es anterior o posterior a otro en el tiempo. Ya no es cierto, pues, que cualquier línea no horizontal pueda ser una línea de universo. Sólo pueden serlo aquellas cuya pendiente no sea inferior a la de los rayos de luz. Puede probarse que cualquier línea por debajo de las de los rayos de luz es la línea de simultaneidad de un observador adecuado. Todo esto puede deducirse del principio de relatividad, pero en lugar de ello vamos a mostrar directamente la conclusión, es decir, veremos que la geometría del plano de Lorentz es consistente con los postulados de la relatividad especial.

Espacio y tiempo Sea X un plano de Lorentz. Su recta infinita tiene dos puntos circulares C_1 y C_2 , que la dividen en dos segmentos. A los puntos de

uno de estos segmentos distintos de sus extremos los llamaremos *puntos infinitos espaciales*, y a los del otro *puntos infinitos temporales*. Con esto las rectas quedan divididas en tres categorías: *isótropas*, *espaciales* y *temporales*, según que su punto infinito sea un punto circular, un punto espacial o un punto temporal. Dos rectas paralelas son del mismo tipo. Un segmento será isótropo, espacial o temporal según lo sea la recta que lo contiene. Un vector no nulo \vec{v} es isótropo, *espacial* o *temporal* si lo son las rectas de dirección \vec{v} . Si r es una recta no isótropa y P es su punto infinito, su conjugado ortogonal es su conjugado armónico respecto a los puntos circulares, luego éstos separan (armónicamente) a P e $I(P)$. Esto significa que P es espacial si y sólo si $I(P)$ es temporal, luego las rectas perpendiculares a rectas espaciales son temporales y viceversa.⁵

Veamos la interpretación de estos conceptos en términos de la geometría afín al tiempo que introducimos una precisión adicional: la distinción entre pasado y futuro e izquierda y derecha.

Definición 10.28 Una *orientación* es un par ordenado (\vec{v}_1, \vec{v}_2) de vectores isótropos linealmente independientes.

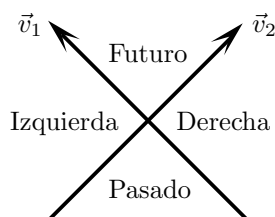
Una orientación así definida divide a los vectores no isótropos en cuatro regiones angulares, según que sus coordenadas en la base \vec{v}_1, \vec{v}_2 tengan signos

$$(+, +), \quad (-, -), \quad (+, -) \quad \text{o} \quad (-, +).$$

Fijado un punto O , el sistema de referencia $(O, \vec{v}_1, \vec{v}_2)$ determina un sistema de coordenadas homogéneas respecto al cual la recta infinita es $z = 0$. Los puntos $(1, 0, 0)$, $(0, 1, 0)$ y $(1, 1, 0)$ determinan un sistema de referencia en la recta infinita respecto al cual las coordenadas de $(x, y, 0)$ son (x, y) . Si un vector tiene coordenadas (a, b) en la base \vec{v}_1, \vec{v}_2 , entonces la recta que pasa por O con dirección dicho vector es $-bx + ay = 0$, y su punto infinito es $(a, b, 0)$, que respecto a la recta infinita es (a, b) y en coordenadas cartesianas es a/b . Así pues, los vectores de las clases $(+, +)$ y $(-, -)$ corresponden a rectas cuyos puntos infinitos tienen la coordenada cartesiana positiva y los de las otras dos clases a puntos cuya coordenada cartesiana es negativa. Estas clases son los dos segmentos que determinan en la recta infinita los puntos $P_0 = (0, 1)$ y $P_\infty = (1, 0)$, es decir, los puntos circulares.

En general hemos tomado como puntos infinitos espaciales a los de uno de estos dos segmentos, arbitrariamente. Respecto de una orientación dada, tomaremos como puntos infinitos espaciales a los del segmento determinado por los vectores de las clases $(+, -)$ y $(-, +)$, de modo que éstos serán precisamente los vectores espaciales, y los vectores temporales serán los de las clases $(+, +)$ y $(-, -)$. Ahora diremos que un vector espacial *apunta a la derecha* o a la *izquierda* según si está en la clase $(-, +)$ o $(+, -)$. Un vector temporal *apunta al futuro* o al *pasado* según si está en la clase $(+, +)$ o $(-, -)$.

⁵Las rectas isótropas serán las líneas de universo de los rayos de luz, las rectas temporales serán las líneas de universo de observadores inerciales y las rectas espaciales serán las líneas de simultaneidad. Más exactamente: las líneas de simultaneidad de un observador serán las perpendiculares a su línea de universo.



Notar que podríamos haber empezado por el concepto de orientación y haber definido a partir de aquí las rectas espaciales y temporales, pero sucede que estos conceptos juegan un papel central en la geometría del plano de Lorentz, mientras que la orientación y la consecuente distinción entre pasado y futuro e izquierda y derecha es secundaria. Por ello hemos de tener presente que la distinción espacial/temporal depende sólo de la distinción entre los dos segmentos en que los puntos circulares dividen a la recta infinita, pero no de una orientación en particular.

Semejanzas e isometrías Veamos en qué grado respetan las distinciones entre espacio y tiempo las distintas clases de aplicaciones del plano de Lorentz. Consideremos un punto O y sea P otro punto tal que la recta OP no sea isótropa. Digamos, por ejemplo, que es temporal, aunque todo vale igual si es espacial. Sabemos que existe una única circunferencia \mathcal{C} de centro O que pasa por P . Ésta será una hipérbola que pasa por los puntos circulares, luego las rectas isótropas que pasan por P serán sus asíntotas. La recta OP es secante a \mathcal{C} , luego \mathcal{C} la divide en dos segmentos, uno formado por puntos exteriores (el que contiene a O) y otro formado por puntos interiores (el que contiene a su punto infinito, que es temporal). Por otro lado, \mathcal{C} también divide a la recta infinita en un segmento de puntos interiores y otro de puntos exteriores, pero estos segmentos tienen por extremos los puntos circulares, luego uno es el de los puntos espaciales y otro el de los puntos temporales. Como acabamos de ver que existe un punto interior temporal, vemos que los puntos infinitos interiores son exactamente los temporales. Así pues, todas las rectas temporales cortan a \mathcal{C} , mientras que las espaciales son exteriores.

Una isometría que fije a O ha de fijar a \mathcal{C} , luego ha de transformar la recta OP en una recta OP' tal que $P' \in \mathcal{C}$, luego la recta OP' ha de ser temporal. Recíprocamente, toda recta temporal es de la forma OP' , con $P' \in \mathcal{C}$, luego existe una isometría que fija a O y transforma P en P' , luego OP en OP' .

Teniendo en cuenta que las traslaciones conservan el carácter espacial y temporal, podemos suprimir la restricción de que las isometrías fijen a O , y así concluimos:

Teorema 10.29 *Dos rectas del plano de Lorentz son del mismo tipo (espaciales, temporales o isótropas) si y sólo si son congruentes.*

En particular, y a diferencia de lo que sucede en un plano euclídeo, no todas las rectas son congruentes. Más en concreto: una recta no es congruente con ninguna de sus perpendiculares. Por otra parte, dadas dos rectas no isótropas

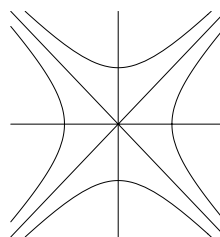
AB y $A'B'$, es claro que la homografía $ABC_1C_2 \xrightarrow{\sim} A'B'C_1C_2$ es una semejanza que transforma una en otra, luego todas las rectas no isótropas son semejantes.

Ejercicio: Probar que en el plano de Lorentz no es cierto que toda semejanza sea composición de una homotecia y una isometría.

Diremos que una circunferencia es *temporal* si las rectas que la cortan y pasan por su centro son temporales o, equivalentemente, si los puntos infinitos interiores a ella son temporales. Similarmente definimos una circunferencia *espacial*. Hemos probado que toda circunferencia es espacial o temporal y es claro que dos circunferencias congruentes han de ser del mismo tipo. En cambio, todas las circunferencias son semejantes.

Una rama de una hipérbola es uno de los arcos en que la dividen los puntos infinitos. Por definición, dos cualesquiera de sus puntos A y B no están separados por los puntos infinitos, luego la recta que los une corta a la recta infinita en un punto exterior. Esto significa que el segmento afín \overline{AB} (el que no contiene al punto infinito) está formado por puntos interiores, luego no contiene puntos de las asíntotas. Esto significa que la rama está contenida en uno de los ángulos determinados por las asíntotas. Un argumento similar prueba que puntos de ramas distintas están en semiplanos distintos respecto a ambas asíntotas, luego concluimos que las ramas de una hipérbola están en dos de los ángulos opuestos por el vértice determinados por sus asíntotas.

De aquí se sigue fácilmente que, respecto de una orientación dada, de las dos ramas de una circunferencia espacial de centro O , una está formada por los puntos P tales que \overrightarrow{OP} apunta hacia la izquierda y otra por los puntos P tales que \overrightarrow{OP} apunta hacia la derecha. Similarmente para circunferencias temporales. La figura muestra dos circunferencias del mismo centro. Una temporal, con su rama futura en la parte superior y su rama pasada en la inferior, y otra espacial, con su rama izquierda a la izquierda y su rama derecha a la derecha.



Veamos ahora el efecto sobre la orientación de las semejanzas e isometrías. Sea \vec{v}_1, \vec{v}_2 una orientación de un plano de Lorentz. Si f es una semejanza, entonces f fija o invierte los puntos circulares, luego $\vec{f}(\vec{v}_1)$ ha de ser un múltiplo de \vec{v}_1 o de \vec{v}_2 y lo mismo sucede con $\vec{f}(\vec{v}_2)$. Supongamos que $\vec{f}(\vec{v}_1) = \alpha \vec{v}_2$ y $\vec{f}(\vec{v}_2) = \beta \vec{v}_1$, con $\alpha > 0$ y $\beta < 0$. Entonces es fácil ver que \vec{f} transforma cada vector que apunta hacia el futuro en un vector que apunta hacia la derecha, cada vector que apunta hacia el pasado en un vector que apunta hacia la izquierda, cada vector que apunta hacia la derecha en un vector que apunta hacia el pasado y cada vector que apunta hacia la izquierda en un vector que apunta hacia el futuro. Podemos analizar así todas las posibilidades, pero lo importante es que en cualquier caso \vec{f} actúa igual sobre todos los vectores de un mismo tipo, es decir, que si \vec{f} transforma un vector que apunta hacia el futuro en un vector que apunta hacia la izquierda, entonces transforma a todos los vectores que apuntan hacia el futuro en vectores que apuntan hacia la izquierda, etc.

Restrinjámonos ahora a isometrías. Si una reflexión f tiene, por ejemplo, eje espacial, es claro que conservará la izquierda y la derecha. En cambio, si \vec{v} es perpendicular al eje, tendremos $f(\vec{v}) = -\vec{v}$, luego los vectores que apuntan hacia el futuro se transforman en vectores que apuntan hacia el pasado y viceversa. Diremos entonces que f conserva la orientación espacial pero no la temporal. Similarmente, las reflexiones con eje temporal conservan la orientación temporal pero no la espacial. En cualquier caso, una reflexión nunca conserva la orientación.

Si componemos dos reflexiones que conserven ambas la orientación espacial o ambas la temporal obtendremos un giro que conserva la orientación, mientras que si componemos una reflexión que conserva la orientación espacial con otra que conserve la orientación temporal obtendremos un giro que no conserva ninguna de las dos orientaciones. A los primeros los llamaremos *giros orientados*. Las isometrías que conservan la orientación son las composiciones de traslaciones con giros orientados, y forman un subgrupo del grupo de los movimientos.

Distancia A continuación nos ocupamos del problema de definir una distancia entre puntos. No podemos definir la proporción entre dos vectores cualesquiera, pues un vector espacial no puede ser congruente con ningún vector temporal. Sin embargo sí podemos comparar dos vectores del mismo tipo. El análogo al teorema 10.23 es ahora el siguiente:

Teorema 10.30 *Si \vec{u} y \vec{v} son dos vectores ambos espaciales o ambos temporales existe un único número real $\alpha > 0$ tal que $\vec{v} \equiv \pm\alpha\vec{u}$.*

DEMOSTRACIÓN: Podemos representar los vectores como $\vec{u} = \overrightarrow{OA}$ y $\vec{v} = \overrightarrow{OB}$. Supongamos, por ejemplo, que ambos son temporales. Sea \mathcal{C} la circunferencia de centro O que pasa por A . Claramente la circunferencia es temporal, luego la recta OB ha de cortarla en un punto A' . Como A y A' están ambos sobre \mathcal{C} , existe una isometría que fija a O y transforma A' en A . Sea B' la imagen de B por dicha isometría, que obviamente está sobre la recta OA . Por consiguiente $\vec{v} \equiv \overrightarrow{OB'} = \alpha\vec{u}$. El teorema 10.21 prueba que podemos cambiar α por $-\alpha$ para garantizar que es positivo, a la vez que nos da la unicidad. ■

Esto nos permite definir la *razón* entre dos segmentos \overline{AB} y \overline{CD} , ambos del mismo tipo, como el único escalar positivo

$$\alpha = \frac{\overline{AB}}{\overline{CD}},$$

tal que $\overrightarrow{AB} \equiv \pm\alpha\overrightarrow{CD}$. Si fijamos *dos* unidades de longitud, una espacial y otra temporal, podemos definir la *longitud* de un segmento arbitrario no isótropo como su razón respecto a la unidad correspondiente, espacial o temporal, y la *norma* de un vector no isótropo $\vec{v} = \overrightarrow{AB}$ como la longitud del segmento \overline{AB} . La representaremos también por $\|\vec{v}\|$. Como en el caso euclídeo se prueba ahora que dos segmentos o vectores ambos espaciales o ambos temporales son congruentes si y sólo si tienen la misma longitud (o norma). Respecto a los

segmentos y vectores isótopos (no nulos), sabemos por el teorema 10.21 que todos son congruentes entre sí. Si convenimos en que los segmentos isótopos tienen todos longitud nula y que los vectores isótopos tienen todos norma nula, podemos extender el resultado anterior al caso isótopo, cuyo enunciado queda como sigue:

Teorema 10.31 *Dos segmentos o (vectores no nulos) del mismo tipo (espaciales, temporales o isótopos) son congruentes si y sólo si tienen la misma longitud (la misma norma).*

Dado un punto O y un $r > 0$, los puntos cuya distancia a O es igual a r (más los puntos circulares) forman dos circunferencias de centro O y radio r . Un punto P está en una u otra según si la recta OP es espacial o temporal. De aquí se sigue a su vez:

Teorema 10.32 *Una biyección afín es una isometría si y sólo si conserva el carácter espacial, temporal o isótopo de los segmentos, así como su longitud.*

DEMOSTRACIÓN: Si conserva los vectores isótopos entonces conserva la involución ortogonal, luego es una semejanza. Si conserva las longitudes de segmentos transforma cada circunferencia \mathcal{C} en otra circunferencia \mathcal{C}' del mismo radio. Si además conserva el carácter de los segmentos entonces \mathcal{C} y \mathcal{C}' son ambas espaciales o ambas temporales. Pero dos circunferencias del mismo radio y del mismo carácter son congruentes (a través de una traslación que haga corresponder sus centros), luego por 10.20 la afinidad es una isometría. ■

Producto escalar Vamos a ver que el plano de Lorentz puede ser descrito en términos algebraicos de forma análoga al plano euclídeo.

Definición 10.33 Un *sistema de referencia ortonormal* es un sistema de referencia $(O; \vec{v}_1, \vec{v}_2)$ tal que $\vec{v}_1 \perp \vec{v}_2$, $\|\vec{v}_1\| = \|\vec{v}_2\| = 1$, \vec{v}_1 es espacial y \vec{v}_2 es temporal. El sistema de referencia está *orientado* si el vector espacial apunta hacia la derecha y el temporal hacia el futuro.

Sabemos que la ortogonalidad de vectores puede expresarse en términos de una forma bilineal simétrica regular F . Su matriz en un sistema de referencia ortonormal será de la forma

$$A = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}.$$

(Podemos exigir un 1 porque la forma bilineal está determinada salvo un factor escalar. Con esta restricción, la forma F es única.) Ahora bien, no todas las matrices de este tipo son posibles. Es necesario que impliquen la existencia de vectores isótopos. Un vector de coordenadas (x, t) será isótopo si cumple

$$(x, t)A \begin{pmatrix} x \\ t \end{pmatrix} = 0,$$

es decir, si $x^2 + at^2 = 0$. Esta ecuación tiene dos soluciones si y sólo si $a < 0$. Por lo tanto la matriz de la forma bilineal ha de ser

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -c^2 \end{pmatrix}, \quad \text{con } c > 0.$$

Entonces los vectores isótropos son los que cumplen $x^2 - c^2t^2 = 0$, es decir $x = \pm ct$. Las rectas isótropas serán las que en el sistema de referencia considerado tienen ecuaciones de la forma $x = \pm ct + x_0$.

Veamos que, como en el caso euclídeo, la forma F también nos permite calcular las normas. Consideremos un vector espacial \vec{v} que en el sistema de referencia dado tenga coordenadas (a, b) . La circunferencia centrada en el origen $(0, 0)$ y que pasa por (a, b) cortará al eje espacial en un punto $(x, 0)$, de modo que $|x|$ es por definición la norma de \vec{v} . El punto $(x, 0)$ puede obtenerse a partir de (a, b) por una reflexión cuyo eje pasa por $(0, 0)$. Dicho eje pasará también por el punto medio entre $(x, 0)$ y (a, b) , que es

$$M = \left(\frac{a+x}{2}, \frac{b}{2} \right).$$

La dirección del eje viene dada por el vector de coordenadas M , y ha de ser perpendicular a la recta que une $(x, 0)$ con (a, b) , cuyo vector director es $(a-x, b)$. La condición $M \perp (a-x, b)$ equivale a $x^2 = a^2 - c^2b^2$, luego

$$\|\vec{v}\| = \sqrt{a^2 - c^2b^2} = \sqrt{F(\vec{v}, \vec{v})}.$$

En particular vemos que una condición necesaria para que un vector \vec{v} sea espacial es que $F(\vec{v}, \vec{v}) = a^2 - c^2b^2 > 0$.

Si partimos de un vector temporal tendremos que buscar el corte de la circunferencia con el eje temporal, luego buscaremos un punto $(0, x)$. Cálculos análogos nos llevan a la ecuación $a^2 - c^2b^2 + c^2x^2 = 0$, luego en tal caso

$$\|\vec{v}\| = \frac{\sqrt{-a^2 + c^2b^2}}{c} = \frac{\sqrt{-F(\vec{v}, \vec{v})}}{c}.$$

Por consiguiente, una condición necesaria para que un vector sea temporal es que $F(\vec{v}, \vec{v}) = a^2 - c^2b^2 < 0$.

Más aún, hemos probado que $\|\vec{v}\| = 1$ equivale a que

$$F(\vec{v}, \vec{v}) = \begin{cases} 1 & \text{si } \vec{v} \text{ es espacial,} \\ -c^2 & \text{si } \vec{v} \text{ es temporal.} \end{cases}$$

Por consiguiente, la matriz de F en *cualquier* base ortonormal será la misma matriz A , con la misma constante c . En resumen:

Teorema 10.34 *Si en un plano de Lorentz fijamos arbitrariamente una unidad de longitud espacial y otra temporal, entonces existe una única constante $c > 0$ de modo que en cualquier sistema de referencia ortonormal se cumple:*

1. Dos vectores de coordenadas (x, t) y (x', t') son ortogonales si y sólo si $xx' - c^2tt' = 0$.
2. Las rectas isótropas son las determinadas por las ecuaciones $x = \pm ct + x_0$.
3. Un vector de coordenadas (x, t) es isótropo si $x^2 - c^2t^2 = 0$, temporal si $x^2 - c^2t^2 < 0$ y espacial si $x^2 - c^2t^2 > 0$.
4. La norma de un vector de coordenadas (x, y) es

$$\|(x, y)\| = \begin{cases} \sqrt{x^2 - c^2t^2} & \text{si } (x, y) \text{ es espacial,} \\ \frac{\sqrt{-x^2 + c^2t^2}}{c} & \text{si } (x, y) \text{ es temporal.} \end{cases}$$

Usando esto podemos probar:

Teorema 10.35 Una biyección afín es una isometría si y sólo si transforma un sistema de referencia ortonormal en un sistema de referencia ortonormal.

DEMOSTRACIÓN: La aplicación lineal asociada \vec{f} transforma una base ortonormal (\vec{u}, \vec{v}) en otra base ortonormal (\vec{u}', \vec{v}') . Dado cualquier $\vec{w} = x\vec{u} + t\vec{v}$, tendremos que $\vec{f}(\vec{w}) = x\vec{u}' + t\vec{v}'$, luego $\|\vec{w}\|$ y $\|\vec{f}(\vec{w})\|$ se calculan ambos a partir de (x, t) por las fórmulas del teorema anterior, luego $\|\vec{w}\| = \|\vec{f}(\vec{w})\|$ y por consiguiente f conserva distancias. Además esto muestra que \vec{f} transforma cada vector en otro del mismo tipo. Concluimos que f es una isometría. ■

Las leyes de la teoría de la relatividad especial se enuncian muy fácilmente en términos de la geometría del plano de Lorentz:

1. Un observador inercial es un sistema de referencia ortonormal orientado.
2. Si dos sucesos tienen coordenadas (x_0, t_0) y (x_1, t_1) respecto a un observador inercial, el tiempo transcurrido entre ellos para dicho observador es $|t_1 - t_0|$ y la distancia que los separa es $|x_1 - x_0|$.
3. Las líneas de universo de los rayos de luz son las rectas isótropas.

El hecho de que las líneas isótropas tengan ecuaciones de la forma $x = \pm ct + x_0$ nos permite interpretar ahora la constante c como la velocidad de la luz. Hemos probado que ésta es la misma para todos los observadores inerciales.

Notemos que la constante c no depende de la estructura parabólica del espacio de Lorentz, sino de la elección de las unidades de longitud. Si cambiamos el vector unitario temporal \vec{v}_2 por $\vec{v}_2' = c^{-1/2} \vec{v}_2$, tendremos que $F(\vec{v}_2', \vec{v}_2') = -1$, es decir, con esta nueva unidad de tiempo $c = 1$. Físicamente esto se interpreta como que si tomamos como unidad de tiempo el año, entonces tomamos como unidad de longitud el año luz, con lo que la velocidad de la luz resulta ser de 1 año luz por año. En resumen, no perdemos generalidad si suponemos $c = 1$, pero pese a ello conservaremos la c para que las fórmulas que obtengamos coincidan con las usuales en la teoría de la relatividad.

Los últimos resultados que hemos obtenido sugieren que adoptemos una serie de convenios que simplifiquen considerablemente la teoría. En primer lugar

definimos el *producto escalar* de dos vectores \vec{v} y \vec{w} , cuyas coordenadas en una base ortonormal sean (x, t) y (x', t') , como

$$\vec{v} \vec{w} = xx' - c^2 tt'.$$

Sabemos que esta definición es independiente de la base elegida. El producto escalar es una forma bilineal simétrica regular. Ahora podemos decir que un vector \vec{v} es isótropo, espacial o temporal según si $\vec{v} \vec{v}$ es igual, mayor o menor que 0. Tenemos que la norma de un vector arbitrario \vec{v} viene dada por

$$\|\vec{v}\|^+ = \begin{cases} \sqrt{\vec{v} \vec{v}} & \text{si } \vec{v} \text{ es espacial o isótropo,} \\ \frac{\sqrt{-\vec{v} \vec{v}}}{c} & \text{si } \vec{v} \text{ es temporal.} \end{cases}$$

Hemos añadido el superíndice + porque en este punto vamos a introducir una nueva norma respecto a la cual todos los resultados se expresan de forma mucho más simple:

$$\|\vec{v}\| = \begin{cases} \sqrt{\vec{v} \vec{v}} & \text{si } \vec{v} \text{ es espacial o isótropo,} \\ \frac{\sqrt{\vec{v} \vec{v}}}{c} & \text{si } \vec{v} \text{ es temporal.} \end{cases}$$

La única diferencia es que hemos eliminado el signo negativo en el caso temporal, de modo que la raíz cuadrada ha de entenderse como un número imaginario puro con parte real positiva. Observar que si $c = 1$ la definición se reduce a $\|\vec{v}\| = \sqrt{\vec{v} \vec{v}}$. La relación entre ambas normas es que la primera es el módulo de la segunda.

Ahora ya no tenemos vectores espaciales y temporales con la misma norma. Al contrario, la norma diferencia a los vectores espaciales (cuya norma es real) de los temporales (cuya norma es imaginaria). Del mismo modo, consideraremos que la longitud de un segmento temporal es imaginaria. Al incorporar a la norma la distinción entre espacio y tiempo, muchos enunciados se simplifican formalmente. El teorema siguiente recoge los más importantes.

Teorema 10.36 *Sea X un espacio de Lorentz en el que hemos fijado unas unidades de longitud y de tiempos. Sea r un número real positivo o imaginario puro con parte real positiva.*

1. *Un segmento es espacial, temporal o isótropo según que su longitud sea real, imaginaria o nula.*
2. *Una biyección afín en X es una isometría si y sólo si conserva las normas (o las distancias).*
3. *Para cada punto O , el conjunto formado por los puntos circulares más todos los puntos P tales que $\overline{OP} = r$ es una circunferencia de centro O . Diremos que su radio es r . Toda circunferencia es de esta forma.*
4. *Una circunferencia es espacial o temporal según si su radio es real o imaginario.*

Además tenemos un perfecto análogo al teorema de Pitágoras:

Teorema 10.37 (Teorema de Pitágoras) Si $c = 1$ y \widehat{ABC} es un triángulo rectángulo con hipotenusa \overline{AB} , entonces

$$\overline{AB}^2 = \overline{CB}^2 + \overline{CA}^2.$$

DEMOSTRACIÓN: La prueba es la misma que en el caso euclídeo. Tenemos que $\overrightarrow{AB} = \overrightarrow{CB} - \overrightarrow{CA}$ y $\overrightarrow{CB} \perp \overrightarrow{CA}$, luego

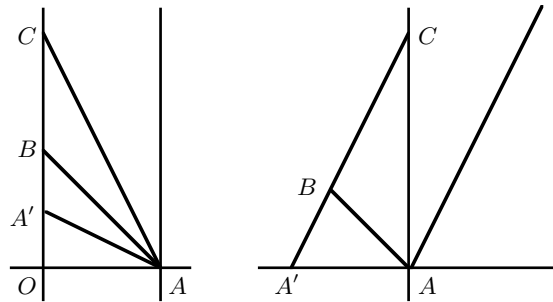
$$\overline{AB}^2 = (\overrightarrow{CB} - \overrightarrow{CA})(\overrightarrow{CB} - \overrightarrow{CA}) = \overline{CB}^2 + \overline{CA}^2 - 2\overrightarrow{CB}\overrightarrow{CA} = \overline{CB}^2 + \overline{CA}^2.$$

■

Ahora podemos caracterizar fácilmente los conceptos de futuro/pasado, izquierda/derecha en términos de coordenadas. Consideremos un sistema de referencia ortonormal y una orientación. Los vectores de coordenadas $(-c, 1)$ y $(c, 1)$ son isótropos linealmente independientes, luego los vectores que determinan la orientación han de ser de la forma $\alpha(-c, 1)$ y $\beta(c, 1)$. Comparando los signos de las coordenadas de los vectores de la base $(1, 0)$ y $(0, 1)$ respecto a estas dos bases de vectores isótropos se concluye que α y β son ambos positivos. Por consiguiente los vectores $(-c, 1)$ y $(c, 1)$ determinan la misma orientación que la dada. Usando estos vectores es fácil probar:

Teorema 10.38 Un vector temporal (espacial) cuyas coordenadas en un sistema de referencia ortonormal orientado sean (x, y) apunta hacia el futuro (hacia la derecha) si y sólo si $t > 0$ (resp. $x > 0$) y apunta hacia el pasado (hacia la izquierda) si $t < 0$ (resp. $x < 0$).

