

Capítulo 1

El plano euclidiano

1.1 La geometría griega

En el principio, la geometría era una colección de reglas de uso común para medir y construir casas y ciudades. No fue hasta el año 300 a.C. que Euclides de Alejandría, en sus *Elementos*, ordenó y escribió todo ese saber, imprimiéndole el sello de rigor lógico que caracteriza y distingue a las matemáticas. Se dio cuenta de que todo razonamiento riguroso (o *demostración*) debe basarse en ciertos principios previamente establecidos ya sea, a su vez, por otra demostración o bien por convención. Pero a final de cuentas, este método conduce a la necesidad ineludible de convenir en que ciertos principios básicos (*postulados* o *axiomas*) son válidos sin necesidad de demostrarlos, que están dados y son incontrovertibles para poder construir sobre ellos el resto de la teoría. Lo que hoy se conoce como Geometría Euclidiana, y hasta hace dos siglos simplemente como Geometría, está basada en los cinco postulados de Euclides:

- I Por cualesquiera dos puntos, se puede trazar el segmento de recta que los une.
- II Dados un punto y una distancia, se puede trazar el círculo con centro en el punto y cuyo radio es la distancia.
- III Un segmento de recta se puede extender en ambas direcciones indefinidamente.
- IV Todos los ángulos rectos son iguales.
- V Dadas dos rectas y una tercera que las corta, si los ángulos internos de un lado suman menos de dos ángulos rectos, entonces las dos rectas se cortan y lo hacen de ese lado.

Obsérvese que en estos postulados se describe el comportamiento y la relación entre ciertos elementos básicos de la geometría, como son “punto”, “trazar”, “segmento”, “distancia”, etc. De alguna manera, se le da la vuelta a su definición haciendo uso de

la noción intuitiva que se tiene de ellos y haciendo explícitas ciertas relaciones básicas que deben cumplir.

De estos postulados o axiomas, el quinto es el más famoso pues se creía que podría deducirse de los otros. Es más sofisticado, y entonces se pensó que debía ser un *teorema*, es decir, ser demostrable. De hecho, un problema muy famoso fue ése: demostrar el quinto postulado usando únicamente los otros cuatro. Y muchísimas generaciones de matemáticos lo atacaron sin éxito. No es sino hasta el siglo XIX, y para sorpresa de todos, que se demuestra que no se puede demostrar; que efectivamente hay que suponerlo como axioma, pues negaciones de él dan origen a “nuevas geometrías”, tan lógicas y bien fundamentadas como la euclidiana. Pero esta historia se verá más adelante (en los capítulos 6 y 8) pues por el momento nos interesa introducir la geometría analítica.

La publicación de la “*Géométrie*” de Descartes marca una revolución en las matemáticas. La introducción del álgebra a la solución de problemas de índole geométrica desarrolló una nueva intuición y, con ésta, un nuevo entender de la naturaleza de las “*lignes courves*”.

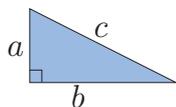
Para comprender lo que hizo René Descartes (1596–1650), se debe tener más familiaridad con el quinto postulado. Además del enunciado original, hay otras dos versiones que son equivalentes:

V.a (El Quinto) Dada una línea recta y un punto fuera de ella, existe una única recta que pasa por el punto y que es paralela a la línea.

V.b Los ángulos interiores de un triángulo suman dos ángulos rectos.

De las tres versiones que hemos dado del quinto postulado de Euclides, usaremos a lo largo de este libro a la **V.a**, a la cual nos referiremos simplemente como “*El Quinto*”.

Antes de entrar de lleno a la Geometría Analítica demostremos, a manera de homenaje a los griegos y con sus métodos, uno de sus teoremas más famosos e importantes.

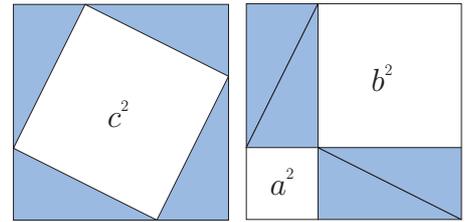


Teorema 1.1 (de Pitágoras) Dado un triángulo rectángulo, si los lados que se encuentran en un ángulo recto (llamados catetos) miden a y b , y el tercero (la hipotenusa) mide c , entonces

$$a^2 + b^2 = c^2.$$

Demostración. Considérese un cuadrado de lado $a + b$, y colóquense en él cuatro copias del triángulo de dos maneras diferentes como en la figura. Las áreas que quedan fuera son iguales. \square

Aunque la demostración anterior (hay otras) parece no usar nada, queda implícito el uso del quinto postulado; pues, por ejemplo, en el primer cuadrado, que el cuadradito interno de lado c tenga ángulo recto usa su versión **V.b**. Además, hace uso de nuevos conceptos como son el área y cómo calcularla; en fin, hace uso de nuestra intuición geométrica, que debemos creer y fomentar. Pero vayamos a lo nuestro.



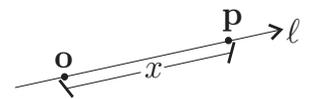
***EJERCICIO 1.1** Demuestra la equivalencia de **V**, **V.a** y **V.b**. Aunque no muy formalmente (como en nuestra demostración del Teorema de Pitágoras), convencerse con dibujos de que tienen todo que ver entre ellos.

EJERCICIO 1.2 Demuestra el Teorema de Pitágoras ajustando (sin que se traslapen) cuatro copias del triángulo en un cuadrado de lado c y usando que $(b-a)^2 = b^2 - 2ab + a^2$ (estamos suponiendo que $a < b$, como en la figura anterior).

1.2 Puntos y parejas de números

Para reinterpretar el razonamiento que hizo Descartes, supongamos por un momento que conocemos bien el *plano euclidiano* definido por los cinco axiomas de los *Elementos*; es el conjunto de puntos que se extiende indefinidamente a semejanza de un pizarrón, un papel, un piso o una pared y viene acompañado de nociones como rectas y distancia, entre otras; y en el cual se pueden demostrar teoremas como el de Pitágoras. Denotaremos por \mathbb{E}^2 a este plano; en este caso el exponente 2 se refiere a la dimensión y no es exponenciación (no es “ \mathbb{E} al cuadrado”, sino que debe leerse “ \mathbb{E} dos”). Da la idea de que puede haber espacios euclidianos de cualquier dimensión \mathbb{E}^n (léase “ \mathbb{E} ene”), aunque sería complicado definirlos axiomáticamente. De hecho los definiremos, pero de una manera más simple, que es con el uso de coordenadas: la idea genial de Descartes. Podríamos ahora, apoyados en el lenguaje moderno de los conjuntos,¹ recrear su razonamiento como sigue.

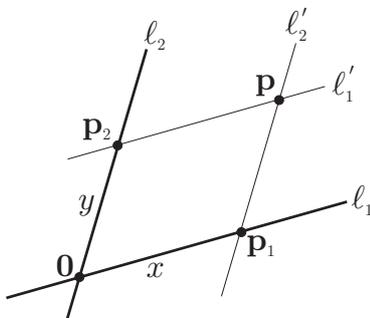
El primer paso es notar que los puntos de una recta $\ell \subset \mathbb{E}^2$ corresponden biunívocamente a los *números reales*, que se denotan por \mathbb{R} . Escogemos un punto $\mathbf{o} \in \ell$ que se llamará el *origen* y que corresponderá al número cero. El origen parte a la recta en dos mitades; a una de ellas se le asocian los números positivos de acuerdo con su distancia al origen, y a la otra mitad, los números negativos. Así, a cada número real x se le asocia un punto $\mathbf{p} \in \ell$, y a cada punto en ℓ le corresponde un número real.



¹La notación básica de conjuntos se establece e introduce brevemente en el Apéndice.

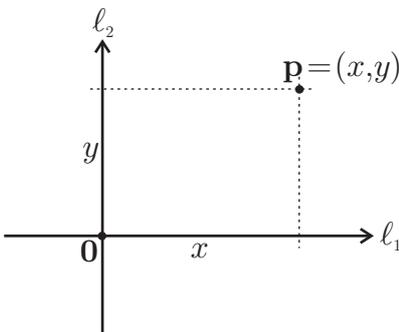
De esta identificación natural surge el nombre de “recta de los números reales” y todo el desarrollo ulterior de la Geometría Analítica, e inclusive del Cálculo. Los números tienen un significado geométrico (los griegos lo sabían bien al entenderlos como distancias), y entonces problemas geométricos (y físicos) pueden atacarse y entenderse manipulando números.

El segundo paso para identificar los puntos del plano euclidiano con parejas de números hace uso esencial del quinto postulado. Tómense dos rectas ℓ_1 y ℓ_2 en el plano \mathbb{E}^2 que se intersecten en un punto \mathbf{o} que será el *origen*. Orientamos las rectas para que sus puntos correspondan a números reales, como arriba. Entonces a cualquier punto $\mathbf{p} \in \mathbb{E}^2$ se le puede hacer corresponder una pareja de números de la siguiente



forma. Por el Quinto, existe una única recta ℓ'_1 que pasa por \mathbf{p} y es paralela a ℓ_1 ; y análogamente, existe una única recta ℓ'_2 que pasa por \mathbf{p} y es paralela a ℓ_2 . Las intersecciones $\ell_1 \cap \ell'_2$ y $\ell_2 \cap \ell'_1$ determinan los puntos $\mathbf{p}_1 \in \ell_1$ y $\mathbf{p}_2 \in \ell_2$, respectivamente, y por tanto, determinan dos números reales x y y ; es decir, una pareja ordenada (x, y) . Y al revés, dada la pareja de números (x, y) , tómense \mathbf{p}_1 como el punto que está sobre ℓ_1 a distancia x de \mathbf{o} , y \mathbf{p}_2 como el punto que está sobre ℓ_2 a distancia y de \mathbf{o} (tomando en cuenta signos, por supuesto). Sea ℓ'_1 la recta que pasa por \mathbf{p}_2 paralela a ℓ_1 y sea

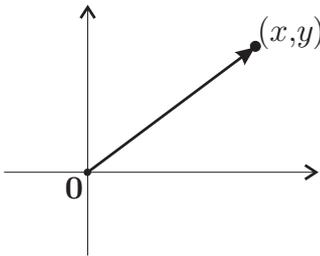
ℓ'_2 la recta que pasa por \mathbf{p}_1 paralela a ℓ_2 (¡cabíamos de usar de nuevo al Quinto!). La intersección $\ell'_1 \cap \ell'_2$ es el punto \mathbf{p} que corresponde a la pareja (x, y) .



La generalidad con que hicimos el razonamiento anterior fue para hacer explícito el uso del quinto postulado en el razonamiento clásico de Descartes, pero no conviene dejarlo tan ambiguo. Es costumbre tomar a las rectas ℓ_1 y ℓ_2 ortogonales: a ℓ_1 horizontal con su dirección positiva a la derecha (como vamos leyendo), conocida tradicionalmente como el *eje de las x*; y a ℓ_2 vertical con dirección positiva hacia arriba, el famoso *eje de las y*. No sólo es por costumbre que se utiliza esta convención, sino también porque simplifica el álgebra asociada a la geometría clásica.

Por ejemplo, permitirá usar el teorema de Pitágoras para calcular fácilmente las distancias a partir de las coordenadas. Pero esto se verá más adelante.

En resumen, al fijar los ejes coordenados, a cada pareja de números (x, y) le corresponde un punto $\mathbf{p} \in \mathbb{E}^2$, pero iremos más allá y los identificaremos diciendo $\mathbf{p} = (x, y)$. Además, se le puede hacer corresponder la flecha (segmento de recta dirigido llamado *vector*) que “nace” en el origen y termina en el punto. Así, las siguientes interpretaciones son equivalentes y se usarán de manera indistinta a lo largo de todo el libro.



- a. Un punto en el plano.
- b. Una pareja ordenada de números reales.
- c. Un vector que va del origen al punto.

Nótese que si bien en este texto las palabras *punto* y *vector* son equivalentes, tienen diferentes connotaciones. Mientras que un punto es pensado como un lugar (una posición) en el espacio, un vector es pensado como un segmento de línea dirigido de un punto a otro.

El conjunto de todas las parejas ordenadas de números se denota por $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. En este caso el exponente 2 sí se refiere a exponenciación, pero de conjuntos, lo que se conoce ahora como el *producto cartesiano* en honor de Descartes (véase el Apéndice A); aunque la usanza común es leerlo “ \mathbb{R} dos”.

1.2.1 Geometría analítica

La idea genialmente simple $\mathbb{R}^2 = \mathbb{E}^2$ (es decir, las parejas ordenadas de números reales se identifican naturalmente con los puntos del plano euclidiano) logra que converjan las aguas del álgebra y la geometría. A partir de Descartes se abre la posibilidad de reconstruir la geometría de los griegos sobre la base de nuestra intuición numérica (básicamente la de sumar y multiplicar). Y a su vez, la luz de la geometría baña con significados y problemas al álgebra. Se hermanan, se apoyan, se entrelazan: nace la geometría analítica.

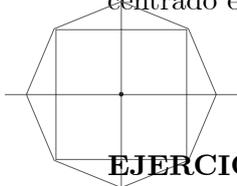
De aquí en adelante, salvo en contadas ocasiones donde sea inevitable y como comentarios, abandonaremos la línea del desarrollo histórico. De hecho, nuestro tratamiento algebraico está muy lejos del original de Descartes, pues usa ideas desarrolladas en el siglo XIX. Pero antes de entrar de lleno a él, vale la pena enfatizar que no estamos abandonando el método axiomático iniciado por Euclides, simplemente cambiaremos de conjunto de axiomas: ahora nos basaremos en los axiomas de los números reales (ver la sección 1.3.1). De tal manera que los objetos primarios de Euclides (línea, segmento, distancia, ángulo, etc.) serán definiciones (por ejemplo, *punto* ya es “pareja ordenada de números” y *plano* es el conjunto de todos los puntos), y los cinco postulados serán teoremas que hay que demostrar. Nuestros objetos primarios básicos serán ahora los números reales y nuestra intuición geométrica primordial es que forman una recta: La Recta Real, \mathbb{R} . Sin embargo, la motivación básica para las definiciones sigue siendo la intuición geométrica desarrollada por los griegos. No queremos hacernos los occisos como si no conociéramos la geometría

griega; simplemente la llevaremos a nuevos horizontes con el apoyo del álgebra —el enfoque analítico— y para ello tendremos que reconstruirla.

EJERCICIO 1.3 Encuentra diferentes puntos en el plano dados por sus coordenadas (por ejemplo, el $(3, 2)$, el $(-1, 5)$, el $(0, -4)$, el $(-1, -2)$, el $(1/2, -3/2)$).

EJERCICIO 1.4 Identifica los *cuadrantes* donde las parejas tienen signos determinados.

EJERCICIO 1.5 ¿Cuáles son las coordenadas de los vértices de un cuadrado de lado 2, centrado en el origen y con lados paralelos a los ejes?



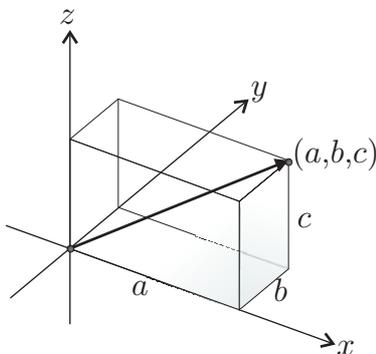
EJERCICIO 1.6 ¿Cuáles son las coordenadas de los vértices del octógono regular que incluye como vértices a los del cuadrado anterior? (*Tienes que usar el Teorema de Pitágoras.*)

EJERCICIO 1.7 ¿Puedes dar las coordenadas de los vértices de un triángulo equilátero centrado en el origen? Dibújalo. (*Usa de nuevo el Teorema de Pitágoras, en un triángulo rectángulo con hipotenusa 2 y un cateto 1.*)

1.2.2 El espacio de dimensión n

La posibilidad de extender con paso firme la geometría euclidiana a más de dos dimensiones es una de las aportaciones de mayor profundidad del método cartesiano.

Aunque nos hayamos puesto formales en la sección anterior para demostrar una correspondencia natural entre \mathbb{E}^2 y \mathbb{R}^2 , intuitivamente es muy simple. Al fijar los



dos ejes coordenados —las dos direcciones preferidas— se llega de manera inequívoca a cualquier punto en el plano dando las distancias que hay que recorrer en esas direcciones para llegar a él. En el espacio, habrá que fijar tres ejes coordenados y dar tres números. Los dos primeros dan un punto en el plano (que podemos pensar horizontal, como el piso), y el tercero nos da la altura (que puede ser positiva o negativa). Si denotamos por \mathbb{R}^3 al conjunto de todas las ternas (x, y, z) de números reales, como éstas corresponden a puntos en el espacio una vez que se fijan los tres ejes (el eje vertical se conoce como *eje z*), podemos definir el espacio

euclidiano de tres dimensiones como \mathbb{R}^3 , sin preocuparnos de axiomas. ¿Y por qué pararse en 3? ¿o en 4?

Dado un número natural $n \in \{1, 2, 3, \dots\}$, denotamos por \mathbb{R}^n al conjunto de todas las n -adas (léase “eneadas”) ordenadas de números reales (x_1, x_2, \dots, x_n) . Formalmente,

$$\mathbb{R}^n := \{(x_1, x_2, \dots, x_n) \mid x_i \in \mathbb{R}, i = 1, 2, \dots, n\}.$$

Para valores pequeños de n , se pueden usar letras x, y, z e inclusive w para denotar las coordenadas; pero para n general esto es imposible y no nos queda más que

usar los subíndices. Así, tenemos un conjunto bien definido, \mathbb{R}^n , al que podemos llamar “*espacio euclidiano de dimensión n*” y hacer geometría en él. A una n -ada $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ se le llama *vector* o *punto*.

El estudiante que se sienta incómodo con esto de “muchas dimensiones”, puede sustituir (pensar en) un 2 o un 3 siempre que se use n y referirse al plano o al espacio tridimensional que habitamos para su intuición. No pretendemos en este libro estudiar la geometría de estos espacios de muchas dimensiones. Nos concentraremos en dimensiones 2 y 3, así que puede pensarse la n como algo que puede tomar valores 2 o 3 y que resulta muy útil para hablar de los dos al mismo tiempo pero en singular. Sin embargo, es importante que el estudiante tenga una visión amplia y pueda preguntarse en cada momento ¿qué pasaría en 4 o más dimensiones? Como verá, y a veces señalaremos explícitamente, algunas cosas son muy fáciles de generalizar, y otras no tanto. Por el momento, basta con volver a enfatizar el amplísimo mundo que abrió el método de las coordenadas cartesianas para las matemáticas y la ciencia.

Vale la pena en este momento puntualizar las convenciones de notación que utilizaremos en este libro. A los números reales los denotamos por letras simples, por ejemplo x , y , z , o bien a , b y c o bien t , r y s ; e inclusive conviene a veces utilizar letras griegas, α (alfa), β (beta) y γ (gamma), o λ (lambda) y μ (mu) (por alguna razón hemos hecho costumbre de utilizarlas siempre en pequeñas familias). Por su parte, a los puntos o vectores los denotamos por letras en negritas, por ejemplo, \mathbf{p} y \mathbf{q} , para señalar que estamos pensando en su carácter de puntos, o bien, usamos \mathbf{u} , \mathbf{v} , \mathbf{w} o \mathbf{d} (acrónimo elegante que usaremos para “dirección”) y \mathbf{n} (acrónimo para “normal”) para subrayar su papel vectorial. Por supuesto, \mathbf{x} , \mathbf{y} y \mathbf{z} siempre son buenos comodines para vectores (puntos o eneadas) variables o incógnitos, que no deben confundirse con x , y , z , que son variables reales o numéricas.

1.3 El espacio vectorial \mathbb{R}^2

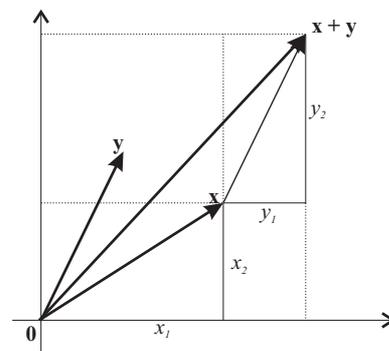
En esta sección se introduce la herramienta algebraica básica para hacer geometría con parejas, ternas o n -adas de números. Y de nuevo la idea central es muy simple: así como los números se pueden sumar y multiplicar, también los vectores tienen sus operacioncitas. Lo único que suponemos de los números reales es que sabemos, o, mejor aún, que “saben ellos” sumarse y multiplicarse; a partir de ello extenderemos estas nociones a los vectores.

Definición 1.3.1 Dados dos vectores $\mathbf{u} = (x, y)$ y $\mathbf{v} = (x', y')$ en \mathbb{R}^2 , definimos su *suma vectorial*, o simplemente su *suma*, como el vector $\mathbf{u} + \mathbf{v}$ que resulta de sumar coordenada a coordenada:

$$\mathbf{u} + \mathbf{v} := (x + x', y + y'),$$

es decir,

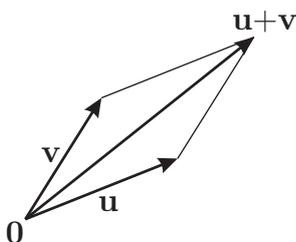
$$(x, y) + (x', y') := (x + x', y + y').$$



Nótese que en cada coordenada, la suma que se usa es la suma usual de números reales. Así que al signo “+” de la suma le hemos ampliado el espectro de “saber sumar números” a “saber sumar vectores”; pero con una receta muy simple: “coordenada a coordenada”. Por ejemplo,

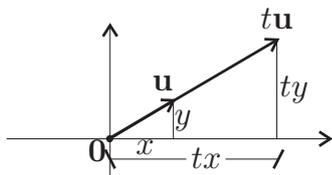
$$(3, 2) + (-1, 1) = (2, 3).$$

La suma vectorial corresponde geoméricamente a la regla del paralelogramo usada para encontrar la resultante de dos vectores. Esto es, se consideran los vectores como segmentos dirigidos que salen del origen, generan entonces un paralelogramo y el vector que va del origen a la otra esquina es la suma. También se puede pensar como la acción de dibujar un vector tras otro, pensando que los vectores son segmentos dirigidos que pueden moverse paralelos a sí mismos.



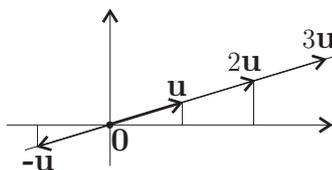
Definición 1.3.2 Dados un vector $\mathbf{u} = (x, y) \in \mathbb{R}^2$ y un número $t \in \mathbb{R}$ se define la *multiplicación escalar* $t\mathbf{u}$ como el vector que resulta de multiplicar cada coordenada del vector por el número:

$$t\mathbf{u} := (tx, ty).$$



Nótese que en cada coordenada, la multiplicación que se usa es la de los números reales.

La multiplicación escalar corresponde a la dilatación o contracción, y posiblemente al cambio de dirección, de un vector. Veamos. Está claro que



$$2\mathbf{u} = (2x, 2y) = (x + x, y + y) = \mathbf{u} + \mathbf{u},$$

así que $2\mathbf{u}$ es el vector “ \mathbf{u} seguido de \mathbf{u} ” o bien, “ \mathbf{u} dilatado a su doble”. De la misma manera que $3\mathbf{u}$ es un vector que apunta en la misma dirección pero de tres veces su tamaño. O bien, es fácil deducir que $(\frac{1}{2})\mathbf{u}$, como punto, está justo a la mitad del camino del origen $\mathbf{0} = (0, 0)$ a \mathbf{u} . Así que $t\mathbf{u}$ para $t > 1$ es, estrictamente hablando, una *dilatación* de \mathbf{u} , y para $0 < t < 1$, una *contracción* del mismo. Por último, para $t < 0$, $t\mathbf{u}$ apunta en la dirección contraria ya que, en particular, $(-1)\mathbf{u} =: -\mathbf{u}$ es el vector que, como segmento dirigido, va del punto \mathbf{u} al $\mathbf{0}$ (puesto que $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$) y el resto se obtiene como dilataciones o contracciones de $-\mathbf{u}$.

No está de más insistir en una diferencia esencial entre las dos operaciones que hemos definido. Si bien la suma vectorial es una operación que de dos ejemplares de la misma especie (vectores) nos da otro de ellos, la multiplicación escalar involucra a dos objetos de diferente índole, por un lado el “escalar”, un número real, y por el otro un vector, lo que da como resultado un nuevo vector. Los vectores no se multiplican (por el momento), sólo los escalares (los números reales) saben multiplicarlos (pegarles, podría decirse) y les cambian con ello su “escala”.

EJERCICIO 1.8 Sean $\mathbf{v}_1 = (2, 3)$, $\mathbf{v}_2 = (-1, 2)$, $\mathbf{v}_3 = (3, -1)$ y $\mathbf{v}_4 = (1, -4)$.

i) Calcula y dibuja: $2\mathbf{v}_1 - 3\mathbf{v}_2$; $2(\mathbf{v}_3 - \mathbf{v}_4) - \mathbf{v}_3 + 2\mathbf{v}_4$; $2\mathbf{v}_1 - 3\mathbf{v}_3 + 2\mathbf{v}_4$.

ii) ¿Qué vector $\mathbf{x} \in \mathbb{R}^2$ cumple que $2\mathbf{v}_1 + \mathbf{x} = 3\mathbf{v}_2$; $3\mathbf{v}_3 - 2\mathbf{x} = \mathbf{v}_4 + \mathbf{x}$?

iii) ¿Puedes encontrar $r, s \in \mathbb{R}$ tales que $r\mathbf{v}_2 + s\mathbf{v}_3 = \mathbf{v}_4$?

EJERCICIO 1.9 Dibuja el origen y tres vectores cualesquiera \mathbf{u} , \mathbf{v} y \mathbf{w} en un papel. Con un par de escuadras encuentra los vectores $\mathbf{u} + \mathbf{v}$, $\mathbf{v} + \mathbf{w}$ y $\mathbf{w} + \mathbf{u}$.

EJERCICIO 1.10 Dibuja el origen y dos vectores cualesquiera \mathbf{u} , \mathbf{v} en un papel. Con regla (escuadras) y compás, encuentra los puntos $(1/2)\mathbf{u} + (1/2)\mathbf{v}$, $2\mathbf{u} - \mathbf{v}$, $3\mathbf{u} - 2\mathbf{v}$, $2\mathbf{v} - \mathbf{u}$. ¿Resultan colineales?

EJERCICIO 1.11 Describe el subconjunto de \mathbb{R}^2 , o *lugar geométrico*, definido por $\mathcal{L}_v := \{t\mathbf{v} \mid t \in \mathbb{R}\}$.

Estas definiciones se extienden a \mathbb{R}^n de manera natural. Dados dos vectores $\mathbf{x} = (x_1, \dots, x_n)$ y $\mathbf{y} = (y_1, \dots, y_n)$ en \mathbb{R}^n y un número real $t \in \mathbb{R}$, la *suma vectorial* (o simplemente la *suma*) y el *producto* (o *multiplicación*) *escalar* se definen como sigue:

$$\mathbf{x} + \mathbf{y} := (x_1 + y_1, \dots, x_n + y_n),$$

$$t\mathbf{x} := (tx_1, \dots, tx_n).$$

Es decir, la suma de dos vectores (con el mismo número de coordenadas) se obtiene sumando coordenada a coordenada y el producto por un escalar (un número) se obtiene multiplicando cada coordenada por ese número.

Las propiedades básicas de la suma vectorial y la multiplicación escalar se reúnen en el siguiente teorema, donde el vector $\mathbf{0} = (0, \dots, 0)$ es llamado *vector cero* que corresponde al *origen*, y, para cada $\mathbf{x} \in \mathbb{R}^n$, el vector $-\mathbf{x} := (-1)\mathbf{x}$ se llama *inverso aditivo* de \mathbf{x} .

Teorema 1.2 Para todos los vectores \mathbf{x} , \mathbf{y} , $\mathbf{z} \in \mathbb{R}^n$ y para todos los números $s, t \in \mathbb{R}$ se cumple que:

- i) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
- ii) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
- iii) $\mathbf{x} + \mathbf{0} = \mathbf{x}$
- iv) $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$
- v) $s(t\mathbf{x}) = (st)\mathbf{x}$
- vi) $1\mathbf{x} = \mathbf{x}$
- vii) $t(\mathbf{x} + \mathbf{y}) = t\mathbf{x} + t\mathbf{y}$
- viii) $(s + t)\mathbf{x} = s\mathbf{x} + t\mathbf{x}$

1.3.1 ¿Teorema o axiomas?

Antes de pensar en demostrar el teorema anterior, vale la pena reflexionar un poco sobre su carácter pues está muy cerca de ser un conjunto de axiomas y es sutil qué

quiere decir eso de demostrarlo.

Primero, debemos observar que \mathbb{R}^n tiene sentido para $n = 1$, y es simplemente una manera rimbombante de referirse a los números reales. Así que del teorema, siendo $n = 1$ (es decir, si \mathbf{x} , \mathbf{y} , \mathbf{z} están en \mathbb{R}), se obtiene parte de los *axiomas de los números reales*. Esto es, cada enunciado es una de las reglas elementales de la suma y la multiplicación que conocemos desde pequeños. Las propiedades **(i)** a **(iv)** son los axiomas que hacen a \mathbb{R} , con la operación suma, lo que se conoce como un “*grupo conmutativo*”: **(i)** dice que la suma es “*asociativa*”, **(ii)** que es “*conmutativa*”, **(iii)** que el $\mathbf{0}$ —o bien, el número 0 — es su “*neutro*” (aditivo) y **(iv)** que todo elemento tiene “*inverso*” (aditivo). Por su parte, **(v)** y **(vi)** nos dicen que la multiplicación es asociativa y que tiene un neutro (multiplicativo), el 1 ; pero nos faltaría que es conmutativa y que tiene inversos (multiplicativos). Entonces tendríamos que añadir:

$$\begin{array}{l} \mathbf{ix)} \quad \mathbf{t s = s t} \\ \mathbf{x)} \quad \text{Si } \mathbf{t} \neq \mathbf{0}, \text{ existe } \mathbf{t}^{-1} \text{ tal que } \mathbf{t t}^{-1} = \mathbf{1} \end{array}$$

para obtener que $\mathbb{R} - \{0\}$ (los reales sin el 0) son un grupo conmutativo con la multiplicación. Finalmente, **(vii)** y **(viii)** ya dicen lo mismo (en virtud de **(ix)**), que las dos operaciones se *distribuyen*. Estos axiomas definen lo que se llama un *campo*.

Obsérvese que en el caso general ($n > 1$) **(ix)** y **(x)** ni siquiera tienen sentido; pues sólo cuando $n = 1$ la multiplicación escalar involucra a seres de la misma especie. En resumen, para el caso $n = 1$, el teorema es un subconjunto de los axiomas que definen las operaciones de los números reales. No hay nada que demostrar, pues son parte de los axiomas básicos que vamos a usar. El resto de los axiomas que definen los números reales se refieren a su orden y, para que el libro quede autocontenido, de una vez los enunciamos. Los números reales tienen una *relación de orden*, denotada por \leq y que se lee “*menor o igual que*” que cumple, para $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbb{R}$, con:

$$\begin{array}{l} \mathbf{Oi)} \quad \mathbf{a \leq b} \quad \text{o} \quad \mathbf{b \leq a} \quad (\text{es un orden total}) \\ \mathbf{Oii)} \quad \mathbf{a \leq b} \quad \text{y} \quad \mathbf{b \leq a} \quad \Leftrightarrow \quad \mathbf{a = b} \\ \mathbf{Oiii)} \quad \mathbf{a \leq b} \quad \text{y} \quad \mathbf{b \leq c} \quad \Rightarrow \quad \mathbf{a \leq c} \quad (\text{es transitivo}) \\ \mathbf{Oiv)} \quad \mathbf{a \leq b} \quad \text{y} \quad \mathbf{c \leq d} \quad \Rightarrow \quad \mathbf{a + c \leq b + d} \\ \mathbf{Ov)} \quad \mathbf{a \leq b} \quad \text{y} \quad \mathbf{0 \leq c} \quad \Rightarrow \quad \mathbf{ac \leq bc} \end{array}$$

Además, los números reales cumplen con el *axioma del supremo* que, intuitivamente, dice que la recta numérica no tiene “hoyos”, que los números reales forman un continuo. Pero este axioma, fundamental para el cálculo pues hace posible formalizar lo “infinitesimal”, no se usa en este libro, así que lo dejamos de lado.

Regresando al Teorema 1.2, para $n > 1$ la cosa es sutilmente diferente. Nosotros definimos la suma vectorial, y al ser algo nuevo sí tenemos que verificar que cumple las propiedades requeridas. Vale la pena introducir a Dios, el de las matemáticas no el de las religiones, para que quede claro. Dios nos da los números reales con la suma y la multiplicación, de alguna manera nos “ilumina” y de golpe: ahí están, con todo y sus axiomas. Ahora, nosotros —simples mortales— osamos definir la suma vectorial

y la multiplicación de escalares a vectores, pero Dios ya no nos asegura nada, él ya hizo lo suyo: nos toca demostrarlo a nosotros.

Demostración. (del Teorema 1.2) Formalmente sólo nos interesa demostrar el teorema para $n = 2, 3$, pero esencialmente es lo mismo para cualquier n . La demostración de cada inciso es muy simple, tanto así que hasta confunde, y consiste en aplicar el axioma correspondiente que cumplen los números reales coordenada a coordenada y la definición de las operaciones involucradas. Sería muy tedioso, y aportaría muy poco al lector, demostrar los ocho incisos, así que sólo demostraremos con detalle el primero para \mathbb{R}^2 , dejando los demás y el caso general como ejercicios.

i) Sean $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^2$, entonces $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2)$ y $\mathbf{z} = (z_1, z_2)$, donde cada x_i, y_i y z_i ($i = 1, 2$) son números reales (nótese que conviene usar subíndices con la misma letra que en negritas denota al vector). Tenemos entonces, a partir de la definición de suma vectorial, que

$$(\mathbf{x} + \mathbf{y}) + \mathbf{z} = (x_1 + y_1, x_2 + y_2) + (z_1, z_2)$$

y usando nuevamente la definición se obtiene que esta última expresión es

$$= ((x_1 + y_1) + z_1, (x_2 + y_2) + z_2).$$

Luego, como *la suma de números reales es asociativa* (el axioma correspondiente de los números reales usado coordenada a coordenada) se sigue

$$= (x_1 + (y_1 + z_1), x_2 + (y_2 + z_2)).$$

Y finalmente, usando dos veces la definición de suma vectorial, se obtiene

$$\begin{aligned} &= (x_1, x_2) + (y_1 + z_1, y_2 + z_2) \\ &= (x_1, x_2) + ((y_1, y_2) + (z_1, z_2)) \\ &= \mathbf{x} + (\mathbf{y} + \mathbf{z}), \end{aligned}$$

lo que demuestra que $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$; y entonces tiene sentido escribir $\mathbf{x} + \mathbf{y} + \mathbf{z}$. \square

Se conoce como *espacio vectorial* a un conjunto en el que están definidas dos operaciones (suma vectorial y multiplicación escalar) que cumplen con las ocho propiedades del Teorema 1.2. De tal manera que este teorema puede parafrasearse “ \mathbb{R}^n es un espacio vectorial”. ¿Podría el lector mencionar un espacio vectorial distinto de \mathbb{R}^n ?

EJERCICIO 1.12 Usando los axiomas de los números reales, así como las definiciones de suma vectorial y producto escalar, demuestra algunos incisos del Teorema 1.2 para el caso $n = 2$ y $n = 3$. ¿Verdad que lo único que cambia es la longitud de los vectores, pero los argumentos son exactamente los mismos? Demuestra (aunque sea mentalmente) el caso general.

EJERCICIO 1.13 Demuestra que si $a \leq 0$ entonces $0 \leq -a$. (Usa el axioma **(Oiv)**.)

EJERCICIO 1.14 Observa que los axiomas de orden no dicen nada sobre cómo están relacionados el 0 y el 1, pero se pueden deducir de ellos. Supón que $1 \leq 0$, usando el ejercicio anterior y los axiomas **(Ov)** y **(Oii)**, demuestra que entonces $-1 = 0$; como esto no es cierto se debe cumplir la otra posibilidad por **(Oi)**, es decir, que $0 \leq 1$.

EJERCICIO 1.15 Demuestra que si $a \leq b$ y $c \leq 0$ entonces $ac \geq bc$ (donde \geq significa “es mayor o igual”).

Ejemplo. Como corolario de este teorema se tiene que el “álgebra” simple a la que estamos acostumbrados con números también vale con los vectores. Por ejemplo, en el ejercicio (1.8.ii.b) se pedía encontrar el vector \mathbf{x} tal que

$$3\mathbf{v}_3 - 2\mathbf{x} = \mathbf{v}_4 + \mathbf{x}.$$

Hagámoslo. Se vale pasar con signo menos del otro lado de la ecuación (sumando el inverso aditivo a ambos lados), y por tanto esta ecuación equivale a

$$-3\mathbf{x} = \mathbf{v}_4 - 3\mathbf{v}_3$$

y multiplicando por $-1/3$ se obtiene que

$$\mathbf{x} = \mathbf{v}_3 - (1/3)\mathbf{v}_4.$$

Ya nada más falta sustituir por los valores dados. Es probable que el estudiante haya hecho algo similar y qué bueno: tenía la intuición correcta. Pero hay que tener cuidado, así como con los números no se puede dividir entre cero, no se le vaya a ocurrir tratar de ¡dividir entre un vector! Aunque a veces se pueden “cancelar” (ver Ej. 1.17). Para tal efecto, el siguiente lema será muy usado y vale la pena verlo en detalle.

Lema 1.3 Si $\mathbf{x} \in \mathbb{R}^2$ y $t \in \mathbb{R}$ son tales que $t\mathbf{x} = \mathbf{0}$ entonces $t = 0$ o $\mathbf{x} = \mathbf{0}$.

Demostración. Suponiendo que $t \neq 0$, hay que demostrar que $\mathbf{x} = \mathbf{0}$ para completar el lema. Pero entonces t tiene inverso multiplicativo y podemos multiplicar por t^{-1} ambos lados de la ecuación $t\mathbf{x} = \mathbf{0}$ para obtener (usando **(v)** y **(vi)**) que $\mathbf{x} = (t^{-1})\mathbf{0} = \mathbf{0}$ por la definición del vector nulo $\mathbf{0} = (0, 0)$. \square

EJERCICIO 1.16 Demuestra que si $\mathbf{x} \in \mathbb{R}^3$ y $t \in \mathbb{R}$ son tales que $t\mathbf{x} = \mathbf{0}$ entonces $t = 0$ o $\mathbf{x} = \mathbf{0}$. ¿Y para \mathbb{R}^n ?

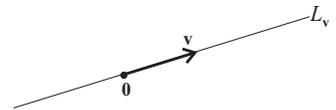
EJERCICIO 1.17 Demuestra que si $\mathbf{x} \in \mathbb{R}^n$ es distinto de $\mathbf{0}$, y $t, s \in \mathbb{R}$ son tales que $t\mathbf{x} = s\mathbf{x}$, entonces $t = s$. (Es decir, si $\mathbf{x} \neq \mathbf{0}$ está permitido “cancelarlo” aunque sea vector.)

1.4 Líneas

En el estudio de la geometría clásica, las líneas o rectas son subconjuntos básicos descritos por los axiomas; se les reconoce intuitivamente y se parte de ellas para construir lo demás. Con el método cartesiano, esto no es necesario. Las líneas se pueden definir o construir, correspondiendo a nuestra noción intuitiva de ellas —que no varía en nada de la de los griegos—, para después ver que efectivamente cumplen con los axiomas que antes se les asociaban. En esta sección, definiremos las rectas y veremos que cumplen los axiomas primero y tercero de Euclides.

Nuestra definición de recta estará basada en la intuición física de una partícula que viaja en movimiento rectilíneo uniforme; es decir, con velocidad fija y en la misma dirección. Estas dos nociones (velocidad como magnitud y dirección) se han amalgamado en el concepto de vector.