Ejemplo Consideremos el caso de la nave que viaja a la mitad de la velocidad de la luz entre dos estrellas a un año luz de distancia. Esta descripción es la correspondiente a un observador situado en cualquiera de las dos estrellas. Veamos ahora la situación desde el punto de vista de un observador en la nave.



Con las unidades que estamos manejando se cumple $c = 1$. Tenemos que $\overline{OA} = 1$, $\overline{OC} = 2i$, luego $\overline{AC} = \sqrt{1^2 + (2i)^2} = \sqrt{3}i$. Por lo tanto el observador de la nave dirá que la duración del viaje ha sido de $\sqrt{3}$ años (aproximadamente un año y nueve meses). Si llamamos AA' a la perpendicular a AC por A , su vector director (a, b) ha de ser

perpendicular a $(-1, 2)$ (tomamos las coordenadas en el primer sistema de referencia), luego $-a - 2b = 0$, o sea, el vector director es $(1, -2)$. Como la recta pasa por $A(0, 1)$ es fácil ver que A' tiene coordenadas $(0, 1/2)$, luego $\overline{AA'} = \sqrt{1^2 + (1/2)^2} = \sqrt{5}/2$, lo que significa que la distancia entre las dos estrellas para el segundo observador es de $\sqrt{5}/2 = 0,87$ años luz. ■

Giros Vamos a investigar las propiedades del plano de Lorentz que explican los resultados del ejemplo anterior. Todo depende de los giros. Sabemos que los giros con un centro común forman un grupo abeliano. En el caso del espacio euclídeo éste es isomorfo a \mathbb{R}/\mathbb{Z} , y un isomorfismo viene dado por

$$\alpha \mapsto \begin{pmatrix} \cos 2\pi\alpha & \sin 2\pi\alpha \\ -\sin 2\pi\alpha & \cos 2\pi\alpha \end{pmatrix}.$$

La situación en el plano de Lorentz es muy distinta. Dado un punto O , consideramos una circunferencia \mathcal{C} de centro O . Un giro de centro O ha de fijar a O y a los dos puntos circulares, luego está unívocamente determinado por la imagen de un punto cualquiera $P_1 \in \mathcal{C}$, que ha de ser otro punto P de \mathcal{C} . Graduémos la cónica tomando como P_0 y P_∞ los puntos circulares. Sabemos que $\mathcal{C} \setminus P_\infty$ tiene estructura de cuerpo isomorfo a \mathbb{R} , luego los puntos finitos de \mathcal{C} tienen estructura de grupo isomorfo a \mathbb{R}^* . Dado cualquier punto finito $P \in \mathcal{C}$, la multiplicación por P transforma P_1 en P . Como fija a los puntos circulares es una semejanza, como fija a la circunferencia es una isometría, y de hecho un giro (de nuevo porque fija a los puntos circulares). Así pues, es el único giro que transforma P_1 en P . Es claro entonces que si un giro f de centro O cumple $f(P_1) = P_\alpha$, la correspondencia $f \mapsto \alpha$ es un isomorfismo entre el grupo de giros y el grupo multiplicativo \mathbb{R}^* .

Ejercicio: Probar que el grupo de giros orientados se corresponde con $]0, +\infty[$.

Veamos esto mismo desde un punto de vista algebraico. Tomemos un sistema de referencia ortonormal \vec{u}_1, \vec{u}_2 y consideremos los vectores \vec{v}_1 y \vec{v}_2 de coordenadas

$$\frac{1}{\sqrt{2}c}(c, 1) \quad \text{y} \quad \frac{1}{\sqrt{2}c}(c, -1).$$

Son dos vectores isótropos linealmente independientes. Además $\vec{v}_1 \vec{v}_2 = 1$. Por consiguiente $\vec{f}(\vec{v}_1) = \alpha \vec{v}_1$ y $\vec{f}(\vec{v}_2) = \beta \vec{v}_2$. Además

$$1 = \vec{v}_1 \vec{v}_2 = \vec{f}(\vec{v}_1) \vec{f}(\vec{v}_2) = \alpha\beta \vec{v}_1 \vec{v}_2 = \alpha\beta.$$

Así pues, la matriz de \vec{f} en la base \vec{v}_1, \vec{v}_2 es

$$M_\alpha = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix},$$

y es claro que la aplicación $\alpha \mapsto M_\alpha$ es un isomorfismo de grupos. De las relaciones

$$\vec{v}_1 = \frac{1}{\sqrt{2}c}(c\vec{u}_1 + \vec{u}_2)$$

$$\vec{v}_2 = \frac{1}{\sqrt{2}c}(c\vec{u}_1 - \vec{u}_2)$$

se sigue fácilmente que la matriz de este mismo giro en la base \vec{u}_1, \vec{u}_2 es:

$$\begin{aligned} & -\frac{1}{\sqrt{2}} \begin{pmatrix} -1 & -1 \\ -c & c \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} \frac{1}{\sqrt{2}c} \begin{pmatrix} c & 1 \\ c & -1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{\alpha + \alpha^{-1}}{2} & \frac{\alpha - \alpha^{-1}}{2c} \\ c \frac{\alpha - \alpha^{-1}}{2} & \frac{\alpha + \alpha^{-1}}{2} \end{pmatrix}. \end{aligned}$$

Así pues, la matriz de un giro en un sistema ortonormal cualquiera es de la forma

$$\pm G_\alpha = \pm \begin{pmatrix} \frac{\alpha + \alpha^{-1}}{2} & \frac{\alpha - \alpha^{-1}}{2c} \\ c \frac{\alpha - \alpha^{-1}}{2} & \frac{\alpha + \alpha^{-1}}{2} \end{pmatrix}, \quad \text{con } \alpha > 0.$$

Notemos que indirectamente hemos probado que $G_\alpha G_\beta = G_{\alpha\beta}$. Es conveniente conservar la notación aditiva que tenemos en el caso euclídeo, así que vamos a componer el isomorfismo $\alpha \mapsto G_\alpha$ con el isomorfismo $\mathbb{R} \rightarrow]0, +\infty[$ determinado por una función exponencial. Ello nos lleva a la definición siguiente:

Definición 10.39 Sea $e > 1$ un número real. Definimos las funciones *seno* y *coseno hiperbólico* (en base e) como las dada por:

$$\sinh x = \frac{e^x - e^{-x}}{2}, \quad \cosh x = \frac{e^x + e^{-x}}{2}.$$

En estos términos hemos probado lo siguiente:

Teorema 10.40 Las matrices de los giros en un sistema de referencia ortonormal son las de la forma

$$\pm G_\alpha = \pm \begin{pmatrix} \cosh \alpha & \frac{1}{c} \sinh \alpha \\ c \sinh \alpha & \cosh \alpha \end{pmatrix}, \quad \text{con } \alpha \in \mathbb{R}.$$

Además se cumple la relación $G_\alpha G_\beta = G_{\alpha+\beta}$.

Ahora conviene investigar las propiedades de las funciones hiperbólicas. Para ello podemos suponer $c = 1$. Fijado un sistema de referencia, la circunferencia de centro $(0,0)$ y radio 1 está formada por los puntos (x,t) que satisfacen la ecuación $x^2 - t^2 = 1$. Uno de sus puntos es $(1,0)$, y todos los demás pueden obtenerse a partir de éste mediante un giro. Por otra parte tenemos que sus puntos son los de la forma

$$\pm(1,0)G_\alpha = \pm(\cosh \alpha, \sinh \alpha), \quad \text{para } \alpha \in \mathbb{R}.$$

Vemos, pues, que del mismo modo que $(\cos \alpha, \sin \alpha)$ recorre una circunferencia en el plano euclídeo, la expresión análoga con las funciones hiperbólicas recorre una circunferencia en el plano de Lorentz o, más exactamente, recorre una rama de la hipérbola. La otra rama se obtiene con el signo negativo.

En particular obtenemos la relación fundamental entre las razones hiperbólicas:

$$\cosh^2 \alpha - \sinh^2 \alpha = 1.$$

De la definición se sigue que $\cosh \alpha > 0$, la relación anterior prueba que de hecho ha de ser $\cosh \alpha \geq 1$. Hemos probado que cualquier par (x, t) con $x > 0$ y tal que $x^2 - t^2 = 1$ es de la forma $(\cosh \alpha, \sinh \alpha)$. De hecho la expresión es única, pues hay un único giro que transforme un punto de una circunferencia en otro dado. Si desarrollamos la igualdad $G_\alpha G_\beta = G_{\alpha+\beta}$ obtenemos las fórmulas para el seno y el coseno hiperbólico de una suma. Las incluimos en este teorema:

Teorema 10.41 *Las funciones hiperbólicas $\sinh : \mathbb{R} \rightarrow \mathbb{R}$ y $\cosh : \mathbb{R} \rightarrow \mathbb{R}$ cumplen las propiedades siguientes:*

1. $\cosh^2 \alpha - \sinh^2 \alpha = 1$.
2. $\cosh(\alpha + \beta) = \cosh \alpha \cosh \beta + \sinh \alpha \sinh \beta$,
 $\sinh(\alpha + \beta) = \cosh \alpha \sinh \beta + \sinh \alpha \cosh \beta$.
3. $\cosh 0 = 1$, $\sinh 0 = 0$, $\cosh(-\alpha) = \cosh \alpha$, $\sinh(-\alpha) = -\sinh \alpha$.
4. Para cada par de números reales (x, t) con $x > 0$ existe un único $\alpha \in \mathbb{R}$ tal que $x = \cosh \alpha$, $t = \sinh \alpha$.

Ejercicio: Probar que si $0 \leq \alpha < \beta$ entonces $\cosh \alpha < \cosh \beta$, $\sinh \alpha < \sinh \beta$. Deducir que $\cosh \alpha$ toma exactamente dos veces cada valor $x \in]1, +\infty[$ y que $\sinh \alpha$ toma una sola vez cada valor real.

Ejercicio: Deducir las fórmulas del ángulo doble y del ángulo mitad.

Ejercicio: Estudiar la función *tangente hiperbólica* $\tanh : \mathbb{R} \rightarrow \mathbb{R}$ dada por

$$\tanh x = \frac{\sinh x}{\cosh x}.$$

Notemos ahora que si un giro tiene matriz G_α en un sistema de referencia ortonormal (con origen en el centro de giro) entonces su matriz en cualquier otro sistema de referencia del mismo origen es $G_{\pm\alpha}$, pues la traza del giro es $2 \cosh \alpha$, y es un invariante. Por consiguiente la expresión $\pm G_\alpha$ asigna a cada giro f un *signo* $\text{sig } f$ (que es -1 si f intercambia las ramas de las circunferencias y 1 si las deja invariantes) y un *ángulo* α determinado salvo signo.

Ejercicio: Probar que la matriz de cambio de base respecto a dos sistemas de referencia ortonormales orientados y con el mismo origen es de la forma G_α . Deducir que un giro tiene la misma matriz en todos los sistemas de referencia ortonormales orientados con origen en su centro.

En términos de las funciones hiperbólicas los puntos finitos de la circunferencia espacial de radio r y centro el origen de coordenadas son los de la forma

$$(\pm r)_\alpha = (\pm r \cosh \alpha, \pm \frac{r}{c} \sinh \alpha), \quad \alpha \in \mathbb{R}.$$

Los puntos finitos de la circunferencia temporal de radio ir son los que tienen coordenadas

$$(\pm ir)_\alpha = (\pm rc \sinh \alpha, \pm r \cosh \alpha), \quad \alpha \in \mathbb{R},$$

es decir, los que se obtienen a partir de $(0, r)$ mediante giros de ángulo α .

Así, respecto a un sistema de referencia dado, todo punto distinto del origen se expresa de forma única como r_α , donde r es un número real o imaginario puro no nulo y α es cualquier número real. Diremos que r y α son las *coordenadas polares* del punto. El número α es el *argumento* del punto. También podemos hablar de las coordenadas polares de un vector no nulo. Así, la acción de un giro de ángulo α sobre un vector es sumarle α a su argumento y multiplicarlo por su signo.

Ejercicio: Definir el ángulo que forman dos rectas, ambas espaciales o ambas temporales. Probar que el ángulo α entre las rectas de vectores \vec{v}_1 y \vec{v}_2 (ambas espaciales o ambas temporales) viene dado por

$$\cosh \alpha = \frac{|\vec{v}_1 \vec{v}_2|}{\|\vec{v}_1\| + \|\vec{v}_2\|}.$$

Consideremos dos observadores inerciales $A \equiv (O; \vec{u}_1, \vec{u}_2)$ y $B \equiv (O', \vec{v}_1, \vec{v}_2)$. La línea de universo de B es la recta que pasa por O' y tiene dirección \vec{v}_2 . Las coordenadas de \vec{v}_2 respecto al primer observador serán de la forma $(c \sinh \alpha, \cosh \alpha)$, para un cierto $\alpha \in \mathbb{R}$ (notar que, como \vec{v}_2 apunta hacia el futuro, su segunda componente es positiva), y la ecuación de dicha línea de universo será de la forma $x = vt + x_0$, donde $v = c \tanh \alpha$. La interpretación física de v es clara: se trata de la velocidad con que A ve alejarse a B . Si $v = 0$ (o equivalentemente, $\alpha = 0$) los observadores están en reposo relativo. Por otra parte, la interpretación geométrica de α también es clara: se trata del ángulo hiperbólico que forman las líneas de universo de ambos observadores. La condición $\alpha = 0$ equivale a que éstas sean paralelas.

De las relaciones trigonométricas se sigue fácilmente que

$$\cosh \alpha = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad \sinh \alpha = \frac{v/c}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

(Notar que el signo del seno es el de v .)

Es claro que al aplicar a \vec{u}_2 el giro de ángulo α obtenemos \vec{v}_2 , y como \vec{u}_1 es unitario, ortogonal a \vec{u}_2 y apunta hacia la derecha, al aplicarle el giro de ángulo α se convierte en \vec{v}_1 . Por consiguiente, si un vector tiene coordenadas (x, t) respecto al observador A , sus coordenadas respecto a B serán $(x', t') = (x, t)G_{-\alpha}$, y según las relaciones anteriores esto equivale a:

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Si el suceso O tiene coordenadas (x'_0, t'_0) respecto al observador B , las coordenadas respecto a B del suceso que en A tiene coordenadas (x, y) son

$$x' = x'_0 + \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad t' = t'_0 + \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Estas fórmulas se conocen como *relaciones de Lorentz*, y muestran claramente las propiedades de la geometría relativista. Veamos cuáles son sus consecuencias. Supongamos que B lleva consigo una regla de longitud L . Los extremos de la regla en un instante dado t' tendrán para B coordenadas (a', b') , $(a' + L, b')$. Despejamos x en la fórmula anterior:

$$x = (x' - x_0) \sqrt{1 - \frac{v^2}{c^2}} + vt,$$

sustituimos $(x', t') = (a', b')$, $(x', t') = (a' + L, b')$ y restamos los valores de x resultantes. Lo que obtenemos es que para el observador A la regla de B mide

$$L' = L \sqrt{1 - \frac{v^2}{c^2}},$$

por lo que A considera que la regla de B es más corta de lo que B afirma. Del mismo modo se prueba que si el reloj de B mide un tiempo T entre dos sucesos, el tiempo transcurrido según A será de

$$T' = T \sqrt{1 - \frac{v^2}{c^2}},$$

lo cual nos indica que, según A , el reloj de B atrasa.

Si $v = 0$ no hay discrepancias. Si los observadores empiezan a moverse, cada cual está legitimado para afirmar que es el otro el que se mueve, mientras él permanece en reposo, por lo que ninguno notará variación alguna en su regla o su reloj, sin embargo cada uno verá menor la regla del otro y verá correr más lentamente el reloj del otro.⁶ La reducción de la regla⁷ y el retraso del reloj serán tanto mayores cuanto mayor sea la velocidad relativa v . Si A pudiera ver a B moverse a la velocidad de la luz, vería su regla con longitud 0 y su reloj parado. Esto es imposible, pero dado un suceso de duración arbitrariamente grande (para A) siempre es posible encontrar un observador B (sin más que exigir que v sea suficientemente grande) de modo que perciba el mismo suceso con duración tan pequeña como se desee.

La interpretación geométrica de todo esto es que la relación entre los sistemas de referencia de dos observadores inerciales es que uno se obtiene del otro aplicándole un giro orientado (más una traslación si no tienen el mismo origen de coordenadas), cuyo ángulo depende únicamente de la velocidad relativa.

Ejemplo En el caso de la nave que viaja entre las dos estrellas, tenemos que la velocidad relativa es $v = 1/2$ (en realidad $v = -1/2$ respecto del observador en A la estrella y $v = 1/2$ respecto del observador B en la nave). El tiempo de viaje ha sido de 2 años para A , mientras que B dice que ha sido de

$$T' = 2\sqrt{1 - (1/2)^2} = \sqrt{3} \text{ años.}$$

⁶Este efecto puede compararse con dos personas que se ven desde lejos. Cada una verá a la otra menor que a sí misma. Sin embargo los efectos de las fórmulas de Lorentz no corresponden a ilusiones ópticas, sino a las magnitudes que cada observador medirá realmente si sus instrumentos son suficientemente exactos. Cuando una partícula subatómica inestable es acelerada a velocidades cercanas a la de la luz, el tiempo que tarda en desintegrarse aumenta considerablemente, y no es que “parezca” que la partícula vive más tiempo, es que la partícula vive más tiempo.

⁷Debemos comentar que en el caso tetradimensional (con un espacio de tres dimensiones) esta reducción afecta sólo a la dirección del movimiento, mientras que las longitudes perpendiculares al movimiento permanecen inalteradas.

Por lo tanto A afirma que el reloj de B ha ido más lento. Sin embargo B considera que el inicio de su viaje (de $\sqrt{3}$ años) fue simultáneo del suceso A' , y A afirma que entre A' y C ha pasado un tiempo de

$$\sqrt{3}\sqrt{1 - (1/2)^2} = 1,5 \text{ años},$$

que es menos de $\sqrt{3}$, luego B afirma que es el reloj de A el que atrasa. ■

Ejercicio: Según comentamos al principio de la sección, si un observador A ve moverse a un observador A' a una velocidad v y A' ve moverse a un observador A'' a una velocidad v' , no es cierto que A vea moverse a A'' a una velocidad $v'' = v + v'$. Si estas velocidades corresponden a ángulos hiperbólicos α , α' y α'' , es fácil ver que $\alpha'' = \alpha + \alpha'$. Usar la fórmula de la tangente hiperbólica de una suma para expresar esta relación en términos de las velocidades.

Semejanzas Para acabar la descripción del plano de Lorentz caracterizaremos las semejanzas en términos similares al caso euclídeo. El resultado básico es el siguiente:

Teorema 10.42 *Una biyección afín f en el plano de Lorentz es una semejanza si y sólo si existe un número real k tal que para todo par de vectores \vec{u} y \vec{v} se cumple $\vec{f}(\vec{u})\vec{f}(\vec{v}) = k\vec{u}\vec{v}$.*

DEMOSTRACIÓN: Es obvio que una aplicación que cumpla esto conserva la ortogonalidad, luego es una semejanza. Recíprocamente, si f conserva la ortogonalidad, consideramos una base ortonormal \vec{u} , \vec{v} . Si α y β son escalares tales que $\alpha\vec{u}\vec{u} + \beta\vec{v}\vec{v} = 0$, entonces $\alpha\vec{u} + \beta\vec{v} \perp \vec{u} + \vec{v}$, luego

$$\alpha\vec{f}(\vec{u}) + \beta\vec{f}(\vec{v}) \perp \vec{f}(\vec{u}) + \vec{f}(\vec{v})$$

y también $\alpha\vec{f}(\vec{u})\vec{f}(\vec{u}) + \beta\vec{f}(\vec{v})\vec{f}(\vec{v}) = 0$.

Esto prueba que los vectores de \mathbb{R}^2 dados por $(\vec{u}\vec{u}, \vec{v}\vec{v})$ y $(\vec{f}(\vec{u})\vec{f}(\vec{u}), \vec{f}(\vec{v})\vec{f}(\vec{v}))$ tienen los mismos vectores ortogonales (α, β) (en el sentido euclídeo usual), luego son linealmente dependientes, es decir, existe un número real k (no nulo) tal que

$$\vec{f}(\vec{u})\vec{f}(\vec{u}) = k\vec{u}\vec{u} \quad \text{y} \quad \vec{f}(\vec{v})\vec{f}(\vec{v}) = k\vec{v}\vec{v}.$$

Usando la bilinealidad del producto es fácil concluir que este k cumple esto mismo con cualquier otro vector. ■

Es claro que el número k que cumple el teorema anterior está unívocamente determinado por f . Llamaremos *razón* de f al número $r = \sqrt{k}$, donde k puede ser real positivo o bien imaginario puro con parte real positiva.

El teorema anterior implica claramente que si f es una semejanza de razón r entonces para todo par de puntos A y B se cumple

$$\overline{f(A)f(B)}^2 = r^2 \overline{AB}^2. \quad (10.1)$$

Por consiguiente $|\overline{f(A)f(B)}| = |r| |\overline{AB}|$. Si r es real el segmento \overline{AB} y su imagen tienen ambos el mismo carácter espacial o temporal, mientras que si r es imaginario el carácter se invierte.

Es evidente entonces que las isometrías son las semejanzas de razón 1. La razón (en este sentido) de una homotecia de radio r es $|r|$. La razón de la composición de una semejanza de razón r con una semejanza de razón r' es $\sqrt{r^2 r'^2}$, que en general no es lo mismo que $r r'$.

Ejercicio: Probar que si una afinidad f cumple 10.1 entonces es una semejanza de razón r .

Ejercicio: Describir las semejanzas de razón i . Probar que toda semejanza es producto de una isometría, una homotecia y (quizá) una semejanza de razón i .

10.4 Propiedades métricas de las cónicas

Ahora estamos en condiciones de caracterizar los conceptos métricos de una cónica, como son sus focos, sus ejes, sus vértices, etc. Vamos a definirlos de nuevo sin tener en cuenta las definiciones dadas en la sección 4.8. Al final de la sección será evidente que las nuevas definiciones son equivalentes a las dadas allí.

Definición 10.43 Sea \mathcal{C} una cónica. Un *eje* de \mathcal{C} es una recta tal que la reflexión respecto a ella deja invariante a \mathcal{C} . Un *foco* F de \mathcal{C} es un punto tal que los pares de rectas conjugadas respecto a \mathcal{C} que pasan por F son perpendiculares. Un *vértice* de \mathcal{C} es un punto de \mathcal{C} tal que su tangente es perpendicular al diámetro que pasa por él.

Salvo que indiquemos lo contrario, cuando hablemos de focos, vértices y ejes de una cónica entenderemos que son reales. Así, todos los diámetros de una circunferencia son ejes, todos sus puntos son vértices y su único foco es su centro. La situación en las demás cónicas es muy distinta.

Teorema 10.44 Una parábola tiene un único eje y un único vértice. Además el eje es el diámetro que pasa por el vértice.

DEMOSTRACIÓN: Sea O el centro de la parábola. Que un punto V sea un vértice significa que la recta OV es perpendicular a la tangente por V , es decir, que ésta pasa por el conjugado ortogonal O' de O . Pero por O' pasan sólo dos tangentes a la parábola. Una es la recta infinita, y la otra toca a la cónica en un punto V que resulta ser el único vértice.

Veamos que OV es un eje. Dado un punto P en la parábola que no sea V , la perpendicular a OV por P pasa por O' , luego no es tangente a la parábola, luego la corta en un segundo punto P' . Como OV es la polar de O' , el teorema 9.39 nos permite concluir que P y P' se corresponden por la reflexión respecto a OV .

Finalmente veamos que OV es el único eje. Si una recta r es un eje entonces la reflexión de eje r ha de fijar a la parábola. Como es una isometría, es claro

que ha de fijar a su único vértice, luego V está en r . La reflexión ha de fijar también a la tangente por V , lo que sólo es posible si dicha tangente t es igual a r o perpendicular a r . No puede ser $t = r$, pues, dado un punto P en la parábola distinto de V , la perpendicular a t por P pasa por O , luego no corta a la parábola en otro punto finito, luego la imagen de P por la reflexión respecto a t , que es otro punto en dicha perpendicular, queda fuera de la cónica. Así pues, r es perpendicular a t y pasa por V , luego es OV . ■

La situación en los demás tipos de cónicas no es tan fácil de estudiar. Recogemos en un teorema lo que podemos obtener por razonamientos similares a los que acabamos de emplear.

Teorema 10.45 *Sea \mathcal{C} una cónica que no sea una parábola. Entonces:*

1. *Una recta es un eje de \mathcal{C} si y sólo si es un diámetro perpendicular a su diámetro conjugado.*
2. *Un punto de \mathcal{C} es un vértice si y sólo si está sobre un eje.*

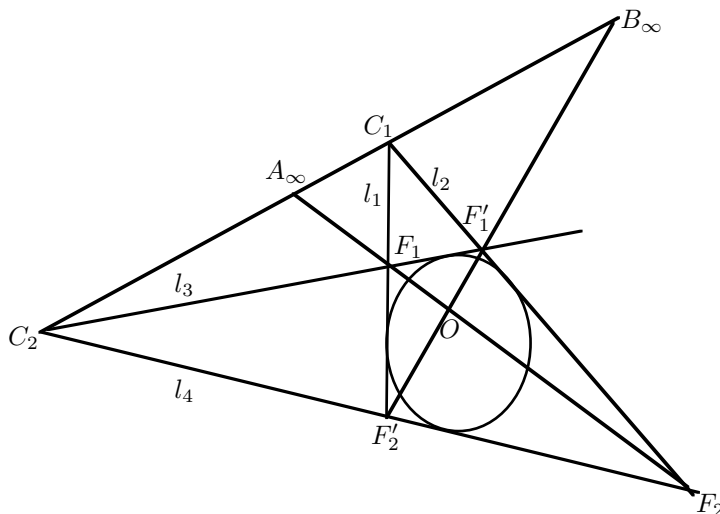
DEMOSTRACIÓN: Si r es un eje de \mathcal{C} , la reflexión de eje r es una isometría que deja fija a \mathcal{C} , luego fija a su centro O , luego O está en r , es decir, r es un diámetro. Al fijar a \mathcal{C} , la reflexión respecto a r transforma pares de rectas conjugadas en pares de rectas conjugadas. Como fija a r , también fija a su diámetro conjugado, pero esto sólo es posible si éste es perpendicular a r .

Recíprocamente, si r es un diámetro perpendicular a su diámetro conjugado r' , entonces el polo de r es el punto infinito de r' , digamos P_∞ . Dado un punto Q en \mathcal{C} , la perpendicular a r por Q es QP_∞ . Si es la tangente por Q entonces Q está en la polar de P_∞ , o sea, en r , luego es fijado por la reflexión respecto de r . Si por el contrario QP_∞ corta a la cónica en otro punto Q' , entonces el teorema 9.39 garantiza que Q y Q' se corresponden por la reflexión respecto a r . En cualquier caso esta reflexión fija a \mathcal{C} .

Si un punto V de \mathcal{C} está en un eje r , entonces el polo de r es el punto infinito de su diámetro conjugado r' , por donde pasa también la tangente por V a \mathcal{C} , luego ésta es perpendicular a r . Recíprocamente, si V es un vértice y r es el diámetro que pasa por V , entonces su polo es el punto infinito de la tangente por V y también el del diámetro conjugado, luego éstas son rectas paralelas y la tangente es perpendicular a r , luego el diámetro conjugado también. ■

Este teorema no afirma nada sobre la existencia de ejes o vértices. Vamos a probar que toda cónica que no sea una parábola ni una circunferencia tiene exactamente dos ejes. El número de vértices dependerá de forma obvia del tipo de cónica. Al mismo tiempo probaremos que existen dos focos.

Sea \mathcal{C} una hipérbola o una elipse real o imaginaria que no sea una circunferencia, es decir, que no pase por ninguno de los puntos circulares C_1 y C_2 . Sean l_1, l_2 las tangentes a \mathcal{C} que pasan por C_1 y l_3, l_4 las tangentes por C_2 . Sean F_1, F_2, F'_1, F'_2 las intersecciones entre estas tangentes, en el orden que indica la figura:



El cuadrilátero completo dual de lados l_1, l_2, l_3, l_4 está circunscrito a \mathcal{C} y sus pares de vértices opuestos son (C_1, C_2) , (F_1, F_2) , (F'_1, F'_2) . Por el dual del teorema 9.38, las tres rectas determinadas por estos pares son los lados de un triángulo autopolar, es decir, tal que cada vértice es el polo del lado opuesto. Sus vértices son los puntos A_∞, B_∞, O , definidos como indica la figura.

Por lo tanto tenemos que O es el polo de la recta infinita C_1C_2 , luego O es el centro de \mathcal{C} . En particular es un punto real. Veremos que las rectas $a = F_1F_2$ y $b = F'_1F'_2$ son los únicos ejes de \mathcal{C} . Por comodidad los llamaremos ya ejes aunque no hayamos probado todavía que lo son. Similarmente, a los puntos F_1, F_2, F'_1, F'_2 los llamaremos focos, aunque probaremos que sólo dos de ellos cumplen la definición que hemos dado, mientras que los otros dos serán puntos imaginarios. De momento tenemos que los ejes son dos diámetros y que hay dos focos sobre cada eje. Además los focos son los vértices de un cuadrilátero completo cuyos puntos diagonales son el centro O y los puntos circulares. Por la propia definición de separación armónica tenemos $\mathcal{H}(O, A_\infty; F_1, F_2)$ y $\mathcal{H}(O, B_\infty; F'_1, F'_2)$. Esto significa que si un foco es real el otro foco del mismo eje también lo es, y ambos se encuentran simétricamente situados respecto al centro.

Consideremos un punto real X fuera de \mathcal{C} y de las cuatro tangentes l_i . Aplicamos el dual del teorema 9.30 al punto X y al cuadrilátero completo dual de lados l_i . Concluimos que las tangentes a \mathcal{C} por X (digamos t_1 y t_2) son conjugadas (en el haz de rectas de centro X) respecto de la involución que hace corresponder los pares (XC_1, XC_2) , (XF_1, XF_2) , (XF'_1, XF'_2) .

Sean x_1 y x_2 las rectas fijadas por esta involución. Tenemos entonces

$$\mathcal{H}(x_1, x_2; t_1, t_2), \quad \mathcal{H}(x_1, x_2; XC_1, XC_2).$$

Ahora bien, XC_1 y XC_2 son las rectas fijas respecto a la involución que induce en el haz de rectas la involución ortogonal, luego x_1 y x_2 son también conjugadas por dicha involución, es decir, son perpendiculares. Similarmente, t_1 y t_2 son las rectas fijas por la involución que \mathcal{C} induce en el haz de rectas, luego

x_1 y x_2 también son conjugadas respecto a \mathcal{C} . Tenemos, pues, que x_1 y x_2 son un par común respecto a las involuciones inducidas en el haz de centro X tanto por la involución ortogonal como por \mathcal{C} . Ambas involuciones son distintas. Para verlo basta tener en cuenta que XC_1 no es tangente a \mathcal{C} , luego no se cumple $p(XC_1) \subset XC_1$, luego la conjugación respecto a \mathcal{C} no fija a XC_1 , mientras que la involución ortogonal sí lo hace. Por consiguiente x_1 y x_2 son el único par común de ambas involuciones, es decir, x_1 y x_2 son el único par de rectas perpendiculares conjugadas que pasan por X . Por el teorema 9.61 (para el caso de involuciones sobre un haz de rectas) podemos concluir que x_1 y x_2 son ambas reales.

Consideremos ahora el caso particular en que $X = O$. Entonces los tres pares que determinan la involución en el haz de rectas son (OC_1, OC_2) , (OF_1, OF_2) , (OF'_1, OF'_2) , pero $OF_1 = OF_2$ y $OF'_1 = OF'_2$, luego las rectas x_1 , x_2 son en este caso los dos ejes. Acabamos de probar que son el único par de diámetros conjugados perpendiculares, o sea, los únicos ejes de la cónica.

Teorema 10.46 *Una hipérbola o una elipse real o imaginaria que no sea una circunferencia tiene exactamente dos ejes.*

Continuando con nuestro argumento, ahora sabemos que los puntos A_∞ , B_∞ son reales. Volvamos al caso en que X es un punto real arbitrario fuera de las tangentes l_i . Supongamos además que no está en los ejes ni en la recta infinita. Entonces las rectas x_1 y x_2 cortan al eje $a = F_1F_2$ en dos puntos reales X_1 y X_2 . Recordemos que $\mathcal{H}(O, A_\infty; F_1, F_2)$. Por otra parte $\mathcal{H}(x_1, x_2; XF_1, XF_2)$, pues las dos últimas rectas son conjugadas respecto a la involución que fija a las dos primeras, luego $\mathcal{H}(X_1, X_2; F_1, F_2)$.

De aquí se siguen varias consecuencias. En primer lugar, si los pares (X_1, X_2) y (O, A_∞) tuvieran un elemento en común, tendrían de hecho los dos elementos en común, pero entonces XO y XA_∞ serían perpendiculares, y como XA_∞ es paralela al eje OA_∞ , tendríamos que XO es perpendicular a dicho eje, luego sería el eje OB_∞ , mientras que hemos tomado X fuera de esta recta.

Así pues, los pares (X_1, X_2) y (O, A_∞) son distintos, y la involución que fija a F_1 y F_2 coincide con la que los hace corresponder, luego es una involución real. Esto implica que F_1 y F_2 son reales o imaginarios conjugados. Lo mismo vale para F'_1 y F'_2 . Ahora bien, puesto que los puntos circulares son imaginarios conjugados, si F_1 y F_2 también lo son, el teorema 9.66 nos da que F'_1 y F'_2 son reales. Recíprocamente, si F_1 y F_2 son reales entonces F'_1 y F'_2 han de ser imaginarios, pues $F_1F_2 \stackrel{F'_1}{\perp} C_1C_2$ y, si F'_2 fuera real, también lo serían C_1 y C_2 . En resumen: dos de los focos son reales y los otros dos son imaginarios conjugados. No perdemos generalidad si suponemos que F_1 y F_2 son reales.

La involución que \mathcal{C} induce en el haz de rectas de centro un foco real F_i fija a las tangentes F_iC_1 y F_iC_2 , y lo mismo sucede con la involución que induce en dicho haz la involución ortogonal, luego ambas son la misma. Por consiguiente, los pares de rectas conjugadas respecto a \mathcal{C} que pasan por F_i son perpendiculares, es decir, F_i es realmente un foco. Recíprocamente, si un punto real P cumple

la definición de foco, las tangentes a \mathcal{C} por P deben pasar por C_1 y C_2 , luego P es uno de F_i . Así pues:

Teorema 10.47 *Una hipérbola o una elipse real o imaginaria que no sea una circunferencia tiene exactamente dos focos. Ambos están sobre un eje y simétricamente situados respecto al centro. El eje que contiene a los focos se llama eje mayor de la cónica, mientras que al otro se le llama eje menor. Si la cónica es real los focos son puntos interiores.*

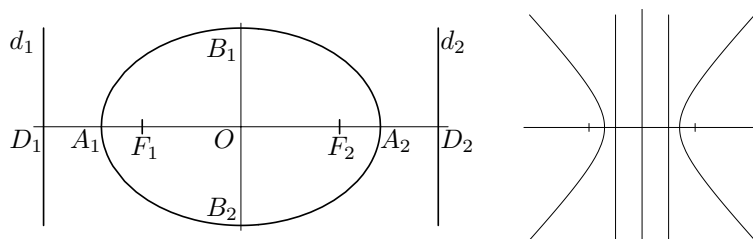
Una forma de ver que los focos son interiores es observar que los focos imaginarios son las intersecciones con el eje menor de las tangentes a la cónica por los focos reales. Como son imaginarios, dichas tangentes también han de serlo, luego por los focos no pasan tangentes reales.

Ejercicio: ¿Por qué en la figura los focos aparecen como puntos exteriores?

Los ejes de \mathcal{C} no pueden ser tangentes (porque pasan por el centro), luego cada uno corta a \mathcal{C} en dos puntos distintos (reales o imaginarios conjugados). Digamos que a corta a \mathcal{C} en A_1, A_2 y que b corta en B_1, B_2 . Llamaremos vértices a estos cuatro puntos, aunque no son necesariamente reales. Obviamente en una elipse imaginaria son todos imaginarios. En una elipse real son todos reales (porque el centro es interior) y en una hipérbola los vértices del eje mayor son reales (porque este eje pasa por los focos, que son interiores), mientras que los del eje menor han de ser imaginarios conjugados, pues dicho eje es la polar de A_∞ , que es un punto interior, luego el eje menor es exterior y no corta a \mathcal{C} .

Definición 10.48 Llamaremos *directrices* de una cónica a las polares de sus focos.

Si llamamos d_1, d_2, d'_1 y d'_2 a las polares de F_1, F_2, F'_1, F'_2 respectivamente, tenemos que d_1 y d_2 son rectas reales exteriores a la cónica (salvo en el caso de una elipse imaginaria, en que esto no tiene sentido). Las directrices d_1 y d_2 pasan por el polo de F_1F_2 , que es B_∞ , luego son paralelas al eje menor, luego perpendiculares al eje mayor. Similarmente, d'_1 y d'_2 pasan por A_∞ y son conjugadas imaginarias.



Veamos ahora una propiedad importante de los focos de una elipse real o de una hipérbola. Tomemos un punto real finito X en \mathcal{C} que no esté en sus ejes. Por el dual del teorema 9.30 tenemos que la tangente t a \mathcal{C} por X es una recta doble de la involución en el haz de rectas de centro X que empareja

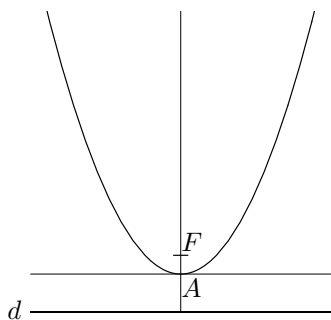
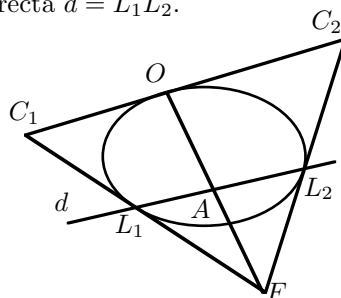
(XC_1, XC_2) , (XF_1, XF_2) , (XF'_1, XF'_2) . La otra recta fijada, digamos n , ha de ser la conjugada harmónica de t respecto a XC_1 y XC_2 , luego la conjugada de t respecto a la involución inducida en el haz por la involución ortogonal. Es decir, t y n son perpendiculares. Se dice que n es la recta *normal* a la cónica por X . Tenemos también que XF_1 y XF_2 son conjugadas armónicas respecto a t y n . Por consiguiente:

Teorema 10.49 *Si \mathcal{C} es una elipse real que no sea una circunferencia (o una hipérbola), entonces la normal (tangente) a \mathcal{C} por uno cualquiera de sus puntos X es la bisectriz del ángulo $\widehat{F_1XF_2}$, donde F_1 y F_2 son sus focos.*

En realidad en la prueba hemos excluido la posibilidad de que X fuera uno de los vértices de la cónica, pero este caso es trivial. El teorema anterior puede interpretarse como que si un rayo de luz parte de un foco de una elipse y se refleja en ella, entonces el rayo reflejado pasará por el otro foco. Si la cónica es una hipérbola el rayo reflejado estará también sobre la recta que pasa por el otro foco, pero se reflejará en sentido contrario. Es esta propiedad la que ha dado lugar al término “foco”.

Volvemos ahora a las parábolas. Sólo nos queda probar que una parábola tiene un único foco. Si \mathcal{C} es una parábola, la recta infinita es tangente a \mathcal{C} . El punto de tangencia es el centro O . Entonces las tangentes a \mathcal{C} por cada punto circular C_i son la propia recta infinita y otra recta l_i . Sea L_i el punto de tangencia y sea $F = l_1 \cap l_2$. La polar de F es la recta $d = L_1L_2$.

Veamos que F y d son reales. La polaridad de la cónica hace corresponder cada punto de la recta infinita con una recta que pasa por O . La involución ortogonal se corresponde con una involución cuyas rectas fijas son OL_1 y OL_2 . A su vez, esta involución induce otra sobre la parábola, cuyos puntos fijos son L_1 y L_2 . Como la parábola es real, éstos son imaginarios conjugados, luego la recta que los une es real, pero ésta es d , luego su polo F también es real, así como la recta OF y el punto A donde éste corta a la parábola.



Como en el caso de las elipses y las hipérbolas se razona fácilmente que F es el único foco de la parábola. Además es un punto interior, luego la directriz es exterior.

Veamos que A es el vértice de la parábola, con lo que OA será su eje. Sea B_∞ el polo de OA , es decir, el punto infinito de la tangente por A . Entonces las rectas FB_∞ y OA son conjugadas y pasan por F , luego son perpendiculares. Por consiguiente B_∞ es el conjugado ortogonal de O . En particular la tangente por A es perpendicular

a OA , lo que prueba que A es el vértice. Puesto que F está en la polar de B_∞ , tenemos además que B_∞ está en d , luego la directriz es perpendicular al eje.

Teorema 10.50 *Una parábola tiene un único foco, el cual es un punto interior situado sobre su eje.*

Sea P un punto real de una hipérbola \mathcal{C} . Aplicamos el dual del teorema 9.31 al triángulo de vértices C_1 , C_2 , F y al punto X , con lo que resulta que la tangente t por X es una recta fija de la involución que en el haz de rectas de centro X determinan los pares XC_1 , XC_2 y XF , XA_∞ . Si n es la otra recta fija, tenemos que XC_1 y XC_2 están separados armónicamente por t y n , luego t y n son conjugadas respecto de la involución que fija a XC_1 y XC_2 , pero ésta es la involución inducida en el haz por la involución ortogonal, luego t y n son perpendiculares y n es la recta normal a \mathcal{C} por X . Con esto llegamos al análogo de 10.49:

Teorema 10.51 *Si \mathcal{C} es una parábola y F es su foco, la tangente y la normal a \mathcal{C} por uno cualquiera de sus puntos X son las bisectrices de los ángulos determinados por XF y la paralela al eje por X .*

Podemos interpretar esto como que un espejo parabólico condensa en su foco todos los rayos de luz que le llegan paralelamente a su eje. Ésta es la razón por la que las antenas parabólicas tienen forma de paraboloide, que es la superficie que se genera al girar una parábola alrededor de su eje.

Ejercicio: Probar que por un punto que no esté sobre una cónica y que no sea un foco pasa exactamente un par de rectas perpendiculares conjugadas respecto a la misma.

Consideremos una cónica real \mathcal{C} que no sea una circunferencia. Sea F uno de sus focos y d la directriz correspondiente. Sea D el punto en que d corta al eje mayor a . Sea H el punto tal que D es el punto medio de F y H . Sea h la paralela a d que pasa por h . Consideremos la homología f de centro F y eje h que transforma D en el punto infinito de DH . Considerando razones dobles se ve enseguida que la imagen del punto infinito ha de ser D (por ser el conjugado armónico de D respecto a F y H). Por otra parte f deja invariante el punto infinito común a d y h , luego $f[d]$ intercambia d con la recta infinita.

Como F es el polo de d respecto a \mathcal{C} , es claro que F será el polo de la recta infinita respecto de $f[\mathcal{C}]$, es decir, F es el centro de \mathcal{C} .

Las tangentes a \mathcal{C} que pasan por F tocan a \mathcal{C} en dos puntos de d , y f los transforma en los puntos infinitos de dichas tangentes, pero por definición de foco éstos son los puntos circulares. Así pues, los puntos circulares están en $f[\mathcal{C}]$, luego $f[\mathcal{C}]$ es una circunferencia.

Sea P' cualquier punto de \mathcal{C} y $P = f(P')$. Sea $X = PP' \cap h$ y P_∞ el punto infinito de PP' . Entonces $XP_\infty PF \bar{\wedge} P_\infty XP'F$, luego $\mathcal{R}(X, P_\infty, P, F) = \mathcal{R}(P_\infty, X, P', F) = \mathcal{R}(X, P_\infty, F, P')$, es decir,

$$\frac{\overrightarrow{XP}}{\overrightarrow{XF}} = \frac{\overrightarrow{XF}}{\overrightarrow{XP'}},$$

Esto vale para las circunferencias si convenimos que su excentricidad es 0.

Por otra parte, dado un punto cualquiera F , una recta cualquiera d que no pase por O y un número $e > 0$, la figura \mathcal{C} de los puntos tales que la razón de las distancias a F y d es e forman una cónica real no degenerada. En efecto, tomando un sistema de referencia ortonormal adecuado podemos suponer que $F = (0, 0)$ y $d \equiv x = D$, con $D > 0$. Entonces la condición equivale a

$$x^2 + y^2 = e^2(D - x)^2.$$

Esta ecuación corresponde a la cónica de matriz

$$A = \begin{pmatrix} 1 - e^2 & 0 & e^2 D \\ 0 & 1 & 0 \\ e^2 D & 0 & -e^2 D^2 \end{pmatrix}. \quad (10.2)$$

Su determinante es $-e^2 D^2$, luego no es degenerada. Además haciendo $y = 0$ obtenemos que la ecuación se cumple para $x = \frac{e}{e+1}D$, luego \mathcal{C} no es imaginaria.

Más aún, F es un foco de \mathcal{C} . Para probarlo notamos primero que $d = p(F)$, luego la intersección de D con \mathcal{C} está formada por los puntos de contacto de las tangentes que pasan por F . Haciendo $x = D$ en la ecuación obtenemos que éstos son $(D, \pm Di, 1)$, luego las tangentes son las rectas que pasan por estos puntos y $(0, 0, 1)$. Es claro que contienen a los puntos $(D, \pm Di, 0) = (1, \pm i, 0)$, que son los puntos circulares, luego F es ciertamente un foco y d la directriz correspondiente. Así pues:

Teorema 10.54 *Dada una recta d , un punto F que no esté en D y un número $e > 0$, la figura formada por todos los puntos tales que la razón de las distancias a F y d sea igual a e es una cónica real con un foco igual a F , directriz d y excentricidad e . Toda cónica real que no sea una circunferencia es de esta forma.*

Esto implica que toda cónica que no sea una circunferencia está completamente determinada por su excentricidad, uno de sus focos y su directriz correspondiente. Las semejanzas conservan las proporciones, por lo que también conservan la excentricidad de las cónicas. Como fijan o invierten los puntos circulares, es claro que conservan los focos (luego también las directrices). Por otra parte, dadas dos cónicas, siempre es posible transformar la primera mediante una semejanza para obtener otra con un foco y una directriz común con la segunda. Si ambas tenían la misma excentricidad, entonces la cónica que obtenemos coincide con la segunda. En resumen:

Teorema 10.55 *Dos cónicas reales son semejantes si y sólo si tienen la misma excentricidad.*

Como siempre, el argumento no incluye las circunferencias, pero el resultado es válido igualmente para ellas. En particular dos parábolas cualesquiera son semejantes.

Algunas relaciones Vamos a probar varias relaciones importantes entre los elementos de una cónica, la mayoría de las cuales las conocemos ya de la sección 4.8. Conviene introducir algunas definiciones:

Llamaremos *semiejes mayores* de una elipse a los segmentos que unen su centro con sus vértices en el eje mayor. También llamaremos semieje mayor a la longitud de estos segmentos, y la representaremos por a . Análogamente se define el semieje menor, que representaremos por b . En una hipérbola sólo tenemos definido el semieje mayor. La *distancia focal* de una elipse o una hipérbola es la distancia de su centro O a cualquiera de sus focos, y la representaremos por c . Para una hipérbola, definimos formalmente $b = \sqrt{c^2 - a^2}$.

Teorema 10.56 *Sea \mathcal{C} una cónica real. Entonces:*

1. Si \mathcal{C} es una elipse, sus puntos son exactamente los que cumplen

$$\overline{PF_1} + \overline{PF_2} = 2a.$$

2. Si \mathcal{C} es una hipérbola, sus puntos son exactamente los que cumplen

$$\overline{PF_1} - \overline{PF_2} = \pm 2a.$$

3. Si \mathcal{C} es una elipse, la distancia de un vértice menor a uno de los focos es a , y se cumple $a^2 = b^2 + c^2$.

4. Si \mathcal{C} es una elipse o una hipérbola, $e = c/a$.

5. Si \mathcal{C} es una hipérbola, el ángulo que forman sus asíntotas es el dado por

$$\cos \theta = 1 - \frac{2}{e^2}.$$

DEMOSTRACIÓN: 1) Dado un punto P de \mathcal{C} , sean D_1 y D_2 los puntos de las directrices donde se alcanza la distancia de éstas a P . Entonces

$$\overline{PF_1} + \overline{PF_2} = e \overline{PD_1} + e \overline{PD_2} = e \overline{D_1 D_2}.$$

Por lo tanto la suma de las distancias a los focos es constante. Evaluándola en un vértice es fácil ver que es exactamente $2a$. Se comprueba sin dificultad que los puntos que cumplen esta ecuación forman una cónica que contiene a \mathcal{C} , luego ha de ser \mathcal{C} .

2) La prueba es similar a la anterior, salvo que ahora $\overline{PD_1} - \overline{PD_2}$ es la distancia entre las dos directrices con signo positivo o negativo según la rama en que esté P , luego la diferencia de distancias sólo es constante en módulo.

3) Si B es uno de los vértices menores, tenemos que $\overline{BF_1} + \overline{BF_2} = 2a$, pero ambos sumandos son iguales. Si A es un vértice mayor, el triángulo $\overline{BAF_1}$ es rectángulo, lo que nos da la relación $a^2 = b^2 + c^2$. En particular vemos que el semieje mayor es ciertamente mayor que el semieje menor.

4) La matriz (10.2) corresponde a una cónica arbitraria con un foco en $(0, 0)$ y directriz $x = D$. Es fácil calcular su centro y su vértice:

$$O\left(-\frac{e^2 D}{1-e^2}\right), \quad A\left(\frac{eD}{1+e}, 0\right),$$

luego

$$a = \frac{ed}{|1-e^2|}, \quad c = \frac{e^2 D}{|1-e^2|},$$

luego ciertamente $e = c/a$.

5) La cónica de matriz A según 10.2 corta a la recta infinita en los puntos $(1, \pm\sqrt{e^2-1}, 0)$, luego sus asíntotas son las polares de estos puntos, cuyas ecuaciones son $(1-e^2)x \pm \sqrt{e^2-1}y + e^2D = 0$. Los vectores directores son $(\pm\sqrt{e^2-1}, 1-e^2)$ y al calcular el ángulo que forman obtenemos la fórmula pedida. ■

Las hipérbolas cuyas asíntotas forman ángulos rectos se llaman *equiláteras*. Según el teorema anterior son exactamente las de excentricidad $e = \sqrt{2}$.

Ejercicio: Si A_1 y A_2 son los vértices mayores de una elipse o una hipérbola y F es uno de sus focos, $\overline{FA_1} \overline{FA_2} = b^2$.

10.5 Espacios de dimensiones superiores

Un espacio parabólico es un espacio afín en cuyo hiperplano infinito se ha fijado una polaridad simétrica. Si no existen puntos autopolares obtenemos el espacio euclídeo usual. En general, fijar una polaridad simétrica en el hiperplano infinito es equivalente a fijar una forma bilineal simétrica en el espacio vectorial asociado al espacio afín (en realidad cada polaridad determina una forma bilineal salvo una constante). El número de geometrías reales parabólicas distintas aumenta con la dimensión. En dimensión 3 todavía tenemos sólo dos, una correspondiente a la polaridad de índice 0 (la geometría euclídea) y otra a la polaridad de índice 1, que es el análogo tridimensional del plano de Lorentz. En dimensión 4 tenemos ya tres espacios, el euclídeo, de índice 0, el espacio de Minkowski, de índice 1, que es el auténtico espacio de la teoría de la relatividad especial, y el de índice 2, que puede verse como suma de dos planos de Lorentz.

Un estudio sintético de los espacios parabólicos puede volverse poco práctico en dimensiones superiores, pero un enfoque algebraico desde el punto de vista de la geometría proyectiva puede ser en ciertos casos preferible al tratamiento puramente afín, que es el más habitual. El lector no debería tener problemas en generalizar a dimensión tres los resultados algebraicos que hemos obtenido en este capítulo. En tal caso el plano infinito puede contener una *cónica circular* en lugar de dos puntos circulares.

El papel análogo al de las cónicas lo juegan las llamadas *cuádricas*, que algebraicamente pueden definirse como las figuras formadas por los puntos que

en un sistema de referencia satisfacen una ecuación matricial similar a la de las cónicas. Entre ellas se encuentran las esferas, que juegan el papel de las circunferencias.

Capítulo XI

La geometría circular

Sabemos que los puntos de una cónica se corresponden con los de una recta a través de las proyecciones perspectivas con centro en un punto de la cónica. Estas proyecciones son homografías, lo que en muchas ocasiones nos ha permitido tratar a las cónicas como rectas o viceversa. En este capítulo vamos a estudiar una relación similar entre figuras con una dimensión más. No hemos desarrollado la teoría de las “cónicas bidimensionales”, las llamadas cuádricas, pero no perdemos generalidad, y de hecho es más natural, si nos centramos en la relación entre el plano y la esfera.

11.1 La proyección estereográfica

Aunque no hemos hablado de esferas hasta ahora, los hechos que necesitamos sobre ellas son todos elementales y requieren poco más que el teorema de Pitágoras. Comencemos con la definición.

Definición 11.1 Una *esfera* de centro un punto O y radio $r > 0$ en un espacio tridimensional euclídeo E es el conjunto de todos los puntos de E cuya distancia a O es igual a r .

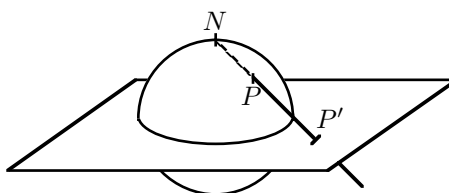
Si fijamos un sistema de referencia con origen en O , es claro que la esfera está formada por los puntos cuyas coordenadas cumplen la ecuación

$$x^2 + y^2 + z^2 = r^2.$$

Si π es un plano, la distancia de O a π se define como la distancia de O al punto donde la perpendicular a π por O corta a π . Es fácil ver que es la menor distancia posible entre O y un punto de π . Se comprueba sin dificultad que si S es la esfera de centro O y radio r y si la distancia de O a π es d , entonces $S \cap \pi$ es vacía si $d > r$, un punto si $d = r$ o una circunferencia de radio $\sqrt{r^2 - d^2}$ si $d < r$. Recíprocamente, toda circunferencia contenida en S es la intersección con S del plano que la contiene.

Es muy práctico referirse a una esfera con el vocabulario propio de la geografía. Fijado un sistema de referencia con origen en su centro, el punto $(0, 0, r)$ es el *polo norte*, el punto $(0, 0, -r)$ es el *polo sur*, la circunferencia $z = 0$ es el *ecuador*, las circunferencias que pasan por ambos polos son los *meridianos* y las circunferencias determinadas por planos paralelos al ecuador son los *paralelos*.

Ahora podemos definir la *proyección estereográfica* de una esfera S en su plano ecuatorial π como la aplicación que a un punto P de S que no sea el polo norte N le hace corresponder el punto P' donde la recta NP corta a π . Observar que las rectas que pasan por N y son paralelas a π forman un plano que dista r del centro de la esfera, luego ninguna de ellas corta a S en otro punto P , luego la proyección está bien definida y es fácil ver que biyecta $S \setminus \{N\}$ con π .