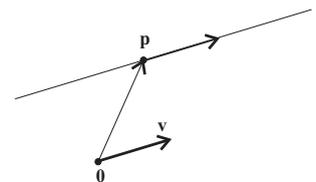
Recordemos que en el Ejercicio 1.11, se pide describir el conjunto $\mathcal{L}_{\mathbf{v}} := \{t\mathbf{v} \in \mathbb{R}^2 \mid t \in \mathbb{R}\}$ donde $\mathbf{v} \in \mathbb{R}^2$. Debía ser claro que si $\mathbf{v} \neq \mathbf{0}$, entonces $\mathcal{L}_{\mathbf{v}}$, al constar de todos los múltiplos escalares del vector \mathbf{v} , se dibuja como una recta que pasa por el origen (pues $\mathbf{0} = 0\mathbf{v}$) con la dirección de \mathbf{v} . Si consideramos la variable t como el tiempo, la función $\varphi(t) = t\mathbf{v}$ describe el movimiento rectilíneo uniforme de una partícula que viene desde tiempo infinito negativo, pasa por el origen $\mathbf{0}$ en el tiempo $t = 0$ (llamada la posición inicial), y seguirá por siempre con velocidad constante \mathbf{v} .



Pero por el momento queremos pensar en las rectas como conjuntos (en el capítulo siguiente estudiaremos más a profundidad las funciones). Los conjuntos $\mathcal{L}_{\mathbf{v}}$, al rotar \mathbf{v} , nos dan las rectas por el origen, y para obtener otras, bastará “empujarlas” fuera del origen (o bien, arrancar el movimiento con otra posición inicial).

Definición 1.4.1 Dados un punto \mathbf{p} y un vector $\mathbf{v} \neq \mathbf{0}$, la *recta que pasa por \mathbf{p}* con dirección \mathbf{v} es el conjunto:

$$\ell := \{\mathbf{p} + t\mathbf{v} \mid t \in \mathbb{R}\}. \quad (1.1)$$



Una *recta* o *línea* en \mathbb{R}^2 es un subconjunto que tiene, para algún \mathbf{p} y $\mathbf{v} \neq \mathbf{0}$, la descripción anterior.

A esta forma de definir una recta se le conoce como *representación o expresión paramétrica*. Esta representación trae consigo una función entre los números reales y la recta:

$$\begin{aligned} \varphi : \mathbb{R} &\rightarrow \mathbb{R}^2 \\ \varphi(t) &= \mathbf{p} + t\mathbf{v}. \end{aligned} \quad (1.2)$$

De hecho, esta función define una biyección entre \mathbb{R} y ℓ . Como ℓ se definió por medio del *parámetro* $t \in \mathbb{R}$, es claro que *es* la imagen de la función φ ; es decir, φ es sobre. Demostremos (aunque de la intuición física parezca obvio) que es uno-a-uno. Supongamos para esto que $\varphi(t) = \varphi(s)$ para algunos $t, s \in \mathbb{R}$ y habrá que concluir

que $t = s$. Por la definición de φ se tiene

$$\mathbf{p} + t\mathbf{v} = \mathbf{p} + s\mathbf{v}.$$

De aquí, al sumar el inverso del lado izquierdo a ambos (que equivale a pasarlo de un lado al otro con signo contrario) se tiene

$$\begin{aligned} \mathbf{0} &= \mathbf{p} + s\mathbf{v} - (\mathbf{p} + t\mathbf{v}) \\ &= \mathbf{p} + s\mathbf{v} - \mathbf{p} - t\mathbf{v} \\ &= (\mathbf{p} - \mathbf{p}) + (s\mathbf{v} - t\mathbf{v}) \\ &= \mathbf{0} + (s - t)\mathbf{v} \\ &= (s - t)\mathbf{v}, \end{aligned}$$

donde hemos usado las propiedades (i), (ii), (iii), (iv) y (viii) del Teorema 1.2.² Del Lema 1.3, se concluye que $s - t = 0$ o bien que $\mathbf{v} = \mathbf{0}$. Pero por hipótesis $\mathbf{v} \neq \mathbf{0}$ (esto es la esencia), así que no queda otra que $s = t$. Esto demuestra que φ es inyectiva, y nuestra intuición se corrobora.

Hemos demostrado que cualquier recta está en biyección natural con \mathbb{R} , la recta modelo, así que todas las líneas son una “copia” de la recta real abstracta \mathbb{R} .

Observación 1.4.1 En la definición 1.4, así como en la demostración, no se usa de manera esencial que \mathbf{p} y \mathbf{v} estén en \mathbb{R}^2 . Sólo se usa que hay una suma vectorial y un producto escalar bien definidos. Así que el punto y el vector podrían estar en \mathbb{R}^3 o en \mathbb{R}^n . Podemos entonces dar por establecida la noción de recta en cualquier dimensión al cambiar 2 por n en la Definición 1.4.

Observación 1.4.2 A la función 1.2 se le conoce como “movimiento rectilíneo uniforme” o “movimiento inercial”, pues, como decíamos al principio de la sección, describe la manera en que se mueve una partícula, o un cuerpo, que no está afectada por ninguna fuerza y tiene posición \mathbf{p} y vector velocidad \mathbf{v} en el tiempo $t = 0$.

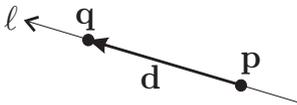
Ya podemos demostrar un resultado clásico, e intuitivamente claro.

Lema 1.4 *Dados dos puntos \mathbf{p} y \mathbf{q} en \mathbb{R}^n , existe una recta que pasa por ellos.*

Demostración. Si tuviéramos que $\mathbf{p} = \mathbf{q}$, como hay muchas rectas que pasan por \mathbf{p} , ya acabamos. Supongamos entonces que $\mathbf{p} \neq \mathbf{q}$, que es el caso interesante.

Si tomamos a \mathbf{p} como punto base para la recta que buscamos, bastará encontrar un vector que nos lleve de \mathbf{p} a \mathbf{q} , para tomarlo como dirección. Éste es la diferencia $\mathbf{q} - \mathbf{p}$, pues claramente

$$\mathbf{p} + (\mathbf{q} - \mathbf{p}) = \mathbf{q},$$



²Ésta es la última vez que haremos mención explícita del uso del Teorema 1.2; de aquí en adelante se aplicarán sus propiedades como si fueran lo más natural del mundo. Pero vale la pena que el estudiante se cuestione por algún tiempo qué es lo que se está usando en las manipulaciones algebraicas.

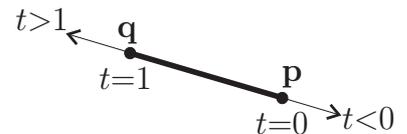
de tal manera que si definimos $\mathbf{d} := \mathbf{q} - \mathbf{p}$, como dirección, la recta

$$\ell = \{\mathbf{p} + t\mathbf{d} \mid t \in \mathbb{R}\},$$

(que sí es una recta pues $\mathbf{d} = \mathbf{q} - \mathbf{p} \neq \mathbf{0}$), es la que funciona. Con $t = 0$ obtenemos que $\mathbf{p} \in \ell$, y con $t = 1$ que $\mathbf{q} \in \ell$. \square

De la demostración del lema, se siguen los postulados I y III de Euclides. Obsérvese que cuando t toma valores entre 0 y 1, se obtienen puntos entre \mathbf{p} y \mathbf{q} (pues el vector direccional $(\mathbf{q} - \mathbf{p})$ se encoge al multiplicarlo por t), así que el *segmento* de \mathbf{p} a \mathbf{q} , que denotaremos por $\overline{\mathbf{p}\mathbf{q}}$, se debe definir como

$$\overline{\mathbf{p}\mathbf{q}} := \{\mathbf{p} + t(\mathbf{q} - \mathbf{p}) \mid 0 \leq t \leq 1\}.$$



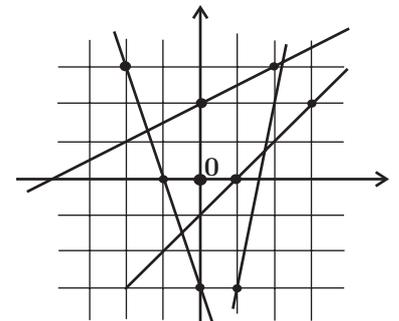
Y la recta ℓ que pasa por \mathbf{p} y \mathbf{q} se extiende “indefinidamente” a ambos lados del segmento $\overline{\mathbf{p}\mathbf{q}}$; para $t > 1$ del lado de \mathbf{q} y para $t < 0$ del lado de \mathbf{p} .

***EJERCICIO 1.18** Aunque lo demostraremos más adelante, es un buen ejercicio demostrar formalmente en este momento que la recta ℓ , cuando $\mathbf{p} \neq \mathbf{q}$, es única.

EJERCICIO 1.19 Encuentra representaciones paramétricas de las rectas en la figura. Observa que su representación paramétrica no es única.

EJERCICIO 1.20 Dibuja las rectas

$$\begin{aligned} &\{(2, 3) + t(1, 1) \mid t \in \mathbb{R}\}; \\ &\{(-1, 0) + s(2, 1) \mid s \in \mathbb{R}\}; \\ &\{(0, -2) + (-r, 2r) \mid r \in \mathbb{R}\}; \\ &\{(t - 1, -2t) \mid t \in \mathbb{R}\}. \end{aligned}$$



EJERCICIO 1.21 Exhibe representaciones paramétricas para las seis rectas que pasan por dos de los siguientes puntos en \mathbb{R}^2 : $\mathbf{p}_1 = (2, 4)$, $\mathbf{p}_2 = (-1, 2)$, $\mathbf{p}_3 = (-1, -1)$ y $\mathbf{p}_4 = (3, -1)$.

EJERCICIO 1.22 Muestra representaciones paramétricas para las tres rectas que genera el triángulo en \mathbb{R}^3 : $\mathbf{q}_1 = (2, 1, 2)$, $\mathbf{q}_2 = (-1, 1, -1)$, $\mathbf{q}_3 = (-1, -2, -1)$.

EJERCICIO 1.23 Si una partícula viaja de $2\mathbf{q}_1$ a $3\mathbf{q}_3$ en movimiento inercial y tarda cuatro unidades de tiempo en llegar. ¿Cuál es su vector velocidad?

EJERCICIO 1.24 Dados dos vectores \mathbf{u} y \mathbf{v} en \mathbb{R}^n , el *paralelogramo* que definen tiene como vértices los puntos $\mathbf{0}$, \mathbf{u} , \mathbf{v} y $\mathbf{u} + \mathbf{v}$ (como en la figura). Demuestra que sus *diagonales*, es decir, los segmentos de $\mathbf{0}$ a $\mathbf{u} + \mathbf{v}$ y de \mathbf{u} a \mathbf{v} se intersectan en su punto medio.

1.4.1 Coordenadas baricéntricas

Todavía podemos exprimirle más información a la demostración del Lema 1.4. Veremos cómo se escriben, en términos de \mathbf{p} y \mathbf{q} , los puntos de su recta, y que esto tiene que ver con la clásica ley de las palancas. Además, de aquí surgirá una demostración muy simple del teorema clásico de concurrencia de medianas en un triángulo.

Supongamos que \mathbf{p} y \mathbf{q} son dos puntos distintos del plano. Para cualquier $t \in \mathbb{R}$, se tiene que

$$\begin{aligned} \mathbf{p} + t(\mathbf{q} - \mathbf{p}) &= \mathbf{p} + t\mathbf{q} - t\mathbf{p} \\ &= (1 - t)\mathbf{p} + t\mathbf{q} \end{aligned}$$

al reagrupar los términos. Y esta última expresión, a su vez, se puede reescribir como

$$s\mathbf{p} + t\mathbf{q} \quad \text{con} \quad s + t = 1,$$

donde hemos introducido la nueva variable $s = 1 - t$. De lo anterior se deduce que la recta ℓ que pasa por \mathbf{p} y \mathbf{q} puede también describirse como

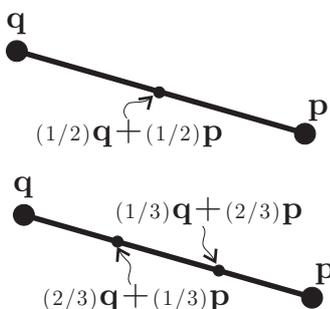
$$\ell = \{s\mathbf{p} + t\mathbf{q} \mid s + t = 1\},$$

que es una *expresión baricéntrica* de ℓ . A los números s, t (con $s + t = 1$, insistimos) se les conoce como *coordenadas baricéntricas* del punto $\mathbf{x} = s\mathbf{p} + t\mathbf{q}$ con respecto a \mathbf{p} y \mathbf{q} .

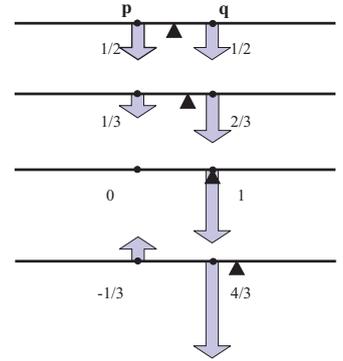
Las coordenadas baricéntricas tienen la ventaja de que ya no distinguen entre los dos puntos. Como lo escribíamos antes —sólo con t —, \mathbf{p} era el punto base (“el que la hace de 0”) y luego \mathbf{q} era hacia donde íbamos (“el que la hace de 1”), y tenían un papel distinto. Ahora no hay preferencia por ninguno de los dos; las coordenadas baricéntricas son “democráticas”. Más aún, al expresar una recta con coordenadas baricéntricas no le damos una dirección preferida. Se usan simultáneamente los parámetros naturales para las dos direcciones (de \mathbf{p} a \mathbf{q} y de \mathbf{q} a \mathbf{p}); pues si $s + t = 1$ entonces

$$\begin{aligned} s\mathbf{p} + t\mathbf{q} &= \mathbf{p} + t(\mathbf{q} - \mathbf{p}) \\ &= \mathbf{q} + s(\mathbf{p} - \mathbf{q}). \end{aligned}$$

Nótese que $\mathbf{x} = s\mathbf{p} + t\mathbf{q}$ está en el segmento entre \mathbf{p} y \mathbf{q} si y sólo si sus dos coordenadas baricéntricas son no negativas, es decir, si y sólo si $t \geq 0$ y $s \geq 0$. Por ejemplo, el punto medio tiene coordenadas baricéntricas $1/2, 1/2$, es $(1/2)\mathbf{p} + (1/2)\mathbf{q}$; y el punto que divide al segmento de \mathbf{p} a \mathbf{q} en la proporción de $2/3$ a $1/3$ se escribe $(1/3)\mathbf{p} + (2/3)\mathbf{q}$, pues se acerca más a \mathbf{q} que a \mathbf{p} . La extensión de la recta más allá de \mathbf{q} tiene coordenadas baricéntricas s, t con $s < 0$ (y por lo tanto $t > 1$); así que los puntos de ℓ fuera del segmento de \mathbf{p} a \mathbf{q} tienen alguna coordenada baricéntrica negativa (la correspondiente al punto más lejano).



En términos físicos, podemos pensar la recta por \mathbf{p} y \mathbf{q} como una barra rígida. Si distribuimos una masa (que podemos pensar unitaria, es decir que vale 1) entre estos dos puntos, el punto de equilibrio tiene las coordenadas baricéntricas correspondientes a las masas. Si, por ejemplo, tienen el mismo peso, su punto de equilibrio está en el punto medio. Las masas negativas pueden pensarse como una fuerza hacia arriba y las correspondientes coordenadas baricéntricas nos dan entonces el punto de equilibrio para las palancas.



Veamos ahora un teorema clásico cuya demostración se simplifica enormemente usando coordenadas baricéntricas. Dado un triángulo con vértices \mathbf{a} , \mathbf{b} , \mathbf{c} , se definen las *medias* como los segmentos que van de un vértice al punto medio del lado opuesto a éste.

Teorema 1.5 Dado un triángulo \mathbf{a} , \mathbf{b} , \mathbf{c} , sus tres medianas “concurrenten” en un punto que las parte en la proporción de 2/3 (del vértice) a 1/3 (del lado opuesto).

Demostración. Como el punto medio del segmento \mathbf{b} , \mathbf{c} es $(\frac{1}{2} \mathbf{b} + \frac{1}{2} \mathbf{c})$, entonces la ...Dibujo mediana por \mathbf{a} es el segmento

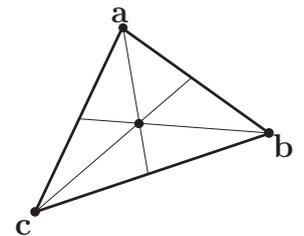
$$\{s \mathbf{a} + t (\frac{1}{2} \mathbf{b} + \frac{1}{2} \mathbf{c}) \mid s + t = 1, s \geq 0, t \geq 0\},$$

y análogamente se describen las otras dos medianas. Por suerte, el enunciado del teorema nos dice dónde buscar la intersección: el punto que describe en la mediana de \mathbf{a} es precisamente

$$\frac{1}{3} \mathbf{a} + \frac{2}{3} (\frac{1}{2} \mathbf{b} + \frac{1}{2} \mathbf{c}) = \frac{1}{3} \mathbf{a} + \frac{1}{3} \mathbf{b} + \frac{1}{3} \mathbf{c}.$$

Entonces, de las igualdades

$$\begin{aligned} \frac{1}{3} \mathbf{a} + \frac{1}{3} \mathbf{b} + \frac{1}{3} \mathbf{c} &= \frac{1}{3} \mathbf{a} + \frac{2}{3} (\frac{1}{2} \mathbf{b} + \frac{1}{2} \mathbf{c}) \\ &= \frac{1}{3} \mathbf{b} + \frac{2}{3} (\frac{1}{2} \mathbf{c} + \frac{1}{2} \mathbf{a}) \\ &= \frac{1}{3} \mathbf{c} + \frac{2}{3} (\frac{1}{2} \mathbf{a} + \frac{1}{2} \mathbf{b}) \end{aligned}$$



se deduce que las tres medianas *concurrenten* en el punto $\frac{1}{3}(\mathbf{a} + \mathbf{b} + \mathbf{c})$, es decir, pasan por él. Este punto se llama el *baricentro* del triángulo, y claramente parte a las medianas en la proporción deseada. De nuevo, el baricentro corresponde al “centro de masa” o “punto de equilibrio” del triángulo.

□

EJERCICIO 1.25 Si tienes una barra rígida de un metro y con una fuerza de 10 kg quieres levantar una masa de 40 kg, ¿de dónde debes colgar la masa, estando el punto de apoyo al extremo de tu barra? Haz un dibujo.

EJERCICIO 1.26 Sean \mathbf{a} , \mathbf{b} y \mathbf{c} tres puntos distintos entre sí. Demuestra que si \mathbf{a} está en la recta que pasa por \mathbf{b} y \mathbf{c} , entonces \mathbf{b} está en la recta por \mathbf{a} y \mathbf{c} .

EJERCICIO 1.27 Encuentra el baricentro del triángulo $\mathbf{a} = (2, 4)$, $\mathbf{b} = (-1, 2)$, $\mathbf{c} = (-1, -1)$. Haz un dibujo.

EJERCICIO 1.28 Observa que en el Teorema 1.5, así como en la discusión que le precede, nunca usamos que estuviéramos en el plano. ¿Podría el lector enunciar y demostrar el teorema análogo para tetraedros en el espacio? (Un tetraedro está determinado por cuatro puntos en el espacio tridimensional. ¿Dónde estará su centro de masa?) ¿Puede generalizarlo a más dimensiones?

1.4.2 Planos en el espacio I

En la demostración del Teorema de las Medianas (1.5) se usó una idea que podemos utilizar para definir planos en el espacio. Dados tres puntos \mathbf{a} , \mathbf{b} y \mathbf{c} en \mathbb{R}^3 (aunque debe observar el lector que nuestra discusión se generaliza a \mathbb{R}^n), ya sabemos describir las tres líneas entre ellos; supongamos que son distintas. Entonces podemos tomar nuevos puntos en alguna de ellas y de estos puntos las nuevas líneas que los unen al vértice restante (como lo hicimos del punto medio de un segmento al vértice opuesto en el teorema). Está claro que la unión de todas estas líneas debe ser el plano que pasa por \mathbf{a} , \mathbf{b} y \mathbf{c} ; ésta es la idea que vamos a desarrollar.

Un punto \mathbf{y} en la línea por \mathbf{a} y \mathbf{b} se escribe como

$$\mathbf{y} = s \mathbf{a} + t \mathbf{b} \quad \text{con} \quad s + t = 1.$$

Y un punto \mathbf{x} en la línea que pasa por \mathbf{y} y \mathbf{c} se escribe entonces como

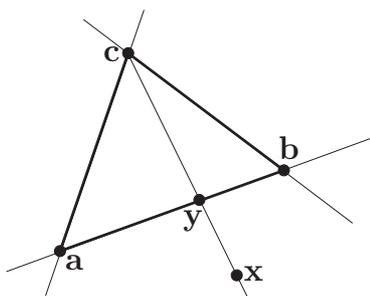
$$\mathbf{x} = r (s \mathbf{a} + t \mathbf{b}) + (1 - r) \mathbf{c},$$

para algún $r \in \mathbb{R}$; que es lo mismo que

$$\mathbf{x} = (rs) \mathbf{a} + (rt) \mathbf{b} + (1 - r) \mathbf{c}.$$

Observemos que los tres coeficientes suman uno:

$$(rs) + (rt) + (1 - r) = r(s + t) + 1 - r = r(1) + 1 - r = 1.$$



Y lo mismo hubiera pasado si en lugar de haber empezado con \mathbf{a} y \mathbf{b} , hubiéramos comenzado con otra de las parejas. De lo que se trata entonces es que cualquier combinación de \mathbf{a} , \mathbf{b} y \mathbf{c} con coeficientes que sumen uno debe estar en su plano. Demostraremos que está en una línea por uno de los vértices y un punto en la línea que pasa por los otros dos.

Sean $\alpha, \beta, \gamma \in \mathbb{R}$ tales que $\alpha + \beta + \gamma = 1$. Consideremos al punto

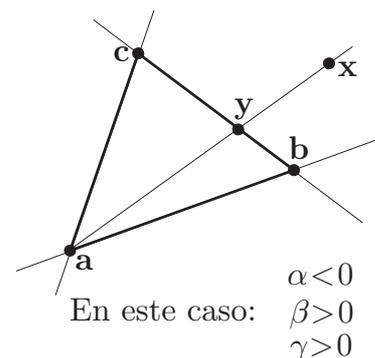
$$\mathbf{x} = \alpha \mathbf{a} + \beta \mathbf{b} + \gamma \mathbf{c}.$$

Como alguno de los coeficientes es distinto de 1 (si no, sumarían 3), podemos suponer sin pérdida de generalidad (esto quiere decir que los otros dos casos son análogos) que $\alpha \neq 1$. Entonces podemos dividir entre $1 - \alpha$ y se tiene

$$\mathbf{x} = \alpha \mathbf{a} + (1 - \alpha) \left(\frac{\beta}{1 - \alpha} \mathbf{b} + \frac{\gamma}{1 - \alpha} \mathbf{c} \right),$$

así que \mathbf{x} está en la recta que pasa por \mathbf{a} y el punto

$$\mathbf{y} = \frac{\beta}{1 - \alpha} \mathbf{b} + \frac{\gamma}{1 - \alpha} \mathbf{c}.$$



Como $\alpha + \beta + \gamma = 1$, entonces $\beta + \gamma = 1 - \alpha$, y los coeficientes de esta última expresión suman uno. Por lo tanto \mathbf{y} está en la recta que pasa por \mathbf{b} y \mathbf{c} . Hemos argumentado la siguiente definición:

Definición 1.4.2 Dados tres puntos \mathbf{a} , \mathbf{b} y \mathbf{c} en \mathbb{R}^3 no colineales (es decir, que no estén en una misma línea), el *plano* que pasa por ellos es el conjunto

$$\Pi = \{ \alpha \mathbf{a} + \beta \mathbf{b} + \gamma \mathbf{c} \mid \alpha, \beta, \gamma \in \mathbb{R}; \alpha + \beta + \gamma = 1 \}.$$

A una expresión de la forma $\alpha \mathbf{a} + \beta \mathbf{b} + \gamma \mathbf{c}$ con $\alpha + \beta + \gamma = 1$ la llamaremos *combinación afín* (o *baricéntrica*) de los puntos \mathbf{a} , \mathbf{b} , \mathbf{c} y a los coeficientes α , β , γ las *coordenadas baricéntricas* del punto $\alpha \mathbf{a} + \beta \mathbf{b} + \gamma \mathbf{c}$.

EJERCICIO 1.29 Considera tres puntos \mathbf{a} , \mathbf{b} y \mathbf{c} no colineales, y sean α , β y γ las correspondientes coordenadas baricéntricas del plano que generan. Observa que cuando una de las coordenadas baricéntricas es cero, entonces el punto correspondiente está en una de las tres rectas por \mathbf{a} , \mathbf{b} y \mathbf{c} . ¿En cuál?, ¿puedes demostrarlo? Haz un dibujo de los tres puntos y sus tres rectas. Éstas parten el plano en pedazos (¿cuántos?); en cada uno de ellos escribe los signos que toman las coordenadas baricéntricas (por ejemplo, en el interior del triángulo se tiene $(+, +, +)$ que corresponden respectivamente a (α, β, γ)).

EJERCICIO 1.30 Dibuja tres puntos \mathbf{a} , \mathbf{b} y \mathbf{c} no colineales. Con regla y compás encuentra los puntos $\mathbf{x} = (1/2)\mathbf{a} + (1/4)\mathbf{b} + (1/4)\mathbf{c}$, $\mathbf{y} = (1/4)\mathbf{a} + (1/2)\mathbf{b} + (1/4)\mathbf{c}$ y $\mathbf{z} = (1/4)\mathbf{a} + (1/4)\mathbf{b} + (1/2)\mathbf{c}$. Describe y argumenta tu construcción. Supón que puedes partir en tres el segmento $\overline{\mathbf{bc}}$ (hazlo midiendo con una regla o a ojo). ¿Puedes construir los puntos $\mathbf{u} = (1/2)\mathbf{a} + (1/3)\mathbf{b} + (1/6)\mathbf{c}$ y $\mathbf{v} = (1/2)\mathbf{a} + (1/6)\mathbf{b} + (1/3)\mathbf{c}$?

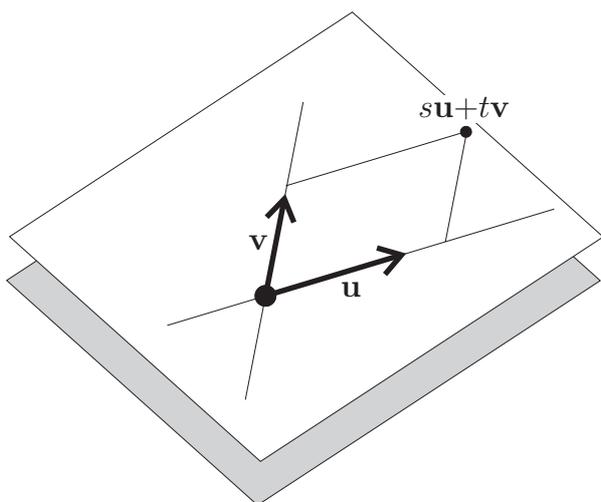
EJERCICIO 1.31 Demuestra que si \mathbf{u} y \mathbf{v} no son colineales con el origen, entonces el conjunto $\{s\mathbf{u} + t\mathbf{v} \mid s, t \in \mathbb{R}\}$ es un plano por el origen. (Usa que $s\mathbf{u} + t\mathbf{v} = r\mathbf{0} + s\mathbf{u} + t\mathbf{v}$ para cualquier $r, s, t \in \mathbb{R}$).

EJERCICIO 1.32 Demuestra que si \mathbf{p} y \mathbf{q} están en el plano Π de la Definición 1.4.2, entonces la recta que pasa por ellos está contenida en Π .

Definición lineal

Emocionado el autor por su demostración del teorema de medianas (1.5) usando coordenadas baricéntricas, se siguió de largo y definió planos en el espacio de una manera quizá no muy intuitiva. Que quede entonces la sección anterior como ejercicio en la manipulación de combinaciones de vectores; si no quedó muy clara puede releerse después de ésta. Pero conviene recapitular para definir los planos de otra manera, para aclarar la definición y así dejarnos la tarea de demostrar una equivalencia de dos definiciones.

La recta que genera un vector $\mathbf{u} \neq \mathbf{0}$ es el conjunto de todos sus “alargamientos” $\mathcal{L}_{\mathbf{u}} = \{s\mathbf{u} \mid s \in \mathbb{R}\}$. Si ahora tomamos un nuevo vector \mathbf{v} que no esté en $\mathcal{L}_{\mathbf{u}}$,



tenemos una nueva recta $\mathcal{L}_{\mathbf{v}} = \{t\mathbf{v} \mid t \in \mathbb{R}\}$ que intersecta a la anterior sólo en el origen. Estas dos rectas generan un plano que consiste en todos los puntos (en \mathbb{R}^3 , digamos, para fijar ideas) a los cuales se puede llegar desde el origen moviéndose únicamente en las direcciones \mathbf{u} y \mathbf{v} . Este plano que pasa por el origen claramente se describe con dos parámetros independientes:

$$\Pi_0 = \{s\mathbf{u} + t\mathbf{v} \mid s, t \in \mathbb{R}\},$$

que por su propia definición está hecho a imagen y semejanza del plano \mathbb{R}^2 (ver Ejercicio 1.33, adelante). Como lo hicimos con

las rectas, podemos definir ahora un plano cualquiera como un plano por el origen (el que acabamos de describir) trasladado por cualquier otro vector constante.

Definición 1.4.3 Un *plano* en \mathbb{R}^3 es un conjunto de la forma

$$\Pi = \{\mathbf{p} + s\mathbf{u} + t\mathbf{v} \mid s, t \in \mathbb{R}\},$$

donde \mathbf{u} y \mathbf{v} son vectores no nulos tales que $\mathcal{L}_{\mathbf{u}} \cap \mathcal{L}_{\mathbf{v}} = \{\mathbf{0}\}$ y \mathbf{p} es cualquier punto. A esta manera de describir un plano la llamaremos *expresión paramétrica*; a \mathbf{u} y \mathbf{v} se les llama *vectores direccionales* del plano Π y a \mathbf{p} el *punto base* de la expresión paramétrica.

Para no confundirnos entre las dos definiciones de plano que hemos dado, llamemos *plano afín* a los que definimos en la sección anterior. Y ahora demostremos que las dos definiciones coinciden.

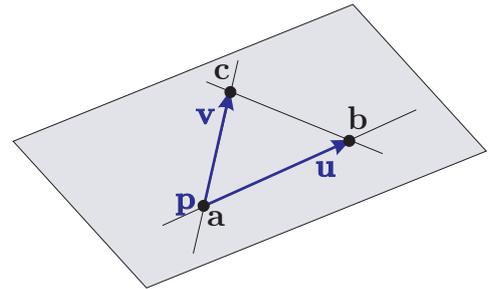
Lema 1.6 *En \mathbb{R}^3 , todo plano es un plano afín y viceversa.*

Demostración. Sea Π como en la definición precedente. Debemos encontrar tres puntos en él y ver que todos los elementos de Π se expresan como combinación baricéntrica de ellos. Los puntos más obvios son

$$\mathbf{a} := \mathbf{p}; \quad \mathbf{b} := \mathbf{p} + \mathbf{u}; \quad \mathbf{c} := \mathbf{p} + \mathbf{v},$$

que claramente están en Π . Obsérvese que entonces se tiene que $\mathbf{u} = \mathbf{b} - \mathbf{a}$ y $\mathbf{v} = \mathbf{c} - \mathbf{a}$, de tal manera que para cualquier $s, t \in \mathbb{R}$ tenemos que

$$\begin{aligned} \mathbf{p} + s\mathbf{u} + t\mathbf{v} &= \mathbf{a} + s(\mathbf{b} - \mathbf{a}) + t(\mathbf{c} - \mathbf{a}) \\ &= \mathbf{a} + s\mathbf{b} - s\mathbf{a} + t\mathbf{c} - t\mathbf{a} \\ &= (1 - s - t)\mathbf{a} + s\mathbf{b} + t\mathbf{c}. \end{aligned}$$



Puesto que los coeficientes de esta última expresión suman 1, esto demuestra que Π está contenido en el plano afín generado por $\mathbf{a}, \mathbf{b}, \mathbf{c}$. E inversamente, cualquier combinación afín de $\mathbf{a}, \mathbf{b}, \mathbf{c}$ tiene esta última expresión, y por la misma igualdad se ve que está en Π . Obsérvese finalmente que si nos dan tres puntos $\mathbf{a}, \mathbf{b}, \mathbf{c}$, se pueden tomar como punto base a \mathbf{a} y como vectores direccionales a $(\mathbf{b} - \mathbf{a})$ y $(\mathbf{c} - \mathbf{a})$ para expresar paramétricamente el plano afín que generan, pues las igualdades anteriores se siguen cumpliendo.

Ver que las condiciones que pusimos en ambas definiciones coinciden se deja como ejercicio. \square

Antes de seguir adelante, conviene establecer cierta **terminología** y **notación** para cosas, nociones y expresiones que estamos usando mucho:

- Dados los vectores $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ en \mathbb{R}^n (para incluir nuestros casos de interés $n = 2, 3$ de una buena vez), a una expresión de la forma

$$s_1\mathbf{u}_1 + s_2\mathbf{u}_2 + \dots + s_k\mathbf{u}_k,$$

donde s_1, s_2, \dots, s_k son números reales (escalares), se le llama una *combinación lineal* de los vectores $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ con *coeficientes* s_1, s_2, \dots, s_k . Obsérvese que toda combinación lineal da como resultado un vector, pero que un mismo vector tiene muchas expresiones tales.

- Como ya vimos, a una combinación lineal cuyos coeficientes suman 1 se le llama *combinación afín*. Y a una combinación afín de dos vectores distintos o de tres no colineales se le llama, además, *baricéntrica*.
- Al conjunto de *todas* las combinaciones lineales de los vectores $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ se le llama el *subespacio generado* por ellos y se le denotará $\langle \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k \rangle$. Es decir,

$$\langle \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k \rangle := \{s_1 \mathbf{u}_1 + s_2 \mathbf{u}_2 + \dots + s_k \mathbf{u}_k \mid s_1, s_2, \dots, s_k \in \mathbb{R}\}.$$

Nótese que entonces, si $\mathbf{u} \neq \mathbf{0}$ se tiene que $\mathcal{L}_{\mathbf{u}} = \langle \mathbf{u} \rangle$ es la recta generada por \mathbf{u} ; y ambas notaciones se seguirán usando indistintamente. Aunque ahora $\langle \mathbf{u} \rangle$ tiene sentido para $\mathbf{u} = \mathbf{0}$, en cuyo caso $\langle \mathbf{u} \rangle = \{\mathbf{0}\}$, y $\mathcal{L}_{\mathbf{u}}$ se usará para destacar que *es* una recta.

- Se dice que dos vectores \mathbf{u} y \mathbf{v} son *linealmente independientes* si son no nulos y tales que $\mathcal{L}_{\mathbf{u}} \cap \mathcal{L}_{\mathbf{v}} = \{\mathbf{0}\}$.
- Dado cualquier subconjunto $A \subset \mathbb{R}^n$, su *traslado* por el vector (o al punto) \mathbf{p} es el conjunto

$$A + \mathbf{p} = \mathbf{p} + A := \{\mathbf{x} + \mathbf{p} \mid \mathbf{x} \in A\}.$$

Podemos resumir entonces nuestras dos definiciones básicas como: una *recta* es un conjunto de la forma $\ell = \mathbf{p} + \langle \mathbf{u} \rangle$ con $\mathbf{u} \neq \mathbf{0}$; y un *plano* es un conjunto de la forma $\Pi = \mathbf{p} + \langle \mathbf{u}, \mathbf{v} \rangle$ con \mathbf{u} y \mathbf{v} linealmente independientes.

EJERCICIO 1.33 Demuestra que si \mathbf{u} y \mathbf{v} son linealmente independientes, entonces la función $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definida por $f(s, t) = s\mathbf{u} + t\mathbf{v}$ es inyectiva.

EJERCICIO 1.34 Demuestra que cualquier plano está en biyección con \mathbb{R}^2 .

EJERCICIO 1.35 Demuestra que tres puntos \mathbf{a} , \mathbf{b} , \mathbf{c} son no colineales si y sólo si los vectores $\mathbf{u} := (\mathbf{b} - \mathbf{a})$ y $\mathbf{v} := (\mathbf{c} - \mathbf{a})$ son linealmente independientes.

EJERCICIO 1.36 Da una expresión paramétrica para el plano que pasa por los puntos $\mathbf{a} = (0, 1, 2)$, $\mathbf{b} = (1, 1, 0)$ y $\mathbf{c} = (-1, 0, 2)$.

EJERCICIO 1.37 Demuestra que $\mathbf{0} \in \langle \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k \rangle$ para cualquier $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ en \mathbb{R}^n .

EJERCICIO 1.38 Demuestra que dos vectores \mathbf{u} y \mathbf{v} son linealmente independientes si y sólo si la única combinación lineal de ellos que da $\mathbf{0}$ es la *trivial* (i.e., con ambos coeficientes cero).

EJERCICIO 1.39 Demuestra que

$$\mathbf{w} \in \langle \mathbf{u}, \mathbf{v} \rangle \iff \mathbf{w} + \langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle \iff \langle \mathbf{u}, \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle.$$

1.5 Medio Quinto

Regresemos al plano. Ya tenemos la noción de recta, y en esta sección veremos que nuestras rectas cumplen con la parte de existencia del quinto postulado de Euclides, que nuestra intuición va correspondiendo a nuestra formalización analítica de la geometría y que al cambiar los axiomas de Euclides por unos más básicos (los de los números reales) obtenemos los anteriores, pero ahora como teoremas demostrables. Ya vimos que por cualquier par de puntos se puede trazar un segmento que se extiende indefinidamente a ambos lados (axiomas I y III), es decir, que por ellos pasa una recta. El otro axioma que involucra rectas es el Quinto e incluye la noción de paralelismo, así que tendremos que empezar por ella.

Hay una definición conjuntista de rectas paralelas, así que formalicémosla. Como las rectas son, por definición, ciertos subconjuntos distinguidos del plano, tiene sentido la siguiente:

Definición 1.5.1 Dos rectas ℓ_1 y ℓ_2 en \mathbb{R}^2 son *paralelas*, que escribiremos $\ell_1 \parallel \ell_2$, si no se intersectan, es decir si

$$\ell_1 \cap \ell_2 = \emptyset,$$

donde \emptyset denota al conjunto vacío (ver Apéndice A).

Pero además de rectas tenemos algo más elemental que son los vectores (segmentos dirigidos) y entre ellos también hay una noción intuitiva de paralelismo que corresponde al “alargamiento” o multiplicación por escalares.

Definición 1.5.2 Dados dos vectores $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ distintos de $\mathbf{0}$, diremos que \mathbf{u} es *paralelo a* \mathbf{v} , lo que escribiremos $\mathbf{u} \parallel \mathbf{v}$, si existe un número $t \in \mathbb{R}$ tal que $\mathbf{u} = t\mathbf{v}$.

Hemos eliminado el vector $\mathbf{0}$, el origen, de la definición para no complicar la situación. Si lo hubiéramos incluido no es cierto el siguiente ejercicio.

EJERCICIO 1.40 Demuestra que la relación “ser paralelo a” es una relación de equivalencia en $\mathbb{R}^2 - \{\mathbf{0}\}$ (en el plano menos el origen llamado el “plano agujerado”). Describe las clases de equivalencia. (En el Apéndice A se define relación de equivalencia.)

EJERCICIO 1.41 Sean \mathbf{u}, \mathbf{v} dos vectores distintos de $\mathbf{0}$. Demuestra que:

$$\mathbf{u} \parallel \mathbf{v} \Leftrightarrow \mathcal{L}_{\mathbf{u}} \cap \mathcal{L}_{\mathbf{v}} \neq \{\mathbf{0}\} \Leftrightarrow \mathcal{L}_{\mathbf{u}} = \mathcal{L}_{\mathbf{v}}.$$

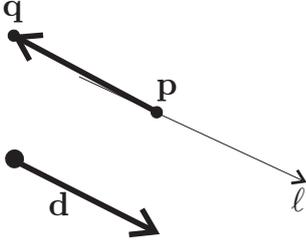
EJERCICIO 1.42 Demuestra que la relación entre rectas “ser paralelo a” es simétrica pero no reflexiva.