Conviene calcular su expresión en coordenadas. Por simplicidad supondremos que la esfera tiene radio 1. Dado uno de sus puntos $P = (x, y, z)$ que no sea el polo norte, la recta que lo une con N es

$$(0, 0, 1) + \lambda(x, y, z - 1), \quad \lambda \in \mathbb{R}.$$

El valor de λ que hace que este punto esté en π es el que cumple $1 + \lambda(z - 1) = 0$, o sea, $\lambda = 1/(1 - z)$. El punto de corte es

$$P' = \left(\frac{x}{1 - z}, \frac{y}{1 - z}, 0 \right).$$

Así pues, si tomamos como sistema de referencia en π el formado por O y los dos primeros vectores de la base del sistema de referencia de E , la expresión en coordenadas de la proyección estereográfica es

$$f(x, y, z) = \left(\frac{x}{1 - z}, \frac{y}{1 - z} \right).$$

También es fácil obtener una expresión para la inversa. Dado un punto $P' = (a, b)$ en π , la recta que pasa por N y P' es

$$(0, 0, 1) + \lambda(a, b, -1), \quad \lambda \in \mathbb{R},$$

y los puntos de esta forma que están en la esfera son los que cumplen

$$(\lambda a)^2 + (\lambda b)^2 + (1 - \lambda)^2 = 1,$$

lo que implica que o bien $\lambda = 0$ (que corresponde al polo norte) o bien

$$\lambda = \frac{2}{a^2 + b^2 + 1},$$

que corresponde al punto

$$g(a, b) = \left(\frac{2a}{a^2 + b^2 + 1}, \frac{2b}{a^2 + b^2 + 1}, \frac{a^2 + b^2 - 1}{a^2 + b^2 + 1} \right). \quad (11.1)$$

La situación que encontramos no es completamente análoga al caso bidimensional. Al proyectar una cónica en una recta podíamos asignar también una imagen al centro de la proyección de modo que toda la cónica se correspondía con toda la recta proyectiva. En cambio, si tratamos de extender la proyección estereográfica a la compleción proyectiva de π nos encontramos con que hay infinitos puntos infinitos a los que tenemos que asignar una antiimagen mientras que sólo nos queda un punto en la esfera al que asignar una imagen.

Por lo tanto, tenemos que los puntos de la esfera no se corresponden con los del plano proyectivo, sino con los del plano completado con un único punto infinito. Llegados aquí la teoría encaja con otra parte de la geometría proyectiva. Podemos identificar los puntos del plano afín real con los de la recta afín compleja, con la diferencia de que la compleción proyectiva de la recta compleja se obtiene adjuntando un solo punto. Así pues, resulta que los puntos de la esfera se corresponden de forma natural con los de la recta proyectiva compleja. Fijado un sistema de referencia en el espacio de la esfera y otro en la recta compleja, una biyección entre esfera y recta es la dada por

$$f(x, y, z) = \frac{x}{1-z} + i \frac{y}{1-z},$$

entendiendo que $f(0, 0, 1) = \infty$. Alternativamente, $f(x, y, z) = (x + yi, 1 - z)$.

Aunque desde un punto de vista estricto la recta proyectiva compleja es toda ella una recta, sin más distinción, lo cierto es que, fijado en ella un sistema de referencia proyectivo, podemos pensar en ella como $\mathbb{R}^2 \cup \{\infty\}$ y hablar de rectas y circunferencias reales. Enseguida veremos que conviene adoptar el convenio siguiente:

Consideraremos que el punto infinito pertenece a todas las rectas de \mathbb{R}^2 , de modo que es colineal con cualquier par de puntos finitos.

Por ejemplo, con este convenio podemos enunciar:

Teorema 11.2 *La proyección estereográfica biyecta las circunferencias de la esfera con las rectas y las circunferencias de $\mathbb{R}^2 \cup \{\infty\}$.*

DEMOSTRACIÓN: Una circunferencia en la esfera S está formada por los puntos que cumplen la ecuación de un plano: $Ax + By + Cz + D = 0$. La expresión (11.1) nos da entonces que las coordenadas (a, b) de sus imágenes han de cumplir

$$2Aa + aBb + C(a^2 + b^2 - 1) + D(a^2 + b^2 + 1) = 0,$$

o equivalentemente,

$$(C + D)(a^2 + b^2) + 2Aa + 2Bb + (D - C) = 0. \quad (11.2)$$

Recíprocamente, todos los puntos que cumplen (11.2) provienen de puntos en la circunferencia dada. Ahora bien, esta última ecuación corresponde a una circunferencia si $C \neq -D$ y a una recta si $C = -D$. Notemos que esto sucede si y sólo si el polo norte $(0, 0, 1)$ está en la circunferencia de partida, luego las circunferencias que pasan por el polo norte son las que se transforman en rectas, las cuales contienen, por convenio, la imagen del polo norte.

Por otra parte es claro que ajustando A , B , C y D se puede hacer que los coeficientes de (11.2) tomen cualquier conjunto de valores, luego todas las circunferencias y todas las rectas del plano son imágenes de circunferencias de la esfera. ■

11.2 Transformaciones circulares

En esta sección describiremos las homografías de la recta proyectiva compleja $P^1(\mathbb{C}) = \mathbb{R}^2 \cup \{\infty\}$ desde el punto de vista de la geometría euclídea de \mathbb{R}^2 . En la sección siguiente las relacionaremos con la esfera a través de la proyección estereográfica.

Sabemos que las homografías de la recta compleja son las aplicaciones de la forma

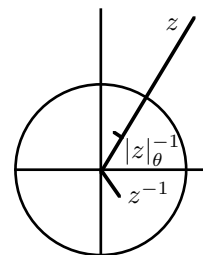
$$f(z) = \frac{az + b}{cz + d}, \quad \text{con } ad - bc \neq 0.$$

Si $c = 0$ tenemos simplemente una aplicación de la forma $f(z) = az + b$. Si $c \neq 0$ podemos expresarla como

$$f(z) = \frac{a}{c} + \frac{b - \frac{ad}{c}}{cz + d},$$

con lo que en cualquier caso f se expresa como composición de aplicaciones de la forma $z \mapsto az + b$ con la inversión $z \mapsto 1/z$. Vamos a estudiar por separado cada una de estas aplicaciones. Comencemos por la última.

La aplicación $z \mapsto 1/z$ transforma 0 en el punto infinito. Si $z \neq 0$ entonces admite una expresión en coordenadas polares como $z = |z|_\theta$, y entonces su inverso es $|z|_\theta^{-1}$. Ahora bien, el punto $|z|_\theta^{-1}$ es el inverso de z respecto a la circunferencia de centro 0 y radio 1 en el sentido de la definición 10.27, luego $z \mapsto z^{-1}$ es la composición de esta inversión con la reflexión respecto al eje real.



Consideremos ahora las aplicaciones de la forma $z \mapsto az + b$. Podemos descomponerlas como composición de un producto $z \mapsto az$ y una suma $z \mapsto b + z$. La suma es geoméricamente una traslación, luego se expresa como producto de dos reflexiones. El producto se descompone a su vez en los productos

$$z \mapsto \frac{a}{|a|} z \quad \text{y} \quad z \mapsto |a|z.$$

El primero es un giro, luego también es un producto de dos reflexiones. El segundo es una homotecia de centro 0. Veamos que puede expresarse como composición de dos inversiones respecto a dos circunferencias de centro 0. En efecto, la inversión respecto a la circunferencia unitaria transforma cada número z de argumento θ en $|z|_\theta^{-1}$, y si ahora aplicamos la inversión respecto a la circunferencia de radio r obtenemos $r^2|z|_\theta$. Si $r = \sqrt{|a|}$ esto es $|a|z$.

Con esto tenemos probada la mitad del teorema siguiente:

Teorema 11.3 *Las homografías de $P^1(\mathbb{C})$ son las aplicaciones expresables como composición de un número par de reflexiones respecto de rectas e inversiones respecto de circunferencias.*

DEMOSTRACIÓN: Hay que probar que un producto de dos reflexiones/inversiones es una homografía. Observar que las inversiones respecto a circunferencias no son biyectivas en el plano proyectivo euclídeo, pues hacen corresponder el centro de la circunferencia con todos los puntos infinitos, pero sí lo son en $P^1(\mathbb{C})$, si convenimos que intercambian el centro con el único punto infinito. Del mismo modo podemos considerar a las biyecciones afines de \mathbb{R}^2 como biyecciones en $P^1(\mathbb{C})$ si convenimos que todas ellas fijan a ∞ .

El producto de dos reflexiones es un giro o una traslación. Las traslaciones son de la forma $z \mapsto z + a$, luego son homografías. Si f es un giro de centro O , entonces $z \mapsto f(z + O) - O$ es un giro de centro O , luego existe un número complejo a de módulo 1 tal que $f(z + O) - O = az$, luego $f(z) = a(z - O) + O$, luego es una homografía.

Consideremos ahora una reflexión f y una inversión g . Supongamos primero que el eje de f pasa por el centro de la circunferencia fijada por g . Consideremos una traslación que lleve dicho centro al punto 0, seguida de un giro de centro 0 que transforme el trasladado del eje de f en el eje real, seguido de una homotecia de centro 0 que transforme la traslación de la circunferencia de g en la circunferencia unidad. Si llamamos h a esta composición, es claro que h es una homografía. Más aún, es una semejanza del plano euclídeo. Es fácil ver que $h^{-1}fh$ es la reflexión respecto al eje real y que $h^{-1}gh$ es la inversión respecto de la circunferencia unidad. Por consiguiente $h^{-1}fgh^{-1}gh$ y $h^{-1}ghh^{-1}fh$ son ambos iguales a la homografía $z \mapsto 1/z$, luego fg y gf también son homografías.

Supongamos ahora que el eje de f no pasa por el centro de la circunferencia de g . Sea f' una reflexión respecto a una recta que pase por dicho centro. Entonces $fg = ff'f'g$ es el producto de dos homografías, luego es una homografía, e igualmente se prueba que gf lo es.

Si tenemos una composición de dos inversiones f y g , tomamos como h la reflexión respecto a la recta que pasa por los centros de sus circunferencias (una cualquiera si ambos coinciden). Entonces $fg = fhgh$ es un producto de dos homografías, luego es una homografía, y lo mismo vale para gf . ■

Definición 11.4 Llamaremos *transformaciones circulares* a las biyecciones de $P^1(\mathbb{C})$ en sí mismo que se expresan como producto de reflexiones e inversiones.

Una transformación circular es *directa* si se expresa como producto de un número par de reflexiones e inversiones. En caso contrario es *inversa*. Llamaremos *grupo circular* al grupo de todas las transformaciones circulares.

Hemos probado que las homografías en \mathbb{C} son las transformaciones circulares directas. Una transformación circular no puede ser a la vez directa e inversa, pues en tal caso podríamos expresar una reflexión o una inversión como producto de un número par de ellas, luego sería una homografía, pero esto es imposible, pues las reflexiones y las inversiones fijan a infinitos puntos.

Es claro que las homografías que fijan a ∞ son las de la forma $z \mapsto az + b$ y hemos visto que éstas se descomponen en producto de un número par de reflexiones (sin que aparezcan inversiones). Más en general, si f es una transformación circular que fija a ∞ , o bien es una homografía y estamos en el caso anterior, o bien es inversa, y entonces podemos tomar cualquier reflexión g y tenemos que $h = fg$ es una homografía que fija a ∞ , luego $f = hg$ es también producto de reflexiones. Puesto que las composiciones de reflexiones son las semejanzas de \mathbb{R}^2 , hemos probado lo siguiente:

Teorema 11.5 *Las transformaciones circulares que fijan a ∞ son exactamente las semejanzas de \mathbb{R}^2 .*

Si una transformación circular f es directa y deja fijos a tres puntos (finitos o no) entonces es la identidad, porque es una homografía. Si f es inversa esto ya no es cierto, pero sólo hay dos posibilidades. En efecto, sea g la reflexión respecto de la recta que pasa por los tres puntos fijos si son colineales o la inversión respecto de la circunferencia que pasa por ellos si no lo son. Entonces fg es una homografía que fija a los tres puntos, luego es la identidad, luego $f = g$. Hemos probado el teorema siguiente:

Teorema 11.6 *Si una transformación circular fija a tres puntos (finitos o no), entonces es la identidad, una reflexión o una inversión.*

De aquí podemos deducir muchas consecuencias interesantes:

Teorema 11.7 *Si f es una transformación circular y g es una reflexión o una inversión, entonces $f^{-1}gf$ es una reflexión o una inversión.*

DEMOSTRACIÓN: Sean A, B, C tres puntos fijados por g . Entonces $f^{-1}gf$ fija a $f(A)$, $f(B)$ y $f(C)$, luego por el teorema anterior es la identidad, una reflexión o una inversión. Como es una transformación inversa no puede ser la identidad. ■

Teorema 11.8 *Las transformaciones circulares transforman cada circunferencia y cada recta en una circunferencia o una recta.*

DEMOSTRACIÓN: Sea f una transformación circular. Dada una recta o circunferencia C , sea g la reflexión o inversión que fija a sus puntos. Entonces $f^{-1}gf$ es también una reflexión o una inversión, y fija exactamente a los puntos de $f[C]$, luego $f[C]$ es una recta o una circunferencia. ■

Teorema 11.9 *Si C y C' son dos rectas/circunferencias, existe una transformación circular que transforma una en otra.*

DEMOSTRACIÓN: Tomemos tres puntos P, Q, R en C y tres puntos P', Q', R' en C' y sea f la homografía $PQR \xrightarrow{\wedge} P'Q'R'$. La imagen de C por f contiene tres puntos de C' , luego ha de ser C' , pues por tres puntos pasa una única recta/circunferencia. ■

En vista de los teoremas anteriores resulta razonable introducir los convenios siguientes:

Definición 11.10 Llamaremos *circunferencias* de $P^1(\mathbb{C})$ a las rectas y a las circunferencias. Llamaremos *inversiones* tanto a las reflexiones como a las inversiones (respecto a circunferencias).

En estos términos las transformaciones circulares transforman circunferencias en circunferencias, y cualquier par de circunferencias están conectadas por una transformación circular. Por tres puntos cualesquiera pasa una única circunferencia.

El teorema fundamental de la geometría afín nos permite caracterizar las transformaciones circulares:

Teorema 11.11 *Una biyección de $P^1(\mathbb{C})$ en sí mismo es una transformación circular si y sólo si transforma circunferencias en circunferencias.*

DEMOSTRACIÓN: Sea f una biyección que transforme circunferencias en circunferencias. Sea $f(\infty) = P$. Sea g cualquier transformación circular que cumpla $g(P) = \infty$. Entonces fg transforma circunferencias en circunferencias y deja fijo a ∞ , luego vista como aplicación en \mathbb{R}^2 transforma rectas en rectas. Por el teorema fundamental de la geometría afín se trata de una biyección afín, luego se extiende a una homografía en \mathbb{R}^2 . Como transforma circunferencias en circunferencias es una semejanza, luego es una transformación circular y f también. ■

Ejercicio: Probar que cada circunferencia C divide al plano en dos partes, de modo que dos puntos están en la misma parte si y sólo si existe un segmento de circunferencia que los contiene y no corta a C . Además una transformación circular hace corresponder dos puntos a un mismo lado de una circunferencia con puntos a un mismo lado de su imagen.

11.3 Homografías en la esfera

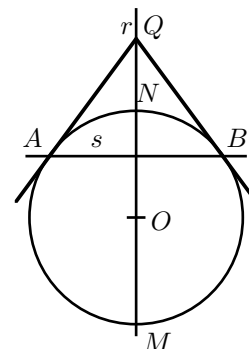
Los resultados de la sección anterior muestran que la geometría circular no distingue las rectas de las circunferencias. Esto resulta completamente natural cuando se relaciona con la geometría de la esfera.

Definición 11.12 Sea S una esfera en un espacio tridimensional euclídeo E . Una *homografía* en S es una biyección $f : S \rightarrow S$ inducida por restricción a partir de una homografía en E que fije a S .

Vamos a probar que la proyección estereográfica hace corresponder las homografías de una esfera con las transformaciones circulares de la recta compleja.

Las homografías de E transforman planos en planos, luego las homografías de una esfera S transforman circunferencias en circunferencias. Cuando componemos una homografía f en S con la proyección estereográfica p obtenemos una biyección $p^{-1}fp$ de la recta compleja en sí misma que claramente transforma circunferencias en circunferencias, luego es una transformación circular. Hemos de probar que toda transformación circular puede obtenerse de este modo. De hecho basta probar que todas las reflexiones de la recta compleja provienen de homografías en la esfera, ahora bien, dada una inversión f , la antiimagen en S del conjunto de puntos fijados por f es una circunferencia en S , luego si probamos que existe una homografía que fija a dicha circunferencia, su correspondiente en $P^1(\mathbb{C})$ será una transformación circular que fija a la misma circunferencia que f , luego será f . En resumen, basta probar que para cada circunferencia C en S existe una homografía en S que no es la identidad pero fija a todos los puntos de C .

Sea π el plano de C , sea r la perpendicular a π que pasa por el centro O de S , sean N y M los puntos donde r corta a S , sea π' cualquier plano que contenga a r , sea \mathcal{C} la circunferencia en que π' corta a S , sea $s = \pi \cap \pi'$, que es una recta perpendicular a r . Sean A y B los puntos donde s corta a r , sea Q el polo de s .



El punto Q es el conjugado armónico de O respecto a M y N , luego no depende de la elección de π' . La involución en \mathcal{C} de centro Q cumple $ABMN \bar{\wedge} ABNM$. Consideremos ahora la homología f de centro Q y eje π que envía M a N . Su restricción a cada plano π' cumple $ABM \bar{\wedge} ABN$, luego es de hecho la involución anterior, luego en particular fija a $S \cap \pi'$, luego f fija a S , con lo que induce una homografía en S que ciertamente fija a cada punto de C .

Con esto tenemos demostrado el teorema que perseguíamos:

Teorema 11.13 *La proyección estereográfica induce un isomorfismo entre el grupo de homografías de una esfera y el grupo de transformaciones circulares de la recta proyectiva compleja.*

Este teorema nos permite traducir a términos de homografías en esferas los resultados que ya conocemos sobre transformaciones circulares. Por ejemplo, dada una circunferencia C en una esfera S , existe una única homografía en S distinta de la identidad que fija a todos sus puntos, a la cual llamaremos *inversión* respecto a C . Toda homografía en una esfera es producto de reflexiones. Si una homografía fija a tres puntos es la identidad o la inversión respecto a la circunferencia que los contiene.

A su vez, el teorema anterior nos aporta un dato importante sobre las transformaciones circulares:

Teorema 11.14 *La restricción de una transformación circular a una circunferencia es una homografía. Más aún, cada homografía entre dos circunferencias se extiende a una única transformación circular directa.*

DEMOSTRACIÓN: La restricción de una transformación circular a una circunferencia C puede descomponerse en producto de tres homografías (elegimos el sistema de referencia de modo que ni C ni su imagen contengan el punto infinito): primero la inversa de la proyección estereográfica, que es una proyección perspectiva (entre dos planos) que transforma C en una circunferencia C' en S , luego la homografía que induce en la esfera la homografía correspondiente a la transformación dada, que transforma C' en otra circunferencia C'' , y por último la restricción a C'' de la proyección estereográfica, que es de nuevo una proyección perspectiva.

Dada una homografía entre dos circunferencias, tomamos tres puntos de la primera y consideramos la transformación circular directa que coincide sobre ellos con la homografía dada. Obviamente ambas han de coincidir sobre toda la circunferencia y la unicidad es clara. ■

11.4 Conservación de ángulos

Vamos a definir el ángulo entre dos circunferencias secantes y probaremos que las transformaciones circulares conservan ángulos. Antes conviene aclarar las posiciones relativas entre dos circunferencias cualesquiera:

Teorema 11.15 *Dos circunferencias distintas tienen cero, uno o dos puntos en común. Según el caso se llaman disjuntas, tangentes o secantes. Dado un punto A en una circunferencia C y un punto B que no esté en C , existe una única circunferencia C' que pasa por B y toca a C en A .*

DEMOSTRACIÓN: Aplicando una transformación circular si es preciso, podemos suponer que una de las circunferencias pasa por ∞ y la otra no, con lo que tenemos una recta y una circunferencia (en el sentido usual), luego efectivamente se cortan en a lo sumo dos puntos.

Respecto a la segunda parte, aplicando una transformación circular podemos suponer que A es el punto infinito, con lo que el resultado equivale a que por un punto exterior a una recta pasa una única paralela. ■

En lo sucesivo convendremos que la tangente a una circunferencia por dos de sus puntos es ella misma. De este modo, si P es un punto de una circunferencia C y Q es otro punto cualquiera (en C o fuera de C), existe una única circunferencia tangente a C que pasa por P y Q . En estos términos la recta tangente a una circunferencia por uno de sus puntos es la circunferencia tangente por dicho punto y por ∞ .

Dos rectas secantes forman dos pares de ángulos opuestos por el vértice. Llamaremos *ángulo* entre ambas a la medida del menor de ellos. El *ángulo* entre

una circunferencia y una recta secante es el ángulo entre la recta y la recta tangente por cualquiera de los puntos de corte. No importa la tangente que se escoja, pues la reflexión respecto a la perpendicular a la recta por el punto medio de los puntos de corte deja invariante a la recta y transforma una tangente en otra. El *ángulo* entre dos circunferencias (usuales) secantes es el ángulo entre sus tangentes por los puntos de corte. También es claro que no importa el punto de corte elegido.

Con esto tenemos definido el ángulo entre dos circunferencias secantes cualesquiera. La definición puede resumirse así: el ángulo entre dos circunferencias es el ángulo entre sus rectas tangentes por ∞ y uno de los puntos de corte.

Teorema 11.16 *Las transformaciones circulares conservan los ángulos entre circunferencias.*

DEMOSTRACIÓN: Sean C_1 y C_2 dos circunferencias secantes y sea P un punto de corte (podemos suponerlo finito). Sea t_i la tangente a C_i por P o bien $t_i = C_i$ si C_i es una recta. Entonces el ángulo entre C_1 y C_2 es por definición el ángulo entre t_1 y t_2 . Una inversión respecto de una circunferencia (usual) de centro P deja invariantes a t_1 y t_2 y transforma cada circunferencia C_i en una recta sin más punto en común con t_i que ∞ (o bien $t_i = C_i$). En cualquier caso transforma a C_i en una recta paralela a t_i , de modo que estas rectas forman entre sí el mismo ángulo que las circunferencias de partida.

Sean ahora C'_i las imágenes de cada C_i por una transformación circular arbitraria. Del mismo modo podemos transformarlas en dos rectas que forman el mismo ángulo que ellas.

El teorema se reduce, pues, a probar que si una transformación circular hace corresponder dos pares de rectas secantes, entonces el ángulo que forman es el mismo.

Sea f una transformación circular que haga corresponder dos rectas secantes r_1 y r_2 con dos rectas secantes r'_1 y r'_2 . Sea O el punto de corte de las primeras y O' el punto de corte de las segundas. Entonces o bien $f(O) = O'$ y $f(\infty) = \infty$ o bien $f(O) = \infty$ y $f(\infty) = O'$. En el segundo caso podemos componer f con una inversión respecto a una circunferencia de centro O' , lo cual deja inalteradas a r'_1 y r'_2 pero hace que ∞ quede fijo. Por lo tanto los dos pares de rectas se corresponden por una semejanza, y las semejanzas conservan los ángulos. ■

Los teoremas anteriores ilustran una potente técnica para obtener resultados sobre circunferencias: reducirlos mediante inversiones a resultados sobre rectas. El ejercicio y los teoremas siguientes son tres ejemplos más.

Ejercicio: Probar que si dos circunferencias son ortogonales, cada una de ellas queda fija por la inversión respecto de la otra.

Teorema 11.17 *Sean C_1 y C_2 dos circunferencias secantes. Sea P un punto de corte y Q un punto cualquiera. Sea C'_1 la tangente a C_1 por P y Q . Entonces el ángulo entre C_1 y C_2 es el mismo que entre C'_1 y C_2 .*

DEMOSTRACIÓN: Basta aplicar una transformación circular f que convierta Q en ∞ . Basta probar que $f[C_1]$ y $f[C_2]$ forman el mismo ángulo que $f[C'_1]$ y $f[C_2]$, pero esto es cierto por definición de ángulo entre circunferencias. ■

Teorema 11.18 Sean P, Q, R y S cuatro puntos colineales y sea C una circunferencia que tenga a PQ por diámetro. Entonces $\mathcal{H}(P, Q; R, S)$ si y sólo si toda circunferencia que pasa por R y S es ortogonal a C .

DEMOSTRACIÓN: Supongamos $\mathcal{H}(P, Q; R, S)$. Aplicando una inversión podemos suponer que P es el punto infinito. Entonces Q es el punto medio de R y S y la circunferencia C es la recta perpendicular a RS por Q . En estas condiciones es fácil ver que toda circunferencia que pasa por R y S es ortogonal a C .

Supongamos ahora que se cumple esta propiedad. Tomemos una circunferencia C' cualquiera que pase por R y S . Por hipótesis corta ortogonalmente a C en dos puntos A y B . Consideremos ahora la circunferencia C'' que pasa por A, R y el conjugado harmónico de R respecto de P y Q . Por la parte anterior C' también es ortogonal a C , ahora bien, por dos puntos A y R pasa una única circunferencia ortogonal a C , luego $C' = C''$, de donde S ha de ser el conjugado harmónico de R respecto de P y Q . ■

Ahora vamos a definir el ángulo entre dos circunferencias en una esfera y probaremos que la proyección estereográfica conserva los ángulos.

Definición 11.19 Si dos circunferencias en una esfera S se cortan en un punto P , definimos el *ángulo* que forman como el ángulo entre sus rectas tangentes por P .

El teorema siguiente prueba en particular que el ángulo entre dos circunferencias secantes en una esfera no depende de cuál de los dos puntos de corte se toma para calcularlo.

Teorema 11.20 La proyección estereográfica conserva ángulos.

DEMOSTRACIÓN: Observemos en primer lugar que si C es una circunferencia en una esfera S y C' es una circunferencia tangente a C por un punto P , entonces ambas tienen la misma recta tangente por P , concretamente, la intersección de los planos de C y C' . Por lo tanto, si tenemos dos circunferencias C_1 y C_2 y C'_1, C'_2 son sus tangentes por P y por otro punto Q , el ángulo entre C_1 y C_2 es el mismo que el ángulo entre C'_1 y C'_2 . Por el teorema 11.17, lo mismo vale para las proyecciones, luego basta probar el teorema para circunferencias que se corten en un par de puntos antípodos P y Q .

Sean p_1 y p_2 las proyecciones estereográficas de una esfera en dos planos π_1 y π_2 . Sea $f : \pi_1 \rightarrow \pi_2$ una isometría entre ellos. Entonces $p_2^{-1}p_1f$ es una biyección de π_2 en sí mismo que transforma circunferencias en circunferencias, luego es una transformación circular, luego conserva ángulos, luego $p_2^{-1}p_1$ también conserva ángulos.

Esto implica que al proyectar dos circunferencias de una esfera en un plano obtenemos dos circunferencias que forman el mismo ángulo independientemente del plano sobre el que proyectemos. En particular, podemos utilizar la proyección respecto al plano perpendicular a la recta PQ .

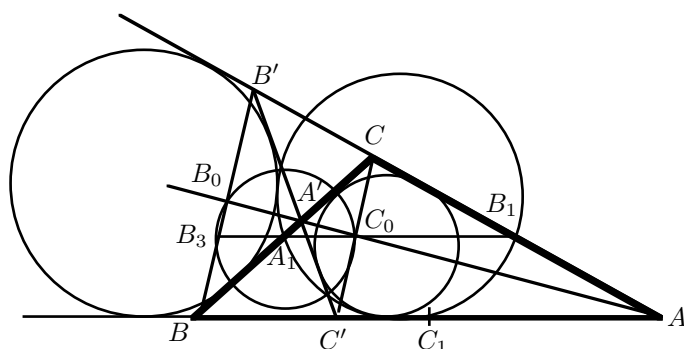
En estas condiciones es fácil ver que las rectas tangentes a C_1 y C_2 tanto por P como por Q son paralelas a las proyecciones de C_1 y C_2 , luego los ángulos son los mismos. ■

11.5 El teorema de Feuerbach

Terminamos el capítulo con una interesante aplicación de la geometría circular a la geometría euclídea:

Teorema 11.21 (Teorema de Feuerbach) *El círculo de los nueve puntos de un triángulo es tangente a sus cuatro circunferencias tritangentes.*

DEMOSTRACIÓN: Consideremos un triángulo \widehat{ABC} . Sean A_1 , B_1 y C_1 los puntos medios de sus lados, etiquetados como indica la figura. El círculo de los nueve puntos es el que pasa por estos tres puntos medios.



Consideremos la circunferencia inscrita \mathcal{C}_1 y una de las circunferencias circunscritas \mathcal{C}_2 , concretamente la que tiene su centro en la bisectriz de \hat{A} . Si el lado BC es perpendicular a esta bisectriz, entonces es fácil ver que el triángulo es isósceles, ambas circunferencias tocan a BC en su punto medio A_1 y el círculo de los nueve puntos tiene también a BC por tangente, con lo que efectivamente toca a \mathcal{C}_1 y \mathcal{C}_2 .

Supongamos que BC no es perpendicular a la bisectriz de \hat{A} . Entonces la reflexión respecto a ella fija a \mathcal{C}_1 y \mathcal{C}_2 , luego transforma a BC en otra recta tangente ambas circunferencias. Digamos que ésta corta a los lados del triángulo en los puntos A' , B' y C' . El punto A' está sobre la bisectriz de \hat{A} .

Las rectas AB , AC , BC y $B'C'$ son los lados de un cuadrilátero completo dual circunscrito tanto a \mathcal{C}_1 como a \mathcal{C}_2 , luego su triángulo diagonal es autopolar para ambas. Sus lados son las rectas AA' , BB' , CC' . Ahora bien, como el lado AA' pasa por los centros de las circunferencias, su polo es infinito, luego las rectas BB' y CC' son paralelas y perpendiculares a AA' . Sean B_0 y C_0 los puntos de corte, es decir, los vértices finitos del triángulo. Notar que el cuadrilátero completo de vértices B , B' , C , C' prueba que $\mathcal{H}(B_0, C_0; A, A')$.

Teniendo en cuenta que A_1 es el punto medio de \overline{BC} , es fácil ver que B_0 y C_0 están a la misma distancia de A_1 , luego podemos trazar una circunferencia \mathcal{C} de centro A_1 que pase por B_0 y C_0 .

Los puntos B_0 y C_0 son conjugados respecto de las circunferencias \mathcal{C}_1 y \mathcal{C}_2 y son colineales con sus centros, luego están armónicamente separados por los dos pares de puntos donde éstas cortan a AA' . Podemos aplicar el teorema 11.18 y concluir que \mathcal{C} es ortogonal tanto a \mathcal{C}_1 como a \mathcal{C}_2 . Por lo tanto la inversión respecto a \mathcal{C} deja fijas a cada una de ellas. Para probar el teorema bastará ver que esta inversión transforma el círculo de los nueve puntos en la recta $B'C'$, pues ésta es tangente a las circunferencias \mathcal{C}_i , luego el círculo de los nueve puntos también lo será.

La recta A_1B_1 es paralela a AB , luego no es paralela a $B'C'$. Sea B_2 el punto de corte. Notemos que como AA' es la bisectriz de \hat{A} , el punto C_0 es el punto medio del segmento $\overline{CC'}$. La recta A_1B_1 pasa por el punto medio de \overline{CB} y es paralela a BC' , luego contiene a C_0 . Por otra parte, del hecho de que BB' y CC' sean paralelas junto con que A_1 es el punto medio de \overline{BC} se sigue que A_1 es el punto medio del segmento que une C_0 con el punto B_3 en que A_1B_1 corta a BB' . Por consiguiente B_3 también está en \mathcal{C} . Ahora observamos que

$$B_0C_0AA' \stackrel{B'}{\underset{\wedge}{\sim}} B_3C_0B_1B_2,$$

y como $\mathcal{H}(B_0, C_0; A, A')$, también $\mathcal{H}(B_3, C_0; B_1, B_2)$, luego la involución que \mathcal{C} induce en la recta A_1B_1 por inversión (que fija a C_0 y B_3) conjugará B_1 con B_2 , es decir, transforma B_1 en un punto de la recta $B'C'$. Simétricamente se prueba que el inverso de C_1 está también sobre $B'C'$ y ciertamente el inverso de A_1 es el punto infinito, luego el inverso del círculo de los nueve puntos es la recta $B'C'$. ■

A causa de este teorema, el círculo de los nueve puntos se conoce también como círculo de Feuerbach.

Capítulo XII

La geometría hiperbólica

La geometría hiperbólica es la geometría no euclídea más próxima a la geometría euclídea, pues satisface todos los axiomas de la geometría absoluta (grupos A, B, C y D) y, por consiguiente, se distingue tan sólo en que incumple el axioma de las paralelas. Aunque podemos definir espacios hiperbólicos de cualquier dimensión, nos limitaremos a estudiar el caso bidimensional por simplicidad. Un plano hiperbólico puede obtenerse a partir de un plano proyectivo por un proceso formalmente similar a como se obtiene un plano euclídeo. Para obtener un plano euclídeo primero seleccionamos una recta, lo que nos permite distinguir entre puntos finitos e infinitos y definir la noción de paralelismo, con lo que tenemos un espacio afín. Después fijamos una involución elíptica en la recta infinita, con lo que podemos definir la perpendicularidad, las semejanzas y las isometrías. Para obtener un plano hiperbólico basta seleccionar una cónica real en lugar de una recta. Esto nos divide los puntos del plano proyectivo en tres clases: puntos finitos (los interiores a la cónica), puntos infinitos (los de la cónica) y puntos ultrainfinitos (los exteriores a la cónica). El plano hiperbólico propiamente dicho está formado por los puntos finitos, de modo que dos rectas que se corten en un punto infinito o ultrainfinito son disjuntas desde el punto de vista hiperbólico. Es fácil ver entonces que por un punto exterior a una recta pasan infinitas paralelas.

Una diferencia importante respecto a la geometría afín es que ahora no necesitamos introducir ninguna estructura adicional para definir la perpendicularidad. El papel de la involución ortogonal lo juega aquí la polaridad de la cónica, y el hecho de que esté determinada por la propia cónica termina traducándose en que los conceptos análogos a los de afinidades, semejanzas e isometrías se confunden en uno solo en el caso hiperbólico. En particular no hay más semejanzas que las isometrías. Pasemos a detallar todas estas ideas.

12.1 El plano hiperbólico

Definición 12.1 Un *plano hiperbólico* es un plano proyectivo real en el que se ha seleccionado una cónica real, a la que llamaremos *cónica infinita*. A los

puntos interiores a la cónica infinita los llamaremos *puntos finitos*, a los de la cónica los llamaremos *infinitos* y a los exteriores *ultrainfinitos*. Llamaremos *grupo hiperbólico* al grupo de las homografías que fijan a la cónica. Las *rectas finitas* serán las que contengan puntos finitos.

Salvo que indiquemos lo contrario, cuando hablemos de puntos y rectas en un plano hiperbólico sobrentenderemos que se trata de puntos y rectas finitas. Es claro que toda recta finita contiene exactamente dos puntos infinitos, los cuales determinan dos segmentos, uno formado por los puntos finitos y otro por los ultrainfinitos. Observar que las aplicaciones del grupo hiperbólico hacen corresponder puntos finitos con puntos finitos, pues éstos se caracterizan por que por ellos no pasan tangentes a la cónica, y esta propiedad se conserva por homografías.

En lugar de definir las rectas paralelas como las que no tienen puntos (finitos) en común, conviene distinguir dos clases de paralelismo:

Diremos que dos rectas son *paralelas* o *ultraparalelas* según si se cortan en un punto infinito o ultrainfinito. De este modo, dos rectas sin puntos (finitos) comunes son paralelas o ultraparalelas. Es claro que por un punto exterior a una recta pasan exactamente dos rectas paralelas, e infinitas ultraparalelas.

Diremos que dos rectas son *perpendiculares* si son conjugadas respecto a la cónica infinita. La prueba del teorema siguiente es elemental:

Teorema 12.2 *Por cada punto pasa una única perpendicular a una recta dada. Dos rectas son ultraparalelas si y sólo si tienen una perpendicular común, y en tal caso es única.*

Diremos que las rectas perpendiculares a una dada forman un *haz* de rectas ultraparalelas, y es claro que tales haces están en correspondencia con los puntos ultrainfinitos.

Tenemos, pues, que un plano hiperbólico es un subconjunto de un plano proyectivo. Si dotamos a éste de estructura euclídea tendremos una representación del plano hiperbólico como subconjunto de un plano euclídeo, con lo que muchos de los teoremas euclídeos se traducirán a teoremas hiperbólicos. Veamos que esto se puede hacer de modo que la relación entre ambos planos sea especialmente simple.

Sea O cualquier punto del plano hiperbólico, sea r su polar, que será una recta ultrainfinita, consideremos el espacio euclídeo que resulta de tomar a r como recta infinita con la involución (elíptica) inducida por la cónica infinita. Entonces dicha cónica se convierte en una circunferencia C de centro O , las rectas hiperbólicas coinciden con los segmentos euclídeos de extremos en C y las relaciones de orden hiperbólicas coinciden con las euclídeas.

Definición 12.3 Se llama *modelo de Klein* del plano hiperbólico al espacio formado por los puntos interiores de una circunferencia euclídea C y tomando como rectas las cuerdas de C .

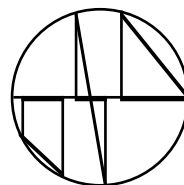
Todos los conceptos referentes al plano hiperbólico se definen en términos de la estructura proyectiva del plano que lo contiene, pero es útil relacionarlos con la geometría euclídea a través del plano de Klein. Por ejemplo, observamos que las perpendiculares hiperbólicas a una recta que pasa por el centro O del plano de Klein coinciden con las euclídeas.

Ejercicio: Describir en términos geométricos sencillos la construcción de la perpendicular a una recta por un punto dado en el plano de Klein.

Ejemplo: Como muestra de la utilidad del plano de Klein probaremos el hecho siguiente:

La proyección ortogonal de una recta en otra que no sea perpendicular a ella tiene por imagen una semirrecta si ambas rectas son paralelas y un segmento en caso contrario.

Podemos tomar uno de los puntos de la recta imagen como centro del plano de Klein, con lo que ésta es un diámetro del mismo. La proyección ortogonal desde la otra recta coincide entonces con la euclídea, pero dicha recta es un segmento, luego su imagen es otro segmento salvo en el caso de que ambas tengan un extremo común, es decir, salvo que se trate de rectas paralelas. ■



Observemos que todas las homografías del grupo hiperbólico conservan la perpendicularidad, luego todas ellas son el equivalente a las semejanzas de la geometría parabólica. Si queremos definir las isometrías como en el caso parabólico deberemos empezar por definir la reflexión respecto a una recta. Es claro que sólo hay una definición razonable: la reflexión respecto de una recta r ha de tener orden 2, luego ha de ser una involución de la cónica infinita, y además ha de fijar los puntos de r , luego ha de ser la involución que fija a los dos puntos de intersección de r con \mathcal{C} . En definitiva, hemos de definir una *reflexión* como una involución hiperbólica de la cónica infinita. Ahora bien, examinando la prueba del apartado 5) del teorema 8.54 observamos que toda involución elíptica se descompone en producto de dos involuciones hiperbólicas, luego de hecho toda homografía del grupo hiperbólico es producto de reflexiones, luego hemos de considerarlas a todas como isometrías.

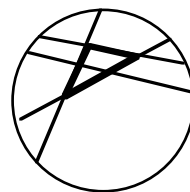
Definición 12.4 A las homografías del grupo hiperbólico las llamaremos también *isometrías*. Diremos que dos figuras son *congruentes* si existe una isometría que transforma una en otra.

Notemos que las isometrías que dejan fijo al centro O del plano de Klein fijan también a su recta polar, que es la recta infinita, luego son biyecciones afines. Puesto que fijan a la circunferencia, de hecho son isometrías euclídeas. El recíproco es obvio, luego tenemos que las isometrías hiperbólicas del plano de Klein que fijan a su centro son exactamente las euclídeas, es decir, los giros de centro O y las reflexiones respecto de rectas que pasan por O . En particular

la reflexión respecto de una recta r que pasa por O fija a todos los puntos de r y, como O puede ser cualquier punto, esto vale para todas las reflexiones hiperbólicas.

Con esto tenemos definidos todos los conceptos que aparecen en los axiomas de la geometría absoluta excepto las relaciones de orden en una recta, pero éstas son fáciles de definir: si a una recta que quitamos un punto infinito o ultrainfinito obtenemos una recta afín con dos ordenaciones naturales, las restricciones de estas ordenaciones al segmento finito no dependen del punto que hemos quitado, por lo que tenemos dos ordenaciones naturales en porción finita de la recta. En el plano de Klein las relaciones de orden hiperbólicas coinciden con las euclídeas.

Ahora es fácil comprobar que el plano hiperbólico cumple los axiomas de la geometría absoluta. Los grupos A y B y el axioma D se comprueban sin dificultad alguna en el plano de Klein. Por ejemplo, un triángulo en el plano de Klein es un triángulo euclídeo contenido en la circunferencia infinita, si una recta hiperbólica corta a uno de sus lados y no pasa por ninguno de los vértices, entonces ha de cortar a otro de los lados, porque esta propiedad se cumple en el plano euclídeo. Esto demuestra el axioma B5 para el plano hiperbólico. Igualmente se prueban los demás axiomas.



Los axiomas de congruencia requieren mayor atención porque la congruencia hiperbólica no coincide con la euclídea, como tampoco lo hacen la perpendicularidad ni mucho menos el paralelismo. Pensemos, por ejemplo, que si el plano de Klein satisface realmente los axiomas del grupo C entonces una recta debe contener infinitos segmentos disjuntos de la misma longitud, pero dicha recta es un segmento euclídeo, y no contiene infinitos segmentos euclídeos de la misma longitud. Lo que sucede es que los segmentos hiperbólicos son más cortos desde el punto de vista euclídeo cuanto más cerca están del borde del círculo finito.

Pese a ello, vamos a reducir también al caso euclídeo la comprobación de los axiomas del grupo C . Sólo necesitamos este hecho adicional: dados dos puntos P y Q existe una isometría hiperbólica que transforma uno en otro. De hecho vamos a ver que existe una que los intercambia. Sea r la recta que los une. Sean A y B sus puntos infinitos. El teorema 9.63 nos da que existe un único par M , N de puntos de r harmónicamente conjugados respecto de P y Q y respecto de A y B . Uno de los dos, digamos M será finito y el otro ultrainfinito. Entonces la polar de N es una recta finita s que pasa por M . La reflexión respecto de s fija a M e intercambia a P y Q . Notar que M es el punto medio hiperbólico del segmento \overline{PQ} .

Ahora vamos a probar los axiomas de congruencia. El axioma C1 es obvio. Para C2 hemos de partir de tres puntos A , B , A' y una semirrecta s de origen A' . Hemos de probar que existe un único punto B' en s tal que $\overline{AB} \equiv \overline{A'B'}$. Aplicando una isometría podemos suponer que A' es el centro del plano de Klein, con lo que la semirrecta es un radio del círculo. Mediante otra isometría transformamos el segmento \overline{AB} en un segmento $\overline{A'B''}$. Ahora podemos aplicar

un giro de centro A' , que es una isometría hiperbólica, para transformar $\overline{A'B''}$ en un segmento $\overline{A'B'}$ con B' en s .

Para probar la unicidad hemos de ver que si $\overline{A'B'} \equiv \overline{A'B''}$ con B'' en s entonces $B' = B''$. Ahora bien, una isometría entre los dos segmentos hace corresponder sus extremos. Componiéndola si es preciso con una isometría que intercambie A' y B'' podemos suponer que fija a A' . Pero si una isometría fija a A' y transforma B' en B'' (ambos en la misma semirrecta) necesariamente $B' = B''$, pues la isometría es también una isometría euclídea.

La prueba de C3 es simple: dados A, B, C colineales y A', B', C' colineales también tales que $\overline{AB} \equiv \overline{A'B'}$ y $\overline{BC} \equiv \overline{B'C'}$, entonces existe una isometría que transforma A en A' y B en B' . La imagen de C ha de ser un punto C'' en la misma semirrecta que C' con origen B' y de modo que $\overline{B'C''} \equiv \overline{BC} \equiv \overline{B'C'}$, luego por C2 ha de ser $C'' = C'$.

El axioma C4 se comprueba exactamente igual que C2: dado un ángulo L , una semirrecta s y un semiplano π cuya frontera sea la prolongación de s , suponemos que el origen de s es el centro del círculo, transportamos el ángulo para que su vértice este en dicho centro y aplicamos un giro para obtener un ángulo de lado s y, si es necesario, una simetría para que quede contenido en π . La unicidad se prueba de forma similar.

Antes de probar C5 hemos de probar que la definición de congruencia de triángulos que dimos en el capítulo I coincide con la que aquí hemos dado. Una implicación es evidente. Supongamos que $\widehat{ABC} \equiv \widehat{A'B'C'}$ en el sentido de que sus lados y ángulos son congruentes. Entonces existe una isometría que transforma A y B en A' y B' respectivamente. Componiéndola con una reflexión respecto a AB podemos suponer que ésta transforma C en un punto C'' en el mismo semiplano que C' respecto de $A'B'$. Por C4 se ha de cumplir $\widehat{B'A'C''} = \widehat{B'A'C'}$ y $\widehat{A'B'C''} = \widehat{A'B'C'}$. Por consiguiente $A'C'' = A'C'$ y $B'C'' = B'C'$, luego $C' = C''$ es el punto de corte de ambas rectas. Así pues, $\widehat{A'B'C'}$ es la imagen de \widehat{ABC} , luego son congruentes en el sentido general.

Ahora la prueba de C5 es completamente análoga a la de C2 y C4 (pero no hace falta probar la unicidad). Así pues:

Teorema 12.5 *El plano hiperbólico cumple los axiomas de la geometría absoluta, pero no cumple el axioma E sobre la unicidad de las paralelas.*

Las diferencias entre la geometría euclídea y la hiperbólica se hacen notables en cuanto hablamos de perpendicularidad y paralelismo. Ya hemos visto que dos rectas paralelas no tienen perpendiculares comunes. Otro hecho insólito es que dos pares cualesquiera de rectas paralelas son congruentes (pues cada uno está determinado por tres puntos infinitos, y existe una isometría que transporta tres puntos infinitos cualesquiera en otros tres).

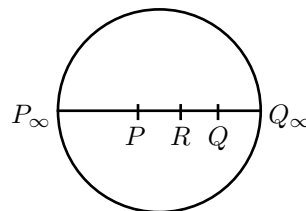
Ejercicio: Probar que dos pares de rectas ultraparalelas correspondientes a un mismo haz son congruentes.

12.2 Medida de segmentos y ángulos

Sabemos que el plano hiperbólico cumple todos los resultados que probamos en el capítulo II. En particular podemos asociar a cada segmento una longitud real, definida salvo un factor constante, de modo que dos segmentos son congruentes si y sólo si tienen la misma longitud. Como ya hemos comentado, aunque representemos el plano hiperbólico como un círculo euclídeo de centro O , la longitud hiperbólica no puede coincidir con la euclídea ni siquiera en el caso de segmentos con un extremo en O .

Vamos a buscar una expresión para la longitud hiperbólica de un segmento. Por el teorema 2.17, cualquier invariante que cumpla $m(u + v) = m(u) + m(v)$ para todo par de segmentos u y v nos da la longitud de los segmentos respecto a cierta unidad de medida. A la hora de buscar un invariante es lógico pensar en la razón doble.

Dados dos puntos P y Q , llamaremos P_∞ y Q_∞ a los puntos infinitos de la recta PQ de modo que el orden de los cuatro sea P_∞, P, Q, Q_∞ . La cantidad $\mathcal{R}(P, Q, Q_\infty, P_\infty)$ es ciertamente invariante por isometrías. Además, si un punto R está entre P y Q es inmediato comprobar la relación



$$\mathcal{R}(P, Q, Q_\infty, P_\infty) = \mathcal{R}(P, R; Q_\infty, P_\infty) \mathcal{R}(R, Q, P_\infty, Q_\infty).$$

Nos interesa una relación de este tipo pero aditiva, no multiplicativa. Para ello basta tomar logaritmos. Notemos que como $PQ \nparallel P_\infty Q_\infty$ la razón doble es positiva. De hecho es fácil ver que es mayor que 1, pues

$$|\mathcal{R}(P, Q, Q_\infty, P_\infty)| = \frac{\overline{PQ_\infty}}{\overline{PP_\infty}} : \frac{\overline{QQ_\infty}}{\overline{QP_\infty}}$$

y ciertamente $\overline{PQ_\infty} \overline{QP_\infty} > \overline{QQ_\infty} \overline{PP_\infty}$. Éste es el motivo por el que hemos considerado esta razón doble y no $\mathcal{R}(P, Q, P_\infty, Q_\infty)$, que es la inversa de la anterior, y por lo tanto menor que 1. Según esto podemos definir

$$d(P, Q) = \frac{1}{2} \log \mathcal{R}(P, Q, Q_\infty, P_\infty)$$

y tenemos que $d(P, Q)$ es un número real positivo invariante por isometrías y cuando P, Q, R son tres puntos colineales con R entre P y Q se cumple $d(P, Q) = d(P, R) + d(R, Q)$. Convenimos además que $d(P, P) = 0$.

Así, $d(P, Q)$ es la *distancia hiperbólica* entre los puntos P y Q . En lugar de la constante $1/2$ podríamos haber elegido cualquier otro número real positivo. Modificar esta constante equivale a cambiar de unidad de longitud. Más adelante justificaremos esta elección particular.

Es clara la analogía entre la fórmula que hemos encontrado para la medida de segmentos hiperbólicos y la fórmula de Cayley que dimos en el capítulo X para la medida de ángulos euclídeos. En el caso de los ángulos la proximidad

es más que formal, pues podemos reducir la medida de ángulos hiperbólicos a la de ángulos euclídeos. En efecto, sabemos que las isometrías que fijan al centro O del plano de Klein son las mismas para la geometría euclídea y para la hiperbólica, de donde se sigue que dos ángulos de vértice O son congruentes para una si y sólo si lo son para la otra. A su vez de aquí se sigue que la suma de ángulos de vértice O es la misma en ambos casos y la medida de ángulos, tal y como la construimos en el capítulo II no depende de nada más, luego la medida euclídea coincide con la hiperbólica en el caso de ángulos de vértice O . Según la fórmula de Cayley, si r_1 y r_2 son dos rectas que pasan por O el ángulo que forman mide

$$\widehat{r_1 r_2} = \frac{1}{2} \arg \mathcal{R}(r_1, r_2, I_1, I_2),$$

donde I_1 y I_2 son las rectas que pasan por O y por los puntos circulares.

Si obtenemos una expresión invariante por isometrías hiperbólicas y que coincida con ésta sobre las rectas que se cortan en O , tendremos una medida de ángulos hiperbólicos. Ahora bien, esto es fácil de conseguir:

Teorema 12.6 *El ángulo que forman dos rectas hiperbólicas secantes r_1 y r_2 viene dado por la expresión¹*

$$\widehat{r_1 r_2} = \frac{1}{2} \arg \mathcal{R}(r_1, r_2, I_1, I_2),$$

donde I_1 e I_2 son las tangentes a la cónica infinita por el punto $r_1 \cap r_2$.

Efectivamente, esta expresión es invariante por isometrías y si el punto de corte es el centro del plano de Klein entonces las tangentes pasan por los puntos circulares, luego el ángulo coincide con el euclídeo.

La fórmula anterior nos da el (menor) ángulo orientado determinado por dos rectas secantes, donde ‘orientado’ quiere decir que cambia de signo si cambiamos el orden de las rectas. Para eliminar esta dependencia del orden podemos tomar el argumento en $[-\pi, \pi]$ y tomar el valor absoluto. Esto nos da la medida (en el sentido del capítulo II) del ángulo agudo (o recto) que determinan las dos rectas. Para calcular el ángulo entre dos semirectas aplicamos la fórmula anterior si calculamos el complementario (restando el resultado de π) si el ángulo que queremos medir es obtuso.

Vamos a deducir unas fórmulas muy útiles para calcular longitudes y ángulos hiperbólicos a partir de las coordenadas homogéneas de los puntos y rectas. Sea A la matriz de la cónica infinita en un sistema de referencia dado. Sea $f(X, Y) = XAY^t$, de modo que la ecuación de la cónica es $f(X, X) = 0$. Sean X e Y las coordenadas homogéneas de dos puntos interiores a la cónica. La recta que los une está formada por los puntos de coordenadas $X + \lambda Y$, donde $\lambda \in \mathbb{R} \cup \{\infty\}$. Esta recta corta a la cónica en los puntos correspondientes a los valores de λ que satisfacen la ecuación

$$0 = f(X + \lambda Y, X + \lambda Y) = \lambda^2 f(Y, Y) + 2f(X, Y)\lambda + f(X, X).$$

¹Como en el caso de la fórmula de Cayley, debemos señalar que la expresión habitual de esta fórmula es $\widehat{r_1 r_2} = -\frac{i}{2} \log \mathcal{R}(r_1, r_2, I_1, I_2)$, donde \log representa al logaritmo complejo.