EJERCICIO 1.43 Demuestra que dos rectas *horizontales*, es decir, con vector direccional $\mathbf{d} = (1, 0)$, o son paralelas o son iguales.

Con la noción de paralelismo de vectores, podemos determinar cuándo un punto está en una recta.

Lema 1.7 Sea ℓ la recta que pasa por \mathbf{p} con dirección \mathbf{d} ($\mathbf{d} \neq \mathbf{0}$), y sea $\mathbf{q} \neq \mathbf{p}$, entonces

$$\mathbf{q} \in \ell \iff (\mathbf{q} - \mathbf{p}) \parallel \mathbf{d}.$$



Demostración. Tenemos que $\mathbf{q} \in \ell$ si y sólo si existe una $t \in \mathbb{R}$ tal que

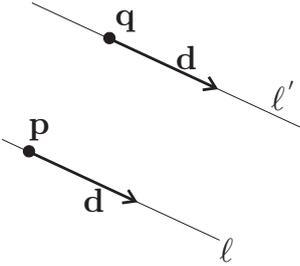
$$\mathbf{q} = \mathbf{p} + t \mathbf{d}.$$

Pero esto es equivalente a que $\mathbf{q} - \mathbf{p} = t \mathbf{d}$, que por definición es que $\mathbf{q} - \mathbf{p}$ es paralelo a \mathbf{d} . \square

Teorema 1.8 (1/2 Quinto) Sean ℓ una recta en \mathbb{R}^2 y \mathbf{q} un punto fuera de ella, entonces existe una recta ℓ' que pasa por \mathbf{q} y es paralela a ℓ .

Demostración. Nuestra definición de recta nos da un punto \mathbf{p} y un vector dirección $\mathbf{d} \neq \mathbf{0}$ para ℓ , de tal manera que

$$\ell = \{\mathbf{p} + t \mathbf{d} \mid t \in \mathbb{R}\}.$$



Es intuitivamente claro —y a la intuición hay que seguirla pues es, de cierta manera, lo que ya sabíamos— que la recta paralela deseada debe tener la misma dirección, así que definamos

$$\ell' = \{\mathbf{q} + s \mathbf{d} \mid s \in \mathbb{R}\}.$$

Como $\mathbf{q} \in \ell'$, nos falta demostrar que $\ell \parallel \ell'$, es decir, que $\ell \cap \ell' = \emptyset$. Para lograrlo, supongamos que no es cierto, es decir, que existe $\mathbf{x} \in \ell \cap \ell'$. Por las expresiones paramétricas de ℓ y ℓ' , se tiene entonces que existen $t \in \mathbb{R}$ y $s \in \mathbb{R}$ para las cuales $\mathbf{x} = \mathbf{p} + t \mathbf{d}$ y $\mathbf{x} = \mathbf{q} + s \mathbf{d}$. De aquí se sigue que

$$\begin{aligned} \mathbf{q} + s \mathbf{d} &= \mathbf{p} + t \mathbf{d} \\ \mathbf{q} - \mathbf{p} &= t \mathbf{d} - s \mathbf{d} \\ \mathbf{q} - \mathbf{p} &= (t - s) \mathbf{d}, \end{aligned}$$

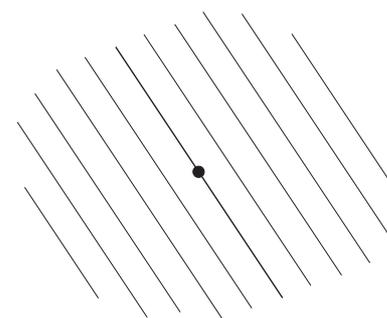
y entonces $\mathbf{q} \in \ell$ por el lema anterior, que es una contradicción a las hipótesis del teorema. Dicho de otra manera, como demostramos que $\ell \cap \ell' \neq \emptyset$ implica que $\mathbf{q} \in \ell$, podemos concluir que si $\mathbf{q} \notin \ell$, como en la hipótesis del teorema, no puede suceder que $\ell \cap \ell'$ sea no vacío, y por lo tanto $\ell \cap \ell' = \emptyset$ y $\ell \parallel \ell'$ por definición. Lo cual concluye con la parte de existencia del Quinto Postulado. \square

La parte que falta demostrar del Quinto es la unicidad, es decir, que cualquier otra recta que pase por \mathbf{q} sí intersecta a ℓ ; que la única paralela es ℓ' . Esto se sigue de que rectas con vectores direccionales no paralelos siempre se intersectan, pero pospondremos la demostración formal de este hecho hasta tener más herramientas conceptuales. En particular, veremos en breve cómo encontrar la intersección de rectas para ejercitar la intuición, pero antes de entrarle, recapitulemos sobre la noción básica de esta sección, el paralelismo.

De la demostración del teorema se sigue que dos rectas con la misma dirección (o, lo que es lo mismo, con vectores direccionales paralelos) o son la misma o no se intersectan. Conviene entonces cambiar nuestra noción conjuntista de paralelismo a una vectorial.

Definición 1.5.3 Dos rectas ℓ_1 y ℓ_2 son *paralelas*, que escribiremos $\ell_1 \parallel \ell_2$, si tienen vectores direccionales paralelos.

Con esta nueva definición (que es la que se mantiene en lo sucesivo) una recta es paralela a sí misma, dos rectas paralelas distintas no se intersectan (son paralelas en el viejo sentido) y la relación es claramente transitiva. Así que es una relación de equivalencia y las clases de equivalencia corresponden a las clases de paralelismo de vectores (las rectas agujeradas por el origen). Llamaremos *haz de rectas paralelas* a una clase de paralelismo de rectas, es decir, al conjunto de todas las rectas paralelas a una dada (todas paralelas entre sí). Hay tantos haces de rectas paralelas como hay rectas por el origen, pues cada haz contiene exactamente una de estas rectas que, quitándole el origen, está formada por los posibles vectores direccionales para las rectas del haz.



Además, nuestra nueva noción de paralelismo tiene la gran ventaja de que se extiende a cualquier espacio vectorial. Nótese primero que la noción de paralelismo entre vectores no nulos se extiende sin problema a \mathbb{R}^3 , pues está en términos del producto por escalares; y luego nuestra noción de paralelismo entre rectas se sigue de la de sus vectores direccionales. Por ejemplo, en el espacio \mathbb{R}^3 corresponde a nuestra noción intuitiva de paralelismo que no es conjuntista. Dos rectas pueden no intersectarse sin tener la misma dirección.

EJERCICIO 1.44 Da la descripción paramétrica de dos rectas en \mathbb{R}^3 que no se intersecten y que no sean paralelas.

EJERCICIO 1.45 (Quinto D3) Demuestra que dados un plano Π en \mathbb{R}^3 y un punto \mathbf{q} fuera de él, existe un plano Π' que pasa por \mathbf{q} y que no intersecta a Π .

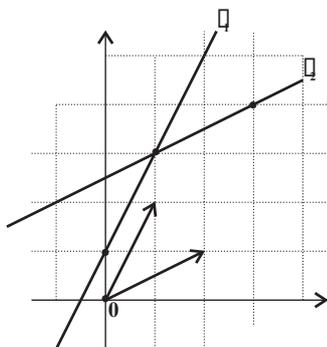
1.6 Intersección de rectas I

En esta sección se analiza el problema de encontrar la intersección de dos rectas. De nuevo, sabemos por experiencia e intuición que dos rectas no paralelas se deben intersectar en un punto. Cómo encontrar ese punto es la pregunta que responderemos en esta sección.

Hagamos un ejemplo. Sean

$$\ell_1 = \{(0, 1) + t(1, 2) \mid t \in \mathbb{R}\}$$

$$\ell_2 = \{(3, 4) + s(2, 1) \mid s \in \mathbb{R}\}.$$



La intersección de estas dos rectas $\ell_1 \cap \ell_2$ es el punto que cumple con las dos descripciones. Es decir, buscamos una $t \in \mathbb{R}$ y una $s \in \mathbb{R}$ que satisfagan

$$(0, 1) + t(1, 2) = (3, 4) + s(2, 1),$$

donde es importante haber dado a los dos parámetros diferente nombre. Esta ecuación vectorial puede reescribirse como

$$t(1, 2) - s(2, 1) = (3, 4) - (0, 1)$$

$$(t - 2s, 2t - s) = (3, 3),$$

que nos da una ecuación lineal con dos incógnitas en cada coordenada. Es decir, debemos resolver el sistema:

$$t - 2s = 3$$

$$2t - s = 3.$$

Si despejamos t en la primera ecuación, se obtiene $t = 3 + 2s$. Y al sustituir en la segunda obtenemos la ecuación lineal en la variable s :

$$2(3 + 2s) - s = 3,$$

que se puede resolver directamente:

$$6 + 4s - s = 3$$

$$3s = -3$$

$$s = -1.$$

Y por lo tanto, al sustituir este valor de s en la descripción de ℓ_2 , el punto

$$(3, 4) - (2, 1) = (1, 3)$$

está en las dos rectas. Nótese que bastó con encontrar el valor de un solo parámetro, pues de la correspondiente descripción paramétrica se obtiene el punto deseado. Pero debemos encontrar el mismo punto si resolvemos primero el otro parámetro. Veámoslo. Otra manera de resolver el sistema es eliminar la variable s . Para esto nos conviene multiplicar por -2 la segunda ecuación y sumarla a la primera, para obtener

$$-3t = -3,$$

de donde $t = 1$. Al sustituir en la parametrización de ℓ_1 , nos da:

$$(0, 1) + (1, 2) = (1, 3),$$

como esperábamos.

Observemos primero que hay muchas maneras de resolver un sistema de dos ecuaciones lineales (se llaman así porque los exponentes de las incógnitas son uno —parece que no hay exponentes—) con dos incógnitas. Cualquiera de ellos es bueno y debe llevar a la misma solución.

Que en este problema particular aparezca un sistema de ecuaciones no es casualidad, pues en general encontrar la intersección de rectas se tiene que resolver así. Ya que si tenemos dos rectas cualesquiera en \mathbb{R}^2 :

$$\begin{aligned}\ell_1 &= \{\mathbf{p} + t\mathbf{v} \mid t \in \mathbb{R}\} \\ \ell_2 &= \{\mathbf{q} + s\mathbf{u} \mid s \in \mathbb{R}\}.\end{aligned}$$

Su intersección está dada por la t y la s que cumplen

$$\mathbf{p} + t\mathbf{v} = \mathbf{q} + s\mathbf{u},$$

que equivale a

$$t\mathbf{v} - s\mathbf{u} = \mathbf{q} - \mathbf{p}. \tag{1.3}$$

Esta ecuación vectorial, a su vez, equivale a una ecuación lineal en cada coordenada. Como estamos en \mathbb{R}^2 , y \mathbf{p} , \mathbf{q} , \mathbf{v} y \mathbf{u} son constantes, nos da entonces un sistema de dos ecuaciones lineales con las dos incógnitas t y s . En cada caso particular el sistema se puede resolver de diferentes maneras; por ejemplo, despejando una incógnita y sustituyendo, o bien tomando múltiplos de las ecuaciones y sumándolas para eliminar una variable y obtener una ecuación lineal con una sola variable (la otra) que se resuelve directamente.

Para hacer la teoría y entender qué pasa con estos sistemas en general, será bueno estudiar las ecuaciones lineales con dos incógnitas de manera geométrica; veremos que también están asociadas a las líneas rectas. Por el momento, de manera “algebraica”, o mejor dicho, mecánica, resolvamos algunos ejemplos.

EJERCICIO 1.46 Encuentra los seis puntos de intersección de las cuatro rectas del Ejercicio 1.20.

EJERCICIO 1.47 Encuentra las intersecciones de las rectas $l_1 = \{(2, 0) + t(1, -2) \mid t \in \mathbb{R}\}$, $l_2 = \{(2, 1) + s(-2, 4) \mid s \in \mathbb{R}\}$ y $l_3 = \{(1, 2) + r(3, -6) \mid r \in \mathbb{R}\}$. Dibújalas para entender qué está pasando.

1.6.1 Sistemas de ecuaciones lineales I

Veamos ahora un método general para resolver sistemas.

Lema 1.9 *El sistema de ecuaciones*

$$\begin{aligned} \mathbf{a} s + \mathbf{b} t &= \mathbf{e} \\ \mathbf{c} s + \mathbf{d} t &= \mathbf{f} \end{aligned} \tag{1.4}$$

en las dos incógnitas s , t tiene solución única si su determinante $\mathbf{ad} - \mathbf{bc}$ es distinto de cero.

Demostración. Hay que intentar resolver el sistema general, seguir un método que funcione en todos los casos independientemente de los valores concretos que puedan tener las constantes \mathbf{a} , \mathbf{b} , \mathbf{c} , \mathbf{d} , \mathbf{e} y \mathbf{f} . Y el método más general es el de “eliminar” una variable para despejar la otra.

Para eliminar la t del sistema (1.4), multiplicamos por \mathbf{d} la primera ecuación, y por $-\mathbf{b}$ la segunda, para obtener

$$\begin{aligned} \mathbf{ad} s + \mathbf{bd} t &= \mathbf{ed} \\ -\mathbf{bc} s - \mathbf{bd} t &= -\mathbf{bf}, \end{aligned} \tag{1.5}$$

de tal manera que al sumar estas dos ecuaciones se tiene

$$(\mathbf{ad} - \mathbf{bc}) s = \mathbf{ed} - \mathbf{bf}. \tag{1.6}$$

Si $\mathbf{ad} - \mathbf{bc} \neq 0$,³ entonces se puede despejar s :

$$s = \frac{\mathbf{ed} - \mathbf{bf}}{\mathbf{ad} - \mathbf{bc}}.$$

Y análogamente (¡hágalo como ejercicio!), o bien, sustituyendo el valor de s en cualquiera de las ecuaciones originales (¡convéznase!), se obtiene t :

$$t = \frac{\mathbf{af} - \mathbf{ec}}{\mathbf{ad} - \mathbf{bc}}.$$

Lo cual demuestra que si el determinante es distinto de cero, la solución es única; es precisamente la de las dos fórmulas anteriores. \square

³Al número $\mathbf{ad} - \mathbf{bc}$ se le llama el *determinante* del sistema, porque determina que en este punto podemos proseguir.

Veamos ahora qué pasa con el sistema (1.4) cuando su determinante es cero. Si $\mathbf{ad} - \mathbf{bc} = 0$, obsérvese que aunque la ecuación (1.6) parezca muy sofisticada porque tiene muchas letras, en realidad es o una contradicción (algo falso) o una trivialidad. El coeficiente de s es 0, así que el lado izquierdo es 0. Entonces, si el lado derecho no es 0 es una contradicción; y si sí es 0, es cierto para cualquier s y hay una infinidad (precisamente un \mathbb{R}) de soluciones.

Lo que sucedió en este caso ($\mathbf{ad} - \mathbf{bc} = 0$) es que al intentar eliminar una de las variables también se eliminó la otra (como debió haberle sucedido al estudiante que hizo el Ejercicio 1.47). Entonces, o las dos ecuaciones son múltiplos y comparten soluciones (cuando $\mathbf{ed} = \mathbf{bf}$ las ecuaciones (1.5) ya son una la negativa de la otra); o bien, sólo los lados izquierdos son múltiplos, pero los derechos no lo son por el mismo factor ($\mathbf{ed} \neq \mathbf{bf}$) y no hay soluciones comunes.

Podemos resumir esto con el siguiente teorema que ya hemos demostrado.

Teorema 1.10 *Un sistema de dos ecuaciones lineales*

$$\begin{aligned} \mathbf{a} s + \mathbf{b} t &= \mathbf{e} \\ \mathbf{c} s + \mathbf{d} t &= \mathbf{f} \end{aligned} \tag{1.7}$$

en las dos incógnitas s , t tiene solución única si y sólo si su determinante $\mathbf{ad} - \mathbf{bc}$ es distinto de cero. Además, si su determinante es cero (si $\mathbf{ad} - \mathbf{bc} = 0$) entonces no tiene solución o tiene una infinidad de ellas (tantas como \mathbb{R}).

Si recordamos que los sistemas de dos ecuaciones lineales con dos incógnitas aparecieron al buscar la intersección de dos rectas, este teorema corresponde a nuestra intuición de las rectas: dos rectas se intersectan en un solo punto o no se intersectan o son la misma. De él dependerá la demostración de la mitad del Quinto que tenemos pendiente, pero conviene entrarle al toro por los cuernos: estudiar y entender primero el significado geométrico de las ecuaciones lineales con dos incógnitas.

EJERCICIO 1.48 Para las rectas de los dos ejercicios anteriores (1.46, 1.47), determina cómo se intersectan las rectas, usando únicamente el determinante.

1.7 Producto interior

En esta sección se introduce un ingrediente central para el estudio moderno de la geometría euclidiana: el “*producto interior*”. Además de darnos la herramienta algebraica para el estudio geométrico de las ecuaciones lineales, en la sección siguiente nos dará mucho más. De él derivaremos después las nociones básicas de distancia y ángulo (de rigidez). Puede decirse que el producto interior es, aunque menos intuitivo, más elemental que las nociones de distancia y ángulo pues éstas se definirán en base a

aquél (aunque se verá también que estas dos juntas lo definen). El producto interior depende tan íntimamente de la idea cartesiana de coordenadas, que no tiene ningún análogo en la geometría griega. Pero tampoco viene de los primeros pininos que hizo la geometría analítica, pues no es sino hasta el siglo XIX cuando se le empezó a dar la importancia debida al desarrollarse las ideas involucradas en la noción general de espacio vectorial. Podría entonces decirse que el uso del producto interior (junto con el lenguaje de espacio vectorial) marca dos épocas en la geometría analítica. Pero es quizá la simpleza de su definición su mejor tarjeta de presentación.

Definición 1.7.1 Dados dos vectores $\mathbf{u} = (u_1, u_2)$ y $\mathbf{v} = (v_1, v_2)$, su *producto interno* (también conocido como producto *escalar* —que preferimos no usar para no confundir con la multiplicación por escalares— o bien como producto *punto*) es el número real:

$$\mathbf{u} \cdot \mathbf{v} = (u_1, u_2) \cdot (v_1, v_2) := u_1v_1 + u_2v_2.$$

En general, en \mathbb{R}^n se define el *producto interior* (o el *producto punto*) de dos vectores $\mathbf{u} = (u_1, \dots, u_n)$ y $\mathbf{v} = (v_1, \dots, v_n)$ como la suma de los productos de sus coordenadas correspondientes, es decir,

$$\mathbf{u} \cdot \mathbf{v} := u_1v_1 + \dots + u_nv_n = \sum_{i=1}^n u_iv_i.$$

Así, por ejemplo:

$$\begin{aligned} (4, 3) \cdot (2, -1) &= 4 \times 2 + 3 \times (-1) = 8 - 3 = 5, \\ (1, 2, 3) \cdot (4, 5, -6) &= 4 + 10 - 18 = -4. \end{aligned}$$

Obsérvese que el producto interior tiene como ingredientes dos vectores ($\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$) y nos da un escalar ($\mathbf{u} \cdot \mathbf{v} \in \mathbb{R}$); no debe confundirse con la multiplicación escalar, que de un escalar y un vector nos da un vector, excepto en el caso $n = 1$, en el que ambos coinciden con la multiplicación de los reales.

Antes de demostrar las propiedades básicas del producto interior, observemos que nos será muy útil para el problema que dejamos pendiente sobre sistemas de ecuaciones. En efecto en una ecuación lineal con dos incógnitas, x y y digamos, aparece una expresión de la forma $\mathbf{a}x + \mathbf{b}y$ con \mathbf{a} y \mathbf{b} constantes. Si tomamos un vector constante $\mathbf{u} = (\mathbf{a}, \mathbf{b})$ y un vector variable $\mathbf{x} = (x, y)$, ahora podemos escribir

$$\mathbf{u} \cdot \mathbf{x} = \mathbf{a}x + \mathbf{b}y$$

y el “paquete básico” de información está en \mathbf{u} . En particular, como la mayoría de los lectores ya deben saber, la ecuación lineal

$$\mathbf{a}x + \mathbf{b}y = c$$

donde c es una nueva constante, define una recta. Con el producto interior esta ecuación se reescribe como

$$\mathbf{u} \cdot \mathbf{x} = c.$$

Veremos cómo en el vector \mathbf{u} y en la constante c se almacena la información geométrica de esa recta. Pero no nos apresuremos.

Como en el caso de la suma vectorial y la multiplicación por escalares, conviene demostrar primero las propiedades elementales del producto interior para manejarlo después con más soltura.

Teorema 1.11 *Para todos los vectores $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, y para todo número $t \in \mathbb{R}$ se cumple que*

$$\begin{aligned} \text{i)} \quad & \mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u} \\ \text{ii)} \quad & \mathbf{u} \cdot (t \mathbf{v}) = t (\mathbf{u} \cdot \mathbf{v}) \\ \text{iii)} \quad & \mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w} \\ \text{iv)} \quad & \mathbf{u} \cdot \mathbf{u} \geq 0 \\ \text{v)} \quad & \mathbf{u} \cdot \mathbf{u} = 0 \Leftrightarrow \mathbf{u} = \mathbf{0} \end{aligned}$$

Demostración. Nos interesa demostrarlo en \mathbb{R}^2 (y como ejercicio en \mathbb{R}^3), aunque el caso general es esencialmente lo mismo. Convendrá usar la notación de la misma letra con subíndices para las coordenadas, es decir, tomar $\mathbf{u} = (u_1, u_2)$, $\mathbf{v} = (v_1, v_2)$ y $\mathbf{w} = (w_1, w_2)$. El enunciado (i) se sigue inmediatamente de las definiciones y la conmutatividad de los números reales; (ii) se obtiene al factorizar:

$$\mathbf{u} \cdot (t \mathbf{v}) = u_1(t v_1) + u_2(t v_2) = t(u_1 v_1 + u_2 v_2) = t(\mathbf{u} \cdot \mathbf{v}).$$

De la distributividad y conmutatividad en los reales se obtiene (iii):

$$\begin{aligned} \mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) &= u_1(v_1 + w_1) + u_2(v_2 + w_2) \\ &= u_1 v_1 + u_1 w_1 + u_2 v_2 + u_2 w_2 \\ &= (u_1 v_1 + u_2 v_2) + (u_1 w_1 + u_2 w_2) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}. \end{aligned}$$

Al tomar el producto interior de un vector consigo mismo se obtiene

$$\mathbf{u} \cdot \mathbf{u} = u_1^2 + u_2^2,$$

que es una suma de cuadrados. Como cada cuadrado es positivo la suma también lo es (iv); y si fuera 0 significa que cada sumando es 0, es decir, que cada coordenada es 0 (v). \square

EJERCICIO 1.49 Demuestra el Teorema 1.11 para $n = 3$.

EJERCICIO 1.50 Calcula el producto interior de algunos vectores del Ejercicio 1.3.

EJERCICIO 1.51 Demuestra sin usar (v), sólo la definición de producto interior, que dado $\mathbf{u} \in \mathbb{R}^2$ se tiene

$$\mathbf{u} \cdot \mathbf{x} = 0, \quad \forall \mathbf{x} \in \mathbb{R}^2 \Leftrightarrow \mathbf{u} = \mathbf{0}.$$

(Observa que sólo hay que usar el lado izquierdo para dos vectores muy sencillos.)

EJERCICIO 1.52 Demuestra el análogo del ejercicio anterior en \mathbb{R}^3 . ¿Y en \mathbb{R}^n ?

1.7.1 El compadre ortogonal

El primer uso geométrico que daremos al producto interior será para detectar la perpendicularidad. Es otra de las nociones básicas en los axiomas de Euclides que incluyen el concepto de ángulo recto.

Fijemos el vector $\mathbf{u} \in \mathbb{R}^2$, y para simplificar la notación digamos que $\mathbf{u} = (a, b) \neq \mathbf{0}$. Si tomamos $\mathbf{x} = (x, y)$ como un vector variable, vamos a ver que las soluciones de la ecuación

$$\mathbf{u} \cdot \mathbf{x} = 0, \quad (1.8)$$

es decir, los puntos en \mathbb{R}^2 cuyas coordenadas (x, y) satisfacen la ecuación

$$ax + by = 0,$$

forman una recta que es perpendicular a \mathbf{u} . Para esto consideraremos una solución particular (una de las más sencillas), $x = -b$, $y = a$; y será tan importante esta solución que le daremos nombre:

Definición 1.7.2 El *compadre ortogonal* del vector $\mathbf{u} = (a, b)$, denotado \mathbf{u}^\perp y que se lee “ \mathbf{u} -perpendicular” o “ \mathbf{u} -ortogonal”, es

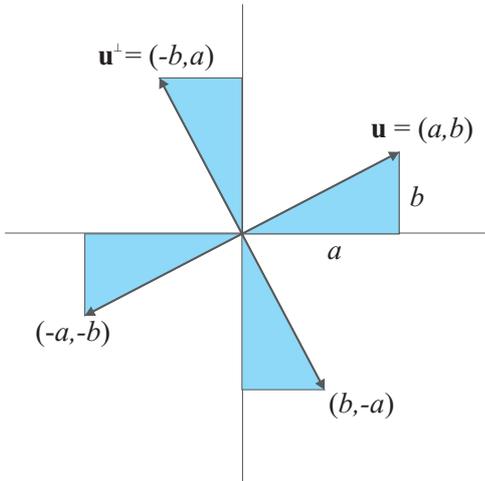
$$\mathbf{u}^\perp = (-b, a).$$

(Se intercambian coordenadas y a la primera se le cambia el signo.)

Se cumple que

$$\mathbf{u} \cdot \mathbf{u}^\perp = a(-b) + ba = 0;$$

y para justificar el nombre hay que mostrar que \mathbf{u}^\perp se obtiene de \mathbf{u} al girarlo 90° en sentido contrario a las manecillas del reloj (alrededor del origen). Para ver esto considérese la figura al margen, donde suponemos que el vector $\mathbf{u} = (a, b)$ está en el primer cuadrante, es decir, que $a > 0$ y $b > 0$. Al rotar el triángulo rectángulo de catetos a y b (e hipotenusa \mathbf{u}) un ángulo de 90° en el origen se obtiene el correspondiente a \mathbf{u}^\perp . Se incluyen además las siguientes dos rotaciones para que quede claro que esto no depende de los signos de a y b .



Podemos concluir entonces que el “compadre ortogonal”, pensado como la función de \mathbb{R}^2 en \mathbb{R}^2 que manda al vector \mathbf{u} en su perpendicular \mathbf{u}^\perp ($\mathbf{u} \mapsto \mathbf{u}^\perp$), es la rotación de 90° en sentido contrario a las manecillas del reloj alrededor del origen. Es fácil comprobar que $(\mathbf{u}^\perp)^\perp = -\mathbf{u}$, ya sea con la fórmula o porque la rotación de 180° (rotar y volver a rotar 90°) corresponde justo a tomar el inverso aditivo.

Como ya es costumbre, veamos algunas propiedades bonitas de la función “compadre ortogonal” respecto de las otras operaciones que hemos definido. La demostración, que consiste en dar coordenadas y aplicar definiciones, se deja al lector.

Lema 1.12 Para todos los vectores $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$, y para todo número $t \in \mathbb{R}$ se cumple que

$$\begin{aligned} i) \quad & (\mathbf{u} + \mathbf{v})^\perp = \mathbf{u}^\perp + \mathbf{v}^\perp \\ ii) \quad & (t\mathbf{u})^\perp = t(\mathbf{u}^\perp) \\ iii) \quad & \mathbf{u}^\perp \cdot \mathbf{v}^\perp = \mathbf{u} \cdot \mathbf{v} \\ iv) \quad & \mathbf{u}^\perp \cdot \mathbf{v} = -(\mathbf{u} \cdot \mathbf{v}^\perp). \end{aligned}$$

EJERCICIO 1.53 Demuestra el lema anterior.

Sistemas de ecuaciones lineales II

Usemos la herramienta del producto interior y el compadre ortogonal para revisar nuestro trabajo previo (sección 1.6.1) sobre sistemas de dos ecuaciones lineales con dos incógnitas. Si pensamos cada ecuación como la coordenada de un vector (y así fue como surgieron), un sistema tal se escribe

$$s\mathbf{u} + t\mathbf{v} = \mathbf{c}, \tag{1.9}$$

donde s y t son las incógnitas y $\mathbf{u}, \mathbf{v}, \mathbf{c} \in \mathbb{R}^2$ están dados. Para que este sistema sea justo el que ya estudiamos, (1.4), denotemos las coordenadas $\mathbf{u} = (\mathbf{a}, \mathbf{c})$, $\mathbf{v} = (\mathbf{b}, \mathbf{d})$; y para las constantes (con acrónimo \mathbf{c}), digamos que $\mathbf{c} = (\mathbf{e}, \mathbf{f})$. Como ésta es una ecuación vectorial, si la multiplicamos (con el producto interior) por un vector, obtendremos una ecuación lineal real. Y para eliminar una variable, digamos s , podemos multiplicar por el compadre ortogonal de \mathbf{u} , para obtener

$$\begin{aligned} \mathbf{u}^\perp \cdot (s\mathbf{u}) + \mathbf{u}^\perp \cdot (t\mathbf{v}) &= \mathbf{u}^\perp \cdot \mathbf{c} \\ s(\mathbf{u}^\perp \cdot \mathbf{u}) + t(\mathbf{u}^\perp \cdot \mathbf{v}) &= \mathbf{u}^\perp \cdot \mathbf{c} \\ s(0) + t(\mathbf{u}^\perp \cdot \mathbf{v}) &= \mathbf{u}^\perp \cdot \mathbf{c} \\ t(\mathbf{u}^\perp \cdot \mathbf{v}) &= \mathbf{u}^\perp \cdot \mathbf{c}. \end{aligned}$$

Definición 1.7.3 El *determinante* de los vectores $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ es el número real

$$\det(\mathbf{u}, \mathbf{v}) := \mathbf{u}^\perp \cdot \mathbf{v}.$$

Así que si suponemos que $\det(\mathbf{u}, \mathbf{v}) \neq 0$, podemos despejar t . De manera análoga podemos despejar s , y entonces concluir que el sistema de ecuaciones tiene solución única:

$$t = \frac{\mathbf{u}^\perp \cdot \mathbf{c}}{\mathbf{u}^\perp \cdot \mathbf{v}}, \quad s = \frac{\mathbf{v}^\perp \cdot \mathbf{c}}{\mathbf{v}^\perp \cdot \mathbf{u}};$$

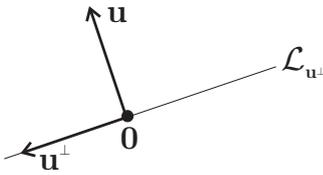
que es justo el Lema 1.9.

EJERCICIO 1.54 Escribe en coordenadas ($\mathbf{u} = (\mathbf{a}, \mathbf{c})$, $\mathbf{v} = (\mathbf{b}, \mathbf{d})$) cada uno de los pasos anteriores para ver que corresponden al método clásico de eliminación. (Quizá te convenga escribir las parejas como columnas, en vez de renglones, para que el sistema de ecuaciones sea claro.)

La ecuación lineal homogénea

Ahora sí, describamos el conjunto de soluciones de la ecuación lineal homogénea:

$$\mathbf{u} \cdot \mathbf{x} = 0.$$



Proposición 1.13 Sea $\mathbf{u} \in \mathbb{R}^2 - \{\mathbf{0}\}$, entonces

$$\{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{u} \cdot \mathbf{x} = 0\} = \mathcal{L}_{\mathbf{u}^\perp},$$

donde, recuérdese, $\mathcal{L}_{\mathbf{u}^\perp} = \{t\mathbf{u}^\perp \mid t \in \mathbb{R}\}$.

Demostración. Tenemos que demostrar la igualdad de dos conjuntos. Esto se hace en dos pasos que corresponden a ver que cada elemento de un conjunto está también en el otro.

La contención más fácil es “ \supseteq ”. Un elemento de $\mathcal{L}_{\mathbf{u}^\perp}$ es de la forma $t\mathbf{u}^\perp$ para algún $t \in \mathbb{R}$. Tenemos que demostrar que $t\mathbf{u}^\perp$ está en el conjunto de la izquierda. Para esto hay que ver que satisface la condición de pertenencia, que se sigue de (ii) en el Lema 1.12, y (ii) en el Teorema 1.11:

$$\mathbf{u} \cdot (t\mathbf{u}^\perp) = \mathbf{u} \cdot t(\mathbf{u}^\perp) = t(\mathbf{u} \cdot \mathbf{u}^\perp) = 0.$$

Para la otra contención, tomamos $\mathbf{x} \in \mathbb{R}^2$ tal que $\mathbf{u} \cdot \mathbf{x} = 0$ y debemos encontrar $t \in \mathbb{R}$ tal que $\mathbf{x} = t\mathbf{u}^\perp$. Sean $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ las coordenadas de \mathbf{u} , es decir $\mathbf{u} = (\mathbf{a}, \mathbf{b})$ y análogamente, sea $\mathbf{x} = (x, y)$. Entonces estamos suponiendo que

$$ax + by = 0.$$

Puesto que $\mathbf{u} \neq \mathbf{0}$ por hipótesis, tenemos que $\mathbf{a} \neq 0$ o $\mathbf{b} \neq 0$; y en cada caso podemos despejar una variable para encontrar el factor deseado:

$$\mathbf{a} \neq 0 \Rightarrow x = -\frac{by}{a},$$

y por lo tanto

$$\mathbf{x} = (x, y) = \frac{y}{a}(-b, a) = \left(\frac{y}{a}\right)\mathbf{u}^\perp;$$

o bien,

$$\mathbf{b} \neq 0 \Rightarrow y = -\frac{ax}{b} \Rightarrow \mathbf{x} = (x, y) = -\frac{x}{b}(-b, a) = \left(-\frac{x}{b}\right)\mathbf{u}^\perp.$$

□

Hemos demostrado que la ecuación lineal *homogénea* (se llama así porque la constante es 0) $\mathbf{a}\mathbf{x} + \mathbf{b}\mathbf{y} = 0$, con \mathbf{a} y \mathbf{b} constantes reales, tiene como soluciones las parejas (\mathbf{x}, \mathbf{y}) que forman una recta por el origen de \mathbb{R}^2 . Si empaquetamos la información de la ecuación en el vector $\mathbf{u} = (\mathbf{a}, \mathbf{b})$, tenemos que $\mathbf{u} \neq \mathbf{0}$, pues de lo contrario no habría tal ecuación lineal (sería la tautología $0 = 0$ de lo que estamos hablando, y no es así). Pero además hemos visto que el vector \mathbf{u} es ortogonal (también llamado *normal*) a la recta en cuestión, pues ésta está generada por el compadre ortogonal \mathbf{u}^\perp .

Se justifica entonces la siguiente definición, que era nuestro primer objetivo con el producto interior (donde, de una vez, estamos extrapolando a todas las dimensiones).

Definición 1.7.4 Se dice que dos vectores \mathbf{u} y \mathbf{v} en \mathbb{R}^n son *perpendiculares* u *ortogonales*, y se escribe $\mathbf{u} \perp \mathbf{v}$, si $\mathbf{u} \cdot \mathbf{v} = 0$.

Podemos reformular entonces la proposición anterior como “dada $\mathbf{u} \in \mathbb{R}^2$, $\mathbf{u} \neq \mathbf{0}$, el conjunto de $\mathbf{x} \in \mathbb{R}^2$ que satisfacen la ecuación $\mathbf{u} \cdot \mathbf{x} = 0$ son la recta *perpendicular* a \mathbf{u} ”.

Antes de estudiar la ecuación general (no necesariamente homogénea), notemos que si el producto interior detecta perpendicularidad, también se puede usar, junto con el compadre ortogonal, para detectar paralelismo en el plano.

Corolario 1.14 Sean \mathbf{u} y \mathbf{v} dos vectores no nulos en \mathbb{R}^2 , entonces

$$\mathbf{u} \parallel \mathbf{v} \iff \det(\mathbf{u}, \mathbf{v}) = \mathbf{u}^\perp \cdot \mathbf{v} = 0.$$

Demostración. Por la Proposición 1.13, $\mathbf{u}^\perp \cdot \mathbf{v} = 0$ si y sólo si \mathbf{v} pertenece a la recta generada por $(\mathbf{u}^\perp)^\perp$, que es la recta generada por \mathbf{u} pues $(\mathbf{u}^\perp)^\perp = -\mathbf{u}$. \square

Hay que hacer notar que si les damos coordenadas a \mathbf{u} y a \mathbf{v} , digamos $\mathbf{u} = (\mathbf{a}, \mathbf{b})$ y $\mathbf{v} = (\mathbf{c}, \mathbf{d})$, entonces el determinante, que detecta paralelismo, es

$$\mathbf{u}^\perp \cdot \mathbf{v} = \mathbf{a}\mathbf{d} - \mathbf{b}\mathbf{c},$$

que ya habíamos encontrado como determinante del sistema de ecuaciones (1.9), pero considerando $\mathbf{u} = (\mathbf{a}, \mathbf{c})$ y $\mathbf{v} = (\mathbf{b}, \mathbf{d})$.

EJERCICIO 1.55 Dibuja la recta definida como $\{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{u} \cdot \mathbf{x} = 0\}$ para \mathbf{u} con cada uno de los siguientes vectores: a) $(1, 0)$, b) $(0, 1)$, c) $(2, 1)$, d) $(1, 2)$, e) $(-1, 1)$.

EJERCICIO 1.56 Describe el lugar geométrico definido como $\{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{u} \cdot \mathbf{x} = 0\}$ para \mathbf{u} con cada uno de los siguientes vectores: a) $(1, 0, 0)$, b) $(0, 1, 0)$, c) $(0, 0, 1)$, d) $(1, 1, 0)$.

EJERCICIO 1.57 Sean $\mathbf{u} = (\mathbf{a}, \mathbf{b})$ y $\mathbf{v} = (\mathbf{c}, \mathbf{d})$ dos vectores no nulos. Sin usar el producto interior ni el compadre ortogonal —es decir, de la pura definición y con “álgebra elemental”— demuestra que \mathbf{u} es paralelo a \mathbf{v} si y sólo si $\mathbf{a}\mathbf{d} - \mathbf{b}\mathbf{c} = 0$. Compara con la última parte de la demostración de la Proposición 1.13.

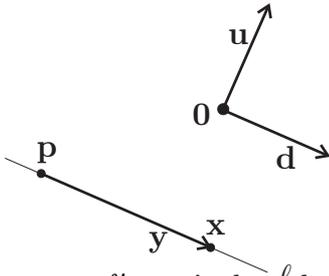
1.8 La ecuación normal de la recta

Demostraremos ahora que todas las rectas de \mathbb{R}^2 se pueden describir con una ecuación

$$\mathbf{u} \cdot \mathbf{x} = c,$$

donde $\mathbf{u} \in \mathbb{R}^2$ es un vector normal a la recta y la constante $c \in \mathbb{R}$ es escogida apropiadamente. Esta ecuación, vista en coordenadas, equivale a una ecuación lineal en dos variables ($a\mathbf{x} + b\mathbf{y} = c$, al tomar $\mathbf{u} = (a, b)$ constante y $\mathbf{x} = (x, y)$ es el vector variable). Históricamente, el hecho de que las rectas tuvieran tal descripción fue una gran motivación en el inicio de la geometría analítica.

Tomemos una recta con dirección $\mathbf{d} \neq \mathbf{0}$



$$l = \{\mathbf{p} + t\mathbf{d} \mid t \in \mathbb{R}\}.$$

Si definimos $\mathbf{u} := \mathbf{d}^\perp$ y $c := \mathbf{u} \cdot \mathbf{p}$, afirmamos, es decir, demostraremos, que

$$l = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{u} \cdot \mathbf{x} = c\}.$$

Llamemos l' a este último conjunto; demostrar la igualdad $l = l'$, equivale a demostrar las dos contenciones $l \subseteq l'$ y $l' \subseteq l$.