Concretamente:

$$\lambda = \frac{-f(X, Y) \pm \sqrt{f(X, Y)^2 - f(X, X)f(Y, Y)}}{f(Y, Y)}.$$

Podemos considerar a λ como la coordenada cartesiana de un punto de la recta XY respecto a un sistema de referencia de origen X y en el que Y es el punto infinito. Si llamamos λ_1 y λ_2 a los dos valores dados por la fórmula anterior y P_1, P_2 a los puntos correspondientes, entonces

$$\mathcal{R}(X, Y, P_1, P_2) = \mathcal{R}(0, \infty, \lambda_1, \lambda_2) = \frac{\lambda_1}{\lambda_2}.$$

Para calcular $d(X, Y)$ hemos de ordenar P_1 y P_2 de modo que la razón doble sea mayor que 1. Si los ponemos en orden equivocado lo único que sucede es que el logaritmo cambia de signo, por lo que en lugar de preocuparnos por determinar el orden correcto en este caso lo que haremos será añadir un valor absoluto a la fórmula. Así, la distancia hiperbólica entre los puntos de coordenadas X e Y es

$$\begin{aligned} d &= \frac{1}{2} \left| \log \frac{f(X, Y) + \sqrt{f(X, Y)^2 - f(X, X)f(Y, Y)}}{f(X, Y) - \sqrt{f(X, Y)^2 - f(X, X)f(Y, Y)}} \right| \\ &= \frac{1}{2} \left| \log \frac{\left(f(X, Y) + \sqrt{f(X, Y)^2 - f(X, X)f(Y, Y)} \right)^2}{f(X, X)f(Y, Y)} \right| \end{aligned}$$

El coseno hiperbólico biyecta los números positivos con los números mayores que 1, luego es equivalente conocer d que $\cosh d$. Sin embargo vamos a ver que la expresión de $\cosh d$ es algo más simple. Calculamos

$$e^{\pm d} = \left| \frac{f(X, Y) + \sqrt{f(X, Y)^2 - f(X, X)f(Y, Y)}}{\sqrt{f(X, X)f(Y, Y)}} \right|,$$

de donde se obtiene fácilmente

$$\cosh d = \frac{e^d + e^{-d}}{2} = \frac{|f(X, Y)|}{\sqrt{f(X, X)f(Y, Y)}}.$$

Ahora particularizamos al caso en que la cónica es la circunferencia de radio 1 y los puntos son $X = (x, y, 1)$, $Y = (x', y', 1)$. Entonces $f(X, Y) = xx' + yy' - 1$, luego en total obtenemos:

Teorema 12.7 *La distancia hiperbólica d en el plano de Klein entre dos puntos cuyas coordenadas cartesianas son (x, y) y (x', y') viene dada por*

$$\cosh d = \frac{1 - xx' - yy'}{\sqrt{(1 - x^2 - y^2)(1 - x'^2 - y'^2)}}.$$

Los mismos cálculos nos dan una expresión similar para el ángulo entre dos rectas. Hemos de partir de la ecuación $F(X, Y) = XA^{-1}Y$ y las coordenadas homogéneas X e Y de dos rectas finitas que se corten en un punto P . Entonces $X + \lambda Y$ recorre todas las rectas que pasan por P y los valores de λ que obtenemos corresponden a las dos tangentes imaginarias por P a la cónica infinita. Si llamamos θ al ángulo entre las dos rectas, llegamos a que

$$\theta = \pm \frac{1}{2} \arg \frac{\left(F(X, Y) + \sqrt{F(X, Y)^2 - F(X, X)F(Y, Y)} \right)^2}{F(X, X)F(Y, Y)}.$$

Más aún, si recordamos la deducción de la fórmula de Cayley veremos que en realidad la razón doble que hemos calculado tiene módulo 1 y así

$$\cos 2\theta \pm i \sin 2\theta = \frac{\left(F(X, Y) + \sqrt{F(X, Y)^2 - F(X, X)F(Y, Y)} \right)^2}{F(X, X)F(Y, Y)}.$$

Ahora usamos que $\cos 2\theta + i \sin 2\theta = (\cos \theta + i \sin \theta)^2$, con lo que simplificamos el cuadrado y tomamos la parte real. El resultado es

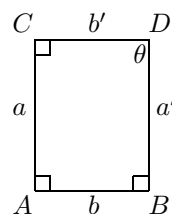
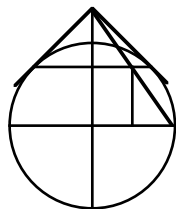
$$\cos \theta = \frac{F(X, Y)}{\sqrt{F(X, X)F(Y, Y)}}.$$

Si lo aplicamos al caso en que la cónica infinita es la circunferencia unidad, la ecuación tangencial es $F(X, Y) = xx' + yy' - zz'$, con lo que llegamos al teorema siguiente:

Teorema 12.8 *El ángulo hiperbólico θ en el plano de Klein entre dos rectas cuyas coordenadas homogéneas son (x, y, z) y (x', y', z') viene dado por*

$$\cos \theta = \frac{xx' + yy' - zz'}{\sqrt{(x^2 + y^2 - z^2)(x'^2 + y'^2 - z'^2)}}.$$

Ejemplo: Consideremos un cuadrilátero hiperbólico con tres ángulos rectos del que conocemos las longitudes a y b de dos de sus lados. Vamos a usar las fórmulas que hemos obtenido para calcular el cuarto ángulo. No perdemos generalidad si suponemos que sus vértices tienen coordenadas $A(0, 0)$, $B(x, 0)$, $C(0, y)$, $D(x, y)$ en el plano de Klein.



Lo primero que conviene observar es que el ángulo θ es agudo. Para ello no hay más que construir la perpendicular a CD por D y observar que forma con CD un ángulo mayor que θ . Por lo tanto, si calculamos el ángulo entre CD y BD estaremos calculando θ y no su complementario. Incidentalmente hemos probado que no existen rectángulos hiperbólicos.

El teorema 12.7 nos da

$$\cosh a = \frac{1}{\sqrt{1-x^2}}, \quad \cosh b = \frac{1}{\sqrt{1-y^2}},$$

de donde podemos despejar $x = \tanh a$, $y = \tanh b$. Por otra parte el teorema 12.8 aplicado a las rectas $CD = (0, 1, -y)$ y $BD = (1, 0, -x)$ nos da

$$\cos \theta = \frac{xy}{\sqrt{(1-y^2)(1-x^2)}},$$

y al sustituir los valores de x e y en función de a y b llegamos a

$$\cos \theta = \sinh a \sinh b.$$

■

Ejercicio: ¿Qué sucede si $\sinh a \sinh b \geq 1$?

Ejercicio: Probar que la hipotenusa de un triángulo rectángulo es mayor que cualquiera de los catetos. Deducir que la distancia más corta entre un punto y una recta se alcanza en el pie de la perpendicular a la recta por el punto.

Ejercicio: Con la notación del ejemplo anterior, probar que $\sinh a' > \sinh a$ y por lo tanto $a' > a$. Deducir que la distancia más corta entre dos puntos de dos rectas ultraparalelas es la que media entre los pies de su única perpendicular común. Probar también que si una recta tiene dos puntos a la misma distancia de otra recta, entonces no pueden ser paralelas.

La distancia entre dos rectas paralelas Veamos otra aplicación del teorema 12.7. Consideremos dos rectas paralelas. No perdemos generalidad si suponemos que son las rectas $y = 0$ e $y = 1 - x$ del plano de Klein. La distancia a la primera del punto (x, y) de la segunda viene dada por

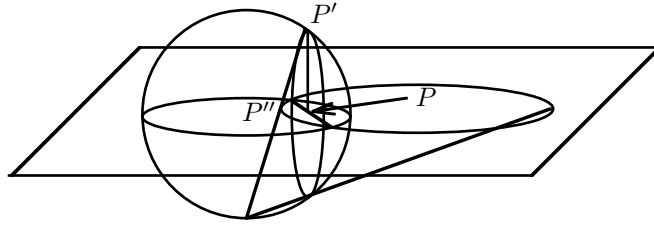
$$\cosh d = \frac{\sqrt{1-x^2}}{\sqrt{1-x^2-(1-x)^2}} = \sqrt{\frac{1+x}{2x}} = \sqrt{\frac{1}{2} \left(1 + \frac{1}{x}\right)}.$$

A la vista de esta fórmula es claro que si r y s son dos rectas paralelas no hay dos puntos en r a la misma distancia de s , así como que la recta r puede orientarse de modo que la distancia a s de un punto P de r se hace arbitrariamente pequeña si P se toma suficientemente a la derecha y arbitrariamente grande si P se toma suficientemente a la izquierda. En particular no hay una mínima distancia entre las dos rectas. ■

12.3 El modelo de Poincaré

Del mismo modo que el plano de Klein relaciona la geometría hiperbólica con la euclídea, hay otro modelo de plano hiperbólico, debido a Poincaré, que relaciona la geometría hiperbólica con la geometría circular.

Consideremos una esfera S y sea D el círculo delimitado por su ecuador C . Sea O su centro. Podemos considerar D como plano de Klein, en cuyo caso lo llamaremos D_K . Sea D_P el mismo círculo D , pero visto ahora como subconjunto de la recta proyectiva compleja. Sea $G : D_K \rightarrow D_P$ la biyección definida como sigue: A cada punto P de D_K le asignamos el punto P' donde la perpendicular a D por P corta a una de las semiesferas de S . A continuación proyectamos P' mediante la proyección estereográfica desde el polo de la semiesfera opuesta, con lo que obtenemos un punto $P'' = G(P)$.



Consideramos a D_P como plano hiperbólico tomando como rectas las imágenes por G de las rectas de D_K , entendiendo que dos rectas $G[r]$ y $G[s]$ son perpendiculares si y sólo si lo son r y s , y que las isometrías de D_P son las aplicaciones de la forma $G^{-1}fG$, donde f es una isometría en D_K . El interés de esta construcción radica en que la geometría de D_P puede describirse de forma natural en términos de la geometría circular de la recta compleja, sin ninguna referencia a la geometría de D_K .

En primer lugar notamos que al levantar las rectas de D_K sobre la semiesfera éstas se transforman en los cortes de S con los planos verticales, que son las circunferencias perpendiculares a C . Más exactamente, las rectas hiperbólicas se transforman en las porciones de estas circunferencias contenidas en la semiesfera elegida. Al aplicar la proyección estereográfica obtenemos que las rectas de D_P son simplemente las intersecciones con D_P de las circunferencias de la recta compleja perpendiculares a C . Entre ellas se encuentran las rectas que pasan por O .

A cada isometría f de D_K le podemos asignar la única transformación circular directa f' que coincide con f en C . Notemos que las transformaciones circulares directas que fijan a C fijan también a los puntos de D_P . Es claro que la correspondencia $f \mapsto f'$ biyecta las isometrías con las transformaciones circulares directas que fijan a C . Vamos a probar que $f' = G^{-1}fG$. Si A y B son dos puntos de C , entonces f transforma la recta AB de D_K en $f(A)f(B)$. Por otra parte $G[AB]$ es el arco de la circunferencia ortogonal a C que pasa por A y por B , luego $f'[G[AB]]$ es el arco de la circunferencia ortogonal a C que pasa por $f(A)$ y $f(B)$, luego $G^{-1}[f'[G[AB]]] = f(A)f(B)$. Así pues, f y $Gf'G^{-1}$ coinciden sobre rectas. Expresando cada punto de D_K como intersección de dos rectas vemos que también coinciden sobre puntos.

Por consiguiente, las isometrías de D_P son las transformaciones circulares directas que fijan a C .

Para acabar de caracterizar la geometría de D_P probaremos que los ángulos entre dos rectas (circunferencias) de D_P son los mismos que en la geometría circular. Esto se debe a que G deja invariantes a las rectas de D_K que pasan por O , luego el ángulo entre dos rectas de D_P que pasan por O coincide con el ángulo entre estas mismas rectas en D_K , que coincide con el euclídeo, que a su vez coincide con el ángulo entre ellas en la geometría circular, luego efectivamente el ángulo entre dos rectas que pasan por O para la geometría de D_P y para la geometría circular es el mismo. Como las isometrías de D_P son transformaciones circulares, conservan ángulos para las dos geometrías, luego la igualdad de ángulos es válida para rectas que se corten en un punto arbitrario.

Por último notamos que, dado que la geometría de D_K está descrita completamente en términos de la geometría circular, no necesitamos suponer que D_K es un círculo euclídeo, sino que todo lo dicho vale en general tomando como D_K uno de los lados (círculo o semiplano) que determina una circunferencia en la recta proyectiva compleja. En resumen:

Definición 12.9 Llamaremos *modelo de Poincaré* del plano hiperbólico al conjunto de los puntos situados a un mismo lado D (círculo o semiplano) de una circunferencia C en la recta proyectiva compleja, tomando como rectas las intersecciones con D de las circunferencias perpendiculares a C . Las isometrías hiperbólicas son entonces las transformaciones circulares directas que fijan a C . Los ángulos hiperbólicos coinciden con los de la geometría circular.

Hemos probado que la geometría del plano de Poincaré es exactamente la misma que la del plano de Klein. Aunque con las nociones de recta, isometría y perpendicular tenemos totalmente caracterizada la geometría hiperbólica, vamos ver que también hay una expresión sencilla para la longitud hiperbólica en el plano de Poincaré en términos de la geometría circular.

Veamos primero un caso particular. Tomemos como C el círculo de centro O y radio 1 en el plano $z = 0$ y consideremos el punto $(x, 0, 0)$ del plano de Klein. Su imagen en el plano de Poincaré se obtiene elevándolo, por ejemplo, a la semiesfera superior, lo que nos da el punto $(x, 0, \sqrt{1-x^2})$, y después calculando la intersección con $z = 0$ de la recta que une este punto con el polo sur $(0, 0, -1)$. Un cálculo sencillo nos da que dicha imagen es el punto $(x', 0)$, donde

$$x' = \frac{x}{1 + \sqrt{1-x^2}}.$$

En términos de números complejos dicha imagen es simplemente el número $x' + 0i = x'$.

Vamos a calcular la distancia hiperbólica en el plano de Poincaré entre los puntos $P(0, 0)$ y $Q(x, 0)$. Los puntos infinitos involucrados son $P_\infty(-1, 0)$ y $Q_\infty(1, 0)$. Entonces

$$\mathcal{R}(P, Q, Q_\infty, P_\infty) = \frac{1+x}{1-x}, \quad \text{luego} \quad d_K(P, Q) = \frac{1}{2} \log \frac{1+x}{1-x}.$$

Ésta ha de ser también la distancia hiperbólica entre $P' = 0$ y $Q' = x'$ en el plano de Poincaré. Vamos a expresarla en términos de 0 , x' y los puntos infinitos de la recta que determinan, que son $P'_\infty = -1$ y $Q'_\infty = 1$. Para ello observamos que

$$\begin{aligned}\mathcal{R}(P', Q', Q'_\infty, P'_\infty)^2 &= \left(\frac{1+x'}{1-x'} \right)^2 = \left(\frac{1+x+\sqrt{1-x^2}}{1-x+\sqrt{1-x^2}} \right)^2 \\ &= \frac{1+x+(1+x)\sqrt{1-x^2}}{1-x+(1-x)\sqrt{1-x^2}} = \frac{(1+x)(1+\sqrt{1-x^2})}{(1-x)(1+\sqrt{1-x^2})} \\ &= \frac{1+x}{1-x} = R(P, Q, Q_\infty, P_\infty).\end{aligned}$$

Así pues, tomando logaritmos queda que la distancia hiperbólica en el plano de Poincaré entre P' y Q' viene dada por

$$d_P(P', Q') = \log \mathcal{R}(P', Q', Q'_\infty, P'_\infty).$$

En principio tenemos probada esta fórmula cuando $P' = 0$ y Q' está sobre el eje real, pero ambos miembros de la igualdad son invariantes por las isometrías del plano de Poincaré (las transformaciones circulares conservan la razón doble), y todo par de puntos puede transformarse mediante una isometría en un par en las condiciones anteriores. Esto prueba el teorema siguiente:

Teorema 12.10 *La distancia hiperbólica entre dos puntos P y Q en el plano de Poincaré viene dada por*

$$d_P(P, Q) = \log \mathcal{R}(P, Q, Q_\infty, P_\infty),$$

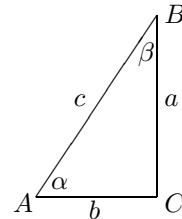
donde P_∞ y Q_∞ son los puntos infinitos de la recta (circular) PQ , de modo que el orden entre los cuatro puntos sea P_∞, P, Q, Q_∞ .

Ejercicio: Transportar la estructura hiperbólica del plano de Poincaré a una semiesfera mediante la proyección estereográfica. Describir las rectas, los ángulos y las isometrías.

12.4 Trigonometría hiperbólica

Los triángulos hiperbólicos verifican relaciones trigonométricas algo más complejas que las de los triángulos euclídeos, pero igualmente manejables.

Consideremos un triángulo rectángulo hiperbólico. Digamos que sus vértices son A, B, C , sus lados a, b, c y sus ángulos α, β, γ , donde $\gamma = \pi/2$. No perdemos generalidad si suponemos que las coordenadas cartesianas de los vértices son $A(0,0)$, $B(x,y)$, $C(x,0)$. Como A es el centro del plano de Klein el ángulo hiperbólico α coincide con el ángulo euclídeo



correspondiente. Llamemos $z = \sqrt{x^2 + y^2}$, es decir, z es la longitud euclídea de la hipotenusa del triángulo. El teorema 12.7 nos da las relaciones

$$\cosh a = \frac{\sqrt{1-x^2}}{\sqrt{1-z^2}}, \quad \cosh b = \frac{1}{\sqrt{1-x^2}}, \quad \cosh c = \frac{1}{\sqrt{1-z^2}}.$$

Es fácil despejar de las dos últimas

$$x = \tanh b, \quad z = \tanh c.$$

De la primera obtenemos

$$(1-x^2) \cosh^2 a - y^2 \cosh^2 a = 1-x^2 \Rightarrow y^2 \cosh^2 a = (1-\tanh^2 b) \sinh^2 a,$$

con lo que

$$y = \frac{\tanh a}{\cosh b}.$$

De aquí se obtienen todas las fórmulas trigonométricas hiperbólicas sobre triángulos rectángulos. Con ellas podemos calcular los cinco elementos del triángulo a partir de dos cualesquiera de ellos. En primer lugar vemos cómo calcular los ángulos si conocemos dos de los lados.

Conocidos un cateto y la hipotenusa Si conocemos b y c podemos aprovechar que el ángulo hiperbólico α coincide con el euclídeo, con lo que

$$\cos \alpha = \frac{x}{z} = \frac{\tanh b}{\tanh c}.$$

Por simetría se ha de cumplir la fórmula análoga con el otro cateto, luego

$$\tanh b = \cos \alpha \tanh c, \quad (12.1)$$

$$\tanh a = \cos \beta \tanh c. \quad (12.2)$$

Conocidos los catetos Si conocemos a y b usamos la tangente de α :

$$\tan \alpha = \frac{y}{x} = \frac{\tanh a}{\sinh b}.$$

Por simetría la misma fórmula ha de valer para el cálculo de β , con lo que

$$\tanh a = \tan \alpha \sinh b, \quad (12.3)$$

$$\tanh b = \tan \beta \sinh a. \quad (12.4)$$

El teorema de Pitágoras hiperbólico Veamos ahora cómo calcular un lado si conocemos los otros dos. Con esto podremos resolver cualquier triángulo rectángulo conocidos dos lados. Sustituimos los valores de x , y , z en la relación pitagórica euclídea $z^2 = x^2 + y^2$, con lo que obtenemos

$$\tanh^2 c = \tanh^2 b + \frac{\tanh^2 a}{\cosh^2 b}.$$

Usando la relación obvia entre cosenos y tangentes hiperbólicas llegamos a

$$1 - \frac{1}{\cosh^2 c} = 1 - \frac{1}{\cosh^2 b} + \frac{1}{\cosh^2 b} \left(1 - \frac{1}{\cosh^2 a} \right),$$

que se reduce a

$$\cosh c = \cosh a \cosh b. \quad (12.5)$$

Conocidos un lado y un ángulo Las fórmulas que hemos dado para los cosenos y las tangentes nos permiten ya resolver este caso. No obstante la fórmula anterior nos da también expresiones para los senos:

$$\operatorname{sen} \alpha = \frac{y}{z} = \frac{\tanh a}{\cosh b \tanh c} = \frac{\operatorname{senh} a \cosh c}{\cosh a \cosh b \operatorname{senh} c} = \frac{\operatorname{senh} a}{\operatorname{senh} c}.$$

Así pues:

$$\operatorname{senh} a = \operatorname{sen} \alpha \operatorname{senh} c, \quad (12.6)$$

$$\operatorname{senh} b = \operatorname{sen} \beta \operatorname{senh} c. \quad (12.7)$$

Conocidos dos ángulos Éste es el punto más peculiar de la trigonometría hiperbólica, y es que nos permite calcular los lados de un triángulo conocidos sus ángulos. De momento lo probamos sólo para triángulos rectángulos, pero luego veremos que vale en general. Dividimos (12.1) entre (12.7) y aplicamos (12.5):

$$\frac{\cos \alpha}{\operatorname{sen} \beta} = \frac{\tanh b \operatorname{senh} c}{\tanh c \operatorname{senh} b} = \frac{\cosh c}{\cosh b} = \cosh a.$$

Similarmente se prueba la fórmula análoga. Multiplicándolas y aplicando (12.5) obtenemos una tercera que nos permite calcular la hipotenusa:

$$\cos \alpha = \operatorname{sen} \beta \cosh a, \quad (12.8)$$

$$\cos \beta = \operatorname{sen} \alpha \cosh b, \quad (12.9)$$

$$1 = \tan \alpha \tan \beta \cosh c. \quad (12.10)$$

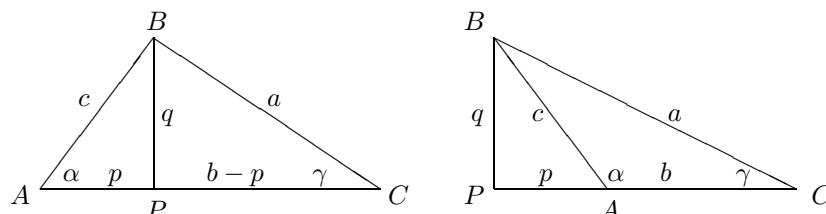
La suma de los ángulos Hay una fórmula de la trigonometría euclídea que ya no es válida en el caso hiperbólico, y es que la suma de los ángulos de un triángulo no es igual a π . De hecho siempre es menor que π . De momento lo probamos para triángulos rectángulos. Llamemos $\bar{\beta}$ al ángulo euclídeo correspondiente a β , de modo que

$$\cos \bar{\beta} = \frac{y}{z} = \frac{\tanh a}{\cosh b \tanh c} = \frac{\cos \beta}{\cosh b} < \cos \beta.$$

Puesto que los ángulos son agudos, se cumple $\beta < \bar{\beta}$, luego

$$\alpha + \beta + \frac{\pi}{2} < \alpha + \bar{\beta} + \frac{\pi}{2} = \pi.$$

Triángulos arbitrarios La resolución de triángulos arbitrarios se consigue descomponiéndolos en dos triángulos rectángulos. Para ello trazamos una de las alturas, y hay que distinguir dos casos según si el pie P queda entre los dos vértices opuestos o no.



De todos modos hay ocasiones en las que las pruebas se pueden reducir al primer caso (figura de la izquierda), pues por el teorema 2.26 sabemos que los ángulos de un triángulo no suman más de π , luego al menos dos de ellos son agudos. Si suponemos que A y C lo son, entonces el pie P de la altura por C está necesariamente entre A y C . Por ejemplo, teniendo en cuenta que los ángulos no rectos de los triángulos rectángulos suman menos de $\pi/2$, basta considerar este primer caso para concluir que los ángulos de cualquier triángulo \widehat{ABC} suman menos de π .

Teorema 12.11 *La suma de los ángulos de un triángulo hiperbólico es siempre menor que π .*

Vamos a ver que tres de los seis elementos de un triángulo determinan los otros tres. Para ello probaremos teoremas análogos hiperbólicos al teorema de los senos y el teorema del coseno.

Teorema 12.12 (Teorema de los senos) *Todo triángulo hiperbólico satisface la relación*

$$\frac{\sin \alpha}{\sinh a} = \frac{\sin \beta}{\sinh b} = \frac{\sin \gamma}{\sinh c}.$$

DEMOSTRACIÓN: Supongamos la situación de la figura de la izquierda. Entonces

$$\sinh q = \sin \gamma \sinh a, \quad \sinh q = \sin \alpha \sinh c,$$

y al igualar obtenemos la relación buscada. Si la situación es la de la figura de la derecha la prueba es similar. ■

En geometría hiperbólica contamos con dos teoremas del coseno, uno para lados y otro para ángulos:

Teorema 12.13 (Teorema del coseno) *Todo triángulo hiperbólico satisface las relaciones*

$$\begin{aligned} \cosh a &= \cosh b \cosh c - \sinh b \sinh c \cos \alpha, \\ \cos \alpha &= -\cos \beta \cos \gamma + \sin \beta \sin \gamma \cosh a. \end{aligned}$$

DEMOSTRACIÓN: Como en el teorema anterior, suponemos que el pie de la altura por B queda entre A y C . El caso contrario es análogo.

$$\begin{aligned}\cosh a &= \cosh q \cosh(b-p) = \cosh q (\cosh b \cosh p - \sinh b \sinh p) \\ &= \cosh b \cosh c - \sinh b \cosh c \tanh p \\ &= \cosh b \cosh c - \sinh b \cosh c \cos \alpha \tanh c \\ &= \cosh b \cosh c - \sinh b \sinh c \cos \alpha.\end{aligned}$$

Sea $\delta = \widehat{PBC}$.

$$\begin{aligned}\cos \alpha &= \cosh q \sin(\beta - \delta) = \cosh q \sin \beta \cos \delta - \cosh q \cos \beta \sin \delta \\ &= \cosh q \cosh(v-p) \sin \gamma \sin \beta - \cos \beta \cos \gamma \\ &= \cosh a \sin \gamma \sin \beta - \cos \beta \cos \gamma.\end{aligned}$$

■

Es fácil ver que con estas fórmulas podemos resolver cualquier triángulo.² En particular, si conocemos los tres ángulos el teorema del coseno para ángulos nos permite calcular los tres lados. Por consiguiente:

Teorema 12.14 *Dos triángulos hiperbólicos son congruentes si y sólo si tienen los tres ángulos iguales.*

He aquí otra consecuencia interesante de los teoremas del coseno:

Teorema 12.15 *Si dos triángulos tienen sus lados proporcionales entonces son iguales.*

DEMOSTRACIÓN: La prueba se basa en que en ningún momento hemos especificado la base e de las razones hiperbólicas. De este modo, si tenemos dos triángulos de lados \widehat{ABC} y $\widehat{A'B'C'}$ de modo que $a' = ka$, $b' = kb$, $c' = kc$, con $k > 0$, observamos que

$$\cosh_e ka = \frac{e^{ka} + e^{-ka}}{2} = \frac{(e^k)^a + (e^k)^{-a}}{2} = \cosh_{ke} a,$$

y lo mismo vale para el seno hiperbólico, por lo que el teorema del coseno aplicado al segundo triángulo (con base e) es

$$\cosh_{ke} a = \cosh_{ke} b \cosh_{ke} c - \sinh_{ke} b \sinh_{ke} c \cos \alpha',$$

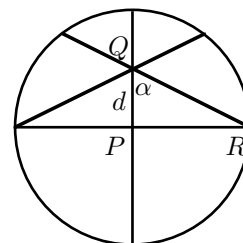
y si lo comparamos con el teorema del coseno para el primer triángulo tomando como base ke concluimos que $\cos \alpha = \cos \alpha'$, luego $\alpha = \alpha'$ y los triángulos tienen sus ángulos iguales, luego son iguales. ■

²Como en el caso euclídeo, si conocemos dos lados y un ángulo distinto del que forman puede haber dos soluciones.

Ejemplo: el ángulo de paralelismo Consideremos una recta cualquiera r y un punto exterior P . Sea s la perpendicular a r por P y sea $Q = r \cap s$. Podemos suponer que Q es el centro del plano de Klein y que r es el eje x . Sea t una de las paralelas a r por P . Sea $R = r \cap t$.

Entonces \widehat{PQR} es lo que se llama un triángulo rectángulo asintótico, con dos lados infinitos y un ángulo nulo. Las mismas técnicas que hemos empleado con los triángulos rectángulos usuales nos dan la relación entre sus dos únicos elementos finitos, el lado $d = \overline{PQ}$ y el ángulo $\alpha = \hat{Q}$, a saber:

$$\cos \alpha = \tanh d.$$



También es fácil ver que α es el mismo para las dos paralelas por P y recibe el nombre de *ángulo de paralelismo* a r por P . Su interpretación es sencilla: las rectas que forman con s un ángulo inferior a α cortan a r , las dos que forman ángulo α son paralelas y las que forman ángulo superior a α son ultraparalelas. Hemos probado que el ángulo de paralelismo a una recta por un punto es agudo y depende únicamente de la distancia d del punto a la recta.

Ejercicio: Probar la relación de Lobachevski para el ángulo de paralelismo:

$$\tan \frac{\alpha}{2} = e^{-d}.$$

12.5 Las isometrías hiperbólicas

Vamos a describir con más detalle los distintos tipos de isometrías hiperbólicas. Los dos primeros los conocemos ya, pero los incluimos por completitud.

Reflexiones Recordemos que hemos definido la reflexión respecto a una recta r como la isometría que induce en la cónica infinita la involución que fija a los puntos infinitos de r . Al representarla en el plano de Klein con centro en un punto de r vemos que la reflexión ha de ser una isometría euclídea, por lo que ha de ser la simetría respecto de r y por consiguiente fija a todos los puntos de r . En el modelo de Poincaré las reflexiones coinciden con las de la geometría circular. Desde un punto de vista puramente hiperbólico la reflexión respecto a la recta r es la isometría que a cada punto P lo lleva al punto situado sobre la perpendicular a r por P y a la misma distancia de ésta que P .