Si $\mathbf{x} = \mathbf{p} + t\mathbf{d} \in l$, entonces (usando las propiedades del producto interior y del compadre ortogonal) tenemos

$$\begin{aligned} \mathbf{u} \cdot \mathbf{x} &= \mathbf{u} \cdot (\mathbf{p} + t\mathbf{d}) = \mathbf{u} \cdot \mathbf{p} + t(\mathbf{u} \cdot \mathbf{d}) \\ &= \mathbf{u} \cdot \mathbf{p} + t(\mathbf{d}^\perp \cdot \mathbf{d}) = \mathbf{u} \cdot \mathbf{p} + 0 = c, \end{aligned}$$

lo cual demuestra que $l \subseteq l'$. Por el otro lado, dado $\mathbf{x} \in l'$ (i.e., tal que $\mathbf{u} \cdot \mathbf{x} = c$), sea $\mathbf{y} := \mathbf{x} - \mathbf{p}$ (obsérvese que entonces $\mathbf{x} = \mathbf{p} + \mathbf{y}$). Se tiene que

$$\mathbf{u} \cdot \mathbf{y} = \mathbf{u} \cdot (\mathbf{x} - \mathbf{p}) = \mathbf{u} \cdot \mathbf{x} - \mathbf{u} \cdot \mathbf{p} = c - c = 0.$$

Por la Proposición 1.13, esto implica que \mathbf{y} es paralelo a $\mathbf{u}^\perp = -\mathbf{d}$; y por tanto que $\mathbf{x} = \mathbf{p} + \mathbf{y} \in l$. Esto demuestra que $l = l'$.

En resumen, hemos demostrado que toda recta puede ser descrita por una ecuación normal:

Teorema 1.15 *Sea $l = \{\mathbf{p} + t\mathbf{d} \mid t \in \mathbb{R}\}$ una recta en \mathbb{R}^2 . Entonces l consta de los vectores $\mathbf{x} \in \mathbb{R}^2$ que cumplen la ecuación $\mathbf{u} \cdot \mathbf{x} = c$, que escribimos*

$$l: \quad \mathbf{u} \cdot \mathbf{x} = c,$$

donde $\mathbf{u} = \mathbf{d}^\perp$ y $c = \mathbf{u} \cdot \mathbf{p}$. □

Este teorema también podría escribirse:

Teorema 1.16 Sea $\mathbf{d} \in \mathbb{R}^2 - \{\mathbf{0}\}$, entonces

$$\{\mathbf{p} + t\mathbf{d} \mid t \in \mathbb{R}\} = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{d}^\perp \cdot \mathbf{x} = \mathbf{d}^\perp \cdot \mathbf{p}\}$$

Estos enunciados nos dicen cómo encontrar una ecuación normal para una recta descrita paramétricamente. Por **ejemplo**, la recta

$$\{(s - 2, 3 - 2s) \mid s \in \mathbb{R}\}$$

tiene como vector direccional al $(1, -2)$. Por tanto tiene vector normal a su compadre ortogonal $(2, 1)$, y como pasa por el punto $(-2, 3)$ entonces está determinada por la ecuación

$$\begin{aligned} (2, 1) \cdot (x, y) &= (2, 1) \cdot (-2, 3) \\ 2x + y &= -1 \end{aligned}$$

(sustituya el lector las coordenadas de la descripción paramétrica en esta última ecuación).

...Dibujo

Para encontrar una representación paramétrica de la recta ℓ dada por una *ecuación normal*

$$\ell: \mathbf{u} \cdot \mathbf{x} = c$$

(léase “ ℓ dada por la ecuación $\mathbf{u} \cdot \mathbf{x} = c$ ”), bastará con encontrar una solución particular \mathbf{p} (que es muy fácil porque se puede dar un valor arbitrario a una variable, 0 es la más conveniente, y despejar la otra), pues sabemos que la dirección es \mathbf{u}^\perp (o cualquier paralelo). Por **ejemplo**, la recta dada por la ecuación normal

$$2x - 3y = 2$$

tiene vector normal $(2, -3)$. Por tanto tiene vector direccional $(3, 2)$; y como pasa por el punto $(1, 0)$, es el conjunto

$$\{(1 + 3t, 2t) \mid t \in \mathbb{R}\}.$$

Esto también se puede obtener directamente de las coordenadas, pero hay que partirlo en casos. Si $\mathbf{u} = (a, b)$, tenemos

$$\ell: ax + by = c.$$

Supongamos que $a \neq 0$. Entonces se puede despejar x :

$$x = \frac{c}{a} - \frac{b}{a}y$$

y todas las soluciones de la ecuación se obtienen dando diferentes valores a y ; es decir, son

$$\left\{ \left(\frac{c}{a} - \frac{b}{a}y, y \right) \mid y \in \mathbb{R} \right\} = \left\{ \left(\frac{c}{a}, 0 \right) + y \left(-\frac{b}{a}, 1 \right) \mid y \in \mathbb{R} \right\},$$

que es una recta con dirección $(-b, a)$. Nos falta ver qué pasa cuando $a = 0$, pero entonces la recta es la horizontal $y = c/b$. Si hubieramos hecho el análisis cuando $b \neq 0$, se nos hubiera confundido la notación del caso clásico que vemos en el siguiente párrafo.

La ecuación “funcional” de la recta

Es bien conocida la ecuación *funcional de la recta*

$$y = mx + b,$$

que se usa para describir rectas como gráficas de una función; aquí m es la *pendiente* y b es la llamada “ordenada al origen” o el “valor inicial”. En nuestros términos, esta ecuación se puede reescribir como

$$-mx + y = b$$

o bien, como

$$(-m, 1) \cdot (x, y) = b.$$

Por lo tanto, tiene vector normal $\mathbf{n} = (-m, 1)$. Podemos escoger a $\mathbf{d} = (1, m) = -\mathbf{n}^\perp$ como su vector direccional y a $\mathbf{p} = (0, b)$ como una solución particular. Así que su parametrización natural es (usando a x como parámetro):

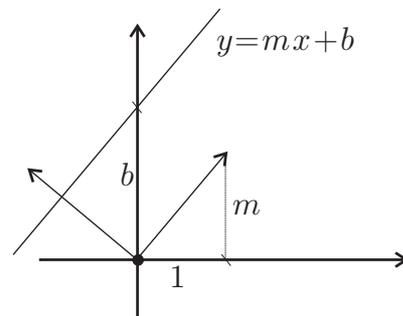
$$\{(0, b) + x(1, m) \mid x \in \mathbb{R}\} = \{(x, mx + b) \mid x \in \mathbb{R}\},$$

que es la gráfica de una función. Obsérvese que todas las rectas excepto las verticales (con dirección $(0, 1)$) se pueden expresar así.

No está de más observar que para cualquier función $f : \mathbb{R} \rightarrow \mathbb{R}$ se obtiene un subconjunto de \mathbb{R}^2 , llamado la *gráfica de la función* f definida paramétricamente como

$$\{(x, f(x)) \mid x \in \mathbb{R}\};$$

y entonces hemos visto que todas las rectas, excepto las verticales, son la gráfica de funciones de la forma $f(x) = mx + b$, que en el capítulo 3 llamaremos funciones afines.



EJERCICIO 1.58 Para las siguientes rectas, encuentra una ecuación normal y, en su caso, su ecuación funcional:

- $\{(2, 3) + t(1, 1) \mid t \in \mathbb{R}\}$
- $\{(-1, 0) + s(2, 1) \mid s \in \mathbb{R}\}$
- $\{(0, -2) + (-r, 2r) \mid r \in \mathbb{R}\}$
- $\{(1, 3) + s(2, 0) \mid s \in \mathbb{R}\}$
- $\{(t - 1, -2t) \mid t \in \mathbb{R}\}$

EJERCICIO 1.59 Da una descripción paramétrica de las rectas dadas por las ecuaciones:

- $2x - 3y = 1$
- $2x - y = 2$
- $2y - 4x = 2$
- $x + 5y = -1$
- $3 - 4y = 2x + 2$

EJERCICIO 1.60 Encuentra una ecuación normal para la recta que pasa por los puntos:

- a) $(2, 1)$ y $(3, 4)$
- b) $(-1, 1)$ y $(2, 2)$
- c) $(1, -3)$ y $(3, 1)$
- d) $(2, 0)$ y $(1, 1)$

EJERCICIO 1.61 Sean \mathbf{p} y \mathbf{q} dos puntos distintos en \mathbb{R}^2 . Demuestra que la recta ℓ que pasa por ellos tiene ecuación normal:

$$\ell : (\mathbf{q} - \mathbf{p})^\perp \cdot \mathbf{x} = \mathbf{q}^\perp \cdot \mathbf{p}.$$

EJERCICIO 1.62 Encuentra la intersección de las rectas (a) y (b) del Ejercicio 1.59.

EJERCICIO 1.63 Encuentra la intersección de la recta (α) del Ejercicio 1.58 con la de la recta (α) del Ejercicio 1.59, donde $\alpha \in \{a, b, c, d\}$.

EJERCICIO 1.64 Dibuja las gráficas de las funciones $f(x) = x^2$ y $g(x) = x^3$.

1.8.1 Intersección de rectas II

Regresemos ahora al problema teórico de encontrar la intersección de rectas, pero con la nueva herramienta de la ecuación normal. Consideremos dos rectas ℓ_1 y ℓ_2 en \mathbb{R}^2 . Sean $\mathbf{u} = (a, b)$ y $\mathbf{v} = (c, d)$ vectores normales a ellas respectivamente, de tal manera que ambos son no nulos y existen constantes $e, f \in \mathbb{R}$ para las cuales

$$\begin{aligned} \ell_1 & : \quad \mathbf{u} \cdot \mathbf{x} = e \\ \ell_2 & : \quad \mathbf{v} \cdot \mathbf{x} = f. \end{aligned}$$

Es claro entonces que $\ell_1 \cap \ell_2$ consta de los puntos $\mathbf{x} \in \mathbb{R}^2$ que satisfacen ambas ecuaciones. Si llamamos, como de costumbre, x y y a las coordenadas de \mathbf{x} (es decir, si hacemos $\mathbf{x} = (x, y)$), las dos ecuaciones anteriores son el sistema

$$\begin{aligned} ax + by & = e \\ cx + dy & = f, \end{aligned}$$

con a, b, c, d, e, f constantes y x, y variables o incógnitas. Éste es precisamente el sistema que estudiamos en la sección 1.6.1; aunque allá hablabamos de los “parámetros s y t ”, son esencialmente lo mismo. Pero ahora tiene el significado geométrico que buscábamos: resolverlo es encontrar el punto de intersección de dos rectas (sus coordenadas tal cual, y no los parámetros para encontrarlo, como antes). Como ya hicimos el trabajo abstracto de resolver el sistema, sólo nos queda por hacer la traducción al lenguaje geométrico. Como ya vimos, el tipo de solución (una única, ninguna o tantas como reales) depende del determinante del sistema $ad - bc$, y éste, a su vez, ya se nos apareció (nótese que $\det(\mathbf{u}, \mathbf{v}) = \mathbf{u}^\perp \cdot \mathbf{v} = ad - bc$) como el “numero mágico” que detecta paralelismo (Corolario 1.14). De tal manera que del Corolario 1.14 y del Teorema 1.10 podemos concluir:

Teorema 1.17 *Dadas las rectas*

$$\begin{aligned} \ell_1 & : \quad \mathbf{u} \cdot \mathbf{x} = e \\ \ell_2 & : \quad \mathbf{v} \cdot \mathbf{x} = f, \end{aligned}$$

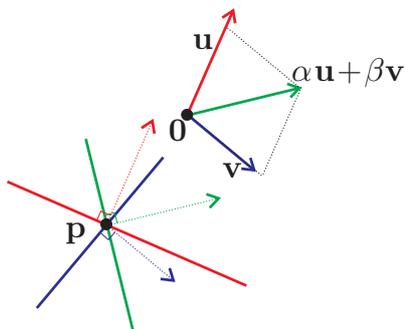
entonces:

- i) $\det(\mathbf{u}, \mathbf{v}) \neq 0 \Leftrightarrow \ell_1 \cap \ell_2$ es un único punto;
- ii) $\det(\mathbf{u}, \mathbf{v}) = 0 \Leftrightarrow \ell_1 \parallel \ell_2 \Leftrightarrow \ell_1 \cap \ell_2 = \emptyset$ o $\ell_1 = \ell_2$.

Consideremos con más detenimiento el segundo caso, cuando el determinante $\mathbf{u}^\perp \cdot \mathbf{v}$ es cero y por tanto los vectores normales (y las rectas) son paralelos. Tenemos entonces que para alguna $t \in \mathbb{R}$, se cumple que $\mathbf{v} = t\mathbf{u}$ (nótese que $t \neq 0$, pues ambos vectores son no nulos). La disyuntiva de si las rectas no se intersectan o son iguales corresponde a que las constantes e y f difieran por el mismo factor, es decir si $f \neq te$ o $f = te$ respectivamente. Podemos sacar dos conclusiones interesantes. Primero, dos ecuaciones normales definen la misma recta si y sólo si las tres constantes que las determinan (en nuestro caso $(\mathbf{a}, \mathbf{b}, e)$ y $(\mathbf{c}, \mathbf{d}, f)$) son vectores paralelos en \mathbb{R}^3 . Y segundo, que al fijar un vector $\mathbf{u} = (\mathbf{a}, \mathbf{b})$ y variar un parámetro $c \in \mathbb{R}$, las ecuaciones $\mathbf{u} \cdot \mathbf{x} = c$ determinan el haz de rectas paralelas que es ortogonal a \mathbf{u} .

Otro punto interesante que vale la pena discutir es el método que usamos para resolver sistemas de ecuaciones. Se consideraron ciertos múltiplos de ellas y luego se sumaron. Es decir, para ciertas α y β en \mathbb{R} se obtiene, de las dos ecuaciones dadas, una nueva de la forma

$$(\alpha \mathbf{u} + \beta \mathbf{v}) \cdot \mathbf{x} = \alpha e + \beta f.$$



Ésta es la ecuación de otra recta. La propiedad importante que cumple es que si \mathbf{p} satisface las dos ecuaciones dadas, entonces también satisface esta última por las propiedades de nuestras operaciones. De tal manera que esta nueva ecuación define una recta que pasa por el punto de intersección \mathbf{p} . El método de eliminar una variable consiste en encontrar la horizontal o la vertical que pasa por el punto

(la mitad del problema). Pero en general, todas las ecuaciones posibles de la forma anterior definen el *haz de rectas concurrentes* por el punto de intersección \mathbf{p} cuando \mathbf{u} y \mathbf{v} no son paralelos, pues los vectores $\alpha \mathbf{u} + \beta \mathbf{v}$ (al variar α y β) son más que suficientes para definir todas las posibles direcciones.

Por último, y aunque ya hayamos demostrado la parte de existencia, concluyamos con el Quinto usando al Teorema 1.17.

Teorema 1.18 *Dada una recta ℓ y un punto \mathbf{p} fuera de ella en el plano, existe una única recta ℓ' que pasa por \mathbf{p} y no intersecta a ℓ .*

Demostración. Existe un vector $\mathbf{u} \neq \mathbf{0}$ y una constante $c \in \mathbb{R}$, tales que ℓ está dada por la ecuación $\mathbf{u} \cdot \mathbf{x} = c$. Sea

$$\ell' : \quad \mathbf{u} \cdot \mathbf{x} = \mathbf{u} \cdot \mathbf{p}.$$

Entonces $\mathbf{p} \in \ell'$ porque satisface la ecuación, $\ell \cap \ell' = \emptyset$ porque $\mathbf{u} \cdot \mathbf{p} \neq c$ (pues $\mathbf{p} \notin \ell$), y cualquier otra recta que pasa por \mathbf{p} intersecta a ℓ pues su vector normal no es paralelo a \mathbf{u} . \square

EJERCICIO 1.65 Encuentra la intersección de las rectas (a), (b), (c) y (d) del Ejercicio 1.59.

EJERCICIO 1.66 Demuestra que dos vectores \mathbf{u} y \mathbf{v} en \mathbb{R}^2 son linealmente independientes si y solo si $\mathbf{u}^\perp \cdot \mathbf{v} \neq 0$.

EJERCICIO 1.67 Demuestra que dos vectores \mathbf{u} y \mathbf{v} en \mathbb{R}^2 son linealmente independientes si y sólo si $\mathbb{R}^2 = \langle \mathbf{u}, \mathbf{v} \rangle$.

EJERCICIO 1.68 A cada terna de números reales $(\mathbf{a}, \mathbf{b}, \mathbf{c})$, asóciate la ecuación $\mathbf{a}x + \mathbf{b}y = \mathbf{c}$.

i) ¿Cuáles son las ternas (como puntos en \mathbb{R}^3) a las que se le asocian rectas en \mathbb{R}^2 por su ecuación normal?

ii) Dada una recta en \mathbb{R}^2 , describe las ternas (como subconjunto de \mathbb{R}^3) que se asocian a ella.

iii) Dado un haz de rectas paralelas en \mathbb{R}^2 , describe las ternas (como subconjunto de \mathbb{R}^3) que se asocian a rectas de ese haz.

iv) Dado un haz de rectas concurrentes en \mathbb{R}^2 , describe las ternas (como subconjunto de \mathbb{R}^3) que se asocian a las rectas de ese haz.

1.8.2 Teoremas de concurrencia

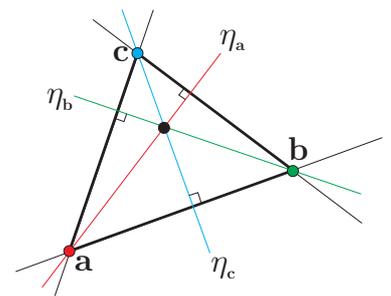
En esta sección, aplicamos la ecuación normal de las rectas para demostrar uno de los teoremas clásicos de concurrencia de líneas y dejamos el otro como ejercicio.

La *altura* de un triángulo es la recta que pasa por uno de sus vértices y es ortogonal al lado opuesto.

Teorema 1.19 *Las alturas de un triángulo son concurrentes.*

Demostración. Dado un triángulo con vértices \mathbf{a} , \mathbf{b} , \mathbf{c} , tenemos que la altura por el vértice \mathbf{a} , llamémosla $\eta_{\mathbf{a}}$ (“eta-sub-a”), está definida por

$$\eta_{\mathbf{a}} : \quad (\mathbf{c} - \mathbf{b}) \cdot \mathbf{x} = (\mathbf{c} - \mathbf{b}) \cdot \mathbf{a},$$



pues es ortogonal al lado que pasa por \mathbf{b} y \mathbf{c} , que tiene dirección $(\mathbf{c} - \mathbf{b})$, y pasa por el punto \mathbf{a} . Análogamente:

$$\begin{aligned}\eta_{\mathbf{b}} &: (\mathbf{a} - \mathbf{c}) \cdot \mathbf{x} = (\mathbf{a} - \mathbf{c}) \cdot \mathbf{b} \\ \eta_{\mathbf{c}} &: (\mathbf{b} - \mathbf{a}) \cdot \mathbf{x} = (\mathbf{b} - \mathbf{a}) \cdot \mathbf{c}.\end{aligned}$$

Obsérvese ahora que la suma (lado a lado) de dos de estas ecuaciones da precisamente el negativo de la tercera. Por ejemplo, sumando las dos primeras obtenemos

$$\begin{aligned}(\mathbf{c} - \mathbf{b}) \cdot \mathbf{x} + (\mathbf{a} - \mathbf{c}) \cdot \mathbf{x} &= (\mathbf{c} - \mathbf{b}) \cdot \mathbf{a} + (\mathbf{a} - \mathbf{c}) \cdot \mathbf{b} \\ (\mathbf{c} - \mathbf{b} + \mathbf{a} - \mathbf{c}) \cdot \mathbf{x} &= \mathbf{c} \cdot \mathbf{a} - \mathbf{b} \cdot \mathbf{a} + \mathbf{a} \cdot \mathbf{b} - \mathbf{c} \cdot \mathbf{b} \\ (\mathbf{a} - \mathbf{b}) \cdot \mathbf{x} &= (\mathbf{a} - \mathbf{b}) \cdot \mathbf{c}.\end{aligned}$$

Así que si $\mathbf{x} \in \eta_{\mathbf{a}} \cap \eta_{\mathbf{b}}$, entonces cumple las dos primeras ecuaciones y por tanto cumple su suma que es “menos” la ecuación de $\eta_{\mathbf{c}}$, y entonces $\mathbf{x} \in \eta_{\mathbf{c}}$. Así que las tres rectas pasan por el mismo punto. \square

EJERCICIO 1.69 Demuestra que las tres mediatrices de un triángulo son concurrentes (el punto en el que concurren se llama *circuncentro*). Donde la mediatriz de un segmento es su ortogonal que pasa por el punto medio.

EJERCICIO 1.70 Encuentra el circuncentro del triángulo con vértices $(1, 1)$, $(1, -1)$ y $(-2, -2)$. Haz el dibujo del triángulo y sus mediatrices.

1.8.3 Planos en el espacio II

Hemos definido las rectas en \mathbb{R}^n por su representación paramétrica y, en \mathbb{R}^2 , acabamos de ver que también tienen una representación normal (con base en una ecuación lineal). Es entonces importante hacer notar que este fenómeno sólo se da en dimensión 2. En \mathbb{R}^3 , que es el otro espacio que nos interesa, los planos (¡no las rectas!) son las que se definen por la ecuación normal. Veamos esto con cuidado.

Dado un vector $\mathbf{n} \in \mathbb{R}^3$ distinto de $\mathbf{0}$; sea, para cualquier $\mathbf{d} \in \mathbb{R}$,

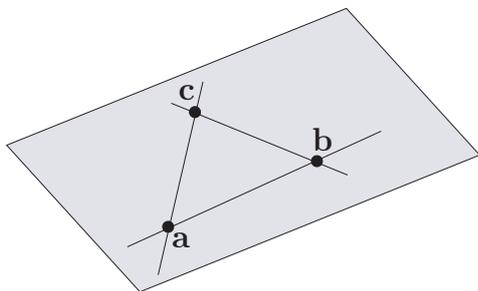
$$\Pi_{\mathbf{d}}: \mathbf{n} \cdot \mathbf{x} = \mathbf{d}$$

es decir, $\Pi_{\mathbf{d}} := \{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{n} \cdot \mathbf{x} = \mathbf{d}\}$. Llamamos a la constante \mathbf{d} pues ahora el vector normal es de la forma $\mathbf{n} = (\mathbf{a}, \mathbf{b}, \mathbf{c})$ y entonces la ecuación anterior se escribe en coordenadas como

$$\mathbf{a}x + \mathbf{b}y + \mathbf{c}z = \mathbf{d}$$

con el vector variable $\mathbf{x} = (x, y, z)$. Afirmamos que $\Pi_{\mathbf{d}}$ es un plano. Recuérdesse que definimos plano como el conjunto de combinaciones afines (o baricéntricas) de tres puntos no colineales.

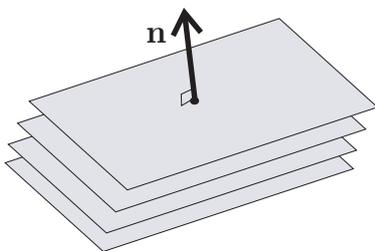
Veamos primero que si tres puntos \mathbf{a} , \mathbf{b} , \mathbf{c} satisfacen la ecuación $\mathbf{n} \cdot \mathbf{x} = d$, entonces los puntos del plano que generan también la satisfacen. Para esto, tomamos $\alpha, \beta, \gamma \in \mathbb{R}$ tales que $\alpha + \beta + \gamma = 1$, y entonces se tiene



$$\begin{aligned} \mathbf{n} \cdot (\alpha \mathbf{a} + \beta \mathbf{b} + \gamma \mathbf{c}) &= \alpha (\mathbf{n} \cdot \mathbf{a}) + \beta (\mathbf{n} \cdot \mathbf{b}) + \gamma (\mathbf{n} \cdot \mathbf{c}) \\ &= \alpha d + \beta d + \gamma d \\ &= (\alpha + \beta + \gamma) d = d. \end{aligned}$$

Esto demuestra que si $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \Pi_d$ entonces el plano (o si son colineales, la recta) que generan está contenido en Π_d . Faltaría ver que en Π_d hay tres puntos no colineales y luego demostrar que no hay más soluciones que las del plano que generan.

Pero mejor veámoslo desde otro punto de vista: el lineal en vez del baricéntrico. Afirmamos que los conjuntos Π_d , al variar d , son la familia de planos normales a \mathbf{n} . Para demostrar esto, hay que ver que Π_0 es un plano por el origen y que Π_d es un trasladado de Π_0 , es decir, Π_0 empujado por algún vector constante.



Esto último ya se hizo en esencia cuando vimos el caso de rectas en el plano, así que lo veremos primero para remarcar lo general de aquella demostración. Supongamos que $\mathbf{p} \in \mathbb{R}^3$ es tal que

$$\mathbf{n} \cdot \mathbf{p} = d,$$

es decir, que es una solución particular. Vamos a demostrar que

$$\Pi_d = \Pi_0 + \mathbf{p} := \{\mathbf{y} + \mathbf{p} \mid \mathbf{y} \in \Pi_0\}; \tag{1.10}$$

es decir, que si $\mathbf{x} \in \Pi_d$ entonces $\mathbf{x} = \mathbf{y} + \mathbf{p}$ para algún $\mathbf{y} \in \Pi_0$; y al revés, que si $\mathbf{y} \in \Pi_0$ entonces $\mathbf{x} = \mathbf{y} + \mathbf{p} \in \Pi_d$.

Dado $\mathbf{x} \in \Pi_d$ (i.e., tal que $\mathbf{n} \cdot \mathbf{x} = d$), sea $\mathbf{y} := \mathbf{x} - \mathbf{p}$. Como

$$\mathbf{n} \cdot \mathbf{y} = \mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = \mathbf{n} \cdot \mathbf{x} - \mathbf{n} \cdot \mathbf{p} = d - d = 0,$$

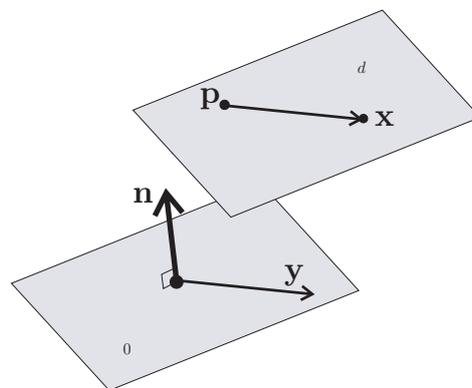
entonces $\mathbf{y} \in \Pi_0$ y es, por definición, tal que $\mathbf{x} = \mathbf{y} + \mathbf{p}$; por lo tanto $\Pi_d \subseteq \Pi_0 + \mathbf{p}$. Y al revés, si $\mathbf{y} \in \Pi_0$ y $\mathbf{x} = \mathbf{y} + \mathbf{p}$, entonces

$$\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot (\mathbf{y} + \mathbf{p}) = \mathbf{n} \cdot \mathbf{y} + \mathbf{n} \cdot \mathbf{p} = 0 + d = d$$

y por lo tanto $\mathbf{x} \in \Pi_d$. Hemos demostrado (1.10).

Nos falta ver que Π_0 es efectivamente un plano por el origen. Puesto que estamos suponiendo que $\mathbf{n} \neq \mathbf{0}$ entonces de la ecuación

$$ax + by + cz = 0$$



puede despejarse alguna variable en términos de las otras dos; sin pérdida de generalidad, digamos que es z , es decir, que $c \neq 0$. De tal forma que al dar valores arbitrarios a x y y , la fórmula nos da un valor de z , y por tanto un punto $\mathbf{x} \in \mathbb{R}^3$ que satisface la condición original $\mathbf{n} \cdot \mathbf{x} = 0$ —esto equivale a parametrizar las soluciones con \mathbb{R}^2 , con dos grados de libertad—. En particular hay una solución \mathbf{u} con $x = 1$ y $y = 0$ (a saber, $\mathbf{u} = (1, 0, -\mathbf{a}/c)$) y una solución \mathbf{v} con $x = 0$ y $y = 1$ (¿cuál es?). Como claramente los vectores \mathbf{u} y \mathbf{v} no son paralelos (tienen ceros en coordenadas distintas), generan linealmente todo un plano de soluciones (pues para cualquier $s, t \in \mathbb{R}$, se tiene que $\mathbf{n} \cdot (s\mathbf{u} + t\mathbf{v}) = s(\mathbf{n} \cdot \mathbf{u}) + t(\mathbf{n} \cdot \mathbf{v}) = 0$). Falta ver que no hay más soluciones que las que hemos descrito, pero esto se lo dejamos al estudiante para el siguiente ejercicio, que hay que comparar con este párrafo para entender cómo se llegó a su elegante planteamiento.

EJERCICIO 1.71 Sea $\mathbf{n} = (a, b, c)$ tal que $a \neq 0$. Demuestra que

$$\{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{n} \cdot \mathbf{x} = 0\} = \{s(-b, a, 0) + t(-c, 0, a) \mid s, t \in \mathbb{R}\}.$$

EJERCICIO 1.72 Describe el plano en \mathbb{R}^3 dado por la ecuación $x + y + z = 1$.

***EJERCICIO 1.73** Sea \mathbf{n} un vector no nulo en \mathbb{R}^n .

a) Demuestra que $V_0 := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{n} \cdot \mathbf{x} = 0\}$ es un *subespacio vectorial* de \mathbb{R}^n ; es decir, que la suma de vectores en V_0 se queda en V_0 y que el “alargamiento” de vectores en V_0 también está en V_0 .

b) Demuestra que para cualquier $d \in \mathbb{R}$, el conjunto $V_d := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{n} \cdot \mathbf{x} = d\}$ es un trasladado de V_0 ; es decir, que existe un $\mathbf{p} \in \mathbb{R}^n$ tal que $V_d = V_0 + \mathbf{p} = \{\mathbf{y} + \mathbf{p} \mid \mathbf{y} \in V_0\}$.

c) ¿Qué dimensión dirías que tiene V_d ? Argumenta un poco tu respuesta.

1.9 Norma y ángulos

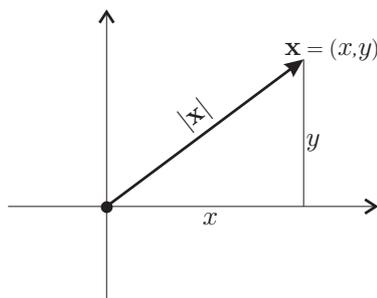
Del producto interior, obtendremos la noción de *norma* o *magnitud* de los vectores, que corresponde a la *distancia* del punto al origen.

Definición 1.9.1 Dado un vector $\mathbf{v} \in \mathbb{R}^n$, su *norma* (o *magnitud*) es el número real:

$$|\mathbf{v}| := \sqrt{\mathbf{v} \cdot \mathbf{v}},$$

de tal manera que la norma es una función de \mathbb{R}^n en \mathbb{R} .

Como $\mathbf{v} \cdot \mathbf{v} \geq 0$ (Teorema 1.11), tiene sentido tomar su raíz cuadrada (que a su vez se define como el número positivo tal que al elevarlo al cuadrado nos da el dado). Se tiene entonces la siguiente fórmula que es una definición equivalente de la norma y que será usada con mucho más frecuencia

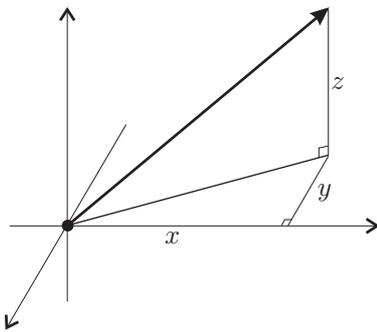


$$|\mathbf{v}|^2 = \mathbf{v} \cdot \mathbf{v}.$$

En \mathbb{R}^2 la norma se escribe, con coordenadas $\mathbf{v} = (x, y)$, como

$$|\mathbf{v}|^2 = x^2 + y^2.$$

Entonces $|\mathbf{v}|$ corresponde a la distancia euclidiana del origen al punto $\mathbf{v} = (x, y)$, pues de acuerdo con el Teorema de Pitágoras, x y y son lo que miden los catetos del triángulo rectángulo con hipotenusa \mathbf{v} . Aquí es donde resulta importante (por primera vez) que los ejes coordenados se tomen ortogonales, pues entonces la fórmula para calcular la distancia euclidiana al origen a partir de las coordenadas se hace sencilla.



En \mathbb{R}^3 , con coordenadas $\mathbf{v} = (x, y, z)$, la norma se escribe

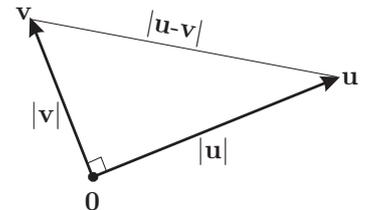
$$|\mathbf{v}|^2 = x^2 + y^2 + z^2,$$

que de nuevo, usando dos veces el Teorema de Pitágoras (en los triángulos de la figura) y que la nueva dirección z es ortogonal al plano x, y , corresponde a la *magnitud* del vector \mathbf{v} .

Demostremos primero el Teorema de Pitágoras para vectores en general. Este teorema fue la motivación básica para la definición de norma; sin embargo su uso ha sido sólo ése: como motivación, pues no lo hemos usado formalmente sino para ver que la norma corresponde a la noción euclidiana de distancia al origen (magnitud de vectores) cuando los ejes se toman ortogonales entre sí. Ahora veremos que al estar en el trasfondo de nuestras definiciones, éstas le hacen honor al hacerlo verdadero.

Teorema 1.20 (Pitágoras vectorial) Sean \mathbf{u} y \mathbf{v} dos vectores en \mathbb{R}^n . Entonces son perpendiculares si y sólo si

$$|\mathbf{u}|^2 + |\mathbf{v}|^2 = |\mathbf{u} - \mathbf{v}|^2.$$



Demostración. De las propiedades del producto interior (y la definición de norma) se obtiene

$$\begin{aligned} |\mathbf{u} - \mathbf{v}|^2 &= (\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v}) \\ &= \mathbf{u} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} - \mathbf{u} \cdot \mathbf{v} - \mathbf{v} \cdot \mathbf{u} \\ &= \mathbf{u} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} - 2(\mathbf{u} \cdot \mathbf{v}) \\ &= |\mathbf{u}|^2 + |\mathbf{v}|^2 - 2(\mathbf{u} \cdot \mathbf{v}). \end{aligned}$$

Por definición, \mathbf{u} y \mathbf{v} son perpendiculares si y sólo si $\mathbf{u} \cdot \mathbf{v} = 0$ y entonces el teorema se sigue de la igualdad anterior. \square

Así que nuestra definición de perpendicularidad (que el producto punto se anule) corresponde a que el Teorema de Pitágoras se vuelva cierto. Y entonces ahora tiene sentido nuestra manera informal de llamar, en \mathbb{R}^3 , a las soluciones de la ecuación $\mathbf{u} \cdot \mathbf{x} = 0$ “el plano normal a \mathbf{u} ”. Pues $\mathbf{u} \cdot \mathbf{x} = 0$ si y solo si los vectores \mathbf{u} y \mathbf{x} son los catetos de un triángulo que cumple el Teorema de Pitágoras y por tanto es rectángulo.

Las propiedades básicas de la norma se resumen en el siguiente teorema.

Teorema 1.21 *Para todos los vectores $\mathbf{v}, \mathbf{u} \in \mathbb{R}^n$ y para todo número real $t \in \mathbb{R}$ se tiene que:*

- i) $|\mathbf{v}| \geq 0$
- ii) $|\mathbf{v}| = 0 \Leftrightarrow \mathbf{v} = \mathbf{0}$
- iii) $|t\mathbf{v}| = |t||\mathbf{v}|$
- iv) $|\mathbf{u}| + |\mathbf{v}| \geq |\mathbf{u} + \mathbf{v}|$
- v) $|\mathbf{u}||\mathbf{v}| \geq |\mathbf{u} \cdot \mathbf{v}|$

Demostración. La primera afirmación es consecuencia inmediata de la definición de raíz cuadrada, y la segunda del Teorema 1.11. La tercera, donde también se usa ese teorema, es muy simple pero con una sutileza:

$$|t\mathbf{v}|^2 = (t\mathbf{v}) \cdot (t\mathbf{v}) = t^2(\mathbf{v} \cdot \mathbf{v}) = t^2 |\mathbf{v}|^2$$

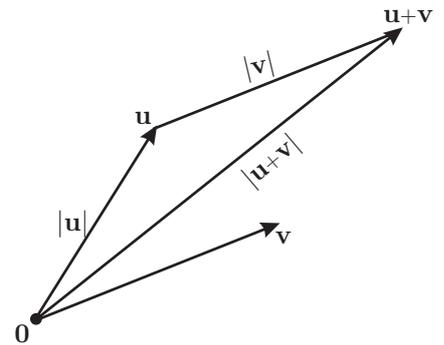
de donde, tomando raíz cuadrada, se deduce $|t\mathbf{v}| = |t||\mathbf{v}|$. Puesto que $\sqrt{t^2}$ no siempre es t sino su *valor absoluto* $|t|$, que es positivo siempre (y que podría definirse $|t| := \sqrt{t^2}$); nótese además que como $|\mathbf{v}| \geq 0$ entonces $\sqrt{|\mathbf{v}|^2} = |\mathbf{v}|$. Obsérvese que para $n = 1$ (es decir, en \mathbb{R}) la norma y el valor absoluto coinciden, así que usar la misma notación para ambos no causa ningún conflicto.

El inciso (iv) se conoce como “**la desigualdad del triángulo**”, pues dice que en el triángulo con vértices $\mathbf{0}$, \mathbf{u} y $\mathbf{u} + \mathbf{v}$, es más corto ir directamente de $\mathbf{0}$ a $\mathbf{u} + \mathbf{v}$ ($|\mathbf{u} + \mathbf{v}|$) que pasar primero por \mathbf{u} ($|\mathbf{u}| + |\mathbf{v}|$). Para demostrarla, nótese primero que como ambos lados de la desigualdad son no negativos, entonces ésta es equivalente a que la misma desigualdad se cumpla para los cuadrados, *i.e.*,

$$|\mathbf{u}| + |\mathbf{v}| \geq |\mathbf{u} + \mathbf{v}| \Leftrightarrow (|\mathbf{u}| + |\mathbf{v}|)^2 \geq |\mathbf{u} + \mathbf{v}|^2.$$

Y esta última desigualdad es equivalente a que $(|\mathbf{u}| + |\mathbf{v}|)^2 - |\mathbf{u} + \mathbf{v}|^2 \geq 0$. Así que debemos desarrollar el lado izquierdo:

$$\begin{aligned} (|\mathbf{u}| + |\mathbf{v}|)^2 - |\mathbf{u} + \mathbf{v}|^2 &= (|\mathbf{u}| + |\mathbf{v}|)^2 - (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) \\ &= |\mathbf{u}|^2 + |\mathbf{v}|^2 + 2|\mathbf{u}||\mathbf{v}| - (\mathbf{u} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} + 2(\mathbf{u} \cdot \mathbf{v})) \\ &= 2(|\mathbf{u}||\mathbf{v}| - (\mathbf{u} \cdot \mathbf{v})). \end{aligned}$$



Queda entonces por demostrar que $|\mathbf{u}||\mathbf{v}| \geq (\mathbf{u} \cdot \mathbf{v})$, pero esto se sigue del inciso (v), que es una afirmación más fuerte que demostraremos a continuación independientemente.

El inciso (v) se conoce como “**la desigualdad de Schwartz**”. Por un razonamiento análogo al del inciso anterior, bastará demostrar que $(|\mathbf{u}||\mathbf{v}|)^2 - |\mathbf{u} \cdot \mathbf{v}|^2$ es positivo. Lo haremos para \mathbb{R}^2 con coordenadas, dejando el caso de \mathbb{R}^3 , y de \mathbb{R}^n , como ejercicios. Supongamos entonces que $\mathbf{u} = (a, b)$ y $\mathbf{v} = (\alpha, \beta)$ para obtener

$$\begin{aligned} |\mathbf{u}|^2 |\mathbf{v}|^2 - |\mathbf{u} \cdot \mathbf{v}|^2 &= (a^2 + b^2)(\alpha^2 + \beta^2) - (a\alpha + b\beta)^2 \\ &= a^2\alpha^2 + b^2\beta^2 + a^2\beta^2 + b^2\alpha^2 \\ &\quad - (a^2\alpha^2 + b^2\beta^2 + 2a\alpha b\beta) \\ &= a^2\beta^2 - 2(a\beta)(b\alpha) + b^2\alpha^2 \\ &= (a\beta - b\alpha)^2 \geq 0; \end{aligned}$$

donde la última desigualdad se debe a que el cuadrado de cualquier número es no negativo. Lo cual demuestra la desigualdad de Schwartz, la del Triángulo y completa el Teorema. \square

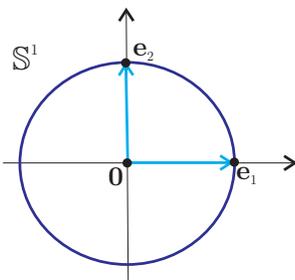
EJERCICIO 1.74 Demuestra que dos vectores \mathbf{u} y \mathbf{v} son perpendiculares si y sólo si $|\mathbf{u} + \mathbf{v}|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2$. Haz el dibujo.

EJERCICIO 1.75 Demuestra que dos vectores \mathbf{u} y \mathbf{v} son perpendiculares si y sólo si $|\mathbf{u} + \mathbf{v}| = |\mathbf{u} - \mathbf{v}|$. Haz el dibujo. Observa que este enunciado corresponde a que un paralelogramo es un rectángulo si y sólo si sus diagonales miden lo mismo.

EJERCICIO 1.76 Sean \mathbf{u} y \mathbf{v} dos vectores no nulos. Demuestra que tienen la misma norma si y sólo si $\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - \mathbf{v}$ son ortogonales.

EJERCICIO 1.77 Demuestra la desigualdad de Schwartz en \mathbb{R}^3 . ¿Puedes dar, o describir, la demostración en \mathbb{R}^n ?

1.9.1 El círculo unitario

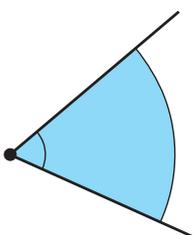
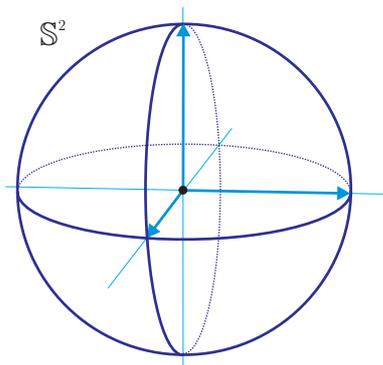


A los vectores que tienen norma igual a uno, se les llama *vectores unitarios*. Y al conjunto de todos los vectores unitarios en \mathbb{R}^2 se le llama el *círculo unitario* y se denota con \mathbb{S}^1 . La notación viene de la palabra *sphere* en inglés, que significa esfera; pues en general se puede definir a la *esfera de dimensión n* como

$$\mathbb{S}^n = \{ \mathbf{x} \in \mathbb{R}^{n+1} \mid |\mathbf{x}| = 1 \}.$$

Nótese que el exponente se refiere a la dimensión de la esfera en sí, y que ésta necesita una dimensión más para “vivir”. Así, la esfera de dimensión 2 vive en \mathbb{R}^3 , y es la representación abstracta de las pompas de jabón. Más adelante la estudiaremos con cuidado. Por lo pronto nos interesa el círculo unitario.

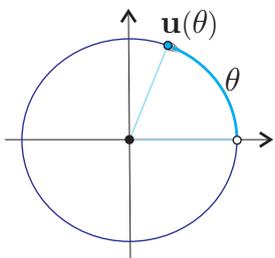
Los puntos de \mathbb{S}^1 corresponden a los *ángulos*, y a estos los mediremos con radianes. La definición formal o muy precisa de ángulo involucra necesariamente nociones de cálculo que no entran en este libro. Sin embargo, es un concepto muy intuitivo y de esa intuición nos valdremos. Un *ángulo* es un sector radial del círculo unitario, aunque se denota gráficamente con un arco pequeño cerca del centro para indicar que no depende realmente del círculo de referencia, sino que es más bien un sector de cualquier círculo concéntrico. Es costumbre partir el círculo completo en 360 sectores iguales llamados *grados*. Resulta conveniente usar el número 360, pues $360 = 2^3 \cdot 3^2 \cdot 5$ tiene muchos divisores, así que el círculo se puede partir en cuatro sectores iguales de 90 grados (denotado 90° y llamado *ángulo recto*) o en 12 de 30° que corresponden a las horas del reloj, etc. Pero la otra manera de medirlos, que no involucra la convención de escoger 360 para la vuelta entera, es por la longitud del arco de círculo que abarcan; a esta medida se le conoce como



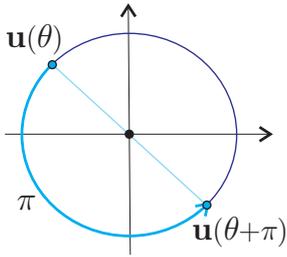
en radianes. Un problema clásico que resolvieron los griegos con admirable precisión fue medir la *circunferencia* del círculo unitario (de radio uno), es decir, cuánto mide un hilo “untado” en el círculo. El resultado, como todos sabemos, es el doble del famoso número $\pi = 3.14159\dots$, donde los puntos suspensivos indican que su expresión decimal sigue infinitamente pues no es racional. Otra manera de entender los radianes es cinemática y es la que usaremos a continuación para establecer notación; es la versión teórica del movimiento de la piedra en una honda justo antes de lanzarla.

Si una partícula viaja dentro de \mathbb{S}^1 a velocidad constante 1, partiendo del punto $\mathbf{e}_1 = (1, 0)$ y en dirección contraria a las manecillas del reloj (es decir, saliendo hacia arriba con vector velocidad $\mathbf{e}_2 = (0, 1)$), entonces en un tiempo θ estará en un punto que llamaremos $\mathbf{u}(\theta)$; en el tiempo $\pi/2$ estará en el $\mathbf{e}_2 = (0, 1)$ (es decir, $\mathbf{u}(\pi/2) = \mathbf{e}_2$), en el tiempo π en el $(-1, 0)$ y en el 2π estará de regreso para empezar de nuevo. El tiempo aquí, que estamos midiendo con el parámetro θ , corresponde precisamente a los *radianes*, pues al viajar a velocidad constante 1 el tiempo es igual a la distancia recorrida. Podemos también dar sentido a $\mathbf{u}(\theta)$ para θ negativa pensando que la partícula viene dando vueltas desde siempre (tiempo infinito negativo) de tal manera que en el tiempo 0 pasa justo por \mathbf{e}_1 , es decir, que

$$\mathbf{u}(0) = (1, 0).$$



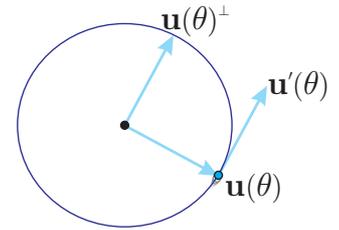
Nuestra suposición básica es que la función $\mathbf{u} : \mathbb{R} \rightarrow \mathbb{S}^1$, que hemos descrito, está bien definida. Se cumple entonces que



$$\mathbf{u}(\theta) = \mathbf{u}(\theta + 2\pi),$$

pues en el tiempo 2π la partícula da justo una vuelta (la longitud del círculo mide lo mismo independientemente de dónde empezemos a medir). Y también que $\mathbf{u}(\theta + \pi) = -\mathbf{u}(\theta)$, pues π es justo el ángulo que da media vuelta.

Observemos ahora que al pedir que la partícula viaje en el círculo unitario con velocidad constante 1 entonces su *vector velocidad*, que incluye ahora dirección además de magnitud, es tangente a \mathbb{S}^1 (la piedra de la honda sale por la tangente), y es fácil ver que entonces es perpendicular al vector de posición $\mathbf{u}(\theta)$, pero más precisamente es justo su compadre ortogonal porque va girando en “su” dirección, “hacia él”. Usando la notación de derivadas, podríamos escribir



$$\mathbf{u}'(\theta) = \mathbf{u}(\theta)^\perp.$$

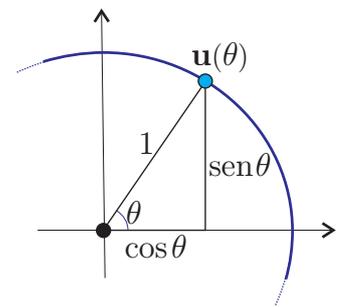
Hay que remarcar que al medir ángulos con radianes, si bien ganamos en naturalidad matemática, es inevitable la ambigüedad de que a los ángulos no corresponda un número único. Pues θ y $\theta + 2n\pi$, para cualquier $n \in \mathbb{Z}$, determinan al mismo ángulo, es decir $\mathbf{u}(\theta) = \mathbf{u}(\theta + 2n\pi)$. Podemos exigir que θ esté en el intervalo entre 0 y 2π , o bien, que resulta más agradable geoméricamente, en el intervalo de $-\pi$ a π , para reducir la ambigüedad sólo a los extremos. Pero al sumar o restar ángulos, inevitablemente nos saldremos de este intervalo y habrá que volver a ajustar.

Funciones trigonométricas

Recordemos ahora que \mathbb{S}^1 es un subconjunto de \mathbb{R}^2 , así que el punto $\mathbf{u}(\theta)$ tiene dos coordenadas precisas. Éstas están dadas por las importantísimas *funciones trigonométricas* coseno y seno. Más precisamente, podemos definir

$$(\cos \theta, \text{sen } \theta) := \mathbf{u}(\theta),$$

es decir, $\cos \theta$ y $\text{sen } \theta$ son las coordenadas cartesianas del punto $\mathbf{u}(\theta) \in \mathbb{S}^1$. Entonces el coseno y el seno corresponden respectivamente al cateto adyacente y al cateto opuesto de un triángulo rectángulo con hipotenusa 1 y ángulo θ , como se ve en trigonometría de secundaria. También pensarse que si damos por establecidas las funciones seno y coseno, entonces la función $\mathbf{u}(\theta)$ está dada por la ecuación anterior. Hay que remarcar que la pertenencia $\mathbf{u}(\theta) \in \mathbb{S}^1$ equivale entonces a la ecuación



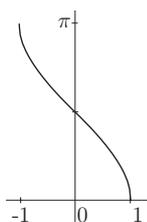
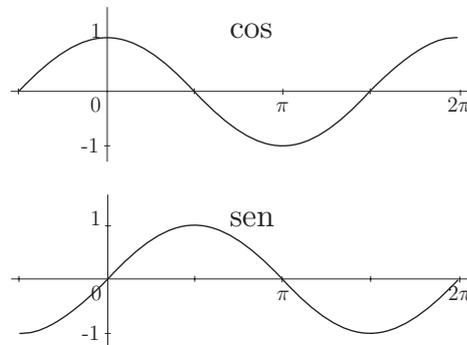
$$\text{sen}^2 \theta + \cos^2 \theta = 1.$$

Las propiedades de la función \mathbf{u} que hemos enumerado, traducidas a sus coordenadas (por ejemplo, la del vector velocidad se traduce a $\cos' \theta = -\operatorname{sen} \theta$ y $\operatorname{sen}' \theta = \cos \theta$) son suficientes para definir a las funciones \cos y sen , y en adelante las daremos por un hecho. Sus bien conocidas gráficas aparecen en la figura.

Observemos, por último, que cualquiera de las funciones coseno o seno casi determinan al ángulo, pues de la ecuación anterior se pueden despejar con la ambigüedad de un signo al tomar raíz cuadrada, es decir,

$$\operatorname{sen} \theta = \pm \sqrt{1 - \cos^2 \theta}.$$

Dicho de otra manera, dado x en el intervalo de -1 a 1 (escrito $x \in [-1, 1]$,⁴ o bien $-1 \leq x \leq 1$) tenemos dos posibles puntos en \mathbb{S}^1 con esa ordenada, a saber $(x, \sqrt{1-x^2})$ y $(x, -\sqrt{1-x^2})$. Entonces con la ambigüedad de un signo podemos determinar el ángulo del cuál x es el coseno. Si escogemos la parte superior del círculo unitario obtenemos una función, llamada *arcocoseno* y leída “el arco cuyo coseno es ...”



$$\arccos : [-1, 1] \rightarrow [0, \pi],$$

y definida por

$$\cos(\arccos(x)) = x.$$

La ambigüedad surge al componer en el otro sentido, pues para $\theta \in [-\pi, \pi]$ se tiene $\arccos(\cos(\theta)) = \pm\theta$.

Obsérvese que entonces la mitad superior del círculo unitario se describe como la gráfica de la función $x \mapsto \operatorname{sen}(\arccos(x))$.

EJERCICIO 1.78 Demuestra (usando que definimos $\mathbf{u}(\theta) \in \mathbb{S}^1$) que para cualquier $\theta \in \mathbb{R}$ se cumplen

$$\begin{aligned} -1 &\leq \cos \theta \leq 1 \\ -1 &\leq \operatorname{sen} \theta \leq 1 \end{aligned}$$

EJERCICIO 1.79 Demuestra que para cualquier $\theta \in \mathbb{R}$ se cumplen

$$\begin{aligned} \cos(\theta + \pi/2) &= -\operatorname{sen} \theta & \cos(\theta + \pi) &= -\cos \theta & \cos(-\theta) &= \cos \theta \\ \operatorname{sen}(\theta + \pi/2) &= \cos \theta & \operatorname{sen}(\theta + \pi) &= -\operatorname{sen} \theta & \operatorname{sen}(-\theta) &= -\operatorname{sen} \theta \end{aligned}$$

EJERCICIO 1.80 Construye una tabla con los valores de las funciones \cos y sen para los ángulos $0, \pi/6, \pi/4, \pi/3, \pi/2, 2\pi/3, 3\pi/4, 5\pi/6, \pi, 5\pi/4, 3\pi/2, -\pi/3, -\pi/4$, y dibuja los correspondientes vectores unitarios.

⁴Dados $a, b \in \mathbb{R}$, con $a \leq b$, denotamos por $[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$ al *intervalo cerrado* entre a y b .

EJERCICIO 1.81 ¿Cuál es el conjunto de números que usan las coordenadas de las horas del reloj?

1.9.2 Coordenadas polares

El círculo unitario \mathbb{S}^1 (cuyos puntos hemos identificado con los ángulos tomando su ángulo con el vector base $\mathbf{e}_1 = (1, 0)$) tiene justo un representante de cada posible dirección en \mathbb{R}^2 , donde ahora una dirección y su opuesta son diferentes (aunque sean, según nuestra definición, paralelas). De tal manera que a cualquier punto de \mathbb{R}^2 que no sea el origen se llega viajando en una (y exactamente una) de estas direcciones. Ésta es la idea central de las coordenadas polares, que en muchas situaciones son más naturales, o útiles, para identificar los puntos del plano. Por ejemplo, son las que intuitivamente usa un cazador: apuntar —decidir una dirección— es la primera coordenada y luego, dependiendo de la distancia a la que vuela el pato —la segunda coordenada— tiene que ajustar el tiro para que las trayectorias de pato y perdigones se intersecten.

Sea \mathbf{x} cualquier vector no nulo en \mathbb{R}^2 . Entonces $|\mathbf{x}| \neq 0$ y tiene sentido tomar el vector $(|\mathbf{x}|^{-1})\mathbf{x}$, que es unitario pues como $|\mathbf{x}| > 0$ implica que $|\mathbf{x}|^{-1} > 0$, entonces tenemos

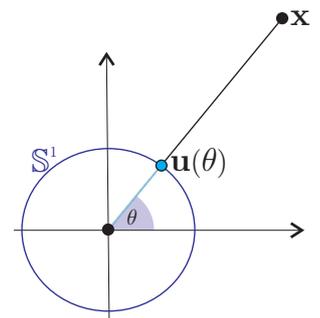
$$| |\mathbf{x}|^{-1} \mathbf{x} | = | |\mathbf{x}|^{-1} | |\mathbf{x}| = |\mathbf{x}|^{-1} |\mathbf{x}| = 1.$$

De aquí que exista un único $\mathbf{u}(\theta) \in \mathbb{S}^1$ (aunque θ no es único como real, sí lo es como ángulo), tal que

$$\mathbf{u}(\theta) = |\mathbf{x}|^{-1} \mathbf{x}$$

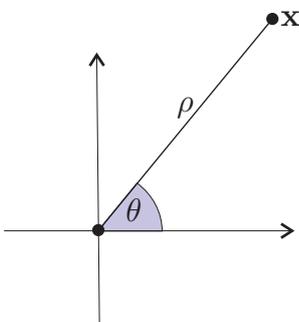
y claramente se cumple que

$$\mathbf{x} = |\mathbf{x}| \mathbf{u}(\theta).$$



A $\mathbf{u}(\theta)$ se le referirá como la *dirección* de \mathbf{x} . A la pareja $(\theta, |\mathbf{x}|)$ se le llama las *coordenadas polares* del vector \mathbf{x} . Se usa el término “coordenadas” pues determinan al vector, es decir, si nos dan una pareja (θ, ρ) con $\rho \geq 0$, obtenemos un vector

$$\mathbf{x} = \rho \mathbf{u}(\theta)$$



con magnitud ρ y dirección $\mathbf{u}(\theta)$, o bien, *ángulo* θ . Obsérvese que sólo con el origen hay ambigüedad: si $\rho = 0$ en la fórmula anterior, para cualquier valor de θ obtenemos el origen; éste no tiene un ángulo definido. Así que al hablar de coordenadas polares supondremos que el vector en cuestión es no nulo. Por nuestra convención de $\mathbf{u}(\theta)$, lo que representa θ es el ángulo respecto al eje x en su dirección positiva.

EJERCICIO 1.82 Da las coordenadas (cartesianas) de los vectores cuyas coordenadas polares son $(\pi/2, 3)$, $(\pi/4, \sqrt{2})$, $(\pi/6, 2)$, $(\pi/6, 1)$.

EJERCICIO 1.83 Si tomamos (θ, ρ) como coordenadas polares, describe los subconjuntos de \mathbb{R}^2 definidos por las ecuaciones $\theta = \text{cte}$ (“ θ igual a una constante”) y $\rho = \text{cte}$.

1.9.3 Ángulo entre vectores

Podemos usar coordenadas polares para definir el ángulo entre vectores en general.

Definición 1.9.2 Dados \mathbf{x} y \mathbf{y} , dos vectores no nulos en \mathbb{R}^2 , sean $(\alpha, |\mathbf{x}|)$ y $(\beta, |\mathbf{y}|)$ sus coordenadas polares respectivamente. El *ángulo de \mathbf{x} a \mathbf{y}* es

$$\overrightarrow{\text{ang}}(\mathbf{x}, \mathbf{y}) = \beta - \alpha.$$

Obsérvese que el ángulo tiene dirección (de ahí que le hayamos puesto la flechita de sombrero), ya que claramente

$$\overrightarrow{\text{ang}}(\mathbf{x}, \mathbf{y}) = -\overrightarrow{\text{ang}}(\mathbf{y}, \mathbf{x}).$$

Pues $\overrightarrow{\text{ang}}(\mathbf{x}, \mathbf{y})$ corresponde al movimiento angular que nos lleva de la dirección de \mathbf{x} a la de \mathbf{y} , y $\overrightarrow{\text{ang}}(\mathbf{y}, \mathbf{x})$ es justo el movimiento inverso. Así que podemos resumir diciendo que las coordenadas polares de un vector \mathbf{x} (recuérdese que al aplicar el término ya suponemos que es no nulo) es la pareja

$$(\overrightarrow{\text{ang}}(\mathbf{e}_1, \mathbf{x}), |\mathbf{x}|).$$

Con mucha frecuencia van a aparecer ángulos no dirigidos, pensados como un sector del círculo unitario sin principio ni fin determinados. Para ellos usaremos la notación $\text{ang}(\mathbf{x}, \mathbf{y})$ (sin la flecha), llamándolo el *ángulo entre \mathbf{x} y \mathbf{y}* . Aunque intuitivamente es claro lo que queremos (piénsese en una porción de una pizza), su definición es un poco más laboriosa. Si tomamos dos vectores unitarios en \mathbb{S}^1 , estos parten el círculo en dos sectores, uno de ellos es menor que medio círculo y ése es el ángulo entre ellos. Así que debemos pedir que

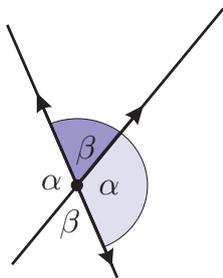
$$0 \leq \text{ang}(\mathbf{x}, \mathbf{y}) \leq \pi.$$

Y es claro que escogiendo a α y a β (los ángulos de \mathbf{x} y \mathbf{y} , respectivamente) de manera adecuada se tiene que $\text{ang}(\mathbf{x}, \mathbf{y}) := |\beta - \alpha| \in [0, \pi]$. Así que, por ejemplo, $\text{ang}(\mathbf{x}, \mathbf{y}) = 0$ significa que \mathbf{x} y \mathbf{y} apuntan en la misma dirección, mientras que $\text{ang}(\mathbf{x}, \mathbf{y}) = \pi$ significa que \mathbf{x} y \mathbf{y} apuntan en direcciones contrarias, y se tiene que $\text{ang}(\mathbf{x}, \mathbf{y}) = \pi/2$ cuando son perpendiculares.

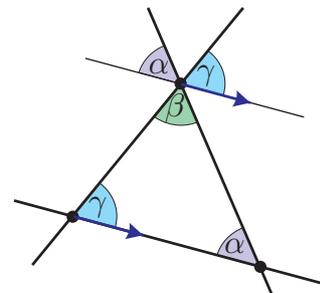
EJERCICIO 1.84 ¿Cuál es el ángulo de \mathbf{x} a \mathbf{y} , cuando a) $\mathbf{x} = (1, 0)$, $\mathbf{y} = (0, -2)$; b) $\mathbf{x} = (1, 1)$, $\mathbf{y} = (0, 1)$; c) $\mathbf{x} = (\sqrt{3}, 2)$, $\mathbf{y} = (1, 0)$.

Suma de ángulos

Si ya definimos ángulo entre vectores, resulta natural definir *ángulo entre dos rectas* como el ángulo entre sus vectores direccionales. Pero hay una ambigüedad pues podemos escoger direcciones opuestas para una misma recta. Esta ambigüedad equivale a que dos rectas que se intersectan definen cuatro sectores o “ángulos” que se agrupan naturalmente en dos parejas opuestas por el vértice que miden lo mismo; y estos dos ángulos son *complementarios*, es decir, suman π . No vale la pena entrar en los detalles formales de esto usando vectores direccionales pues sólo se confunde lo obvio, que desde muy temprana edad sabemos. Pero sí podemos delinear rápidamente la demostración de que la suma de los ángulos de un triángulo siempre es π , que para Euclides tuvo que ser axioma (disfrazado o no, pues recuérdese que es equivalente al Quinto).



Un triángulo son tres rectas que se intersectan dos a dos (pero no las tres), y en cada vértice (la intersección de dos de las rectas) define un ángulo interno; llamémoslos α , β y γ . Al trazar la paralela a un lado del triángulo que pasa por el vértice opuesto (usando el Quinto), podemos medir ahí los tres ángulos, pues por nuestra demostración del Quinto esta paralela tiene el mismo vector direccional que el lado original, y nuestra definición de ángulo entre líneas usa la de ángulo entre vectores direccionales. Finalmente, hay que observar que si en ese vértice cambiamos uno de los ángulos por su opuesto, los tres ángulos se juntan para formar el ángulo de un vector a su opuesto, es decir, suman π . Hemos delineado la demostración del siguiente teorema clásico.

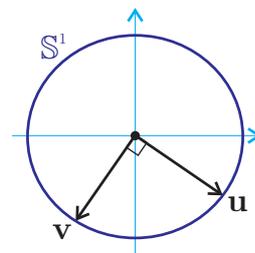


Teorema 1.22 *Los ángulos internos de un triángulo suman π .* □

1.10 Bases ortonormales

Definición 1.10.1 Una *base ortonormal* de \mathbb{R}^2 es una pareja de vectores unitarios perpendiculares.

Por ejemplo, la *base canónica* $\mathbf{e}_1 = (1, 0)$, $\mathbf{e}_2 = (0, 1)$ es una base ortonormal, y para cualquier $\mathbf{u} \in \mathbb{S}^1$ se tiene que $\mathbf{u}, \mathbf{u}^\perp$ es una base ortonormal, pues también $|\mathbf{u}^\perp| = 1$. Obsérvese además que si \mathbf{u} y \mathbf{v} son una base ortonormal entonces ambos están en \mathbb{S}^1 y además $\mathbf{v} = \mathbf{u}^\perp$ o $\mathbf{v} = -\mathbf{u}^\perp$. Por el momento, su importancia radica en que es muy fácil escribir cualquier vector como combinación lineal de una base ortonormal:



Teorema 1.23 Sean \mathbf{u} y \mathbf{v} una base ortonormal de \mathbb{R}^2 , entonces para cualquier $\mathbf{x} \in \mathbb{R}^2$ se tiene que

$$\mathbf{x} = (\mathbf{x} \cdot \mathbf{u}) \mathbf{u} + (\mathbf{x} \cdot \mathbf{v}) \mathbf{v}.$$

Demostración. Suponemos que \mathbf{x} está dado, y vamos a resolver el sistema de ecuaciones

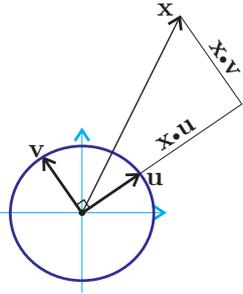
$$s \mathbf{u} + t \mathbf{v} = \mathbf{x} \tag{1.11}$$

con incógnitas t, s . Tomando el producto interior con \mathbf{u} en la ecuación anterior obtenemos

$$s (\mathbf{u} \cdot \mathbf{u}) + t (\mathbf{v} \cdot \mathbf{u}) = \mathbf{x} \cdot \mathbf{u}.$$

Pero $\mathbf{u} \cdot \mathbf{u} = 1$ pues \mathbf{u} es unitario y $\mathbf{v} \cdot \mathbf{u} = 0$ ya que \mathbf{u} y \mathbf{v} son perpendiculares, así que

$$s = \mathbf{x} \cdot \mathbf{u}.$$



Análogamente, tomando el producto interior por \mathbf{v} , se obtiene que $t = \mathbf{x} \cdot \mathbf{v}$. De tal manera que el sistema (1.11) sí tiene solución y es la que asegura el teorema. \square

Vale la pena observar que el truco que acabamos de usar para resolver un sistema de dos ecuaciones con dos incógnitas no es nuevo, lo usamos en la Sección 1.7.1 al tomar el producto interior con el compadre ortogonal; la diferencia es que aquí el otro vector de la base ortonormal está dado.

Como corolario de este teorema obtenemos la interpretación geométrica del producto interior de vectores unitarios.

Corolario 1.24 Si $\mathbf{u}, \mathbf{v} \in \mathbb{S}^1$ y α es el ángulo entre ellos, i.e., $\alpha = \text{ang}(\mathbf{u}, \mathbf{v})$, entonces

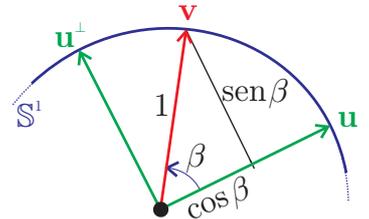
$$\mathbf{u} \cdot \mathbf{v} = \cos \alpha.$$

Demostración. Sea $\beta = \overrightarrow{\text{ang}}(\mathbf{u}, \mathbf{v})$ el ángulo de \mathbf{u} a \mathbf{v} , de tal manera que, escogiéndolos adecuadamente, se tiene que $\alpha = \pm\beta$. Usando el teorema anterior con la base ortonormal $\mathbf{u}, \mathbf{u}^\perp$ y el vector \mathbf{v} (en vez de \mathbf{x}) se obtiene

$$\mathbf{v} = (\mathbf{v} \cdot \mathbf{u}) \mathbf{u} + (\mathbf{v} \cdot \mathbf{u}^\perp) \mathbf{u}^\perp.$$

Pero también es claro geoméricamente que

$$\mathbf{v} = \cos \beta \mathbf{u} + \sin \beta \mathbf{u}^\perp.$$



Puesto que $\cos \beta = \cos \alpha$, pues $\alpha = \pm\beta$, el corolario se sigue de que la solución de un sistema con determinante no cero es única (en nuestro caso el determinante del sistema $\mathbf{v} = s \mathbf{u} + t \mathbf{u}^\perp$ es $\det(\mathbf{u}, \mathbf{u}^\perp) = \mathbf{u}^\perp \cdot \mathbf{u} = 1$). \square

EJERCICIO 1.85 Demuestra que si \mathbf{u} , \mathbf{v} y \mathbf{w} son una *base ortonormal* de \mathbb{R}^3 (es decir, si los tres son unitarios y dos a dos son perpendiculares), entonces para cualquier $\mathbf{x} \in \mathbb{R}^3$ se tiene que

$$\mathbf{x} = (\mathbf{x} \cdot \mathbf{u}) \mathbf{u} + (\mathbf{x} \cdot \mathbf{v}) \mathbf{v} + (\mathbf{x} \cdot \mathbf{w}) \mathbf{w}.$$

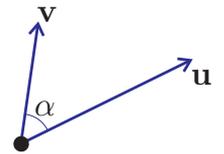
EJERCICIO 1.86 Escribe los vectores $(1, 0)$, $(0, 1)$, $(2, 1)$ y $(-1, 3)$ como combinación lineal de $\mathbf{u} = (3/5, 4/5)$ y $\mathbf{v} = (4/5, -3/5)$, es decir, como $t\mathbf{u} + s\mathbf{v}$ con t y s apropiadas.

1.10.1 Fórmula geométrica del producto interior

Como consecuencia del corolario anterior obtenemos un resultado importante que nos dice que el producto interior puede definirse en términos de la norma y el ángulo entre vectores.

Teorema 1.25 Sean \mathbf{u} y \mathbf{v} dos vectores en \mathbb{R}^2 y α el ángulo entre ellos, entonces

$$\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}| |\mathbf{v}| \cos \alpha.$$



Demostración. Si $\mathbf{u} = \mathbf{0}$ (o $\mathbf{v} = \mathbf{0}$) la igualdad anterior se cumple (pues ambos lados son 0); podemos suponer entonces que ambos son distintos de cero. Ahora podemos reescalar a \mathbf{u} (y a \mathbf{v}) para ser unitarios (pues tiene sentido $(|\mathbf{u}|^{-1})\mathbf{u}$ que es un vector unitario) manteniendo el ángulo entre ellos. Y por el Corolario 1.24,

$$\cos \alpha = \left(\frac{1}{|\mathbf{u}|}\right)\mathbf{u} \cdot \left(\frac{1}{|\mathbf{v}|}\right)\mathbf{v} = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}| |\mathbf{v}|},$$

de donde se sigue inmediatamente el teorema. □

De esta fórmula geométrica para el producto interior, se obtiene que $|\mathbf{u} \cdot \mathbf{v}| = |\cos \alpha| |\mathbf{u}| |\mathbf{v}|$, así que la desigualdad de Schwartz ($|\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}| |\mathbf{v}|$) es equivalente a $|\cos \alpha| \leq 1$.

EJERCICIO 1.87 Demuestra la *ley de los cosenos* para vectores, es decir, que dados dos vectores $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ con ángulo α entre ellos, se cumple que

$$|\mathbf{u} - \mathbf{v}|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2 - 2|\mathbf{u}| |\mathbf{v}| \cos \alpha.$$

Observa que el Teorema de Pitágoras es un caso especial.

EJERCICIO 1.88 Encuentra la distancia de los puntos $\mathbf{p}_1 = (0, 5)$, $\mathbf{p}_2 = (4, 0)$ y $\mathbf{p}_3 = (-1, 1)$ a la recta $\ell = \{t(4, 3) \mid t \in \mathbb{R}\}$.

1.10.2 El caso general

Puesto que el Teorema 1.23 y su Corolario 1.24 se demostraron para \mathbb{R}^2 , puede dudarse de la validez de los resultados en el caso general de \mathbb{R}^n , pero nos interesa establecer el caso de $n = 3$. Vale la pena entonces hacer unas aclaraciones y ajustes a manera de repaso. Lo que realmente se usó del Teorema 1.23 fue que cuando \mathbf{u} es un vector unitario (en cualquier dimensión ahora) entonces $\mathbf{x} \cdot \mathbf{u}$ tiene un significado geométrico muy preciso: es lo que hay que viajar en la dirección \mathbf{u} para que de ahí, \mathbf{x} quede en dirección ortogonal a \mathbf{u} . Podemos precisarlo de la siguiente manera.

Lema 1.26 *Sea $\mathbf{u} \in \mathbb{R}^n$ un vector unitario (es decir, tal que $|\mathbf{u}| = 1$). Entonces para cualquier $\mathbf{x} \in \mathbb{R}^n$ existe un vector \mathbf{y} perpendicular a \mathbf{u} y tal que*

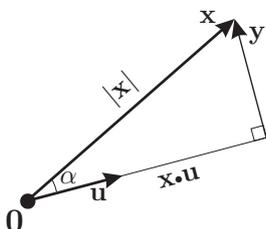
$$\mathbf{x} = (\mathbf{x} \cdot \mathbf{u}) \mathbf{u} + \mathbf{y}.$$

Demostración. Es muy fácil, declaramos $\mathbf{y} := \mathbf{x} - (\mathbf{x} \cdot \mathbf{u}) \mathbf{u}$ y basta ver que $\mathbf{y} \cdot \mathbf{u} = 0$:

$$\begin{aligned} \mathbf{y} \cdot \mathbf{u} &= (\mathbf{x} - (\mathbf{x} \cdot \mathbf{u}) \mathbf{u}) \cdot \mathbf{u} \\ &= \mathbf{x} \cdot \mathbf{u} - (\mathbf{x} \cdot \mathbf{u})(\mathbf{u} \cdot \mathbf{u}) = \mathbf{x} \cdot \mathbf{u} - \mathbf{x} \cdot \mathbf{u} = 0. \end{aligned}$$

□

Si pensamos ahora en el plano generado por \mathbf{u} y \mathbf{x} (en el que también se encuentra \mathbf{y}) y dentro de él en el triángulo rectángulo $\mathbf{0}, (\mathbf{x} \cdot \mathbf{u}) \mathbf{u}, \mathbf{x}$, se obtiene que



$$\cos \alpha = \frac{\mathbf{x} \cdot \mathbf{u}}{|\mathbf{x}|},$$

donde α es el ángulo entre los vectores \mathbf{u} y \mathbf{x} . Y de aquí, el Teorema 1.25 para \mathbb{R}^n se sigue inmediatamente al permitir que \mathbf{u} sea no necesariamente unitario.

Si nos ponemos quisquillosos con la formalidad, se notará una falla en lo anterior (pasamos un *strike* al lector), pues no hemos definido “ángulo entre vectores” más allá de \mathbb{R}^2 . Tómese entonces como la motivación intuitiva para la definición general que hará que todo cuadre bien.

Definición 1.10.2 Dados dos vectores no nulos \mathbf{u} y \mathbf{v} en \mathbb{R}^n el *ángulo entre ellos* es

$$\text{ang}(\mathbf{u}, \mathbf{v}) = \arccos \left(\frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}| |\mathbf{v}|} \right).$$

Nótese que por la desigualdad de Schwartz (dejada al lector para \mathbb{R}^3 en el Ejercicio 1.77) el argumento de la función arccos está en el intervalo $[-1, 1]$, así que está bien definido el ángulo, y queda en el intervalo $[0, \pi]$.

Hay que remarcar que a partir de \mathbb{R}^3 ya no tiene sentido hablar de ángulos dirigidos; esto sólo se puede hacer en el plano. Pues si tomamos dos vectores no nulos

\mathbf{u} y \mathbf{v} en \mathbb{R}^3 , no hay manera de decir si el viaje angular de \mathbf{u} hacia \mathbf{v} es positivo o negativo. Esto dependería de escoger un lado del plano para “verlo” y ver si va contra o con el reloj, pero cualquiera de los dos lados son “iguales” y desde cada uno se “ve” lo contrario del otro; no hay una manera coherente de decidir. Lo que sí se puede decir es cuándo *tres* vectores están orientados positivamente, pero esto se verá mucho más adelante. De tal manera que a partir de \mathbb{R}^3 , sólo el ángulo *entre* vectores está definido, y tiene valores entre 0 y π ; donde los extremos corresponden a paralelismo pero con la misma dirección (ángulo 0) o la contraria (ángulo π).

EJERCICIO 1.89 Sean \mathbf{u} y \mathbf{v} dos vectores unitarios y ortogonales en \mathbb{R}^3 . Demuestra que para cualquier $\mathbf{x} \in \mathbb{R}^3$ existe un único $\mathbf{w} \in \langle \mathbf{u}, \mathbf{v} \rangle$ (el plano generado por \mathbf{u} y \mathbf{v}) tal que $\mathbf{x} - \mathbf{w}$ es ortogonal a \mathbf{u} y a \mathbf{v} ; al punto \mathbf{w} se le llama la *proyección ortogonal* de \mathbf{x} al plano $\langle \mathbf{u}, \mathbf{v} \rangle$. ¿Puedes concluir que la distancia del punto \mathbf{x} al plano $\langle \mathbf{u}, \mathbf{v} \rangle$ es

$$\sqrt{|\mathbf{x}|^2 - (\mathbf{x} \cdot \mathbf{u})^2 - (\mathbf{x} \cdot \mathbf{v})^2} ?$$

1.11 Distancia

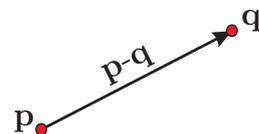
Esta sección cierra el capítulo con una breve discusión sobre el concepto euclidiano de distancia, que se deduce naturalmente de la norma

Dados dos puntos $\mathbf{p}, \mathbf{q} \in \mathbb{R}^n$ se puede definir su *distancia euclidiana*, o simplemente su *distancia*, $d(\mathbf{p}, \mathbf{q})$, como la norma de su diferencia, es decir, como la magnitud del vector que lleva a uno en otro, *i.e.*,

$$d(\mathbf{p}, \mathbf{q}) = |\mathbf{p} - \mathbf{q}|,$$

que explícitamente en coordenadas da la fórmula (en \mathbb{R}^2)

$$d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$



Las propiedades básicas de la distancia se reúnen en el siguiente:

Teorema 1.27 Para todos los vectores $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ se cumple que

- i) $d(\mathbf{x}, \mathbf{y}) \geq 0$
- ii) $d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$
- iii) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$
- iv) $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$.

El *espacio métrico* es un conjunto en el que está definida una distancia que cumple con las cuatro propiedades del Teorema 1.27.

EJERCICIO 1.90 Da explícitamente con coordenadas la fórmula de la distancia en \mathbb{R}^3 y en \mathbb{R}^n .

EJERCICIO 1.91 Demuestra el Teorema 1.27.

EJERCICIO 1.92 ¿Podrías mencionar algún espacio métrico distinto del que se definió arriba?

1.11.1 El espacio euclidiano (primera misión cumplida)

Cuando \mathbb{R}^n se considera junto con la distancia antes definida (llamada la *distancia euclidiana*), se dice que es el espacio euclidiano de dimensión n —que a veces se denota \mathbb{E}^n , y que formalmente podríamos definir $\mathbb{E}^n := (\mathbb{R}^n, \mathbf{d})$ — ya que éste cumple con los postulados de Euclides para $n = 2$. Veamos.

Hemos demostrado que \mathbb{R}^2 cumple los tres postulados que se refieren a las rectas, y aún nos falta analizar dos. El II se refiere a “trazar” círculos. Ya hemos trabajado con el círculo unitario y claramente se generaliza. Dado un punto \mathbf{p} y una distancia r (un número positivo) definimos el *círculo con centro \mathbf{p} y radio r* como el conjunto de puntos a distancia r de \mathbf{p} (que se estudiará en el siguiente capítulo). Se trata de cierto subconjunto de \mathbb{R}^2 , que además es no vacío (es fácil dar explícitamente un punto en él usando coordenadas); y nuestra noción de “definir” un conjunto es la análoga de “trazar” para los griegos. Se cumple entonces el axioma II.