Giros Un *giro* de centro O es la composición de dos reflexiones respecto de dos rectas que se cortan en O . Tomando a O como centro del plano de Klein vemos que los giros así definidos coinciden con los euclídeos, luego los giros hiperbólicos con un mismo centro O forman un grupo isomorfo a \mathbb{R}/\mathbb{Z} . Las imágenes de un punto P por todos los giros de centro O forman una circunferencia de centro O , donde una *circunferencia* hiperbólica de centro O se define igual que una euclídea, es decir, como el conjunto de puntos situados a una misma distancia r de su centro. Existe cierta tradición de llamar *ciclos* a las circunferencias

hiperbólicas. En el plano de Klein de centro O las circunferencias hiperbólicas de centro O coinciden con las euclídeas. En general las circunferencias hiperbólicas de centro O pueden caracterizarse como las cónicas formadas por puntos finitos y que cortan a la cónica infinita en los puntos (imaginarios) por donde pasan las tangentes a ésta desde O . Sabemos que toda isometría que fije a un punto es una reflexión o un giro.

Giros infinitos Definimos un *giro* alrededor de un punto infinito O como la composición de dos reflexiones cuyos ejes pasen por O , es decir, los giros infinitos son la composición de dos reflexiones respecto a dos rectas paralelas. La mejor representación de un giro infinito nos la da el modelo de Poincaré, tomando como conjunto de puntos finitos un hiperplano. No perdemos generalidad si suponemos que el centro del giro O es el punto infinito en el sentido de la geometría circular. Entonces las rectas que pasan por O son las rectas verticales (suponiendo que representamos la frontera del semiplano como una recta horizontal), y la composición de dos reflexiones respecto a dos de estas rectas es simplemente una traslación (euclídea) en dirección horizontal. Vemos, pues, que los giros hiperbólicos alrededor de un punto infinito O forman un grupo isomorfo a \mathbb{R} .

Las imágenes de un punto por todos los giros de centro un punto infinito O constituyen una figura llamada *horociclo* de centro O . En el caso concreto del semiplano de Poincaré, los horociclos de centro ∞ son las rectas horizontales. De aquí deducimos varios hechos que son invariantes por isometrías, y por lo tanto válidos para horociclos con centro arbitrario:

- Por cada punto finito pasa un único horociclo de centro un punto infinito. Dos horociclos con el mismo centro son disjuntos.
- Si P es un punto de un horociclo H de centro O , las únicas rectas cuyo único punto en común con H es P son OP y su perpendicular por P .
- Dados dos puntos de un horociclo de centro O , existe un único giro de centro O que transforma uno en el otro.

Consideremos una transformación circular que convierta el semiplano de Poincaré en un círculo D con circunferencia C de centro P . Sea O la imagen de ∞ . Entonces el diámetro OP es la imagen de una recta vertical, luego los horociclos de centro ∞ se transforman en las circunferencias contenidas en D que pasan por O y perpendiculares a OP , luego todas ellas tienen compartan con C la recta tangente por O . En general, dos cónicas se dicen *osculantes* (lat. ‘que se besan’) si comparten un punto y su tangente por el mismo. Puesto que sólo hay una circunferencia en D osculante a C por P que pase por un punto dado de OP , concluimos que todas ellas son horociclos. Así pues:

Los horociclos en el plano de Poincaré circular son las circunferencias osculantes a la circunferencia infinita.

El paso del plano de Poincaré al plano de Klein puede verse como una composición de dos proyecciones perspectivas, luego los horociclos en el plano de Klein son cónicas osculantes a la cónica infinita, pero esto sólo no las caracteriza. Si H es un horociclo de centro O en el plano de Klein y P es uno de sus puntos, entonces la tangente a H por O es una recta que corta a H únicamente en P , luego sabemos que ha de ser perpendicular a OP . A su vez esto se traduce en que el polo de OP respecto a H coincide con su polo respecto a la cónica infinita C .

Definición 12.16 Se dice que dos cónicas son *hiperosculantes* en un punto P si ambas pasan por P con la misma tangente t y los puntos de t tienen la misma polar respecto de ambas cónicas.

En estos términos tenemos que los horociclos son cónicas hiperosculantes a la circunferencia infinita. El teorema siguiente prueba que esto los caracteriza.

Teorema 12.17 *Dada una cónica \mathcal{C} , un punto P en \mathcal{C} y un punto Q que no esté en \mathcal{C} ni sobre su tangente t por P , existe una única cónica hiperosculante a \mathcal{C} por P que pasa por Q .*

DEMOSTRACIÓN: Sea P' el otro punto donde la recta PQ corta a \mathcal{C} . Consideramos la única elación de centro P y eje t que transforma P' en Q . Obviamente transforma \mathcal{C} en una cónica \mathcal{C}' que pasa por P con tangente t . Además si R es cualquier punto de t distinto de P , la elación transforma su polar respecto de \mathcal{C} en su polar respecto de \mathcal{C}' , pero como dicha polar pasa por P , en realidad queda invariante, luego ciertamente \mathcal{C}' es hiperosculante.

Dada una cónica cualquiera \mathcal{C}' hiperosculante a \mathcal{C} por P y que pase por Q , consideramos una recta cualquiera r que pase por Q y que corte a t en un punto R . La polar de R respecto de \mathcal{C} lo ha de ser también respecto de \mathcal{C}' . Sea R' el punto donde corta a r . Entonces \mathcal{C}' ha de contener al conjugado harmónico de P respecto de (R, R') . De este modo podemos obtener tantos puntos de \mathcal{C}' como queramos. Por consiguiente \mathcal{C}' es única. ■

Así pues, podemos afirmar:

Los horociclos del plano proyectivo son las cónicas hiperosculantes a la cónica infinita y formadas por puntos finitos (salvo el punto de contacto).

En el caso concreto del plano de Klein podemos dar una caracterización mucho más simple. Sea H un horociclo de centro O , sea P el centro del plano de Klein y sea Q el punto de H situado sobre OP . Sea d la distancia (euclídea) entre O y Q . Teniendo en cuenta que H ha de ser la proyección de una circunferencia desde la esfera que tiene por ecuador a la circunferencia infinita, es fácil ver que el diámetro de esta circunferencia ha de ser $\sqrt{2d}$ (suponemos que la circunferencia infinita tiene radio 1), que a su vez ha de coincidir con el eje menor de H . Así pues,

Los horociclos del plano de Klein son las elipses osculantes a la circunferencia infinita cuyos ejes miden d y $\sqrt{2d}$, con $0 < d < 2$.

Ejercicio: Probar que los horociclos de centro O quedan fijos por las reflexiones respecto de rectas que pasan por O .

Ejercicio: Probar que si dos puntos P y Q están en un mismo horociclo de centro O , entonces la distancia de P a OQ es igual a la distancia de Q a OP . Deducir de aquí una caracterización de los horociclos en términos de la distancia hiperbólica.

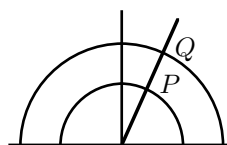
Hagamos una última observación sobre los horociclos: la transformación de una recta horizontal del semiplano de Poincaré en un horociclo en el plano hiperbólico visto como subconjunto del plano proyectivo es una composición de proyecciones perspectivas, luego es una homografía. Los giros infinitos de centro O son traslaciones horizontales en el semiplano de Poincaré, luego inducen sobre los horociclos las homografías de la forma $P_x \mapsto P_{x+y}$, luego lo mismo vale en el modelo proyectivo.

Traslaciones Definimos una *traslación* hiperbólica como la composición de dos reflexiones respecto a rectas hiperparalelas. El modelo que muestra más claramente las traslaciones es el semiplano de Poincaré. Si lo identificamos concretamente con el semiplano $y > 0$ en \mathbb{R}^2 , no perdemos generalidad si suponemos que la perpendicular común a los ejes de las dos reflexiones es el semieje $x = 0$. Entonces estos ejes son dos semicircunferencias de centro $(0, 0)$ y es fácil ver que la composición de las reflexiones respecto a ambas es una homotecia $(x, y) \mapsto (\alpha x, \alpha y)$ con $\alpha > 0$. Más aún, toda homotecia de centro $(0, 0)$ y razón positiva es una traslación hiperbólica.

También es claro que la única recta fijada es el semieje vertical. En general, cada traslación hiperbólica fija a una única recta, a la que llamaremos *eje* de la traslación. Las traslaciones con un eje dado forman un grupo isomorfo a \mathbb{R}^* .

Las imágenes de un punto P a través de todas las traslaciones de eje r forman una figura llamada *curva equidistante* a r por P . En el modelo de traslación en el semiplano de Poincaré, las equidistantes al semieje $x = 0$ son las semirrectas que pasan por el origen.

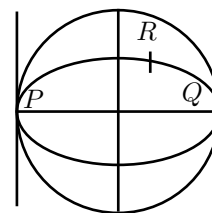
Su nombre se debe a que si P no está en r , la equidistante a r por P está formada por todos los puntos que están en el mismo semiplano que P respecto a r y cuya distancia a r es la misma que la de P . La razón es que, si r es la semirrecta $x = 0$ en el semiplano de Poincaré, la semicircunferencia de centro $(0, 0)$ que pasa por un punto Q de una equidistante es la perpendicular a r por Q , y lo mismo vale para P , y la traslación que lleva P a Q es una isometría que transforma una en la otra.



Dada una recta r y un número $d > 0$, existen dos equidistantes a una distancia d de r , una en cada uno de los semiplanos determinados por r . Si pasamos a un círculo las dos equidistantes a una recta se convierten en dos arcos de circunferencia que pasan por sus puntos infinitos.

Al pasar al modelo proyectivo nos encontramos con que las dos equidistantes se convierten en dos arcos de una misma cónica.

En efecto: No perdemos generalidad si suponemos que r es un diámetro (digamos el horizontal) del plano de Klein. Por concretar supondremos que la circunferencia infinita C es la circunferencia de centro $(0, 0)$ y radio 1 y que r es el eje $y = 0$. Sean P y Q los puntos infinitos de r . Las tangentes a C por P y Q son las rectas verticales $x = \pm 1$. Dado cualquier punto R es claro que existe una única cónica \mathcal{C}' osculante a C por P y Q y que pasa por R .



Es fácil obtenerla como imagen de C por una afinidad que fija a los ejes $x = 0$ e $y = 0$, por lo que \mathcal{C} es una elipse con estas rectas como ejes. Una reflexión respecto a una recta vertical intercambia los puntos P y Q así como las tangentes a C (luego también a \mathcal{C}) por estos puntos. Además fija a dos puntos de \mathcal{C} , luego de hecho fija a \mathcal{C} . Por consiguiente las traslaciones de eje r , que son composición de dos de estas reflexiones, también fijan a \mathcal{C} . Esto prueba que la equidistante a r por R está contenida en \mathcal{C} . Más concretamente en el arco de \mathcal{C} que está en el mismo semiplano que R respecto a r . Por otra parte, cada recta vertical ha de cortar a la equidistante a r por R y sólo corta a dicho arco de \mathcal{C} en un punto, luego tenemos la igualdad.

Falta probar que el arco opuesto de \mathcal{C} es la otra equidistante correspondiente a la misma distancia, pero esto es evidente, pues la reflexión respecto a r fija a \mathcal{C} (porque una elipse es simétrica respecto de sus ejes), luego cada punto del arco opuesto dista de r lo mismo que su simétrico. En resumen:

Las equidistantes a una recta r son las cónicas osculantes a la cónica infinita por los puntos infinitos de r . Más exactamente, son los dos arcos en que estos puntos dividen a la cónica.

Reflexión con traslación Una reflexión con traslación es simplemente la composición de una traslación con la reflexión respecto de su eje.

Teorema 12.18 *Toda isometría del plano proyectivo distinta de la identidad es de uno de los cinco tipos anteriores.*

DEMOSTRACIÓN: Recordemos del capítulo IX que toda homografía en una cónica determina un centro y un eje. De la propia definición se sigue que el eje pasa por los puntos fijos de la cónica si es que los hay. Por lo tanto, el eje es secante, tangente o exterior a la cónica según si la homografía es hiperbólica, parabólica o elíptica. En el último caso el eje corta a la cónica en dos puntos imaginarios, fijados por la extensión de la homografía a la complexificación. Es claro entonces que la homografía ha de fijar al eje en cualquier caso, y por consiguiente también al centro, que es su polo. Recordemos también el teorema 9.53, según el cual el centro de la composición de dos involuciones es la intersección de los ejes de éstas.

Sea, pues, una isometría hiperbólica f de centro O y eje r . Si O es un punto finito entonces podemos tomarlo como centro del plano de Klein y concluir que f ha de ser un giro o una reflexión. Como la restricción de f a la circunferencia

infinita es elíptica (el eje es exterior) de hecho ha de ser un giro. Así pues, los giros son las isometrías hiperbólicas con centro finito. El centro del giro coincide con el centro en el sentido general.

Si O es infinito, podemos descomponer f en producto de dos involuciones, y sus ejes pasarán por O y no pueden ser tangentes (porque no hay involuciones parabólicas) luego son rectas finitas y por consiguiente f es un giro infinito de centro O . Concluimos que los giros infinitos son las isometrías hiperbólicas de centro infinito. El centro del giro coincide con el centro en sentido general.

Si O es ultrainfinito entonces el eje r es una recta finita. Si f es una involución entonces es la reflexión respecto a r . En caso contrario expresamos f como composición de dos involuciones. La prueba del teorema 8.54 muestra que podemos exigir que una de ellas sea hiperbólica, es decir, una reflexión. Si ambas son reflexiones entonces f es una traslación, pues los ejes de las reflexiones se cortan en el punto ultrainfinito O . Notamos también que el eje de la traslación coincide con el eje en sentido general.

Si una de las involuciones sea elíptica el determinante de f ha de ser negativo. Consideramos la reflexión g respecto a r . Entonces fg fija a los puntos infinitos de r , luego sigue teniendo eje r , pero tiene determinante positivo. Así pues, ha de descomponerse en producto de dos reflexiones y es, por consiguiente, una traslación. Por consiguiente, $f = (fg)g$ es una simetría con traslación. Es fácil ver que $fg = gf$. ■

Finalmente probamos que las isometrías pueden caracterizarse como en el caso euclídeo como las biyecciones que conservan las distancias. Más aún, de hecho toda semejanza es una isometría.

Teorema 12.19 *Una biyección f de (el conjunto de puntos finitos del) plano hiperbólico en sí mismo que cumpla $d(f(P), f(Q)) = k d(P, Q)$, para todo par de puntos P y Q , con $k > 0$, es una isometría. En particular ha de ser $k = 1$.*

DEMOSTRACIÓN: Es fácil ver que tres puntos P , Q , R son colineales si y sólo si ordenados adecuadamente cumplen $d(P, Q) = d(P, R) + d(R, Q)$. Por lo tanto f transforma puntos colineales en puntos colineales, luego transforma rectas en rectas. Más aún, f transforma un triángulo en otro con sus lados proporcionales, luego ha de ser $k = 1$ y además f conserva ángulos. Sean P y Q dos puntos cualesquiera. Existe una isometría g que coincide con f sobre P y Q . Basta probar que fg^{-1} es una isometría, y esta aplicación tiene las mismas propiedades que f además de que fija a P y Q . Podemos suponer, pues, que f fija a estos puntos. De aquí se sigue que f fija a cada punto de la recta PQ .

Es fácil ver que f fija o intercambia los semiplanos respecto de PQ . Componiéndola si es preciso con la reflexión respecto de PQ podemos suponer que los fija. Dado cualquier punto R exterior a PQ sea $R' = f(R)$. Los triángulos \overline{PQR} y $\overline{PQR'}$ son congruentes y están en el mismo semiplano respecto a PQ , luego han de ser iguales, luego $f(R) = R$, luego f es la identidad. ■

Capítulo XIII

La geometría elíptica

Hasta ahora hemos estudiado dos ejemplos de geometrías métricas, es decir, geometrías donde está definida la distancia entre dos puntos y la perpendicularidad entre rectas. Una es la geometría de los planos parabólicos, que se obtienen eliminando una recta de un plano proyectivo; la segunda es la geometría hiperbólica, en la que nos quedamos sólo con los puntos interiores de una cónica. La geometría elíptica, que vamos a estudiar ahora, es la geometría métrica del plano proyectivo completo, considerando a todos los puntos finitos. Así como la geometría hiperbólica incumple el axioma *E* de Euclides porque por un punto exterior a una recta pasa más de una paralela, la geometría elíptica lo incumple en sentido contrario: como bien sabemos, en un plano proyectivo no existen paralelas.

Formalmente, la geometría elíptica se puede desarrollar de forma completamente análoga a la hiperbólica. El único cambio es que vamos a tomar como cónica infinita una cónica imaginaria, de modo que no dé lugar a ninguna división de los puntos del plano. Al igual que en los otros casos, podríamos definir espacios elípticos de cualquier dimensión, pero nosotros nos limitaremos al caso plano porque es completamente representativo y más simple en la práctica.

13.1 El plano elíptico

Definición 13.1 Un *plano elíptico* es un plano proyectivo real en el que hemos seleccionado una cónica imaginaria (cónica infinita). Diremos que dos rectas son *perpendiculares* si son conjugadas respecto a la cónica infinita. Llamaremos *grupo elíptico* o *grupo de isometrías* al grupo de las homografías que fijan a la cónica infinita. Dos figuras son *congruentes* si existe una isometría que transforma una en otra.

Vamos a introducir un modelo que nos permita relacionar la geometría elíptica con la euclídea. El plano proyectivo es de la forma $P(V)$, donde V es un espacio vectorial tridimensional real. Existe un sistema de referencia proyectivo respecto al cual la cónica infinita tiene ecuación $x^2 + y^2 + z^2 = 0$. Este

sistema de referencia se corresponde con una base de V . Dotamos a V del producto escalar que en dicha base tiene matriz identidad, es decir, el producto respecto al cual dicha base es ortonormal. Más en general, podemos considerar a V como un espacio afín euclídeo en el que hemos fijado un sistema de referencia ortonormal de modo que los puntos del plano proyectivo son las rectas que pasan por el origen O y las coordenadas homogéneas de un punto son las coordenadas en el sistema de referencia fijado de uno de los puntos (que no sea O) de la recta correspondiente.

Dos puntos de coordenadas X e Y son conjugados respecto de la cónica infinita si y sólo si $XY^t = 0$, es decir, si y sólo si vistos como rectas de V son ortogonales.

Las rectas proyectivas están formadas por los puntos que cumplen una ecuación $ax + by + cz = 0$, es decir, son los planos de V que pasan por O . También es claro que un punto y una recta son conjugados si y sólo si, vistos como recta y plano, uno es perpendicular al otro.

Todas estas relaciones se pueden simplificar si escogemos representantes concretos en V para los puntos proyectivos. Sea S la esfera en V de centro O y radio 1. Cada recta que pasa por O corta a S en dos puntos antípodos. En términos proyectivos, cada punto tiene dos ternas de coordenadas homogéneas sujetas a la restricción $x^2 + y^2 + z^2 = 1$. Podemos, pues, identificar el plano proyectivo con el conjunto de los pares de puntos antípodos de S . De ahora en adelante y salvo que indiquemos lo contrario, cuando hablemos de un punto de S nos referiremos a un par de puntos antípodos considerados como un solo objeto. Las rectas proyectivas se corresponden con las intersecciones con S de los planos que pasan por su centro, es decir, son los llamados *círculos máximos* de la esfera, las circunferencias cuyo radio coincide con el de la esfera. Dos circunferencias son ortogonales (en el sentido de la geometría elíptica) si la perpendicular por el origen del plano de una pasa por la otra. En otros términos, que si tomamos a una como ecuador de la esfera, sus perpendiculares son las circunferencias que pasan por los polos. Es fácil ver que esto equivale a que sean ortogonales en el sentido del capítulo 11.

Consideremos ahora una isometría elíptica f . Podemos verla también como un automorfismo de V (visto como espacio vectorial). Sean X e Y las coordenadas de dos puntos y sean X' e Y' las coordenadas de sus imágenes. Puesto que f ha de conservar la polaridad de la cónica infinita, los puntos serán conjugados si y sólo si lo son sus imágenes, es decir, $XY^t = 0$ es equivalente a $X'Y'^t = 0$. Esto significa que f conserva la ortogonalidad en V , luego es una semejanza, luego $f = kg$, donde k es un número real no nulo y g es una isometría lineal de V . Ahora bien, es claro que f y g inducen la misma homografía en el plano proyectivo, luego podemos concluir que las isometrías elípticas están inducidas por las isometrías lineales de V . El recíproco es cierto: si f es una isometría lineal y A es su matriz en el sistema de referencia que estamos considerando, un punto está en la cónica infinita si sus coordenadas (imaginarias) cumplen $XX^t = 0$. Su imagen por f tiene coordenadas XA y, como A es una matriz ortogonal, $XAA^tX^t = XX^t = 0$, luego la imagen sigue en la cónica. En re-

sumen: las isometrías elípticas son exactamente las homografías inducidas por las isometrías lineales de V o, equivalentemente, por las isometrías de V (visto como espacio afín) que fijan a O .

Definición 13.2 Llamaremos *modelo esférico* de la geometría elíptica al conjunto S formado por los pares de puntos antípodas de una esfera euclídea. Las rectas elípticas son los círculos máximos. Dos rectas son perpendiculares si lo son sus tangentes en los puntos de corte. Una isometría elíptica es la restricción a S de una isometría del espacio euclídeo que fija al centro de S .

Además hemos visto que el polo de una recta en el modelo esférico es el par de polos (en el sentido geográfico) cuando consideramos a la recta como ecuador.

Las isometrías del modelo esférico son en particular homografías, luego conservan ángulos en el sentido del capítulo 11. Podemos, pues, definir el ángulo entre dos rectas elípticas en la esfera como el ángulo euclídeo entre sus tangentes por los puntos de corte y es claro entonces que dos pares de rectas son congruentes si y sólo si forman el mismo ángulo. Veamos que el ángulo entre dos rectas puede definirse sin hacer referencia al modelo esférico exactamente como en la geometría hiperbólica, es decir,

$$\widehat{r_1 r_2} = \frac{1}{2} |\arg \mathcal{R}(r_1, r_2, I_1, I_2)|,$$

donde I_1 e I_2 son las tangentes a la cónica infinita por el punto $r_1 \cap r_2$ y el argumento se toma en $] -\pi, \pi[$.

En efecto, el ángulo definido así es sin duda invariante por isometrías. Basta probar que ambas definiciones coinciden sobre rectas que pasan por el polo norte $(0, 0, 1)$. Sus coordenadas homogéneas serán de la forma $(a, b, 0)$ y $(a', b', 0)$. Sus tangentes por el polo norte serán las intersecciones de los planos $ax + by = 0$ y $a'x + b'y = 0$ con el plano $z = 1$. Ahora bien, podemos identificar el plano $z = 1$ con el plano afín que resulta de eliminar la recta $z = 0$ en el plano proyectivo de partida. Entonces las tangentes cuyo ángulo hemos de medir son las propias rectas iniciales (sin su punto infinito). Con más detalle: en este plano afín tenemos ahora una estructura euclídea, de modo que las coordenadas cartesianas (x, y) correspondientes a las coordenadas homogéneas $(x, y, 1)$ están asociadas a un sistema de referencia ortonormal. Hasta aquí tenemos que el ángulo entre las rectas definido en el modelo esférico corresponde con el ángulo entre ellas mismas respecto a esta estructura euclídea. Este ángulo puede calcularse por la fórmula de Cayley, que es formalmente idéntica a la que queremos probar. Sólo hemos de asegurarnos de que las tangentes a la cónica infinita por el punto de corte $(0, 0, 1)$ coinciden con las rectas isotropas por dicho punto, pero esto es obvio, pues ambas son las rectas $x \pm iy = 0$. En efecto, ambas pasan por los puntos circulares $(1, \pm i, 0)$ y sus coordenadas homogéneas cumplen $x^2 + y^2 + z^2 = 0$.

Podemos plantearnos la fórmula análoga a la distancia hiperbólica entre dos puntos P y Q en términos de la razón doble $\mathcal{R}(P, Q, P_\infty, Q_\infty)$, donde P_∞ y Q_∞ son los puntos de corte entre la cónica infinita y la recta PQ . Ahora bien, si r_1

y r_2 son las polares de P y Q y se cortan en un punto R , la recta polar de R es PQ , luego las tangentes I_1 e I_2 a la cónica infinita que pasan por R cortan a PQ en P_∞ y Q_∞ , luego $\mathcal{R}(P, Q, P_\infty, Q_\infty) = \mathcal{R}(r_1, r_2, I_1, I_2)$. Por lo tanto, esta razón doble es un número complejo de módulo 1 cuyo argumento (dividido entre 2) es el ángulo entre dichas polares. Esto nos lleva a la definición siguiente, donde recogemos también la definición de ángulo que hemos discutido anteriormente.

Definición 13.3 Llamaremos *distancia elíptica* entre dos puntos P y Q como

$$d(P, Q) = \frac{1}{2} |\arg \mathcal{R}(P, Q, P_\infty, Q_\infty)|,$$

donde P_∞ y Q_∞ son los puntos en que la recta PQ corta a la cónica infinita y el argumento se toma en $] -\pi, \pi[$.

El *ángulo elíptico* entre dos rectas r_1 y r_2 es

$$\widehat{r_1 r_2} = \frac{1}{2} |\arg \mathcal{R}(r_1, r_2, I_1, I_2)|,$$

donde I_1 e I_2 son las tangentes a la cónica infinita por el punto $r_1 \cap r_2$ y el argumento se toma en $] -\pi, \pi[$.

Ambas magnitudes pueden variar en $[0, \pi/2[$, con el convenio de que son nulas si los puntos o las rectas son iguales. Ambas se conservan por isometrías. Hemos probado que la distancia entre dos puntos es el ángulo entre sus polares. En el modelo esférico la distancia entre dos puntos P y Q coincide con el ángulo entre las rectas OP y OQ , donde O es el centro de la esfera. El ángulo entre dos rectas es el definido en el capítulo 11.

También es claro que dos pares de puntos/rectas son congruentes si y sólo si su distancia/ángulo es el mismo.

Exactamente el mismo argumento que en el caso de la geometría hiperbólica nos lleva a las fórmulas:

$$\cos d = \frac{xx' + yy' + zz'}{\sqrt{x^2 + y^2 + z^2} \sqrt{x'^2 + y'^2 + z'^2}},$$

$$\cos \theta = \frac{xx' + yy' + zz'}{\sqrt{x^2 + y^2 + z^2} \sqrt{x'^2 + y'^2 + z'^2}},$$

donde (x, y, z) y (x', y', z') son las coordenadas homogéneas de dos puntos/rectas. No obstante estas fórmulas se siguen inmediatamente del modelo esférico usando las propiedades del producto escalar en V .

13.2 Biláteros y triángulos

La ordenación circular de la geometría elíptica confiere a plano propiedades muy distintas de las que se siguen de la ordenación lineal de las geometrías euclídea e hiperbólica. Para empezar hemos de tener presente que dos puntos no

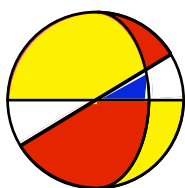
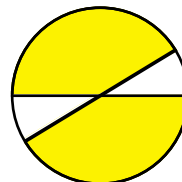
determinan un segmento, sino dos segmentos complementarios. De este modo, cuando hablemos de un segmento \overline{AB} habremos de entender que se trata de uno de los dos segmentos de extremos A y B . Recordemos que habíamos adoptado el convenio de llamar \overline{ACB} al segmento de extremos A y B que contiene a C .

Es fácil ver que un segmento y su complementario no son congruentes salvo que los extremos disten $\pi/2$. Por ello no podemos definir la longitud de un segmento como la distancia entre sus extremos, ya que en tal caso estaríamos asignando la misma longitud a segmentos complementarios no congruentes. Para definir adecuadamente la longitud de un segmento nos basaremos en el hecho siguiente, que se comprueba sin dificultad en el modelo esférico:

Cada segmento elíptico \overline{AB} contiene un único punto C con la propiedad de que $d(A, C) = d(C, B)$. Lo llamaremos punto medio del segmento.