Nos falta únicamente discutir el IV, “todos los ángulos rectos son iguales”. Esta afirmación es sutil. Tenemos, desde hace un buen rato, la noción de ángulo recto, perpendicularidad u ortogonalidad, pero ¿a qué se refiere la palabra “igualdad” en el contexto griego? Si fuera a la igualdad del numerito que mide los ángulos, la afirmación es obvia, casi vacua —“ $\pi/2$ es igual a $\pi/2$ ”—, y no habría necesidad de enunciarla como axioma. Entonces se refiere a algo mucho más profundo. En este postulado está implícita la noción de *movimiento*; se entiende la “igualdad” como que podemos *mover* el plano para llevar un ángulo recto formado por dos rectas a cualquier otro. Esto es claro al ver otros teoremas clásicos como “dos triángulos son *iguales* si sus lados miden lo mismo”: no se puede referir a la igualdad estricta de conjuntos, quiere decir que se puede *llevar* a uno sobre el otro para que, entonces sí, coincidan como conjuntos. Y este *llevar* es el *movimiento* implícito en el uso de la palabra “igualdad”. En términos modernos, esta noción de *movimiento* del plano se formaliza con la noción de función. Nosotros lo haremos en el capítulo 3, y corresponderá formalmente a la noción de *isometría* (una función de \mathbb{R}^2 en \mathbb{R}^2 que preserva distancias). Pero ya no queremos distraernos con la discusión de la axiomática griega. Creamos de buena fe que al desarrollar formalmente los conceptos necesarios el Postulado IV será cierto, o tomémoslo en su acepción numérica trivial, para concluir esta discusión con la que arrancamos el libro.

A partir de los axiomas de los números reales construimos el Plano Euclidiano,

\mathbb{E}^2 , con todo y sus nociones básicas que cumplen los postulados de Euclides. Se tiene entonces que para estudiarlo son igualmente válidos el método *sintético* (el que usaban los griegos pues sus axiomas ya son ciertos) o el *analítico* (que inauguró Descartes y estamos siguiendo). En este último método siempre hay elementos del primero, no siempre es más directo irse a las coordenadas; no hay un divorcio y en general es muy difícil trazar la línea que los separa; además no vale la pena. No hay que preocuparse de eso, en cada momento toma uno lo que le conviene. Pero sí hay que hacer énfasis en la enorme ventaja que dio el método cartesiano, al construir de golpe (y sólo con un poco de esfuerzo extra) espacios euclidianos para todas las dimensiones. Es un método que permitió generalizar y abrir, por tanto, nuevos horizontes. Y no sólo en cuestión de dimensiones.

Como se verá en capítulos posteriores, siguiendo el método analítico pueden construirse espacios de dimensión 2 con formas “raras” de medir ángulos y distancias que los hacen cumplir todos los axiomas menos el Quinto, y haciéndolos entonces espacios *no euclidianos*. Pero esto vendrá a su tiempo; por el momento y para cerrar el círculo de esta discusión, reescribiremos con la noción de distancia el Teorema 1.1, cuya demostración se deja como ejercicio, y donde, nótese, tenemos el extra del “si y sólo si”.

Se puede definir ángulo en tercias ordenadas de puntos. De nuevo, refiriéndonos al ángulo que ya definimos entre vectores (haciendo que el punto de enmedio sea como el origen):

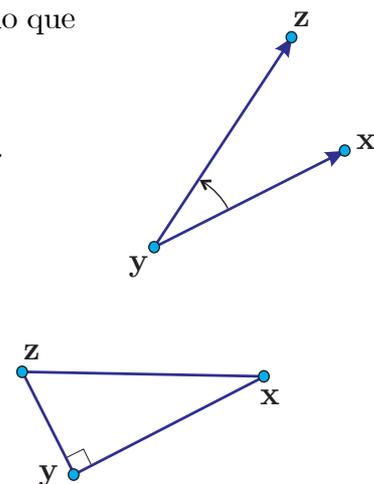
$$\angle \mathbf{x}\mathbf{y}\mathbf{z} := \text{ang}((\mathbf{x} - \mathbf{y}), (\mathbf{z} - \mathbf{y})) = \arccos \frac{(\mathbf{x} - \mathbf{y}) \cdot (\mathbf{z} - \mathbf{y})}{|\mathbf{x} - \mathbf{y}| |\mathbf{z} - \mathbf{y}|}.$$

Y entonces tenemos:

Teorema 1.28 (Pitágoras) *Dados tres puntos \mathbf{x} , \mathbf{y} , \mathbf{z} en \mathbb{E}^n , se tiene que*

$$d(\mathbf{x}, \mathbf{z})^2 = d(\mathbf{x}, \mathbf{y})^2 + d(\mathbf{y}, \mathbf{z})^2 \quad \Leftrightarrow \quad \angle \mathbf{x}\mathbf{y}\mathbf{z} = \pi/2.$$

□



Para retomar el hilo de la corriente principal del texto, en los siguientes ejercicios vemos cómo con la noción de distancia se pueden reconstruir objetos básicos como segmentos y rayos, que juntos dan las líneas; y delineamos un ejemplo de un espacio métrico que da un plano con una geometría “rara”.

EJERCICIO 1.93 Demuestra esta versión del Teorema de Pitágoras.

EJERCICIO 1.94 Demuestra que el segmento de \mathbf{x} a \mathbf{y} es el conjunto

$$\overline{\mathbf{x}\mathbf{y}} = \{\mathbf{z} \mid d(\mathbf{x}, \mathbf{y}) = d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})\}.$$

EJERCICIO 1.95 Demuestra que el *rayo que parte de \mathbf{y} desde \mathbf{x}* (es decir la continuación de la recta por \mathbf{x} y \mathbf{y} más allá de \mathbf{y}) es el conjunto $\{\mathbf{z} \mid d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) = d(\mathbf{x}, \mathbf{z})\}$.

***EJERCICIO 1.96** (Un ejemplo cotidiano de otro espacio métrico). El plano con la métrica de “*Manhattan*” (en honor de la famosísima y bien cuadrículada ciudad) se define como \mathbb{R}^2 con la función de distancia

$$d_m((x_1, x_2), (y_1, y_2)) := |y_1 - x_1| + |y_2 - x_2|;$$

pensando que para ir de un punto a otro sólo se puede viajar en dirección horizontal o vertical (como en las calles de una ciudad).

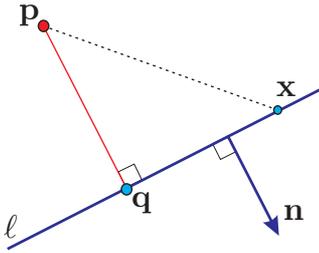
- Demuestra que la métrica de Manhattan cumple el Teorema 1.27.
- Dibuja el conjunto de puntos con distancia 1 al origen; ¿cómo son los círculos?
- ¿Cómo son los segmentos (definidos por la distancia como en el ejercicio anterior)?
- ¿Cómo son los rayos?

1.11.2 Distancia de un punto a una recta

Consideremos el siguiente problema. Nos dan una recta

$$\ell : \mathbf{n} \cdot \mathbf{x} = c$$

y un punto \mathbf{p} (que puede estar fuera o dentro de ella); y se nos pregunta cuál es la distancia de \mathbf{p} a ℓ , que podemos denotar con $d(\mathbf{p}, \ell)$.



Es más fácil resolverlo si lo pensamos junto con un problema aparentemente más complicado: ¿cuál es el punto de ℓ más cercano a \mathbf{p} ? Llamémoslo $\mathbf{q} \in \ell$, aunque sea incógnito. El meollo del asunto es que la recta que pasa por \mathbf{p} y \mathbf{q} debe ser ortogonal a ℓ . Es decir, si el segmento $\overline{\mathbf{p}\mathbf{q}}$ es ortogonal a ℓ entonces \mathbf{q} es el punto en ℓ más cercano a \mathbf{p} . Para demostrarlo considérese cualquier otro punto $\mathbf{x} \in \ell$ y el triángulo rectángulo $\mathbf{x}, \mathbf{q}, \mathbf{p}$ y aplíquese el Teorema de Pitágoras para ver que $d(\mathbf{x}, \mathbf{p}) \geq d(\mathbf{q}, \mathbf{p})$. Como la dirección ortogonal a ℓ es precisamente \mathbf{n} , entonces debe existir $t \in \mathbb{R}$ tal que $\mathbf{q} - \mathbf{p} = t\mathbf{n}$. La nueva incógnita t nos servirá para medir la distancia de \mathbf{p} a ℓ . Tenemos entonces que nuestras incógnitas \mathbf{q} y t cumplen las siguientes ecuaciones:

$$\begin{aligned} \mathbf{n} \cdot \mathbf{q} &= c \\ \mathbf{q} - \mathbf{p} &= t\mathbf{n}, \end{aligned}$$

donde la primera dice que $\mathbf{q} \in \ell$ y la segunda que el segmento $\overline{\mathbf{p}\mathbf{q}}$ es ortogonal a ℓ .

Tomando el producto interior de \mathbf{n} con la segunda ecuación, obtenemos

$$\mathbf{n} \cdot \mathbf{q} - \mathbf{n} \cdot \mathbf{p} = t(\mathbf{n} \cdot \mathbf{n}).$$

De donde, al sustituir la primera ecuación, se puede despejar t (pues $\mathbf{n} \neq \mathbf{0}$) en puros términos conocidos:

$$\begin{aligned} c - (\mathbf{n} \cdot \mathbf{p}) &= t(\mathbf{n} \cdot \mathbf{n}) \\ t &= \frac{c - (\mathbf{n} \cdot \mathbf{p})}{\mathbf{n} \cdot \mathbf{n}}. \end{aligned}$$

De aquí se deduce que

$$\begin{aligned} \mathbf{q} &= \mathbf{p} + \left(\frac{c - (\mathbf{n} \cdot \mathbf{p})}{\mathbf{n} \cdot \mathbf{n}} \right) \mathbf{n} \\ d(\mathbf{p}, \ell) &= |t \mathbf{n}| = |t| |\mathbf{n}| \\ &= \frac{|c - (\mathbf{n} \cdot \mathbf{p})|}{|\mathbf{n} \cdot \mathbf{n}|} |\mathbf{n}| = \frac{|c - (\mathbf{n} \cdot \mathbf{p})|}{|\mathbf{n}|}. \end{aligned}$$

Puesto que sólo hemos usado las propiedades del producto interior y la norma, sin ninguna referencia a las coordenadas, podemos generalizar a \mathbb{R}^3 ; o bien a \mathbb{R}^n pensando que un *hiperplano* (definido por una ecuación lineal $\mathbf{n} \cdot \mathbf{x} = c$) es una recta cuando $n = 2$ y es un plano cuando $n = 3$.

Proposición 1.29 *Sea Π un hiperplano en \mathbb{R}^n dado por la ecuación $\mathbf{n} \cdot \mathbf{x} = c$, y sea $\mathbf{p} \in \mathbb{R}^n$ cualquier punto. Entonces*

$$d(\mathbf{p}, \Pi) = \frac{|c - (\mathbf{n} \cdot \mathbf{p})|}{|\mathbf{n}|},$$

y la proyección ortogonal de \mathbf{p} sobre Π (es decir, el punto más cercano a \mathbf{p} en Π) es

$$\mathbf{q} = \mathbf{p} + \left(\frac{c - \mathbf{n} \cdot \mathbf{p}}{\mathbf{n} \cdot \mathbf{n}} \right) \mathbf{n}.$$

EJERCICIO 1.97 Evalúa las dos fórmulas de la Proposición 1.29 cuando $\mathbf{p} \in \Pi$.

EJERCICIO 1.98 Encuentra una expresión para $d(\mathbf{p}, \ell)$ cuando $\ell = \{\mathbf{q} + t\mathbf{v} \mid t \in \mathbb{R}\}$.

EJERCICIO 1.99 Demuestra que la fórmula de la distancia de un punto a un plano obtenida en el Ejercicio 1.89 coincide con la de la Proposición 1.29.

EJERCICIO 1.100 Encuentra la distancia del punto ... a la recta ... y su proyección ortogonal.

1.11.3 El determinante como área dirigida

No estamos en posición de desarrollar una teoría general de la noción de área; para ello se requieren ideas de cálculo. Pero sí podemos trabajarla para figuras simples

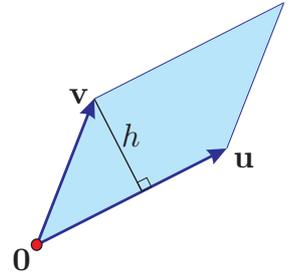
como triángulos y paralelogramos. Se define el *área* de un paralelogramo de manera euclidiana (como en secundaria), es decir, $(\text{base}) \times (\text{altura})$. Veremos que se calcula con un viejo conocido.

Nos encontramos primero el determinante asociado a sistemas de dos ecuaciones lineales con dos incógnitas. Después, definimos el determinante de dos vectores en \mathbb{R}^2 , otra vez por sistemas de ecuaciones y vimos que detecta (cuando es cero) el paralelismo. Ahora podemos darle una interpretación geométrica en todos los casos.

Teorema 1.30 Sean \mathbf{u} y \mathbf{v} dos vectores cualesquiera en \mathbb{R}^2 . El área del paralelogramo que definen \mathbf{u} y \mathbf{v} es $|\det(\mathbf{u}, \mathbf{v})|$. Además $\det(\mathbf{u}, \mathbf{v}) \geq 0$ cuando el movimiento angular más corto de \mathbf{u} a \mathbf{v} es positivo ($\overrightarrow{\text{ang}}(\mathbf{u}, \mathbf{v}) \in [0, \pi]$) y $\det(\mathbf{u}, \mathbf{v}) \leq 0$ cuando $\overrightarrow{\text{ang}}(\mathbf{u}, \mathbf{v}) \in [-\pi, 0]$.

Demostración. Podemos tomar como base del paralelogramo en cuestión al vector \mathbf{u} , es decir, (base) en la fórmula $(\text{base}) \times (\text{altura})$ es $|\mathbf{u}|$, y suponemos que $\mathbf{u} \neq \mathbf{0}$. La altura, h digamos, es entonces la distancia de \mathbf{v} , pensado como punto, a la recta ℓ generada por \mathbf{u} . Esta recta tiene ecuación normal $\mathbf{u}^\perp \cdot \mathbf{x} = 0$. Entonces, por la Proposición 1.29 se tiene que

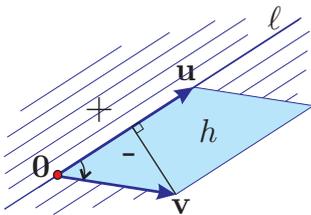
$$h = \frac{|0 - (\mathbf{u}^\perp \cdot \mathbf{v})|}{|\mathbf{u}^\perp|} = \frac{|\mathbf{u}^\perp \cdot \mathbf{v}|}{|\mathbf{u}|}.$$



Al multiplicar por la base se obtiene que el área del paralelogramo es el valor absoluto del determinante

$$\frac{|\mathbf{u}^\perp \cdot \mathbf{v}|}{|\mathbf{u}|} |\mathbf{u}| = |\mathbf{u}^\perp \cdot \mathbf{v}| = |\det(\mathbf{u}, \mathbf{v})|.$$

Para ver lo que significa el signo de $\det(\mathbf{u}, \mathbf{v})$, obsérvese que la recta ℓ , generada por \mathbf{u} , parte el plano en dos *semiplanos*. En uno de ellos se encuentra su compadre ortogonal \mathbf{u}^\perp , éste es el lado *positivo*, pues al tomar $c \geq 0$ las ecuaciones $\mathbf{u}^\perp \cdot \mathbf{x} = c$ definen rectas paralelas a ℓ , y una de ellas (cuando $c = |\mathbf{u}^\perp|^2 > 0$) pasa por \mathbf{u}^\perp ; como estas rectas llenan (“ashuran”) todo ese semiplano, entonces $\mathbf{u}^\perp \cdot \mathbf{v} \geq 0$ si y sólo si \mathbf{v} está en el lado positivo de ℓ . Análogamente, $\mathbf{u}^\perp \cdot \mathbf{v} < 0$ si y sólo si \mathbf{v} está en el mismo lado de ℓ que $-\mathbf{u}^\perp$, que llamamos el lado *negativo*.

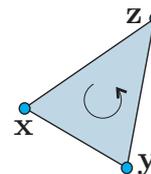


el lado *negativo*.

□

De esta manera, el determinante es un “área signada”: no sólo nos da el área sino que también nos dice si los vectores están orientados positiva o negativamente. Obsérvese que nuestra manera de detectar el paralelismo corresponde entonces a que el área es cero (que el paralelogramo en cuestión es unidimensional).

EJERCICIO 1.101 El área signada de un triángulo *orientado* $\mathbf{x}, \mathbf{y}, \mathbf{z}$ se debe definir como $\det((\mathbf{y} - \mathbf{x}), (\mathbf{z} - \mathbf{x}))$. Demuestra que sólo depende del orden cíclico en que aparecen los vértices, es decir, que da lo mismo si tomamos como ternas a $\mathbf{y}, \mathbf{z}, \mathbf{x}$ o a $\mathbf{z}, \mathbf{x}, \mathbf{y}$. Pero que es su inverso aditivo si tomamos la otra orientación $\mathbf{z}, \mathbf{y}, \mathbf{x}$ para el triángulo.



1.11.4 La mediatriz

Cuando empecemos a estudiar las curvas cónicas, las definiremos como *lugares geométricos* de puntos que cumplen cierta propiedad que involucra distancias. Así que para terminar este capítulo veremos los ejemplos más sencillos.

Dados dos puntos \mathbf{p} y \mathbf{q} ($\in \mathbb{R}^2$), ¿cuál es el lugar geométrico de los puntos que equidistan de ellos? Es decir, hay que describir al conjunto

$$\{\mathbf{x} \in \mathbb{R}^2 \mid d(\mathbf{p}, \mathbf{x}) = d(\mathbf{q}, \mathbf{x})\}.$$

Al desarrollar la ecuación $d(\mathbf{p}, \mathbf{x})^2 = d(\mathbf{q}, \mathbf{x})^2$, que equivale a la anterior pues las distancias son positivas, tenemos

$$\begin{aligned} (\mathbf{p} - \mathbf{x}) \cdot (\mathbf{p} - \mathbf{x}) &= (\mathbf{q} - \mathbf{x}) \cdot (\mathbf{q} - \mathbf{x}) \\ \mathbf{p} \cdot \mathbf{p} - 2(\mathbf{p} \cdot \mathbf{x}) + \mathbf{x} \cdot \mathbf{x} &= \mathbf{q} \cdot \mathbf{q} - 2(\mathbf{q} \cdot \mathbf{x}) + \mathbf{x} \cdot \mathbf{x} \\ \mathbf{p} \cdot \mathbf{p} - \mathbf{q} \cdot \mathbf{q} &= 2(\mathbf{p} - \mathbf{q}) \cdot \mathbf{x} \end{aligned}$$

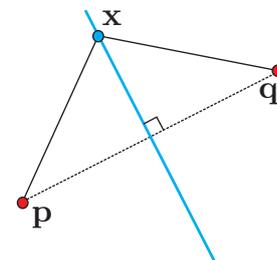
que se puede reescribir como

$$(\mathbf{p} - \mathbf{q}) \cdot \mathbf{x} = (\mathbf{p} - \mathbf{q}) \cdot \left(\frac{1}{2}(\mathbf{p} + \mathbf{q})\right).$$

Y ésta es la ecuación normal de la recta ortogonal al segmento $\overline{\mathbf{p}\mathbf{q}}$ y que lo intersecta en su punto medio, llamada la *mediatriz* de \mathbf{p} y \mathbf{q} .

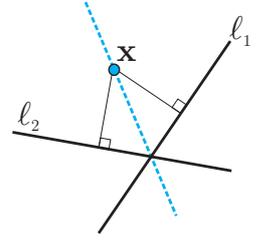
EJERCICIO 1.102 ¿Cuál es el lugar geométrico de los puntos en \mathbb{R}^3 que equidistan de dos puntos? Describe la demostración.

1.11.5 Bisectrices y ecuaciones unitarias



Preguntamos ahora cuál es el lugar geométrico de los puntos que equidistan de dos rectas ℓ_1 y ℓ_2 . De todas las ecuaciones normales que definen a estas rectas, conviene escoger una donde el vector normal es unitario para que las distancias sean más fáciles de expresar. Tal ecuación se llama *unitaria* y se obtiene de cualquier ecuación normal dividiendo —ambos lados de la ecuación, por supuesto— entre la norma del vector normal. Sean entonces $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{S}^1$ y $c_1, c_2 \in \mathbb{R}$ tales que para $i = 1, 2$

$$\ell_i : \mathbf{u}_i \cdot \mathbf{x} = c_i.$$



Por la fórmula de la sección 1.11.2 se tiene entonces que los puntos \mathbf{x} que equidistan de ℓ_1 y ℓ_2 son precisamente los que cumplen con la ecuación

$$|\mathbf{u}_1 \cdot \mathbf{x} - c_1| = |\mathbf{u}_2 \cdot \mathbf{x} - c_2|.$$

Ahora bien, si el valor absoluto de dos números reales coincide entonces o son iguales o son inversos aditivos, lo cual se expresa elegantemente de la siguiente manera:

$$|a| = |b| \Leftrightarrow (a + b)(a - b) = 0,$$

que, aplicado a nuestra ecuación anterior, nos dice que el lugar geométrico que estamos buscando está determinado por la ecuación

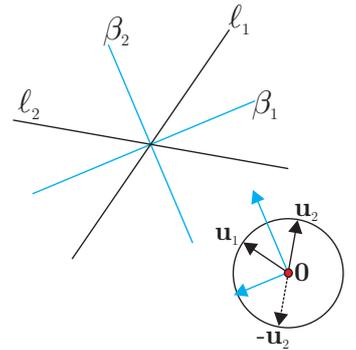
$$\begin{aligned} (\mathbf{u}_1 \cdot \mathbf{x} - c_1 + \mathbf{u}_2 \cdot \mathbf{x} - c_2)(\mathbf{u}_1 \cdot \mathbf{x} - c_1 - \mathbf{u}_2 \cdot \mathbf{x} + c_2) &= 0 \\ ((\mathbf{u}_1 + \mathbf{u}_2) \cdot \mathbf{x} - (c_1 + c_2))((\mathbf{u}_1 - \mathbf{u}_2) \cdot \mathbf{x} - (c_1 - c_2)) &= 0. \end{aligned}$$

Como cuando el producto de dos números es cero alguno de ellos lo es, entonces nuestro lugar geométrico es la unión de las dos rectas:

$$\begin{aligned} \beta_1 : (\mathbf{u}_1 + \mathbf{u}_2) \cdot \mathbf{x} &= (c_1 + c_2) \\ \beta_2 : (\mathbf{u}_1 - \mathbf{u}_2) \cdot \mathbf{x} &= (c_1 - c_2), \end{aligned} \quad (1.12)$$

que son las *bisectrices* de las líneas ℓ_1 y ℓ_2 . Observemos que son ortogonales pues

$$\begin{aligned} (\mathbf{u}_1 + \mathbf{u}_2) \cdot (\mathbf{u}_1 - \mathbf{u}_2) &= |\mathbf{u}_1|^2 - |\mathbf{u}_2|^2 \\ &= 1 - 1 \\ &= 0, \end{aligned}$$

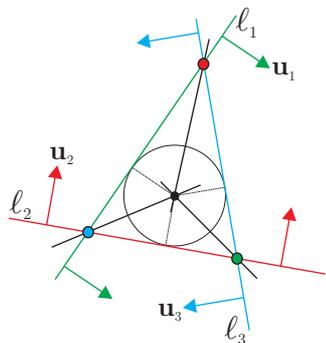


y, geoméricamente, bisectan a los dos sectores o ángulos que forman las rectas; también pasan por el punto de intersección pues sus ecuaciones se obtienen sumando y restando las ecuaciones originales (el caso en que sean paralelas se deja como ejercicio). En el círculo unitario también es claro que la suma y la diferencia de dos vectores (aunque ya no sean necesariamente unitarios) tienen los ángulos adecuados.

Como corolario, obtenemos otro de los teoremas clásicos de concurrencia.

Teorema 1.31 *Las bisectrices (internas) de un triángulo son concurrentes.*

Demostración. Sólo hay que tener cuidado con cuáles ecuaciones unitarias trabajamos, pues hay dos posibles para cada recta. Si escogemos los vectores unitarios de las tres rectas, ℓ_1 , ℓ_2 y ℓ_3 digamos, apuntando hacia adentro del triángulo, \mathbf{u}_1 , \mathbf{u}_2 y \mathbf{u}_3 respectivamente, entonces está claro que sus sumas son normales a las bisectrices exteriores. El teorema habla entonces de las bisectrices que se obtienen como diferencias, que son, numerándolas por el índice de la recta opuesta:



$$\begin{aligned}\beta_1 &: (\mathbf{u}_2 - \mathbf{u}_3) \cdot \mathbf{x} = (c_2 - c_3) \\ \beta_2 &: (\mathbf{u}_3 - \mathbf{u}_1) \cdot \mathbf{x} = (c_3 - c_1) \\ \beta_3 &: (\mathbf{u}_1 - \mathbf{u}_2) \cdot \mathbf{x} = (c_1 - c_2).\end{aligned}$$

El negativo de cualquiera de ellas se obtiene sumando las otras dos.

□

EJERCICIO 1.103 Un argumento aparentemente más elemental para el teorema anterior es que si un punto está en la intersección de dos bisectrices, entonces ya equidista de las tres rectas y por tanto también está en la bisectriz del vértice restante. ¿Qué implica este argumento cuando tomamos bisectrices exteriores? ¿En dónde concurren dos bisectrices exteriores? ¿Una interior y una exterior? Da enunciados y demuéstralos con ecuaciones normales. Completa el dibujo.

EJERCICIO 1.104 Demuestra que el lugar geométrico de los puntos que equidistan a dos rectas paralelas es la paralela a ambas que pasa por el punto medio de sus intersecciones con una recta no paralela a ellas, a partir de las ecuaciones (1.12).

EJERCICIO 1.105 ¿Cuál es el lugar geométrico de los puntos que equidistan de dos planos en \mathbb{R}^3 ? Describe una demostración de tu aseveración.

1.12 *Los espacios de rectas en el plano

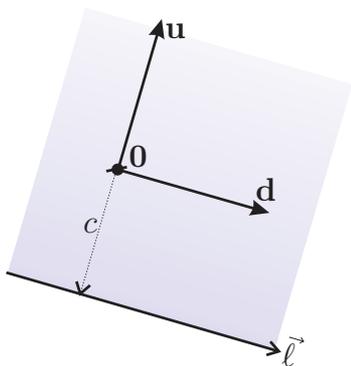
Esta sección “asterisco” o extra es, hasta cierto punto, independiente del texto principal; será relevante como motivación hasta el capítulo 6, donde se estudia el plano proyectivo y cristaliza en la llamada “dualidad”. Estudiaremos dos ejemplos que ya están a la mano de cómo la noción de *el Espacio*, que por milenios se consideró única y amorfa —“el espacio en que vivimos”—, se pluraliza *los espacios*. Ya vimos que el Espacio teórico que consideraron los griegos es sólo la instancia en dimensión 3

de la familia de espacios euclidianos \mathbb{E}^n ; pero además la *pluralización* (pasar de uno a muchos) se da en otro sentido. En matemáticas, y en particular en la geometría, surgen naturalmente otros conjuntos, además de los \mathbb{E}^n , con plenos derechos de ser llamados “espacios”; en la física se les llama “espacios fase”. Los dos “espacios”⁵ que nos ocupan ahora son los de rectas en el plano. Consideraremos cada recta como un punto de un nuevo “espacio” y, puesto que intuitivamente podemos decir si dos rectas son cercanas o no, en este nuevo “espacio”, sabremos si dos puntos lo están. La idea clave está en que así como Descartes asoció parejas de números a los puntos del plano, a las rectas del plano también hemos asociado números (los involucrados en sus ecuaciones) de tal manera que si esos números varían un poco la recta asociada también se mueve poco; dan lugar a una noción de “continuidad” que es la noción básica de la topología.

En el Ejercicio 1.68 de la sección 1.8, se hizo notar que a cada terna de números $(\mathbf{a}, \mathbf{b}, c) \in \mathbb{R}^3$, con $(\mathbf{a}, \mathbf{b}) \neq (0, 0)$, se le puede asociar una recta en \mathbb{R}^2 , a saber, la definida por la ecuación $\mathbf{a}x + \mathbf{b}y = c$; el estudiante debió haber demostrado que las rectas por el origen en \mathbb{R}^3 dan la misma recta en \mathbb{R}^2 y que los planos por el origen en \mathbb{R}^3 corresponden a los haces de rectas en el plano. Dando esto por entendido, en esta última sección abundaremos en tales ideas. Se puede decir que la pregunta que nos guía es: ¿cuántas rectas hay en \mathbb{R}^2 ?, pero el término “cuántas” es demasiado impreciso. Veremos que las rectas de \mathbb{R}^2 forman algo que llamaremos el “espacio” de rectas (un caso especial de las llamadas “Grassmannianas” pues quien primero desarrolló estas ideas fue Grassmann), y para esto les daremos “coordenadas” explícitas, muy a semejanza de las coordenadas polares para los puntos.

1.12.1 Rectas orientadas

Consideremos primero las rectas orientadas en \mathbb{R}^2 , pues a ellas se les puede asociar un vector normal (y por tanto una ecuación) de manera canónica. Una *recta orientada* es una recta $\vec{\ell}$ junto con una orientación preferida, que nos dice la manera *positiva* de viajar en ella. Entonces, de todos sus vectores direccionales podemos escoger el unitario \mathbf{d} con la orientación positiva y, como vector normal, escogemos el compadre ortogonal de éste; es decir, el vector unitario tal que después de la orientación de la recta nos da la orientación positiva del plano. Así, cada recta orientada está definida por una ecuación *unitaria* única

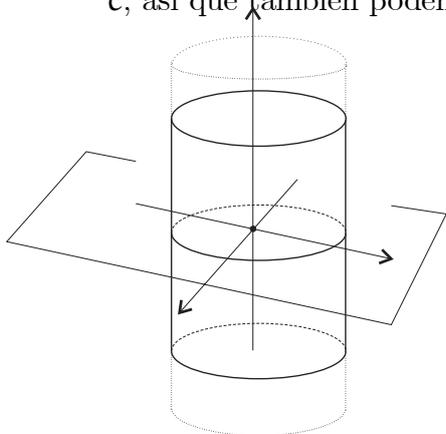


$$\mathbf{u} \cdot \mathbf{x} = c,$$

⁵Mantenemos las comillas pues la noción aún no está bien definida: va en la dirección de lo que ahora se llama topología.

donde $\mathbf{u} \in \mathbb{S}^1$, y $c \in \mathbb{R}$ es su distancia orientada al origen.

Nótese que a cada recta orientada se le puede asociar su semiplano positivo $\mathbf{u} \cdot \mathbf{x} \geq c$; así que también podemos pensar a las rectas orientadas como semiplanos.



Puesto que esta ecuación unitaria es única (por la orientación) entonces hay tantas rectas orientadas como parejas

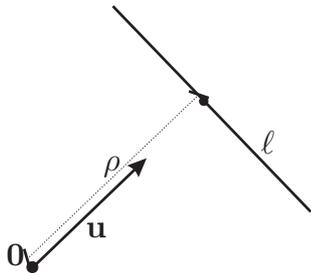
$$(\mathbf{u}, c) \in \mathbb{S}^1 \times \mathbb{R} \subset \mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^3,$$

y este conjunto es claramente un cilindro vértical en \mathbb{R}^3 . Los haces de líneas orientadas paralelas corresponden a líneas verticales en el cilindro, pues en estos haces el vector normal está fijo. Y los haces de líneas (orientadas) concurrentes corresponden a lo que corta un plano por el origen al cilindro $\mathbb{S}^1 \times \mathbb{R}$ (que, como veremos en el capítulo siguiente, se llama elipse).

EJERCICIO 1.106 Demuestra la última afirmación.

1.12.2 Rectas no orientadas

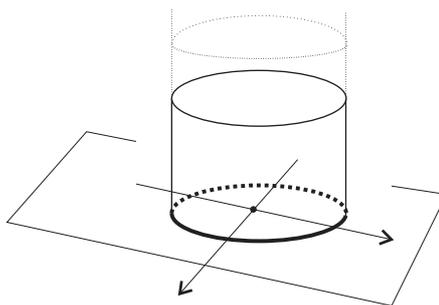
Las dos orientaciones de una recta ℓ tienen ecuaciones unitarias $\mathbf{u} \cdot \mathbf{x} = c$ y $(-\mathbf{u}) \cdot \mathbf{x} = -c$. Si la recta no pasa por el origen (si $\mathbf{0} \notin \ell$ y por tanto $c \neq 0$), de estas dos ecuaciones unitarias podemos tomar la ecuación con constante estrictamente positiva, $\mathbf{u} \cdot \mathbf{x} = \rho$, con $\rho > 0$, digamos. Así que las “coordenadas” (\mathbf{u}, ρ) corresponden a las coordenadas polares del punto en la recta ℓ más cercano al origen; y viceversa, a cada punto $\mathbf{x} \in \mathbb{R}^2 - \{\mathbf{0}\}$ se le asocia la recta que pasa por el punto \mathbf{x} y es normal al vector \mathbf{x} . O dicho de otra manera, si la recta ℓ no contiene al origen, podemos escoger el semiplano que define que no contiene a $\mathbf{0}$.



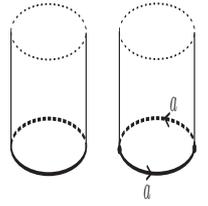
Nos queda una ambigüedad en el origen: para las rectas que pasan por él tenemos dos vectores unitarios normales que definen la misma recta. Podemos resumir que las rectas en \mathbb{R}^2 corresponden a los puntos de un cilindro “semicerrado”

$$\mathbb{S}^1 \times \mathbb{R}_+,$$

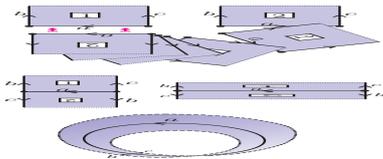
(donde $\mathbb{R}_+ = \{\rho \in \mathbb{R} \mid \rho \geq 0\}$), donde hay que identificar su frontera por antípodas, es decir, en este conjunto $(\mathbf{u}, 0)$ y $(-\mathbf{u}, 0)$ aún representan la misma recta. Si identificamos $(\mathbf{u}, 0)$ con $(-\mathbf{u}, 0)$ en el cilindro $\mathbb{S}^1 \times \mathbb{R}_+$, se obtiene una banda de Moebius abierta. Veámos:



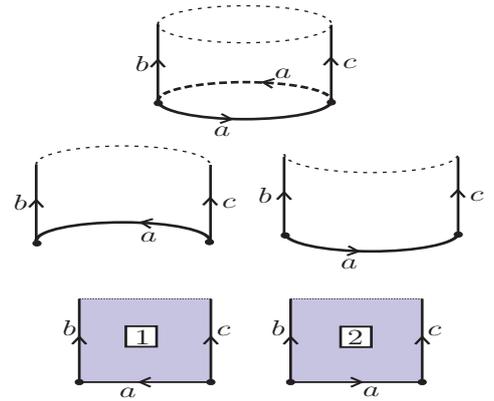
Hay que pensar que el cilindro $\mathbb{S}^1 \times \mathbb{R}_+$ está hecho de un material flexible (pensarlo como “espacio topológico”, se dice actualmente), y hacer muy pequeño el factor \mathbb{R}_+ , para tener un cilindrito con una de sus fronteras inexistente (la que corresponde al infinito), indicada en las figuras con línea punteada. En la otra frontera (que corresponde al 0) debemos pegar puntos antípodos. Llamemos \mathbf{a} al segmento dirigido que es la mitad de ese círculo y entonces la otra mitad, con la misma orientación del círculo, vuelve a ser \mathbf{a} .



Si tratamos de pegar directamente los segmentos \mathbf{a} , la superficie del cilindro nos estorba y no se ve qué pasa con claridad (de hecho no se puede lograr en \mathbb{R}^3 sin que la superficie se autointersece). Conviene entonces cortar el cilindro por los segmentos verticales, denotados con \mathbf{b} y \mathbf{c} en la figura, sobre el principio y fin de \mathbf{a} , para obtener dos pequeños cuadrados



(que hemos bautizado 1 y 2). Los segmentos \mathbf{b} y \mathbf{c} están dirigidos y debemos recordar que hay que volverlos a pegar cuando



se pueda.

Ahora sí podemos pegar los segmentos \mathbf{a} recordando su dirección. Basta voltear el cuadrado 2 de cabeza para lograrlo, y obtener un rectángulo. Si alargamos el rectángulo en su dirección horizontal, obtenemos una banda en la cual hay que identificar las fronteras verticales hechas, cada una, con los pares de segmentos \mathbf{b} y \mathbf{c} (la frontera horizontal sigue siendo punteada). Y esta identificación, según indican nuestros segmentos dirigidos, es invirtiendo la orientación;

según indican nuestros segmentos dirigidos, es invirtiendo la orientación; hay que dar media vuelta

para pegarlos, así que el resultado es una banda de Moebius a la que se le quita su frontera, y ya

no queda nada por identificar.

Hemos demostrado que las rectas del plano corresponden biyectivamente a los puntos de una banda de Moebius abierta, y que esta correspondencia preserva “cercanía” o “continuidad”.

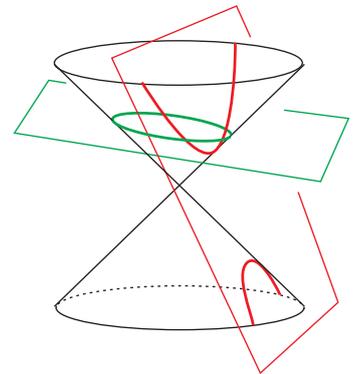
Obsérvese además que las rectas que pasan por el origen (el segmento \mathbf{a} con sus extremos identificados) quedan en el “corazón” (el círculo central) de la banda de Moebius; los haces paralelos corresponden a los intervalos abiertos transversales al corazón que van de la frontera a sí misma (el punto de intersección con el corazón es su representante en el origen). ¿A qué corresponden los haces concurrentes?

EJERCICIO 1.107 Demuestra que dos haces de rectas, no ambos haces paralelos, tienen una única recta en común.

Capítulo 2

Cónicas I (presentación)

Las cónicas son una familia de curvas famosas que definieron los griegos. De ellas, la que vemos cientos de veces al día es la elipse. Al ver la boca de un vaso, una taza o una vasija, o bien, la llanta de un coche, vemos una *elipse* e inmediatamente intuimos que en realidad se trata de un círculo, sólo que lo estamos viendo de lado. Únicamente vemos un círculo cuando la proyección ortogonal de nuestro ojo al plano en el que vive el círculo es justo su centro; así que los cientos de círculos que vemos a diario son nuestra deducción cerebral automática de las elipses que en realidad percibimos. Para fijar ideas, pensemos que vemos con un solo ojo una taza sobre una mesa; cada punto de su borde define una recta a nuestro ojo por donde viaja la luz que percibimos; todas estas rectas forman un *cono* con *ápice* × en el punto ideal de nuestro ojo; pero no es un cono circular sino un cono elíptico: más apachurrado entre más cerca esté nuestro ojo al plano de la boca de la taza, y más circular en cuanto nuestro ojo se acerca a verlo desde arriba. Resulta entonces que si a este cono elíptico lo cortamos con el plano de la boca de la taza nos da un círculo (precisamente, la boca de la taza). Los griegos intuyeron entonces que, al revés, la elipse se puede definir cortando conos circulares (definidos con base en el círculo que ya conocían) con planos inclinados.



Hoy día lo podemos decir de otra manera: si prendemos una lámpara de mano, de ella emana un cono circular de luz; es decir, si la apuntamos ortogonalmente a una pared se ilumina un círculo (con todo y su interior); al girar un poco la lámpara, el círculo se deforma en una elipse que es un plano (la pared) que corta un cono circular (el haz de luz). Entre más inclinada esté la lámpara respecto a la pared, la elipse se alarga (como boca de taza con el ojo casi al ras) hasta que algo de la luz sale de la pared. Pensemos entonces que la pared es infinita (un plano) y que nuestra lámpara es infinitamente potente (un cono perfecto); al girar más la lámpara, la elipse se sigue alargando y alargando hasta el momento en que uno de los rayos del cono no toca la pared (es paralelo a ella). Justo en ese instante, el borde de lo iluminado en la pared es

una *parábola*, y de allí en adelante, al continuar el giro, es una *rama de hipérbola*. La *hipérbola* completa se obtiene al pensar que los rayos del cono son rectas completas, que se continúan hacia atrás de la lámpara, y entonces algunas de estas rectas (cuyos rayos de luz ya no tocan la pared por adelante) la intersectan por atrás en la otra rama de la hipérbola. Si continuamos el giro, pasamos de nuevo por una parábola y luego por elipses que, viniendo de atrás, se redondean hasta el círculo con el que empezamos, pero ahora dibujado por el haz de rayos “hacia atrás” de la lámpara que giró 180° . La definición clásica de los griegos, que es equivalente, es fijando el cono y moviendo el plano, y de allí que las hayan llamado *secciones cónicas* o simplemente *cónicas*: la intersección de planos con conos circulares (completos).

Los geómetras griegos estudiaron con profundidad estas curvas. Lo hicieron por amor al arte, al conocimiento abstracto y puro, por la intuición matemática de su propia belleza y naturalidad; y dieron a la humanidad una lección fundamental de la importancia que tiene la ciencia básica. Pues resultó, casi dos mil años después, que estas curvas teóricas se expresaban en la naturaleza como las órbitas de los planetas (elipses descritas por Kepler) o en las trayectorias de los proyectiles (parábolas descritas por Newton y predecesores). Sus propiedades focales, que también descubrieron los griegos, se usan hoy cotidianamente en las antenas parabólicas, en los telescopios y en el diseño de reflectores o de las lentes de anteojos o cámaras. Su teoría, que parecía ser un divertimento totalmente abstracto, resultó importantísima en la vida diaria y en el entendimiento de la naturaleza.

Los geómetras griegos demostraron que a las cónicas también se les puede definir intrínsecamente como *lugares geométricos* en el plano, es decir, como subconjuntos de puntos que cumplen cierta propiedad; a saber, una propiedad que se expresa en términos de distancias. Ésta será la definición formal que usaremos en el libro y en la última sección de este capítulo veremos que coincide con la de secciones cónicas. En las primeras secciones veremos que su definición como lugares geométricos expresados en términos de distancias es equivalente a su descripción como el lugar geométrico de puntos en \mathbb{R}^2 cuyas coordenadas satisfacen cierta ecuación. Resulta que estas ecuaciones son ecuaciones cuadráticas en dos variables, x y y (las coordenadas). Este hecho, que observó el propio Descartes, fue una gran motivación para el surgimiento de su geometría analítica, pues relacionaba las clásicas secciones cónicas con los polinomios de segundo grado, y en general el álgebra, que habían desarrollado los árabes.

2.1 Círculos

Ya hemos definido y usado extensamente el círculo unitario \mathbb{S}^1 , determinado por la ecuación

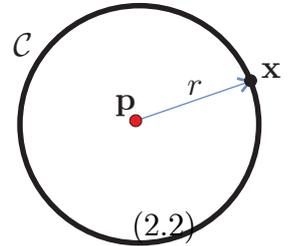
$$x^2 + y^2 = 1.$$

Consideremos ahora cualquier otro círculo \mathcal{C} . Tiene un *centro* $\mathbf{p} = (h, k)$, un *radio* $r > 0$, y es el lugar geométrico de los puntos cuya distancia a \mathbf{p} es r . Es decir, $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^2 \mid d(\mathbf{x}, \mathbf{p}) = r\}$; o bien, \mathcal{C} está definido por la ecuación

$$d(\mathbf{x}, \mathbf{p}) = r. \quad (2.1)$$

Puesto que ambos lados de esta ecuación son positivos, es equivalente a la igualdad de sus cuadrados que en coordenadas toma la forma

$$(x - h)^2 + (y - k)^2 = r^2. \quad (2.2)$$



De tal manera que todos los círculos de \mathbb{R}^2 están determinados por una ecuación cuadrática en las variables x y y . Cuando la ecuación tiene la forma anterior, podemos “leer” toda la información geométrica (el centro y el radio). Pero en general viene disfrazada como

$$x^2 + y^2 - 2hx - 2ky = (r^2 - h^2 - k^2)$$

que es, claramente, su expresión desarrollada. Veamos un ejemplo.

Consideremos la ecuación

$$x^2 + y^2 - 6x + 2y = -6. \quad (2.3)$$

Para ver si define un círculo, y cuál, hay que tratar de escribirla en la forma (2.2). Para esto hay que “completar” los cuadrados sumando (en ambos lados) las constantes que faltan

$$\begin{aligned} (x^2 - 6x + 9) + (y^2 + 2y + 1) &= -6 + 9 + 1 \\ (x - 3)^2 + (y + 1)^2 &= 4. \end{aligned}$$

Es claro que al desarrollar esta ecuación obtenemos la original, de tal manera que aquella define el círculo con centro en $(3, -1)$ y radio 2.

También podemos expresar la ecuación de un círculo de manera vectorial, pues el círculo \mathcal{C} con centro \mathbf{p} y radio r está claramente definido por la ecuación

$$(\mathbf{x} - \mathbf{p}) \cdot (\mathbf{x} - \mathbf{p}) = r^2, \quad (2.4)$$

que es el cuadrado de (2.1). Esta ecuación, que llamaremos la *ecuación vectorial del círculo*, se puede también reescribir como

$$\mathbf{x} \cdot \mathbf{x} - 2\mathbf{p} \cdot \mathbf{x} + \mathbf{p} \cdot \mathbf{p} = r^2.$$

Parece superfluo pero tiene grandes ventajas. Por ejemplo, en algunos de los ejercicios siguientes puede ayudar, pero además, y mucho más profundo, al no hacer referencia a las coordenadas tiene sentido en cualquier dimensión. Sirve entonces para definir esferas en \mathbb{R}^3 y, como veremos en la siguiente sección (en la que vale la pena

que el estudiante tenga en mente cómo traducir a \mathbb{R}^3 , esferas y planos), de la ecuación vectorial se puede extraer mucha información geométrica interesante.

Debemos también remarcar que en las siguientes secciones, donde empezamos a estudiar las otras cónicas, no aparece el equivalente de esta ecuación vectorial del círculo. Para lograrlo en general (en el capítulo 9) nos falta desarrollar mucha más teoría.

EJERCICIO 2.1 ¿Cuáles de las siguientes ecuaciones son ecuaciones de un círculo? Y en su caso, ¿de cuál?

- a) $x^2 - 6x + y^2 - 4y = 12$
- b) $x^2 + 4x + y^2 + 2y = 11$
- c) $2x^2 + 8x + 2y^2 - 4y = -8$
- d) $x^2 - 4x + y^2 - 2y = -6$
- e) $4x^2 + 4x + y^2 - 2y = 4$.

EJERCICIO 2.2 ¿Cuál es el lugar geométrico de los centros de los círculos que pasan por dos puntos (distintos) \mathbf{a} y \mathbf{b} ?

EJERCICIO 2.3 Sean \mathbf{a} y \mathbf{b} dos puntos (distintos) en el plano; y sean \mathcal{C}_1 el círculo con centro \mathbf{a} y radio r_1 y \mathcal{C}_2 el círculo con centro \mathbf{b} y radio r_2 . Demuestra que \mathcal{C}_1 y \mathcal{C}_2 son *tangentes* (i.e. se intersectan en un único punto) si y sólo si $r_1 + r_2 = d(\mathbf{a}, \mathbf{b})$ o $r_1 + d(\mathbf{a}, \mathbf{b}) = r_2$ o $r_2 + d(\mathbf{a}, \mathbf{b}) = r_1$. ¿A qué corresponden geoméricamente las tres condiciones anteriores? Haz un dibujo de cada caso.

EJERCICIO 2.4 Dados los círculos del ejercicio anterior, da condiciones sobre sus radios y la distancia entre sus centros para que:

- a) el círculo \mathcal{C}_1 esté contenido en el interior del círculo \mathcal{C}_2 ;
- b) el círculo \mathcal{C}_1 esté contenido en el exterior del círculo \mathcal{C}_2 (y viceversa);
- c) el círculo \mathcal{C}_1 y el círculo \mathcal{C}_2 se intersecten en dos puntos.

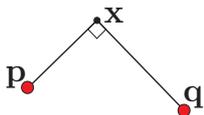
EJERCICIO 2.5 Demuestra que por tres puntos no colineales \mathbf{a}, \mathbf{b} y \mathbf{c} pasa un único círculo, llamado el *circuncírculo* del triángulo $\mathbf{a}, \mathbf{b}, \mathbf{c}$.

EJERCICIO 2.6 Demuestra que dadas tres líneas no concurrentes y no paralelas dos a dos existen exactamente cuatro círculos tangentes a las tres; al que está contenido en el interior del triángulo se le llama su *incírculo*.

EJERCICIO 2.7 Sean \mathbf{p} y \mathbf{q} dos puntos distintos en el plano (o en \mathbb{R}^3). Demuestra que el conjunto de puntos \mathbf{x} cuyas líneas a \mathbf{p} y a \mathbf{q} son ortogonales, es un círculo (o una esfera) que “quiere contener” a \mathbf{p} y a \mathbf{q} . Más precisamente, que el conjunto definido por la ecuación

$$(\mathbf{x} - \mathbf{p}) \cdot (\mathbf{x} - \mathbf{q}) = 0$$

es el círculo con el segmento $\overline{\mathbf{p}\mathbf{q}}$ como diámetro. (Compara esto con el ejercicio siguiente).



EJERCICIO 2.8 Sean \mathbf{p} y \mathbf{q} dos puntos distintos en el plano. ¿Para cuáles números reales c se tiene que la ecuación

$$(\mathbf{x} - \mathbf{p}) \cdot (\mathbf{x} - \mathbf{q}) = c$$

define un círculo? En su caso, ¿cuál es el radio y dónde está su centro?

2.1.1 Tangentes y polares

Observemos primero que las líneas *tangentes* a un círculo son las normales a los *radios* (los segmentos del centro a sus puntos). Efectivamente, si \mathbf{a} es un punto del círculo \mathcal{C} dado por la ecuación vectorial

$$(\mathbf{x} - \mathbf{p}) \cdot (\mathbf{x} - \mathbf{p}) = r^2, \quad (2.4)$$

entonces su tangente es la recta ℓ normal a $(\mathbf{a} - \mathbf{p})$ y que pasa por \mathbf{a} . Pues \mathbf{a} es el punto más cercano a \mathbf{p} en esta recta, de tal manera que para cualquier otro punto $\mathbf{x} \in \ell$ se tiene que $d(\mathbf{x}, \mathbf{p}) > r$. Como el círculo \mathcal{C} parte el plano en dos pedazos (el *interior* donde $d(\mathbf{x}, \mathbf{p}) < r$, y el *exterior* donde $d(\mathbf{x}, \mathbf{p}) > r$), entonces ℓ está contenida en el exterior salvo por el punto $\mathbf{a} \in \mathcal{C}$; ésta será nuestra definición de *tangente*.

Claramente, la recta ℓ está dada por la ecuación

$$\mathbf{x} \cdot (\mathbf{a} - \mathbf{p}) = \mathbf{a} \cdot (\mathbf{a} - \mathbf{p}),$$

que tiene una manera mucho más interesante de escribirse. Restando $\mathbf{p} \cdot (\mathbf{a} - \mathbf{p})$ a ambos lados se obtiene

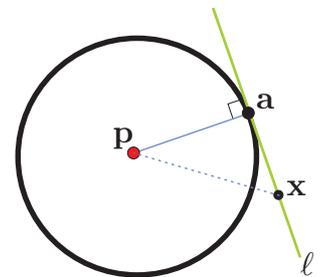
$$\begin{aligned} \mathbf{x} \cdot (\mathbf{a} - \mathbf{p}) - \mathbf{p} \cdot (\mathbf{a} - \mathbf{p}) &= \mathbf{a} \cdot (\mathbf{a} - \mathbf{p}) - \mathbf{p} \cdot (\mathbf{a} - \mathbf{p}) \\ (\mathbf{x} - \mathbf{p}) \cdot (\mathbf{a} - \mathbf{p}) &= (\mathbf{a} - \mathbf{p}) \cdot (\mathbf{a} - \mathbf{p}) \\ (\mathbf{x} - \mathbf{p}) \cdot (\mathbf{a} - \mathbf{p}) &= r^2, \end{aligned}$$

pues $\mathbf{a} \in \mathcal{C}$. La ecuación de la tangente se obtiene entonces al sustituir el punto en una de las dos apariciones de la variable \mathbf{x} en la ecuación vectorial (2.4).

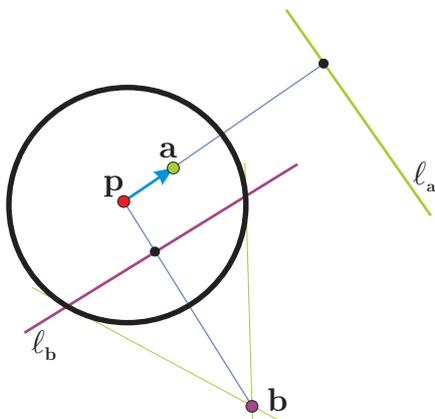
Para cualquier otro punto en el plano, \mathbf{a} digamos, que no sea el centro ($\mathbf{a} \neq \mathbf{p}$), el mismo proceso algebraico —sustituir \mathbf{a} en una de las instancias de \mathbf{x} en la ecuación vectorial— nos da la ecuación de una recta, llamada *la polar* de \mathbf{a} respecto al círculo \mathcal{C} , y que denotaremos $\ell_{\mathbf{a}}$; es decir,

$$\text{sea } \ell_{\mathbf{a}}: (\mathbf{x} - \mathbf{p}) \cdot (\mathbf{a} - \mathbf{p}) = r^2. \quad (2.5)$$

Ya hemos visto que cuando $\mathbf{a} \in \mathcal{C}$, su polar $\ell_{\mathbf{a}}$ es su tangente. Ahora veremos que si \mathbf{a} está en el interior del círculo, entonces $\ell_{\mathbf{a}}$ no lo interseca (está totalmente



contenida en el exterior), y que si está en el exterior (el punto \mathbf{b} en la figura siguiente), entonces lo corta, y además lo corta en los dos puntos de \mathcal{C} a los cuales se pueden trazar tangentes.



Para esto, expresemos la ecuación de ℓ_a en su forma normal; desarrollando (2.5) se obtiene:

$$\mathbf{x} \cdot (\mathbf{a} - \mathbf{p}) = r^2 + \mathbf{p} \cdot (\mathbf{a} - \mathbf{p}).$$

Esto indica que ℓ_a es perpendicular al vector que va de \mathbf{p} a \mathbf{a} . Ahora veamos cuál es su punto de intersección con la recta que pasa por \mathbf{p} y \mathbf{a} . Parametricemos esta última recta con \mathbf{p} de cero y \mathbf{a} de uno (es decir como $\mathbf{p} + t(\mathbf{a} - \mathbf{p})$) y podemos despejar t al sustituir en la variable \mathbf{x} de la ecuación anterior (o bien, esto se ve más directo al sustituir en la (2.5)), para obtener

$$t = \frac{r^2}{(\mathbf{a} - \mathbf{p}) \cdot (\mathbf{a} - \mathbf{p})} = \frac{r^2}{d(\mathbf{p}, \mathbf{a})^2}.$$

Entonces la distancia de \mathbf{p} a ℓ_a es

$$d(\mathbf{p}, \ell_a) = t d(\mathbf{p}, \mathbf{a}) = \frac{r^2}{d(\mathbf{p}, \mathbf{a})} = \left(\frac{r}{d(\mathbf{p}, \mathbf{a})} \right) r \quad (2.6)$$

y tenemos lo primero que queríamos: si $d(\mathbf{p}, \mathbf{a}) < r$ entonces $d(\mathbf{p}, \ell_a) > r$; y al revés, si $d(\mathbf{p}, \mathbf{a}) > r$ entonces $d(\mathbf{p}, \ell_a) < r$. Si el punto \mathbf{a} está muy cerca de \mathbf{p} , su polar está muy lejos, y al revés, sus distancias al centro \mathbf{p} se comportan como inversos pero “alrededor de r ” (si $r = 1$ son precisamente inversos).

Supongamos ahora que $d(\mathbf{p}, \mathbf{a}) > r$, y sea \mathbf{c} un punto en $\ell_a \cap \mathcal{C}$ (que sabemos que existe pues ℓ_a pasa por el interior de \mathcal{C}). Puesto que $\mathbf{c} \in \ell_a$, se cumple la ecuación

$$(\mathbf{c} - \mathbf{p}) \cdot (\mathbf{a} - \mathbf{p}) = r^2.$$

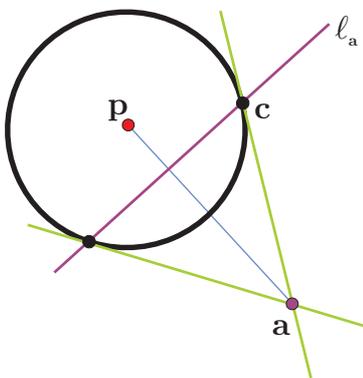
Pero entonces \mathbf{a} cumple la ecuación de ℓ_c que es la tangente a \mathcal{C} en \mathbf{c} ; es decir, la línea de \mathbf{a} a \mathbf{c} es tangente al círculo.

Nótese que el argumento anterior es mucho más general: demuestra que para cualesquiera dos puntos \mathbf{a} y \mathbf{b} (distintos de \mathbf{p}) se tiene que

$$\mathbf{a} \in \ell_b \Leftrightarrow \mathbf{b} \in \ell_a. \quad (2.7)$$

Y los puntos del círculo son los únicos para los cuales se cumple $\mathbf{a} \in \ell_a$.

En particular, hemos aprendido a calcular los puntos de tangencia a un círculo desde un punto exterior \mathbf{a} . A saber, de la ecuación lineal de su polar, ℓ_a , se despeja alguna de las dos variables y se sustituye en la ecuación del círculo. Esto nos da una ecuación de segundo grado en la otra variable que se puede resolver, y nos da dos raíces. Sustituyéndolas de nuevo en la ecuación de la polar se obtiene el otro par de coordenadas.



Por ejemplo, la ecuación del círculo (2.3) tiene la expresión vectorial

$$((x, y) - (3, -1)) \cdot ((x, y) - (3, -1)) = 4.$$

Si queremos conocer los puntos de tangencia desde el punto $\mathbf{a} = (1, 3)$, encontramos primero la ecuación de su línea polar:

$$\begin{aligned} ((x, y) - (3, -1)) \cdot ((1, 3) - (3, -1)) &= 4 \\ (x - 3, y + 1) \cdot (-2, 4) &= 4 \\ -2x + 4y + 10 &= 4 \\ x - 2y &= 3. \end{aligned}$$

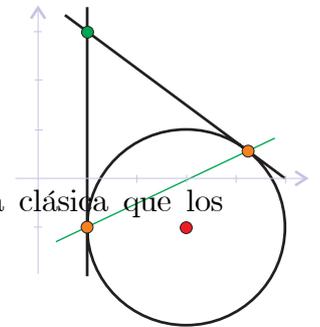
De aquí, para encontrar $\ell_{\mathbf{a}} \cap \mathcal{C}$, conviene sustituir $x = 3 + 2y$ en la ecuación original del círculo (2.3) para obtener

$$\begin{aligned} (3 + 2y)^2 + y^2 - 6(3 + 2y) + 2y &= -6 \\ 5y^2 + 2y - 3 &= 0. \end{aligned}$$

Las raíces de esta ecuación cuadrática se obtienen por la fórmula clásica que los estudiantes de la ciudad de México llaman “del chicharronero”:¹

$$y = \frac{-2 \pm \sqrt{4 + 60}}{10} = \frac{-2 \pm 8}{10},$$

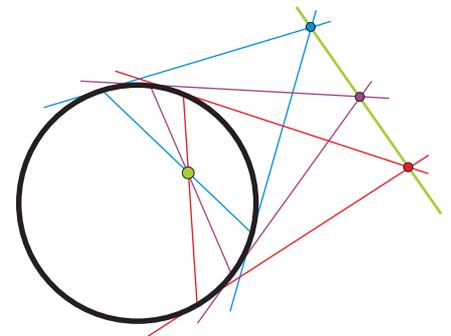
que nos da los valores $y_0 = -1$ y $y_1 = 3/5$. Y estos, al sustituir de nuevo en la ecuación de la polar nos dan los puntos de tangencia de \mathbf{a} ; que son $(1, -1)$ y $(1/5, 3)$. El estudiante puede verificar que satisfacen la ecuación del círculo, aunque los hayamos obtenido (al final) sustituyendo en la ecuación lineal de la polar, y que efectivamente sus tangentes pasan por \mathbf{a} .



EJERCICIO 2.9 Encuentra los puntos de tangencia:

- al círculo $x^2 - 2x + y^2 - 4y = -3$ desde el punto $(-1, 2)$;
- al círculo $x^2 + y^2 = 1$ desde el punto $(2, 2)$.

EJERCICIO 2.10 Sea σ una secante de un círculo \mathcal{C} , es decir un segmento con extremos en el círculo. Si trazamos las tangentes por sus extremos, éstas se intersectan en un punto, que llamaremos el *polar* de σ . Demuestra que tres secantes son concurrentes si y sólo si sus puntos polares son colineales.



¹El polinomio $ax^2 + bx + c$ tiene como raíces a $(-b \pm \sqrt{b^2 - 4ac})/2a$.

EJERCICIO 2.11 El estudiante que en el ejercicio anterior observó la necesidad formal de eliminar los diámetros en la definición de *punto polar* de un segmento, hizo muy bien; pero el que no se puso quisquilloso, y ni cuenta se dio, hizo mejor. Pensó proyectivamente. Demuestra que si permitimos que el centro \mathbf{p} entre a nuestras consideraciones anteriores y añadimos puntos “polares” a los diámetros que juntos forman la nueva línea $\ell_{\mathbf{p}}$ polar con \mathbf{p} , entonces la polaridad se extiende a todas las líneas (falta definirla en las tangentes y exteriores) y que es una biyección entre líneas y puntos que cambia concurrencia por colinearidad. Observa además que los nuevos puntos (polares de los diámetros) tienen una dirección (no orientada) bien definida, aquella en la que se alejan los puntos polares de una secante que se acerca al diámetro correspondiente.

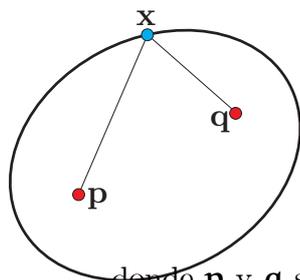
EJERCICIO 2.12 Demuestra que si \mathbf{c} es un punto exterior (al círculo \mathcal{C} con centro \mathbf{p}) entonces su recta a \mathbf{p} bisecta sus dos tangentes a \mathcal{C} . Y además que las distancias a sus *pies* en \mathcal{C} (es decir, a los puntos de tangencia) son iguales.

EJERCICIO 2.13 Observa que toda nuestra discusión sobre líneas polares se basó en la ecuación vectorial (2.4) que tiene sentido en \mathbb{R}^3 donde define esferas. Define la noción de plano polar a un punto respecto de una esfera en \mathbb{R}^3 . Demuestra que un punto \mathbf{c} en el exterior tiene como plano polar a uno que intersecta a la esfera en un círculo formado por los puntos de la esfera cuyas líneas a \mathbf{c} son tangentes (a la unión de estas líneas se le llama el *cono* de la esfera con ápice \mathbf{c}).

2.2 Elipses

Si amarramos un burro hambriento a una estaca en un prado, se irá comiendo el pasto deambulando al azar, pero terminará por dejar pelón el interior de un círculo (con centro la estaca y radio la longitud del mecate). Si amarramos los extremos del mecate a dos estacas y le engarzamos al burro, lo que dibujará es una elipse; por eso, a fijar dos tachuelas en un papel, amarrar holgadamente un hilo entre ellas y luego tensar con un lápiz y girarlo, se le conoce como “*el método del jardinero*” para

trazar elipses. Formalmente, las *elipses* son el lugar geométrico de los puntos (posiciones de la boca del burro cuando tensa el mecate) cuya suma de distancias a dos puntos fijos llamados *focos* (las estacas) es constante (la longitud del mecate). De tal manera que una elipse \mathcal{E} queda determinada por la ecuación



$$d(\mathbf{x}, \mathbf{p}) + d(\mathbf{x}, \mathbf{q}) = 2a, \quad (2.8)$$

donde \mathbf{p} y \mathbf{q} son los focos y a es una constante positiva, llamada *semieje mayor*, tal que $2a > d(\mathbf{p}, \mathbf{q})$.

Incluimos el coeficiente 2 en la constante para que quede claro que si los focos coinciden, $\mathbf{p} = \mathbf{q}$; entonces se obtiene un círculo de radio a y centro en el foco; así que los círculos son elipses muy especiales. Esta ecuación, poniéndole coordenadas a

los focos, incluye raíces cuadradas por las distancias, y es aún muy “fea”. Veremos ahora que en un caso especial es equivalente a una ecuación cuadrática “bonita”.

Supongamos que el *centro* de la elipse \mathcal{E} , i.e., el punto medio entre los focos, está en el origen y que además los focos están en el eje x . Entonces tenemos que $\mathbf{p} = (c, 0)$ y $\mathbf{q} = (-c, 0)$ para alguna c tal que $0 < c < a$ (donde ya suponemos que la elipse no es un círculo al pedir $0 < c$). Es fácil ver entonces que la intersección de \mathcal{E} con el eje x consiste en los puntos $(a, 0)$ y $(-a, 0)$, pues la ecuación (2.8) para puntos $(x, 0)$ es

$$|x - c| + |x + c| = 2a,$$

que sólo tiene las soluciones $x = a$ y $x = -a$ (véase el Ejercicio 2.15), y de aquí el nombre de *semieje mayor* para la constante a . Como el eje y es ahora la mediatriz de los focos, en él, es decir en los puntos $(0, y)$, la ecuación (2.8) se vuelve

$$\sqrt{c^2 + y^2} = a,$$

que tiene soluciones $y = \pm b$, donde $b > 0$, llamado el *semieje menor* de la elipse \mathcal{E} , es tal que

$$b^2 = a^2 - c^2.$$

Ahora sí, consideremos la ecuación (2.8), que con $\mathbf{x} = (x, y)$ y la definición de nuestros focos, se expresa

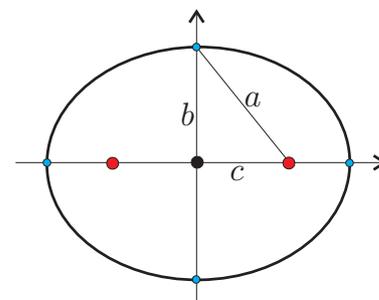
$$\sqrt{(x - c)^2 + y^2} + \sqrt{(x + c)^2 + y^2} = 2a. \quad (2.9)$$

Si elevamos al cuadrado directamente esta ecuación, en el lado izquierdo nos quedaría un incómodo término con raíces. Así que conviene pasar una de las dos raíces al otro lado, para obtener

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

Al elevar al cuadrado, se tiene

$$\begin{aligned} (x - c)^2 + y^2 &= 4a^2 - 4a\sqrt{(x + c)^2 + y^2} + (x + c)^2 + y^2 & (2.10) \\ x^2 - 2cx + c^2 &= 4a^2 - 4a\sqrt{(x + c)^2 + y^2} + x^2 + 2cx + c^2 \\ 4a\sqrt{(x + c)^2 + y^2} &= 4a^2 + 4cx \\ a\sqrt{(x + c)^2 + y^2} &= a^2 + cx \end{aligned}$$



Elevando de nuevo al cuadrado, nos deshacemos de la raíz, y después, al agrupar términos, obtenemos

$$\begin{aligned} a^2 \left((x+c)^2 + y^2 \right) &= a^4 + 2a^2cx + c^2x^2 & (2.11) \\ a^2x^2 + 2a^2cx + a^2c^2 + a^2y^2 &= a^4 + 2a^2cx + c^2x^2 \\ (a^2 - c^2)x^2 + a^2y^2 &= a^2(a^2 - c^2) \\ b^2x^2 + a^2y^2 &= a^2b^2, \end{aligned}$$

que, al dividir entre a^2b^2 , se escribe finalmente como

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad (2.12)$$

llamada la *ecuación canónica de la elipse*.

Hemos demostrado que si $\mathbf{x} = (x, y)$ satisface (2.8), entonces satisface (2.12). Lo inverso, aunque también es cierto, no es tan automático, y requiere más argumentación pues en dos ocasiones (al pasar a las ecuaciones (2.10) y (2.11)) elevamos al cuadrado. Es cierto que si dos números son iguales entonces también sus cuadrados lo son (que fue lo que usamos); pero si $r^2 = t^2$ no necesariamente se tiene que $r = t$, pues podría ser que $r = -t$. Éste es el problema conceptual conocido como “la raíz cuadrada tiene dos ramas”, entonces los pasos de las ecuaciones (2.10) y (2.11) hacia arriba requieren argumentación sobre los signos (todos los demás pasos hacia arriba sí son automáticos). A manera de verificación, demostremos el inverso directamente.

Supongamos que $\mathbf{x} = (x, y)$ satisface la ecuación canónica de la elipse (2.12); veremos que entonces satisface la ecuación (2.9). Como en (2.12) se tiene una suma de números positivos que dan 1, entonces ambos sumandos están entre 0 y 1. Por lo tanto $|x| \leq |a| = a$, y esto habrá que usarse al sacar una raíz. De la ecuación (2.12) se puede despejar y^2 y sustituirla en las expresiones dentro de las raíces de (2.9), para obtener, en la primera,

$$\begin{aligned} (x-c)^2 + y^2 &= (x-c)^2 + b^2 - \frac{b^2}{a^2}x^2 \\ &= \left(1 - \frac{b^2}{a^2}\right)x^2 - 2cx + c^2 + b^2 \\ &= \left(\frac{a^2 - b^2}{a^2}\right)x^2 - 2cx + (c^2 + b^2) \\ &= \frac{c^2}{a^2}x^2 - 2cx + a^2 = \left(\frac{c}{a}x - a\right)^2. \end{aligned}$$

Ahora bien, como $|x| \leq |a| = a$ implica que $x \leq a$, y además $0 < c < a$, entonces $cx < a^2$ y por tanto $\frac{c}{a}x < a$. Así que la expresión que pusimos arriba es negativa y entonces la raíz es la otra posible rama; es decir, hemos demostrado que

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

De manera análoga, pero usando ahora que $-x \leq a$, se demuestra que

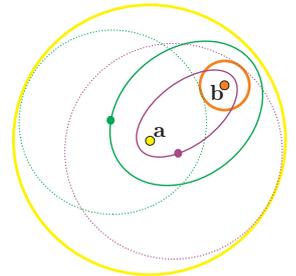
$$\sqrt{(x+c)^2 + y^2} = a + \frac{c}{a}x,$$

de donde se sigue la ecuación de las distancias (2.9).

EJERCICIO 2.14 Dibuja algunas elipses usando el método del jardinero. ¿Verdad que no salen de un solo trazo?

EJERCICIO 2.15 Sea $c > 0$. Dibuja la gráfica de la función $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por $f(x) = |x - c| + |x + c|$. Concluye que si $a > c$ entonces $f(x) = 2a$ si y sólo si $|x| = a$.

EJERCICIO 2.16 Sean \mathcal{C}_1 y \mathcal{C}_2 dos círculos con centros \mathbf{a} y \mathbf{b} respectivamente, tales que \mathcal{C}_2 está contenido en el interior de \mathcal{C}_1 . Demuestra que el conjunto de puntos que son centro de un círculo \mathcal{C}_3 tangente tanto a \mathcal{C}_1 como a \mathcal{C}_2 es un par de elipses con focos \mathbf{a} y \mathbf{b} . Observa que en el caso límite en que \mathcal{C}_1 es tangente a \mathcal{C}_2 , una de las elipses se degenera en un segmento.

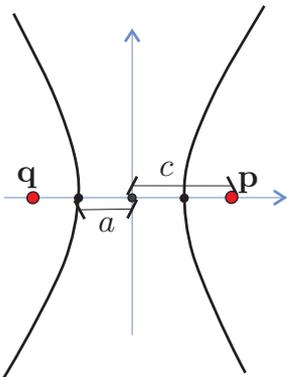


2.3 Hipérbolas

La *hipérbola* está definida como el lugar geométrico de los puntos cuya diferencia (en valor absoluto) de sus distancias a dos puntos fijos \mathbf{p} y \mathbf{q} , llamados focos, es constante. Entonces, una hipérbola \mathcal{H} está definida por la ecuación

$$|d(\mathbf{x}, \mathbf{p}) - d(\mathbf{x}, \mathbf{q})| = 2a, \tag{2.13}$$

donde $a > 0$, y además $2a < d(\mathbf{p}, \mathbf{q}) =: 2c$ (véase el Ejercicio 2.17). Si tomamos como focos a $\mathbf{p} = (c, 0)$ y $\mathbf{q} = (-c, 0)$, esta ecuación toma la forma

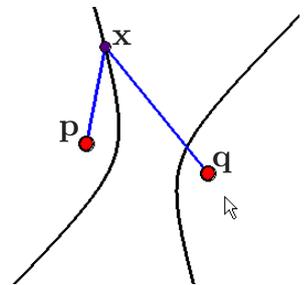


$$\left| \sqrt{(x-c)^2 + y^2} - \sqrt{(x+c)^2 + y^2} \right| = 2a. \tag{2.14}$$

Veremos a continuación que es equivalente a

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \tag{2.15}$$

donde $b > 0$ está definida por $a^2 + b^2 = c^2$. A esta última ecuación se le llama la *ecuación canónica de la hipérbola*.



Como las dos primeras ecuaciones involucran valores absolutos, entonces tienen dos posibilidades que corresponden a las dos ramas de la hipérbola. En una de ellas la distancia a uno de los focos es mayor y en la otra se invierten los papeles. De la ecuación (2.14) se tienen dos posibilidades:

$$\begin{aligned}\sqrt{(x-c)^2 + y^2} &= 2a + \sqrt{(x+c)^2 + y^2} \\ \sqrt{(x+c)^2 + y^2} &= 2a + \sqrt{(x-c)^2 + y^2};\end{aligned}\tag{2.16}$$

la primera corresponde a la rama donde $x < 0$ y la segunda a $x > 0$. Como se hizo en el caso de la elipse (elevando al cuadrado, simplificando para aislar la raíz que queda y luego volviendo a elevar al cuadrado y simplificando, véase Ejercicio 2.18) se obtiene, de cualquiera de las dos ecuaciones anteriores, la ecuación

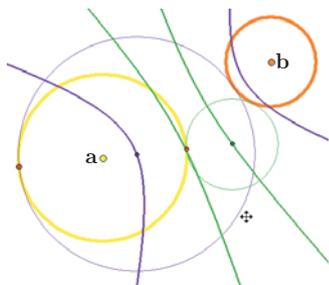
$$(a^2 - c^2)x^2 + a^2y^2 = a^2(a^2 - c^2).\tag{2.17}$$

Y de aquí, al sustituir $-b^2 = a^2 - c^2$, y dividir entre $-a^2b^2$ se obtiene la ecuación canónica.

EJERCICIO 2.17 Demuestra que la ecuación (2.13) define la mediatriz de \mathbf{p} y \mathbf{q} para $a = 0$; los rayos complementarios del segmento $\overline{\mathbf{p}\mathbf{q}}$ para $a = c$, y el conjunto vacío para $a > c$.

EJERCICIO 2.18 Demuestra que la ecuación (2.17) se sigue de cualquiera de las dos anteriores (2.16).

EJERCICIO 2.19 Demuestra que si (x, y) satisface la ecuación (2.15) entonces $|x| \geq a$. Concluye (sustituyendo en las raíces, como en el caso de la elipse) que, entonces, alguna de las dos ecuaciones (2.16) se satisfacen.



EJERCICIO 2.20 Sean \mathcal{C}_1 y \mathcal{C}_2 dos círculos con centros \mathbf{a} y \mathbf{b} respectivamente, y sea \mathcal{X} el conjunto de los puntos que son centro de un círculo \mathcal{C}_3 tangente tanto a \mathcal{C}_1 como a \mathcal{C}_2 .

a) Demuestra que si \mathcal{C}_1 y \mathcal{C}_2 están fuera uno del otro y sus radios son distintos, entonces \mathcal{X} consta de un par de hipérbolas con focos \mathbf{a} y \mathbf{b} .

b) Demuestra que si \mathcal{C}_1 y \mathcal{C}_2 se intersectan pero no son tangentes y sus radios son distintos, entonces \mathcal{X} consta de un hipérbola y una elipse con focos \mathbf{a} y \mathbf{b} .

c) Discute los casos límite entre los casos anteriores y el del Ejercicio 2.16 y el de radios iguales.

2.4 Parábolas

Una *parábola* es el lugar geométrico de los puntos que equidistan de un punto \mathbf{p} (llamado su *foco*) y una recta ℓ , llamada su *directriz*, donde $\mathbf{p} \notin \ell$. Es decir, está

definida por la ecuación

$$d(\mathbf{x}, \mathbf{p}) = d(\mathbf{x}, \ell).$$

Tomemos un ejemplo sencillo con el foco en el eje \mathbf{y} , la directriz paralela al eje \mathbf{x} , y que además pase por el origen. Tenemos entonces $\mathbf{p} = (0, c)$, donde $c > 0$ digamos, y $\ell: \mathbf{y} = -c$; de tal manera que la parábola queda determinada por la ecuación

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

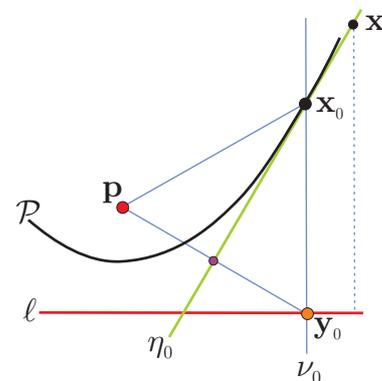
Como ambos lados de la ecuación son positivos, ésta es equivalente a la igualdad de sus cuadrados que da

$$\begin{aligned} x^2 + y^2 - 2cy + c^2 &= y^2 + 2cy + c^2 \\ x^2 &= 4cy. \end{aligned}$$

De tal manera que la gráfica de la función x^2 ($y = x^2$) es una parábola con foco $(0, 1/4)$ y directriz $y = -1/4$. Veremos ahora que cualquier parábola cumple “la propiedad de ser gráfica de una función” respecto de su directriz.

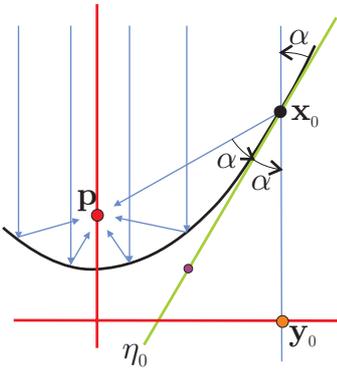
Sea \mathcal{P} la parábola con foco \mathbf{p} y directriz ℓ (donde $\mathbf{p} \notin \ell$). Dado un punto $\mathbf{y}_0 \in \ell$, es claro que los puntos del plano cuya distancia a ℓ coincide con (o se mide por) su distancia a \mathbf{y}_0 son precisamente los de la normal a ℓ que pasa por \mathbf{y}_0 , llamémosla ν_0 . Por otro lado, la mediatriz entre \mathbf{y}_0 y \mathbf{p} , llamémosla η_0 , consta de los puntos cuyas distancias a \mathbf{p} y a \mathbf{y}_0 coinciden. Por lo tanto la intersección de η_0 y ν_0 está en la parábola \mathcal{P} , es decir, $\mathbf{x}_0 = \nu_0 \cap \eta_0 \in \mathcal{P}$, pues $d(\mathbf{x}_0, \ell) = d(\mathbf{x}_0, \mathbf{y}_0) = d(\mathbf{x}_0, \mathbf{p})$. Pero además \mathbf{x}_0 es el único punto en la normal ν_0 que está en \mathcal{P} . Ésta es “la propiedad de la gráfica” a la que nos referíamos.

Podemos concluir aún más: que la mediatriz η_0 es la tangente a \mathcal{P} en \mathbf{x}_0 . Pues para cualquier otro punto $\mathbf{x} \in \eta_0$ se tiene que su distancia a ℓ es menor que su distancia a \mathbf{y}_0 que es su distancia a \mathbf{p} ($d(\mathbf{x}, \ell) < d(\mathbf{x}, \mathbf{y}_0) = d(\mathbf{x}, \mathbf{p})$), y entonces $\mathbf{x} \notin \mathcal{P}$. De hecho, la parábola \mathcal{P} parte el plano en dos pedazos, los puntos más cerca de \mathbf{p} que de ℓ (lo de adentro, digamos, definido por la desigualdad $d(\mathbf{x}, \mathbf{p}) \leq d(\mathbf{x}, \ell)$) y los que están más cerca de ℓ que de \mathbf{p} (lo de afuera, dado por $d(\mathbf{x}, \ell) \leq d(\mathbf{x}, \mathbf{p})$ en donde está η_0), que comparten la frontera donde estas distancias coinciden (la parábola \mathcal{P}). Así que η_0 pasa *tangente* a \mathcal{P} en \mathbf{x}_0 , pues $\mathbf{x}_0 \in \eta_0 \cap \mathcal{P}$ y además η_0 se queda de un lado de \mathcal{P} .