Ahora definimos la *longitud* de un segmento como el doble de la distancia de sus extremos a su punto medio. La idea es que en el modelo esférico un segmento es un arco de circunferencia de amplitud menor que π (en realidad es un par de arcos antípodos de amplitud menor que π), por lo que la mitad de un segmento es un arco de amplitud menor que $\pi/2$ y así la amplitud de este arco sí coincide con la distancia entre sus extremos, luego la longitud del segmento completo así definida coincide con su amplitud como arco. Ahora sí podemos afirmar que dos segmentos son congruentes si y sólo si tienen la misma longitud. Es fácil probar que la longitud de un segmento y la de su complementario suman π . Podemos expresar esto diciendo que las rectas tienen longitud π .

Una recta no divide al plano en dos partes. Dos rectas no lo dividen en cuatro ángulos, sino en dos regiones que podemos llamar *biláteros*. Un bilátero tiene un vértice y dos lados. Dos puntos están en el mismo bilátero si existe un segmento que los une sin cortar a los lados. Una forma fácil de probar estos hechos es observar que los biláteros son simplemente los semiplanos que determina una recta en el espacio afín que resulta de tomar a la otra como recta infinita. Podemos asignar una medida al ángulo de cada bilátero del mismo modo que hemos hecho con los segmentos. La suma de los ángulos de dos biláteros complementarios es π .



Consideremos tres puntos no colineales A , B , C . Tomemos AB como recta infinita. Entonces los biláteros de vértice A y lados AB y AC se convierten en los dos semiplanos determinados por AC . La recta BC es secante a AC , luego divide al plano en cuatro ángulos. Dos puntos están en el mismo ángulo si y sólo si el segmento (afín) que los une no corta a BC ni a AC , o equivalentemente, si y sólo si existe un segmento que los une y no corta a ninguna de las rectas AB , AC , BC .

Esta caracterización es simétrica y no depende de la recta que hemos tomado como infinita ni del orden de los puntos, luego podemos afirmar que tres puntos no colineales A , B , C dividen al plano en cuatro partes que llamaremos

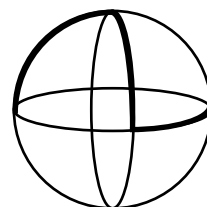
triángulos de vértices A, B, C . Es claro por las definiciones que dos puntos están en el mismo triángulo de vértices A, B, C si y sólo si están en el mismo bilátero de vértice A y lados AB, AC , en el mismo bilátero de vértice B y lados BA, BC y en el mismo bilátero de vértice C y lados CA y CB . Llamaremos \hat{A}, \hat{B} y \hat{C} para un triángulo dado T a los biláteros de vértices A, B, C que contienen a T , y los llamaremos *ángulos* de T .

Sea T un triángulo de vértices A, B, C . Podemos tomar una recta que pase por A pero no esté contenida en \hat{A} . Si la convertimos en recta infinita las rectas AB y AC se vuelven paralelas y T se convierte en el conjunto de los puntos comprendidos entre ambas y que están en uno de los semiplanos determinados por BC . Vemos entonces que de los dos segmentos que B y C determinan en BC , uno está contenido en el ángulo \hat{A} y el otro es exterior a él.

En general, llamaremos *lado* opuesto al vértice A al segmento de BC contenido en \hat{A} . Lo mismo vale para los otros vértices. Dado un triángulo T , cada uno de los otros tres triángulos con sus mismos vértices comparte con T un ángulo y su lado opuesto. Los otros dos lados y ángulos son los complementarios de los de T .

Si, al igual que antes, representamos a un triángulo T como el conjunto de puntos afines comprendidos entre dos rectas paralelas y situados en un mismo semiplano respecto a una tercera recta r , veremos claramente que existe una recta que no corta a ninguno de sus lados (ni pasa por sus vértices). Basta tomar una recta paralela a r en el semiplano opuesto a T . Tenemos, pues, que siempre existe una recta que no corta a ninguno de los lados de un triángulo. Si tomamos esta recta como infinita T se convierte en un triángulo afín ordinario. De aquí podemos traducir al plano elíptico algunos teoremas sobre la ordenación lineal. Por ejemplo, una recta no puede cortar a los tres lados de un triángulo. Sin embargo, el hecho de que exista una recta que no corta a los vértices y lados de un triángulo implica que corta a los complementarios de los lados, luego estos segmentos no son los lados de ningún triángulo.

Esto muestra que tres segmentos pueden conectar tres puntos y pese a ello no ser los lados de ningún triángulo. Con más detalle: digamos que un triángulo tiene lados a, b, c y llamemos $-a, -b, -c$ a los complementarios. Entonces los otros triángulos tienen lados $(a, -b, -c)$, $(-a, b, -c)$ y $(-a, -b, c)$. Las otras cuatro combinaciones son imposibles. Vemos entonces que si en una combinación posible de segmentos cambiamos uno por su complementario obtenemos una combinación imposible y viceversa.

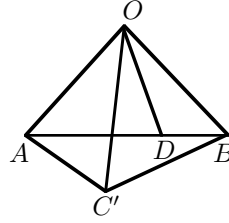


Teorema 13.4 *Un lado de un triángulo es menor que la suma de los otros dos.*

DEMOSTRACIÓN: Consideremos un triángulo elíptico \widehat{ABC} . Supongamos que \overline{AB} es el lado mayor. Si el triángulo tiene dos lados iguales mayores o iguales que el tercero la desigualdad es obvia. Bastará probar que $\overline{AB} < \overline{AC} + \overline{CB}$. Trabajamos en el modelo esférico. Sea O el centro de la esfera. En cuanto sigue \overline{AB} representará el segmento euclídeo de extremos A y B y en lugar del segmento elíptico correspondiente.

Sea D el punto sobre \overline{AB} que hace $\widehat{AOD} = \widehat{AOC}$. Sea C' el punto sobre la semirrecta \overrightarrow{OC} tal que $\overline{OC'} = \overline{OD}$. Entonces $\widehat{AOC'} = \widehat{AOD}$, pues tienen dos lados iguales y el ángulo entre ellos, luego $\overline{AC'} = \overline{AD}$.

Ahora, $\overline{BC'} > \overline{AB} - \overline{AC'} = \overline{AB} - \overline{AD} = \overline{DB}$. Los triángulos $\widehat{OBC'}$ y \widehat{OBD} tienen dos lados iguales y un lado desigual $\overline{BC'} > \overline{DB}$, luego $\widehat{BOC'} > \widehat{BOD} = \widehat{AOB} - \widehat{AOD}$, es decir, $\widehat{BOC'} > \widehat{AOB} - \widehat{AOC}$, que en términos del triángulo elíptico original es $\overline{AB} < \overline{AC} + \overline{CB}$, como había que probar. ■



Ejercicio: Probar que un lado de un triángulo elíptico es siempre mayor que la diferencia de los otros dos.

Teorema 13.5 *La suma de los lados de un triángulo es menor que 2π .*

DEMOSTRACIÓN: Si el triángulo tiene lados a, b, c , consideramos uno de los triángulos adyacentes, digamos el de lados $\pi - a, \pi - b, c$. Por el teorema anterior $c < \pi - a + \pi - b$, luego $a + b + c < 2\pi$. ■

Al aplicar la polaridad de la cónica infinita a las rectas que pasan por el vértice de un bilátero y están contenidas en él obtenemos un segmento. Es fácil ver que si el ángulo del bilátero mide a lo sumo $\pi/2$ lo mismo le ocurre al segmento, y en tal caso el ángulo del bilátero coincide con el que forman sus lados y la longitud del segmento coincide con la distancia entre sus extremos, y además sabemos que ambas coinciden. Es decir, la polaridad transforma biláteros rectos y agudos en segmentos de la misma longitud. Pasando por los complementarios vemos que también es cierto para biláteros obtusos.

En particular la polaridad transforma los tres ángulos de un triángulo T en tres segmentos que conectan los polos de sus lados. La longitud de cada segmento es igual a la amplitud del ángulo del que proviene. Sin embargo estos segmentos no son los lados de ningún triángulo. Para probarlo observamos que un punto P del triángulo está contenido en una recta contenida en cada ángulo, luego su recta polar pasa por un punto de cada segmento. Por consiguiente, los complementarios de estos segmentos sí son los lados de un triángulo (al tomar complementarios tres veces pasamos de una combinación imposible a otra posible).

Definición 13.6 Llamaremos *triángulo polar* de un triángulo elíptico T al triángulo que tiene por vértices los polos de las prolongaciones de sus lados y por lados los complementarios de los segmentos formados por los polos de las rectas contenidas en sus ángulos.

Hemos probado que si T' es el triángulo polar de T y los ángulos de T miden α, β y γ , entonces los lados de T' miden $\pi - \alpha, \pi - \beta, \pi - \gamma$.

Veamos que los ángulos de T' son los complementarios de los biláteros formados por las polares de los puntos de los lados de T . Digamos que los vértices

de T son A, B, C y sus lados opuestos son a, b, c . Tomemos un punto Q de la recta BC . Entonces Q estará en a si y sólo si la recta AQ está en \hat{A} , si y sólo si $p(AQ)$ no está en el lado de T' de extremos $p(AB), p(AC)$, si y sólo si $p(Q)$ no está en el bilátero de vértice $p(BC)$, si y sólo si $p(Q)$ está en el complementario de dicho bilátero.

Con esto es inmediato el teorema siguiente:

Teorema 13.7 *Si T es un triángulo elíptico y T' es su triángulo polar, entonces T es el triángulo polar de T' . Si los ángulos de T miden α, β, γ y sus lados miden a, b, c , entonces los lados de T' miden $\pi - \alpha, \pi - \beta, \pi - \gamma$ y sus ángulos miden $\pi - a, \pi - b, \pi - c$.*

Este teorema nos permite extender a la geometría elíptica el principio de dualidad de la geometría proyectiva. Veamos una primera aplicación:

Teorema 13.8 *La suma de los ángulos de un triángulo es mayor que π y menor que 3π .*

DEMOSTRACIÓN: Si los ángulos del triángulo son α, β, γ , entonces el teorema anterior aplicado al triángulo polar nos da $0 < \pi - \alpha + \pi - \beta + \pi - \gamma < 2\pi$, luego $\pi < \alpha + \beta + \gamma < 3\pi$. ■

Ejercicio: Probar que un triángulo coincide con su polar si y sólo si todos sus lados y ángulos miden $\pi/2$.

13.3 Isometrías elípticas

Las isometrías elípticas son mucho más simples que las euclídeas o las hiperbólicas. Consideremos el modelo esférico. Una isometría elíptica está inducida por una isometría euclídea que fija a la esfera, y sabemos que en un sistema de referencia adecuado su matriz ha de ser de la forma

$$\begin{pmatrix} \cos \theta & \operatorname{sen} \theta & 0 \\ -\operatorname{sen} \theta & \cos \theta & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}.$$

Si multiplicamos todos los coeficientes por -1 obtenemos otra isometría euclídea que induce la misma isometría elíptica (pues estamos componiéndola con la isometría que intercambia los puntos antípodas). Por consiguiente podemos suponer que la última entrada es un $+1$. En definitiva, cada isometría elíptica está inducida por un único giro en \mathbb{R}^3 .

El eje del giro corta al plano elíptico en un único punto O . Es claro que O es un punto fijo, al igual que lo es su recta polar r . Si el ángulo θ es distinto de π entonces O es el único punto fijo y r la única recta fija. Si $\theta = \pi$ entonces la isometría fija también a todos los puntos de r y a todas las rectas que pasan por O . En definitiva:

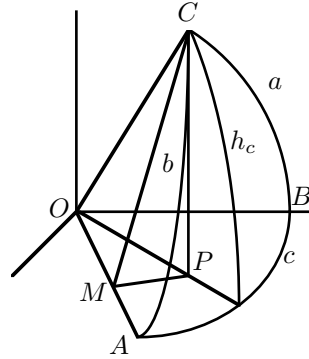
Teorema 13.9 Si f es una isometría del plano elíptico distinta de la identidad, existen un punto O y una recta r unívocamente determinados, a los que llamaremos centro y eje de f de modo que los únicos puntos fijados por f son O y quizá los puntos de r , y las únicas rectas fijadas por f son r y quizá las rectas que pasan por O . Además r es la recta polar de O .

Por razones obvias las isometrías elípticas se llaman también *giros*.

13.4 Trigonometría elíptica

La trigonometría elíptica es completamente análoga a la hiperbólica. Consideremos un triángulo de vértices A, B, C , lados a, b, c y ángulos α, β, γ , relacionados como es habitual. Comenzaremos demostrando el teorema del coseno.

Supongamos primeramente que α y b son agudos. Sea P la proyección ortogonal de C sobre el plano AOB y sea M su proyección ortogonal sobre OA . El plano CPM es perpendicular a OA , pues contiene las rectas CP y CM , ambas perpendiculares a OA . Por consiguiente, las rectas MP y MC son paralelas a las tangentes a la esfera por M contenidas en los planos OAB y OAC respectivamente, es decir, $\widehat{CMP} = \alpha$. Claramente se tiene además que $\overline{OM} = \cos b$, $\overline{MC} = \sin b$, $\overline{MP} = \sin b \cos \alpha$, $\overline{PC} = \sin b \sin \alpha$.



Consideremos la suma de vectores $\overrightarrow{OM} + \overrightarrow{MP} + \overrightarrow{PC} = \overrightarrow{OC}$ y proyectémosla ortogonalmente sobre la recta OB . Si llamamos p a la proyección es fácil comprobar que

$$\begin{aligned} p(\overrightarrow{OM}) &= \cos b \cos c, \\ p(\overrightarrow{MP}) &= \sin b \cos \alpha \cos(\pi/2 - c) = \sin b \sin c \cos \alpha, \\ p(\overrightarrow{PC}) &= 0, \\ p(\overrightarrow{OC}) &= \cos a. \end{aligned}$$

Por lo tanto

$$\cos a = \cos b \cos c + \sin b \sin c \cos \alpha. \quad (13.1)$$

Si $\alpha > \pi/2$ entonces cambia el sentido de \overrightarrow{MP} , pero también cambia el signo de $\cos \alpha$, luego la fórmula sigue siendo válida. Si $b > \pi/2$ entonces aplicamos la fórmula al triángulo adyacente de lados $a, \pi - b, \pi - c$ y el mismo ángulo α y obtenemos la misma fórmula para el triángulo dado. Si $\alpha = \pi/2$ entonces $M = P$ y si $b = \pi/2$ entonces $O = M$. En ambos casos la fórmula se deduce sin dificultad.

Volviendo al caso en que α y b son agudos, consideremos ahora la proyección sobre el eje perpendicular al plano AOB . Entonces

$$\begin{aligned} p(\overrightarrow{OM}) &= 0, \\ p(\overrightarrow{MP}) &= 0, \\ p(\overrightarrow{PC}) &= \text{sen } b \text{ sen } \alpha, \\ p(\overrightarrow{OC}) &= \text{sen } h_c, \end{aligned}$$

donde h_c es la altura del triángulo que une C con c . Tenemos, pues,

$$\text{sen } h_c = \text{sen } b \text{ sen } \alpha,$$

y como antes se comprueba que esta relación es válida en el caso general. Aplicándola a β obtenemos $\text{sen } h_c = \text{sen } a \text{ sen } \beta$. Despejando $\text{sen } h_c$ e igualando obtenemos el teorema de los senos. Así pues,

$$\frac{\text{sen } a}{\text{sen } \alpha} = \frac{\text{sen } b}{\text{sen } \beta} = \frac{\text{sen } c}{\text{sen } \gamma}. \quad (13.2)$$

Las fórmulas (13.1) y (13.2) se conocen como *fórmulas de Bessel*.

La fórmula (13.1) muestra que si dos triángulos tienen los lados iguales entonces también tienen los ángulos iguales, y de aquí se sigue fácilmente que son congruentes. Por dualidad, si dos triángulos tienen los ángulos iguales también son congruentes. De hecho, aplicando la fórmula (13.1) al triángulo polar obtenemos el teorema del coseno para un lado:

$$\cos \alpha = -\cos \beta \cos \gamma + \text{sen } \beta \text{ sen } \gamma \cos a. \quad (13.3)$$

Dejamos a cargo del lector la comprobación de que todas las fórmulas de la trigonometría hiperbólica son válidas para triángulos elípticos cuando las razones hiperbólicas se sustituyen por las correspondientes circulares. Todas las comprobaciones son fáciles a partir de las fórmulas que ya tenemos. Por ejemplo, el teorema del coseno cuando $\alpha = \pi/2$ se convierte en el teorema de Pitágoras elíptico:

$$\cos a = \cos b \cos c.$$

Cada relación tiene una dual, resultante de aplicarla al triángulo polar. Por ejemplo, todo triángulo rectilátero (con un lado recto, digamos a) cumple

$$\cos \alpha = -\cos \beta \cos \gamma.$$

Bibliografía

- [1] Artin, E. *Algèbre géométrique*. Gautier Villars, Paris, 1972
- [2] Baker, A. *Breve introducción a la teoría de números*. Alianza Ed., Madrid, 1986.
- [3] Bastida, J.R. *Field extensions and Galois theory*. Addison-Wesley P.C., California, 1984.
- [4] Busemann, H., Kelly, P.J. *Projective Geometry and Projective Metrics*. Academic Press, New York, 1953.
- [5] Coxeter H.S.M., Greitzer, S.L *Geometry revisited*. L. W. Singer Company, 1967
- [6] Edwards, H.M. *Fermat's last theorem*. Springer, New York, 1977.
- [7] Hungerford, T.W. *Algebra*. Springer, New York, 1974.
- [8] Jennings, G.A. *Modern Geometry with Applications*. Springer, New York, 1994.
- [9] Nieto Vales, J.M. *Curso de trigonometría esférica*. Servicio de Publicaciones de la Universidad de Cádiz, 1996.
- [10] Reventós, A. *Geometria axiomàtica*. Institut d'Estudis Catalans, Barcelona, 1993.
- [11] Seidenberg, A. *Lectures in Projective Geometry*. Van Nostrand, Princeton, 1968.
- [12] Snapper, E., Troyer, R.J. *Metric Affine Geometry*. Dover, New York, 1971.
- [13] Stillwell, J. *Geometry of Surfaces*. Springer, New York, 1992.
- [14] Tisseron, C. *Géométries affine, projective et euclidienne*. Hermann, Paris, 1988.
- [15] Veblen, O, Young, J.W. *Projective Geometry*. Blaisdell, New York, 1938 (vol I), 1946 (vol II).

Índice de Materias

- adyacentes (ángulos), 7
- afín
 - aplicación, 83
 - espacio, 79
- afinidad, 83
- afínmente independientes, 80
- agudo (ángulo), 18
- altura, 57
- amplitud, 102
 - de un arco, 47
 - de un ángulo, 45
- ángulo, 6
 - agudo, 18
 - de un triángulo, 8
 - entre circunferencias, 357, 359
 - entre dos vectores, 89
 - inscrito, 59
 - llano, 14
 - obtuso, 18
 - recto, 18
 - semiinscrito, 60
 - suplementario, 14
- ángulos
 - adyacentes, 7
 - alternos, 51
 - internos/externos, 50
 - opuestos por el vértice, 7
- antisimétrica (forma), 232
- aplicación afín, 83
- arco, 290
 - abarcado, 59
- área (de un triángulo), 58
- arguesiano (plano), 198
- argumento, 99
- Arquímedes (propiedad de), 29, 202
- artiniano (plano), 317
- asíntota, 266
- baricentro, 69, 87
- Bessel (fórmulas), 396
- bilátero, 391
- bisectriz, 17
- cateto, 18
- Cayley (fórmula), 315
- centro, 266
 - de perspectiva, 234
 - de una homografía en una cónica, 287
 - de una semejanza, 151
- cevianas, 67
- circulares (puntos), 300
- círculo, 21
 - de los nueve puntos, 181
- circuncentro, 63
- circunferencia, 21, 302, 380
 - circunscrita, 63, 124
 - inscrita, 65
- circunradio, 63
- circunscrita (figura), 276
- colineales (puntos), 2
- compleción proyectiva, 207, 212
- complemento ortogonal, 94
- complexificación, 294
- congruencia, 9, 95, 307, 308, 387
 - de matrices, 169, 251
- cónica, 264, 271
 - afín, 256
 - de rectas, 278
 - euclídea, 110
 - proyectiva, 257
- conjugación, 116, 136
 - armónica, 239

- respecto a una involución, 243
- conjunto cuadrangular, 239
- constructibilidad, 121
- convexo, 6
- coordenadas
 - baricéntricas, 87
 - cartesianas, 84
 - homogéneas, 216, 226, 284
 - polares, 99, 333
- coplanares (puntos), 2
- correlación, 231
 - proyectiva, 231
 - simétrica, 232
- coseno, 62, 99
 - hiperbólico, 331
- cruce (de rectas), 3
- cuadrado, 53
- cuadrilátero
 - completo, 238
 - dual, 281
- cuerda, 21, 46
- dilatación, 222
- dirección, 188
 - de un vector, 74
- directriz, 340
- distancia, 34, 90
 - elíptica, 390
 - hiperbólica, 368
- divisores elementales, 153, 160
- diámetro, 21
- dual, 229
- dualidad, 233
- ecuación de clases, 137
- eje
 - de perspectiva, 234
 - de una cónica, 336
 - de una homografía en una cónica, 287
 - radical, 106
- elación, 220
- elipse, 106, 296
- entre, 4, 14
- equilátero (triángulo), 12
- equipolentes (vectores), 74
- equivalencia (de matrices), 153
- escaleno, 12
- esfera, 349
- espacial (punto, etc.), 321
- espacio, 2
 - afín, 79, 184
 - euclídeo, 90
 - ordenado, 201
- director, 79
- euclídeo, 90
- fundamental, 167, 168
- proyectivo, 211, 224
- soporte, 211
- estereográfica (proyección), 350
- euclídeo
 - espacio, 90
 - espacio afín, 90
- excentricidad, 343
- excentro, 67
- extradio, 67
- exterior (punto), 2
- exterior/interior, 290
- factores invariantes, 153, 155, 160, 163
- foco, 336
- forma bilineal
 - antisimétrica, 232
 - ortosimétrica, 232
 - simétrica, 232
- forma canónica, 156, 161, 172
- frontera
 - de un círculo, 21
 - de un semiplano, 5
 - de un triángulo, 8
 - de un ángulo, 6
- función
 - exponencial, 42
 - logarítmica, 42
- geometría, 2
 - euclídea, 50
 - métrica, 9
 - ordenada, 3
- giro, 99, 173, 306, 380
- infinito, 381

- orientado, 324
- grupo
 - afín, 139
 - circular, 354
 - de traslaciones, 140
 - elíptico, 387
 - hiperbólico, 364
 - lineal, 139
 - ortogonal, 149
 - proyectivo, 214
- haz de rectas, 186
- hexágono, 198
- hiperosculantes, 382
- hiperplano, 79
- hipotenusa, 18
- hipérbola, 109, 296
- homografía, 214, 286
 - elíptica, parabólica, hiperbólica, 248
 - en una esfera, 355
 - entre cónicas, 282
 - entre haces de rectas, 269
- homología, 220
- homotecia, 141, 186
 - lineal, 141, 190
- imaginarios conjugados, 297
- incentro, 65
- índice, 258
 - reducido, 262
- inradio, 65
- inscrita (figura), 276
- intervalo, 39
- inversión, 356
- inversos (puntos), 313
- involución, 243
 - inducida por una cónica, 282
 - ortogonal, 299
- isometría, 93, 303, 365, 387
 - directa/indirecta, 178
 - lineal, 94
- isósceles (triángulo), 12
- isótropas (rectas), 300
- isótropo, 308
- Klein (plano de), 364
- llano (ángulo), 14
- longitud, 34
- Lorentz
 - plano, 317
 - relaciones, 334
- matriz
 - de coordenadas, 217
 - ortogonal, 149
- medianas, 69
- mediatriz, 21
- módulo, 38, 116
- movimiento, 178, 306
- norma, 74, 90, 324
- número
 - complejo, 115
 - constructible, 121
 - real, 31
- obtuso (ángulo), 18
- ordenación, 200
 - débil, 200
- orientación (del plano de Lorentz), 321
- ortocentro, 70
- ortogonal
 - grupo, 149
 - matriz, 149
- ortogonalidad, 89, 90, 300
- ortonormal, 90
- ortosimétrica (forma), 232
- osculante, 381
- paralelismo, 3, 80, 184, 364
- paralelogramo, 52
- parte entera, 38
- parte real/imaginaria, 116
- parábola, 111, 296
- perpendicular, 18, 20, 300, 364, 387
- perspectividad, 220
- pi, 45
- plano, 2, 79
 - arguesiano, 229
 - artiniano, 317
 - de Lorentz, 317
 - elíptico, 387

- euclídeo, 309
- hiperbólico, 363
- parabólico, 299
- proyectivo, 207
- Poincaré (plano de), 374
- polaridad, 233
 - de una cónica, 265
- polinomio
 - característico, 165
 - mínimo, 158, 160, 163
- polígono regular, 124
- potencia, 106
- primos de Fermat, 126
- principio de dualidad, 231
- producto escalar, 89, 90, 328
- prolongación
 - de un segmento, 4
 - de un semiplano, 6
 - de una semirrecta, 5
- proporción, 34
- proyección
 - paralela, 142, 201
 - perspectiva, 219, 270, 286
- proyectivamente independientes, 212
- punto, 2, 79
 - de tangencia, 278
 - finito, 207
 - infinito, 207, 364
 - medio, 16, 174
 - ultrainfinito, 364
- rango, 156
 - reducido, 262
- razón, 324
 - de una semejanza, 335
- razón doble, 235
 - entre rectas, 314
- recta, 2, 79
 - de Euler, 180
 - graduada, 30
 - orientada, 73
- rectángulo, 53
- reflexión, 175, 302, 365
- relatividad especial, 317
- rombo, 53
- secantes
 - a una circunferencia, 23
 - planos, 3
 - rectas, 3
- segmento, 4, 245
- semejantes (triángulos), 53
- semejanza, 150, 299, 307
 - de matrices, 157
 - lineal, 150
- semicircunferencia, 46
- semihaz de semirrectas, 14
- semiplano, 5
 - complementario, 6
- semirrecta, 5
 - complementaria, 5
- seno, 62, 99
 - hiperbólico, 331
- sentido (de un vector), 74
- separación, 247
 - harmónica, 238
- signatura, 253, 254
- simetría, 175, 306
 - puntual, axial, especular, 175
- simétrica (forma), 232
- sistema de referencia, 84
 - canónico, 86
 - orientado, 325
 - ortonormal, 325
 - proyectivo, 215
- soporte
 - de un círculo, 21
 - de un ángulo, 6
- suma
 - de segmentos, 10
 - de ángulos, 15
 - ortogonal, 94
- suplementarias (variedades), 142
- suplementarios (ángulos), 14
- Tales (posición de), 53
- tangente, 104
 - a una circunferencia, 23
 - a una cónica, 273
 - hiperbólica, 332
- temporal (punto, etc.), 321
- Teorema
 - de Brianchon, 279

- de Cayley, 166
- de Ceva, 242
- de Desargues, 145, 194, 227
 - sobre cónicas, 277
- de Feuerbach, 360
- de Gram-Schmidt, 91
- de la potencia, 105
- de los senos, 63, 104
- de Menelao, 241
- de Papos-Pascal, 198
- de Pitágoras, 56, 329
- de Steiner, 270, 275
- de Sylow, 118
- de Sylvester, 253
- de Tales, 55, 143
- de Wedderburn, 199
- del coseno, 63, 378
- fund. de la geom. afín, 146
- fund. de la geom. proyectiva, 218
- fundamental del álgebra, 118
- torsión, 153
- transformación circular, 353
- transvección, 222
- traslación, 140, 186, 383
- traza, 167, 188
- triángulo, 8, 226
 - elíptico, 392
 - equilátero, 12
 - escaleno, 12
 - isósceles, 12
 - medial, 179
 - órtico, 70
 - polar, 393
- ultraparalelismo, 364
- valor absoluto, 38
- valor propio, 167
- variedad
 - inducida, 211
 - lineal, 79
 - proyectiva, 211
- vector
 - de posición, 84
 - director, 75
 - fijo, 74
 - libre, 75
 - propio, 167
- vectores ortogonales, 89, 90
- vértice
 - de un triángulo, 8
 - de un ángulo, 6
 - de una cónica, 336