2.5 Propiedades focales

2.5.1 De la parábola



Del análisis geométrico anterior es fácil deducir la *propiedad focal* de la parábola que la ha hecho tan importante en la tecnología de los siglos recientes, pero que se conoce desde que Apolonio de Pergas (c. 255-170 A.C.) escribió “*Sobre las cónicas*”.

Como la mediatriz η_0 es bisectriz del ángulo $\mathbf{p}x_0\mathbf{y}_0$, que llamaremos 2α , entonces un fotón (partícula de luz) que viaja por ν_0 (la normal a ℓ) hacia ℓ y “dentro” de la parábola incide en ella con ángulo α (este ángulo se mide infinitesimalmente, es decir, con la tangente η_0); como el ángulo de incidencia es igual al de reflexión, entonces rebota en el rayo que va directo al foco \mathbf{p} .

Llamemos *eje* de la parábola a la perpendicular a la directriz que pasa por el foco. Se tiene entonces que un haz de fotones que viaja en rayos paralelos al eje se va reflejando en la parábola para convertirse en un haz que converge en el foco \mathbf{p} .

Una antena parabólica casera (que se obtiene girando en su eje a una parábola) apuntada a un satélite, refleja entonces la onda de radio que viene de éste (el satélite está tan lejos que esta onda ya “es” paralela) y la hace converger en el receptor (puesto en el foco): “medio metro” cuadrado de la onda se concentra simultáneo en el receptor puntual; tal amplificación hace fácil captarla. El proceso inverso se usa en lámparas o en antenas transmisoras, y reflejan un haz que emana del foco en un haz paralelo al eje.

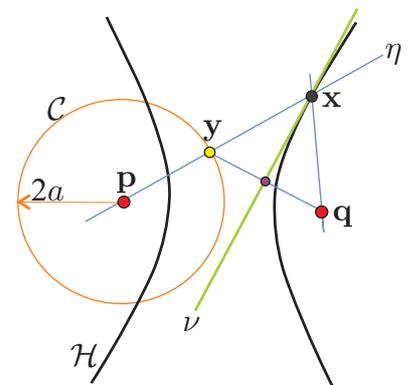
2.5.2 De la hipérbola

¿Qué pasa si plateamos (hacemos espejo) el otro lado de la parábola? Es fácil ver que entonces el haz paralelo al eje que incide en ella (por fuera) se refleja en un haz que emana (o simula emanar) del foco. Si pensamos que un haz paralelo viene de un punto al infinito y queremos reproducir este fenómeno para un haz puntual (que emana de un punto) hay que usar la hipérbola.

Sea \mathcal{H} una hipérbola con focos \mathbf{p} y \mathbf{q} ; el haz que emana del foco \mathbf{p} se refleja en la rama de \mathcal{H} cercana a \mathbf{q} , en un haz que emana de \mathbf{q} . Entonces el uso tecnológico de la hipérbola es con refracción: si el fotón cruza la hipérbola con ángulo de incidencia igual al de refracción, entonces el haz que emana de \mathbf{p} se refracta en uno confluyente en \mathbf{q} (algo así se usa en el diseño de las lentes). Demostrar esto equivale a demostrar que la tangente a la hipérbola \mathcal{H} en un punto \mathbf{x} es bisectriz de sus líneas a los focos \mathbf{p} y \mathbf{q} . Para demostrarlo usaremos una construcción análoga a la de la parábola pero cambiando el uso de las normales a la directriz (que “vienen” del infinito) por el haz de rayos que sale de \mathbf{p} .

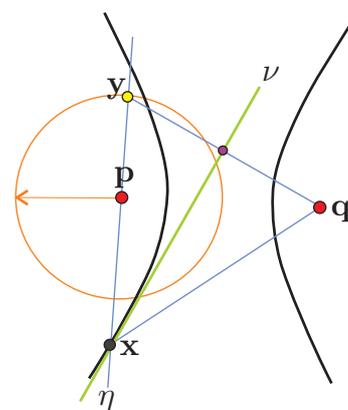
Supongamos que \mathcal{H} está acabada de determinar por la constante $2a$, es decir, que

$$\mathcal{H}: \quad |d(\mathbf{x}, \mathbf{p}) - d(\mathbf{x}, \mathbf{q})| = 2a,$$

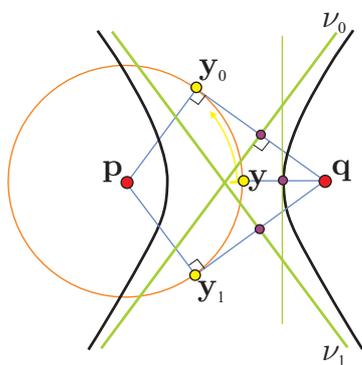


donde $0 < 2a < d(\mathbf{q}, \mathbf{p})$. Sea \mathcal{C} el círculo con centro en \mathbf{p} y radio $2a$, de tal manera que para cada $\mathbf{y} \in \mathcal{C}$ tenemos un rayo que sale de \mathbf{p} . Para fijar ideas, tomemos una $\mathbf{y} \in \mathcal{C}$ que apunta cerca de \mathbf{q} (pero no justo a \mathbf{q}). Los puntos en el rayo $\vec{\eta}_{\mathbf{y}}$ que sale de \mathbf{p} hacia \mathbf{y} usan este rayo para medir su distancia a \mathbf{p} . Si tomamos a $\nu_{\mathbf{y}}$ como la mediatriz de \mathbf{y} y el otro foco \mathbf{q} , entonces (como en la parábola) se tiene que $\mathbf{x} := \vec{\eta}_{\mathbf{y}} \cap \nu_{\mathbf{y}} \in \mathcal{H}$ pues $d(\mathbf{x}, \mathbf{q}) = d(\mathbf{x}, \mathbf{y})$ (porque $\mathbf{x} \in \nu_{\mathbf{y}}$) y $d(\mathbf{x}, \mathbf{p}) = d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{p}) = d(\mathbf{x}, \mathbf{y}) + 2a$ (donde estamos suponiendo que \mathbf{x} está en el rayo que sale de \mathbf{p} y pasa por \mathbf{y} como en la figura), de tal manera que de las dos condiciones obtenemos que $d(\mathbf{x}, \mathbf{p}) - d(\mathbf{x}, \mathbf{q}) = 2a$.

Antes de seguir adelante, consideremos el otro caso, a saber, cuando el punto de intersección \mathbf{x} de las rectas $\eta_{\mathbf{y}}$ (por \mathbf{p} y $\mathbf{y} \in \mathcal{C}$) y $\nu_{\mathbf{y}}$, la mediatriz de \mathbf{y} y \mathbf{q} , está del otro lado de \mathbf{y} (es decir, cuando \mathbf{p} está en el segmento $\overline{\mathbf{x}\mathbf{y}}$). Se tiene entonces que $d(\mathbf{x}, \mathbf{q}) = d(\mathbf{x}, \mathbf{y}) = d(\mathbf{x}, \mathbf{p}) + d(\mathbf{p}, \mathbf{y}) = d(\mathbf{x}, \mathbf{p}) + 2a$, de tal manera que \mathbf{x} está en la otra rama de la hipérbola.

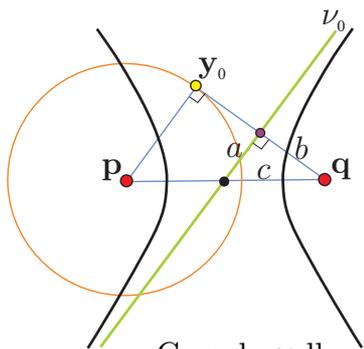


Hemos demostrado que si $\mathbf{y} \in \mathcal{C}$ es tal que su recta por él y \mathbf{p} , $\eta_{\mathbf{y}}$ y su mediatriz con \mathbf{q} , $\nu_{\mathbf{y}}$, se intersectan en un punto \mathbf{x} , entonces $\mathbf{x} \in \mathcal{H}$. Como en el caso de la parábola, la mediatriz $\nu_{\mathbf{y}}$ es ahora el candidato a ser la tangente a \mathcal{H} en el punto \mathbf{x} . Esto se demuestra viendo que $|d(\mathbf{x}', \mathbf{p}) - d(\mathbf{x}', \mathbf{q})| < 2a$ para cualquier otro punto $\mathbf{x}' \in \nu_{\mathbf{y}}$ y se deja como ejercicio (hay que usar dos veces la desigualdad del triángulo). De aquí también se sigue la propiedad focal de la hipérbola.



mediatrices con \mathbf{q} , ν_0 y ν_1 , son las *asíntotas* de \mathcal{H} .

Pero además, se sigue otro punto clave. La existencia de las asíntotas. Hay un par de puntos en el círculo \mathcal{C} , para los cuales la mediatriz con \mathbf{q} es paralela a su recta por \mathbf{p} . Para que esto suceda para un punto $\mathbf{y} \in \mathcal{C}$, como la mediatriz con \mathbf{q} , $\nu_{\mathbf{y}}$, es ortogonal a la recta $\mathbf{q}\mathbf{y}$, entonces es necesario (y suficiente) que el radio $\mathbf{p}\mathbf{y}$ sea también ortogonal a $\mathbf{q}\mathbf{y}$; es decir, que $\mathbf{q}\mathbf{y}$ sea la tangente al círculo \mathcal{C} en el punto \mathbf{y} . Entonces las dos tangentes al círculo \mathcal{C} que se pueden trazar desde \mathbf{q} , definen dos puntos, \mathbf{y}_0 y \mathbf{y}_1 digamos, cuyas



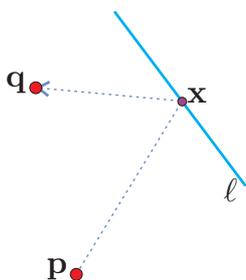
Para fijar ideas consideremos a los focos \mathbf{p} y \mathbf{q} en una línea horizontal y con \mathbf{p} a la izquierda, como en las figuras. Si movemos el punto $\mathbf{y} \in \mathcal{C}$, empezando con el más cercano a \mathbf{q} , hacia arriba (contra el reloj), el correspondiente punto $\mathbf{x} \in \mathcal{H}$ viaja por la rama de \mathbf{q} alejándose arriba a la derecha. Su tangente (que empieza vertical) es mediatriz de un segmento que gira en \mathbf{q} siguiendo a su otro extremo \mathbf{y} ; entonces la tangente gira hacia la derecha, como el reloj.

Cuando \mathbf{y} llega a \mathbf{y}_0 , el punto \mathbf{x} (que viaja en \mathcal{H}) se ha “ido al infinito” pero su tangente llega plácidamente a ν_0 , la asíntota correspondiente. Al mover otro poco a \mathbf{y} , la tangente $\nu_{\mathbf{y}}$ empieza a girar contra el reloj (pues el segmento \mathbf{qy} empieza a bajar) y el correspondiente punto $\mathbf{x} \in \mathcal{H}$ reaparece pero en la otra rama y justo por el lado opuesto. Podemos seguir girando a \mathbf{y} sin problemas, mientras \mathbf{x} recorre toda la rama de \mathbf{p} , hasta llegar al otro punto de tangencia \mathbf{y}_1 (ahora abajo) y tenemos que \mathbf{x} desaparece (en el infinito arriba), justo cuando la tangente está en la otra asíntota. Al seguir girando, \mathbf{x} vuelve a cambiar de rama y la tangente de dirección de giro para regresar, junto con \mathbf{y} , a su lugar de origen.

Además, obtenemos la interpretación geométrica de los tres parámetros \mathbf{a} , \mathbf{b} , \mathbf{c} en el triángulo rectángulo que se forma con el centro, el eje que une los dos focos y una de las asíntotas.

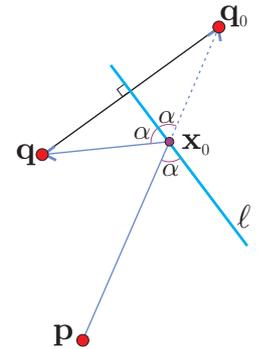
2.5.3 De la elipse

La propiedad focal de la elipse es que cualquier fotón que emana de un foco se refleja en la elipse para llegar al otro foco. De tal manera que si susurramos en uno de los focos de un cuarto elíptico, el sonido rebota en la pared y confluye estruendosamente en el otro foco (pues, además, cada onda sonora que emana de un foco llega simultáneamente al otro foco, sin distorsión o *delay*, pues en todas sus trayectorias ha recorrido la misma distancia). Una manera de ver esto es análoga a la que usamos para la hipérbola: tomar el círculo de radio $2\mathbf{a}$ centrado en el foco \mathbf{p} —este círculo contiene el otro foco \mathbf{q} , puesto que ahora $2\mathbf{a} > \mathbf{d}(\mathbf{p}, \mathbf{q})$ —, y luego ver que para los puntos de este círculo, su mediatriz con \mathbf{q} es tangente a la elipse \mathcal{E} . Pero para variarlo, se la dejamos al estudiante y usamos otro método basado en el clásico “problema del bombero”.

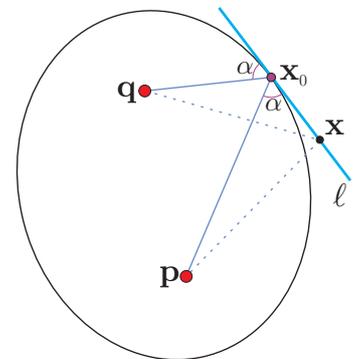


Supongamos que un bombero está parado en el punto \mathbf{p} y hay un fuego —que tiene que apagar— en el punto \mathbf{q} . Pero tiene su cubeta vacía, y entonces tiene que pasar primero a llenarla a un río cuyo borde es la recta ℓ . El problema es ¿cuál es la trayectoria óptima que debe seguir el bombero? Es decir, para cuál punto $\mathbf{x} \in \ell$ se tiene que $\mathbf{d}(\mathbf{p}, \mathbf{x}) + \mathbf{d}(\mathbf{x}, \mathbf{q})$ es mínima. En la vida real, y con muchos bomberos en fila, se aproximarían a la solución poco a poco. Pero desde nuestra cómoda banca de la abstracción, hay una solución muy elegante.

No hemos especificado de qué lado del río está el fuego. Si estuviera del otro lado que el bombero, cualquier trayectoria al fuego tiene que pasar por ℓ y entonces debe irse por la línea recta de \mathbf{p} a \mathbf{q} y tomar agua en $\mathbf{x}_0 = \ell \cap \overline{\mathbf{p}\mathbf{q}}$. Entonces, si fuego y bombero están del mismo lado del río ℓ , podemos pensar en un “fuego virtual”, que es el reflejado de \mathbf{q} en ℓ , llamémosle \mathbf{q}_0 , que cumple que $d(\mathbf{x}, \mathbf{q}) = d(\mathbf{x}, \mathbf{q}_0)$ para todo $\mathbf{x} \in \ell$, (para \mathbf{q} y \mathbf{q}_0 , ℓ es su mediatriz). La solución es, por el caso anterior, $\mathbf{x}_0 = \ell \cap \overline{\mathbf{p}\mathbf{q}_0}$. Pero obsérvese además que el ángulo α con el que llega el bombero a ℓ es igual al ángulo “de reflexión” con el que sale corriendo al fuego (ya con la cubeta llena), e igual al ángulo con el que seguiría su trayecto al fuego virtual, y que esta propiedad determina al punto de mínimo recorrido \mathbf{x}_0 ; es fácil convencerse de que para cualquier otro punto de ℓ los ángulos de llegada y de salida son distintos. Si los bomberos fueran fotones que emanan de \mathbf{p} y ℓ es un espejo, el único que llega a \mathbf{q} es el fotón que sigue el recorrido mínimo.

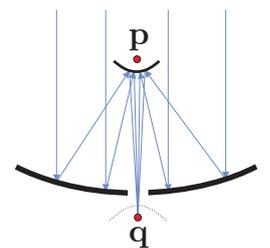


Pensemos ahora que \mathbf{p} y \mathbf{q} son los focos de una elipse y \mathbf{x}_0 un punto en ella. Sea ℓ la recta que pasa por \mathbf{x}_0 y bisecta (por fuera) los segmentos de \mathbf{p} y \mathbf{q} a \mathbf{x}_0 . Por construcción y la solución al problema del bombero, cualquier otro punto $\mathbf{x} \in \ell$ tiene mayor suma de distancias a los focos y por tanto está fuera de la elipse. Esto demuestra que ℓ es la tangente a la elipse en el punto \mathbf{x}_0 y, por tanto, su propiedad focal con la que empezamos esta sección.



2.5.4 Telescopios

Por último, veamos brevemente cómo funciona un telescopio moderno. El haz de fotones que viene de una galaxia, digamos, es “paralelo” pues está tan lejos que desde la Tierra es imposible medir variación de ángulos (mucho más que los satélites respecto a la superficie de las antenas parabólicas). Un gran espejo parabólico hace confluir este haz en su foco \mathbf{p} . Pero \mathbf{p} está del otro lado del espejo. Esto se resuelve poniendo un pequeño espejo secundario elíptico y con el mismo foco \mathbf{p} , que vuelve a reflejar el haz. Éste pasa por un pequeño agujero en el centro del primario y vuelve a confluir en su otro foco \mathbf{q} , que ahora sí está detrás del gran espejo primario, y ahí se captura en el detector. El hoyo que se hace al primario, así como lo que tapa el secundario o la estructura, cuenta poco pues la superficie crece como el radio al cuadrado.



Otra manera de resolver el problema de “regresar el haz” hacia atrás del espejo primario, que es la más usada, es poniendo un secundario hiperbólico en vez de uno elíptico, con la superficie reflejante por fuera y otra vez con el mismo foco \mathbf{p} . De nuevo hace reflejar el haz en uno que confluye en su otro foco \mathbf{q} , cómodamente situado detrás

del hoyito del primario.

Lo que cuenta es qué tantos fotones podemos capturar en el sistema óptico; a simple vista es la superficie de la pupila, los telescopios modernos recolectan los fotones que chocarían en decenas de metros cuadrados en una dirección dada y los hacen pasar, después de dos reflexiones adecuadas, por la pupila de algún instrumento detector.

2.6 *Armonía y excentricidad

Hemos definido elipses e hipérbolas como lugares geométricos con base en las propiedades aditivas de distancias. Ahora estudiaremos las cónicas respecto a propiedades multiplicativas. Pues resulta que las tres cónicas (elipses, parábolas e hipérbolas) se definen también como el lugar geométrico de los puntos \mathbf{x} cuya razón de sus distancias a un foco \mathbf{p} y a una recta ℓ , llamada *directriz*, es una constante fija e , llamada su *excentricidad*. Es decir, están dadas por la ecuación

$$d(\mathbf{x}, \mathbf{p}) = e d(\mathbf{x}, \ell).$$

Para $e = 1$, ésta es la ecuación de la parábola que ya estudiamos. Veremos que para $e < 1$ se define una elipse y para $e > 1$ una hipérbola, de tal manera que las tres cónicas quedan en una sola familia. Pero notemos que no apareció el círculo. Éste tiene excentricidad 0 que no cabe en la definición anterior (más que como límite). Sin embargo, aparece en una instancia (en apariencia) más simple de esta ecuación que consiste en tomar otro punto en vez de la línea, y veremos que ligado a esto surge naturalmente el concepto clásico de los puntos armónicos.

2.6.1 *Puntos armónicos y círculos de Apolonio

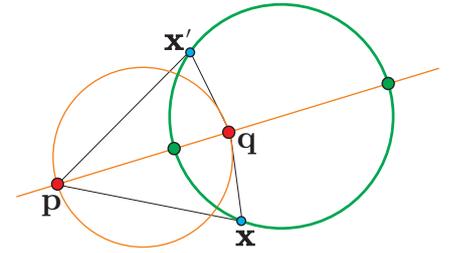
Sean \mathbf{p} y \mathbf{q} dos puntos distintos en el plano. Consideremos la ecuación

$$d(\mathbf{x}, \mathbf{p}) = h d(\mathbf{x}, \mathbf{q}), \tag{2.18}$$

donde $h > 0$. Para $h = 1$ ésta es la ecuación de la mediatriz que estudiamos en el capítulo anterior, pero a excepción de $h = 1$ donde da una recta, veremos que esta ecuación define círculos: los llamados “*círculos de Apolonio*”.²

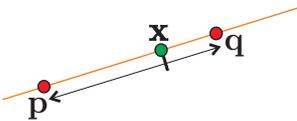
²También a un círculo tangente a tres dados se le llama “*círculo de Apolonio*”.

Notemos primero que el lugar geométrico \mathcal{C} definido por (2.18) es simétrico respecto a la línea ℓ que pasa por \mathbf{p} y \mathbf{q} , pues si un punto \mathbf{x} satisface la ecuación también lo hace su reflejado \mathbf{x}' respecto a la recta ℓ . De tal manera que la recta ℓ debe ser un diámetro y el círculo debe estar determinado por los puntos de intersección $\ell \cap \mathcal{C}$. Encontrémoslos.



Sabemos que los puntos de ℓ se expresan en coordenadas baricéntricas como $\mathbf{x} = \alpha\mathbf{p} + \beta\mathbf{q}$ con $\alpha + \beta = 1$, donde, si damos una dirección a ℓ , de \mathbf{p} hacia \mathbf{q} digamos, y denotamos a las distancias dirigidas con \vec{d} , se tiene que

$$\alpha = \frac{\vec{d}(\mathbf{x}, \mathbf{q})}{\vec{d}(\mathbf{p}, \mathbf{q})} \quad \text{y} \quad \beta = \frac{\vec{d}(\mathbf{x}, \mathbf{p})}{\vec{d}(\mathbf{q}, \mathbf{p})} = \frac{\vec{d}(\mathbf{p}, \mathbf{x})}{\vec{d}(\mathbf{p}, \mathbf{q})}.$$



De tal manera que si $\mathbf{x} = \alpha\mathbf{p} + \beta\mathbf{q} \in \ell$, entonces

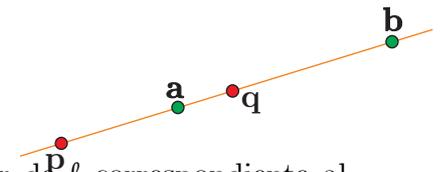
$$\frac{\beta}{\alpha} = \frac{\vec{d}(\mathbf{p}, \mathbf{x})}{\vec{d}(\mathbf{x}, \mathbf{q})}.$$

Por tanto, para encontrar las soluciones de (2.18) en ℓ basta encontrar las soluciones de $|\beta/\alpha| = h$ con $\alpha + \beta = 1$: al sustituir $\beta = h\alpha$ en esta última ecuación, encontramos una solución dentro del segmento $\overline{\mathbf{p}\mathbf{q}}$ (con α y β positivas) dada por

$$\alpha_h := \frac{1}{h+1} \quad \text{y} \quad \beta_h := \frac{h}{h+1}.$$

Llamemos \mathbf{a} al punto correspondiente, es decir, sea $\mathbf{a} = \alpha_h\mathbf{p} + \beta_h\mathbf{q}$. Obsérvese que \mathbf{a} es el punto medio cuando $h = 1$; que \mathbf{a} está más cerca de \mathbf{p} para $h < 1$ y que se aproxima a \mathbf{q} conforme h crece. Para $h \neq 1$, también tenemos otra solución fuera del segmento (que viene de tomar $\beta = -h\alpha$), a saber:

$$\alpha'_h = \frac{1}{1-h} = \frac{\alpha_h}{\alpha_h - \beta_h} \quad \text{y} \quad \beta'_h = \frac{h}{h-1} = \frac{\beta_h}{\beta_h - \alpha_h}.$$

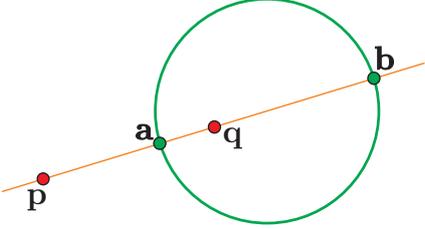


Llamemos \mathbf{b} a esta otra solución de (2.18) en ℓ (es decir, $\mathbf{b} = \alpha'_h\mathbf{p} + \beta'_h\mathbf{q}$), y nótese que \mathbf{b} está en el rayo exterior de ℓ correspondiente al punto, \mathbf{p} o \mathbf{q} , más cercano a \mathbf{a} ($h < 1$ o $h > 1$, respectivamente). (Y nótese también que la expresión de α_h y β_h en términos de α'_h y β'_h es exactamente la misma.)

Clásicamente, se dice que la pareja \mathbf{a}, \mathbf{b} es *armónica* respecto a la pareja \mathbf{p}, \mathbf{q} (o que \mathbf{a} y \mathbf{b} son *conjugados armónicos* respecto a \mathbf{p} y \mathbf{q}) pues (como reza la definición general) los cuatro puntos están alineados y cumplen la relación

$$\frac{d(\mathbf{p}, \mathbf{a})}{d(\mathbf{a}, \mathbf{q})} = \frac{d(\mathbf{p}, \mathbf{b})}{d(\mathbf{b}, \mathbf{q})}, \tag{2.19}$$

que, en nuestro caso, es la “razón de armonía” h . Se tiene claramente que la relación de armonía es simétrica (pues \mathbf{p}, \mathbf{q} son armónicos respecto a \mathbf{a}, \mathbf{b}). Y para aclarar otro poco su relación geométrica, hay que observar que si en la recta real consideramos a \mathbf{p} y \mathbf{q} como los puntos -1 y 1 , entonces el conjugado armónico de x es x^{-1} , y el 0 queda como conjugado armónico del infinito (así como el punto medio de \mathbf{p} y \mathbf{q} , $h = 1$, fue un caso especial).



Ahora sí, demostremos que el círculo cuyo diámetro es el segmento $\overline{\mathbf{a}\mathbf{b}}$ está dado por la ecuación (2.18), que es, recuérdese, $d(\mathbf{x}, \mathbf{p}) = h d(\mathbf{x}, \mathbf{q})$. En el Ejercicio 2.7 se debió demostrar que el círculo con diámetro $\overline{\mathbf{a}\mathbf{b}}$ está dado por la ecuación $(\mathbf{x} - \mathbf{a}) \cdot (\mathbf{x} - \mathbf{b}) = 0$ (si no se hizo, compruébese ahora). Desarrollemos entonces esta expresión, basándonos en que los coeficientes son coordenadas baricéntricas (*i.e.*, que suman 1):

$$\begin{aligned}
 (\mathbf{x} - \mathbf{a}) \cdot (\mathbf{x} - \mathbf{b}) &= (\mathbf{x} - \alpha_h \mathbf{p} - \beta_h \mathbf{q}) \cdot (\mathbf{x} - \alpha'_h \mathbf{p} - \beta'_h \mathbf{q}) \\
 &= (\alpha_h (\mathbf{x} - \mathbf{p}) + \beta_h (\mathbf{x} - \mathbf{q})) \cdot (\alpha'_h (\mathbf{x} - \mathbf{p}) + \beta'_h (\mathbf{x} - \mathbf{q})) \\
 &= \alpha_h \alpha'_h (\mathbf{x} - \mathbf{p}) \cdot (\mathbf{x} - \mathbf{p}) + \beta_h \beta'_h (\mathbf{x} - \mathbf{q}) \cdot (\mathbf{x} - \mathbf{q}) \\
 &\quad + (\alpha_h \beta'_h + \beta_h \alpha'_h) (\mathbf{x} - \mathbf{p}) \cdot (\mathbf{x} - \mathbf{q}) \tag{2.20} \\
 &= \left(\frac{1}{1 - h^2} \right) ((\mathbf{x} - \mathbf{p}) \cdot (\mathbf{x} - \mathbf{p}) - h^2 (\mathbf{x} - \mathbf{q}) \cdot (\mathbf{x} - \mathbf{q})),
 \end{aligned}$$

donde hemos usado las definiciones —en términos de $h \neq 1$ — de nuestros coeficientes (véase el Ejercicio 2.21). Se tiene entonces que $(\mathbf{x} - \mathbf{a}) \cdot (\mathbf{x} - \mathbf{b}) = 0$ si y sólo si $(\mathbf{x} - \mathbf{p}) \cdot (\mathbf{x} - \mathbf{p}) = h^2 (\mathbf{x} - \mathbf{q}) \cdot (\mathbf{x} - \mathbf{q})$, que es justo el cuadrado de la ecuación (2.18).

Además, tenemos al menos otras tres maneras interesantes de detectar armonía.

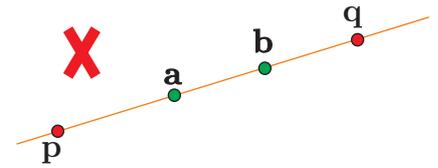
Teorema 2.1 Sean $\mathbf{p}, \mathbf{q}, \mathbf{a}, \mathbf{b}$ cuatro puntos colineales, y sean \mathcal{C}_1 y \mathcal{C}_2 los círculos con diámetro $\overline{\mathbf{p}\mathbf{q}}$ y $\overline{\mathbf{a}\mathbf{b}}$ respectivamente. Entonces son equivalentes:

- i) Las parejas \mathbf{p}, \mathbf{q} y \mathbf{a}, \mathbf{b} son armónicas.
- ii) Se cumple que $(\mathbf{a} - \mathbf{p}) \cdot (\mathbf{b} - \mathbf{q}) = -(\mathbf{a} - \mathbf{q}) \cdot (\mathbf{b} - \mathbf{p})$.
- iii) El punto \mathbf{b} está en la línea polar de \mathbf{a} respecto a \mathcal{C}_1 .
- iv) Los círculos \mathcal{C}_1 y \mathcal{C}_2 se intersectan ortogonalmente.

Demostración. (i) \Leftrightarrow (ii). La condición de armonía (2.19) claramente se puede reescribir, usando normas para expresar distancias, como

$$|\mathbf{a} - \mathbf{p}| |\mathbf{b} - \mathbf{q}| = |\mathbf{a} - \mathbf{q}| |\mathbf{b} - \mathbf{p}|. \tag{i}$$

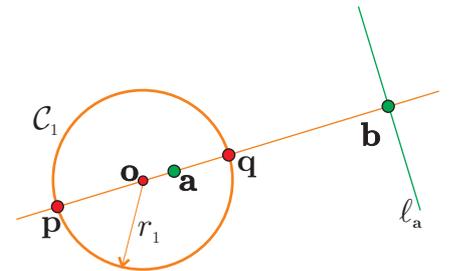
La equivalencia de (i) y (ii) se sigue del hecho de que si dos vectores \mathbf{u} y \mathbf{v} son paralelos entonces $|\mathbf{u}||\mathbf{v}| = \pm (\mathbf{u} \cdot \mathbf{v})$, dependiendo de si apuntan en la misma dirección (+) o en direcciones opuestas (-). Esto da inmediatamente la implicación (ii) \Rightarrow (i), pues por la hipótesis general del teorema los cuatro puntos están alineados y entonces sus diferencias son paralelas. Para el regreso sólo hay que tener cuidado con los signos. Si se cumple (i) es fácil ver que el orden lineal en que aparecen los puntos en la línea es alternado (entre las parejas), pues si, por ejemplo, \mathbf{p} y \mathbf{q} fueran los extremos, un lado de (i) es estrictamente menor que el otro. Entonces —renombrando entre las parejas si es necesario— podemos suponer que aparecen en el orden \mathbf{p} , \mathbf{a} , \mathbf{q} , \mathbf{b} . En este caso (i) toma la expresión (ii).



(i) \Rightarrow (iii). Supongamos que \mathbf{p} , \mathbf{q} y \mathbf{a} , \mathbf{b} son los de los párrafos anteriores con razón de armonía $h \neq 1$. Sean \mathbf{o} (ojo, no necesariamente es el origen $\mathbf{0}$) el centro y r_1 el radio del círculo \mathcal{C}_1 , es decir, $\mathbf{o} = (1/2)(\mathbf{p} + \mathbf{q})$ y $r_1 = |\mathbf{p} - \mathbf{q}|/2$. De tal manera que $\mathcal{C}_1 : (\mathbf{x} - \mathbf{o}) \cdot (\mathbf{x} - \mathbf{o}) = r_1^2$.

Por la sección 1.1, ver que \mathbf{b} está en la línea polar de \mathbf{a} respecto a \mathcal{C}_1 equivale a demostrar que $(\mathbf{a} - \mathbf{o}) \cdot (\mathbf{b} - \mathbf{o}) = r_1^2$. Pero ya evaluamos $(\mathbf{x} - \mathbf{a}) \cdot (\mathbf{x} - \mathbf{b}) = (\mathbf{a} - \mathbf{x}) \cdot (\mathbf{b} - \mathbf{x})$ para cualquier \mathbf{x} ; entonces hay que sustituir \mathbf{o} en (2.20), para obtener

$$\begin{aligned} (\mathbf{a} - \mathbf{o}) \cdot (\mathbf{b} - \mathbf{o}) &= \left(\frac{1}{1 - h^2} \right) ((\mathbf{o} - \mathbf{p}) \cdot (\mathbf{o} - \mathbf{p}) \\ &\quad - h^2 (\mathbf{o} - \mathbf{q}) \cdot (\mathbf{o} - \mathbf{q})) \\ &= \left(\frac{1}{1 - h^2} \right) (r_1^2 - h^2 r_1^2) = r_1^2. \end{aligned}$$



Por el papel que desempeñó (y desempeñará) el punto \mathbf{o} , conviene pensar que sí es el origen. Si no lo fuera, lo trasladamos al origen junto con el resto de los puntos y las propiedades que nos interesan se preservan. Puede pensarse también que en lo que sigue cada letra (\mathbf{b} digamos) representa al punto original menos \mathbf{o} (a $\mathbf{b} - \mathbf{o}$ en el ejemplo). Tenemos entonces que $\mathbf{p} = -\mathbf{q}$ (pues el origen es ahora su punto medio) y que $r_1^2 = \mathbf{p} \cdot \mathbf{p}$. Ahora la condición (iii) equivale a

$$\mathbf{a} \cdot \mathbf{b} = r_1^2, \tag{iii}$$

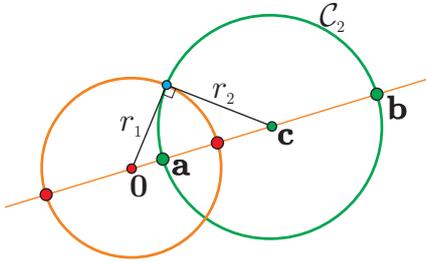
y podemos demostrar la equivalencia de (ii) y (iii) de un solo golpe (aunque un lado ya esté demostrado).

(ii) \Leftrightarrow (iii). Usando que $\mathbf{q} = -\mathbf{p}$, claramente se tiene

$$(\mathbf{a} - \mathbf{p}) \cdot (\mathbf{b} - \mathbf{q}) + (\mathbf{a} - \mathbf{q}) \cdot (\mathbf{b} - \mathbf{p}) = 2(\mathbf{a} \cdot \mathbf{b} - \mathbf{p} \cdot \mathbf{p}),$$

que concluye la demostración ((ii) es que el lado izquierdo sea cero y (iii) que el lado derecho lo es).

(iii) ⇔ (iv). Sean $\mathbf{c} = (1/2)(\mathbf{a} + \mathbf{b})$ el centro de \mathcal{C}_2 , y $r_2 = |\mathbf{b} - \mathbf{a}|/2$ su radio. Que \mathcal{C}_1 y \mathcal{C}_2 sean *ortogonales* quiere decir que sus tangentes en el punto de intersección lo son; y esto equivale a que sus radios al punto de intersección son ortogonales (pues las tangentes son ortogonales a los radios). Por el Teorema de Pitágoras esto equivale a que



$$r_1^2 + r_2^2 = |\mathbf{c}|^2 \tag{iv}$$

(pues \mathcal{C}_1 está centrado en el origen, así que $|\mathbf{c}|$ es la distancia entre los centros), que es lo que tomamos en adelante como (iv). Sabemos, por hipótesis, que \mathbf{a} y \mathbf{b} son paralelos y hay que observar que ambas condiciones ($\mathbf{a} \cdot \mathbf{b} = r_1^2 > 0$ y $r_1^2 + r_2^2 = |\mathbf{c}|^2$) implican que están del mismo lado del origen. Podemos suponer entonces que \mathbf{a} es “más pequeño” que \mathbf{b} , y entonces

$$\begin{aligned} |\mathbf{c}|^2 &= (|\mathbf{a}| + r_2)^2 \\ &= |\mathbf{a}|(|\mathbf{a}| + 2r_2) + r_2^2 \\ &= |\mathbf{a}||\mathbf{b}| + r_2^2 \\ &= \mathbf{a} \cdot \mathbf{b} + r_2^2. \end{aligned}$$

Lo cual demuestra la equivalencia (iii) ⇔ (iv) y concluye el Teorema. □

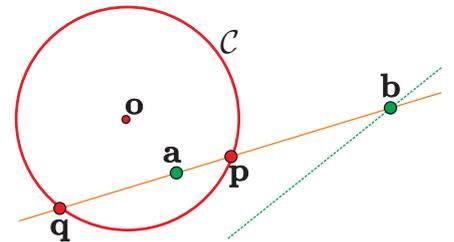
En resumen, los círculos de Apolonio de un par de puntos \mathbf{p} y \mathbf{q} son los ortogonales al círculo que los tiene como diámetro y cuyo centro está en esa línea, y además la cortan en sus conjugados armónicos; teniendo como límites los “círculos de radio cero” en \mathbf{p} y en \mathbf{q} y su línea mediatriz.

EJERCICIO 2.21 Demuestra que $\alpha_h \beta'_h = -\beta_h \alpha'_h$ (usando la definición de los correspondientes primados sin referencia a \mathbf{h}); y usando las definiciones con base en \mathbf{h} , demuestra que

$$\alpha_h \alpha'_h = \frac{1}{1 - h^2} \quad \text{y} \quad \beta_h \beta'_h = \frac{h^2}{h^2 - 1}$$

son las coordenadas baricéntricas del centro del círculo (2.18) respecto a \mathbf{p} y \mathbf{q} .

EJERCICIO 2.22 Sean \mathbf{p} y \mathbf{q} dos puntos en un círculo \mathcal{C} con centro \mathbf{o} ; y sea \mathbf{a} un punto en el segmento $\overline{\mathbf{p}\mathbf{q}}$ distinto del punto medio, es decir, $\mathbf{a} = \alpha\mathbf{p} + \beta\mathbf{q}$ con $\alpha + \beta = 1$, $0 < \alpha < 1$ y $\alpha \neq \beta$. Demuestra que el punto armónico de \mathbf{a} respecto a \mathbf{p} y \mathbf{q} (llamémoslo \mathbf{b}) está en la línea polar de \mathbf{a} respecto al círculo \mathcal{C} . (Expresa \mathbf{b} en coordenadas baricéntricas respecto a \mathbf{p} y \mathbf{q} , y sustituye en la ecuación de la polar de \mathbf{a}). Concluye que la línea polar de un punto \mathbf{a} en el interior de un círculo consiste de todos sus armónicos respecto a la intersección con el círculo de las líneas que pasan por el.



2.6.2 *Excentricidad

Demostraremos ahora que elipses e hipérbolas se definen por la ecuación

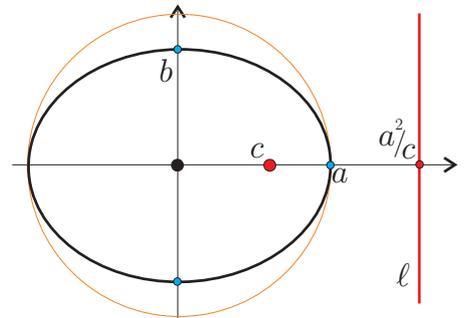
$$d(\mathbf{x}, \mathbf{p}) = e d(\mathbf{x}, \ell), \quad (2.21)$$

con $e < 1$ y $e > 1$ respectivamente, donde \mathbf{p} es uno de sus focos, y a ℓ la llamaremos su *directriz* correspondiente. La constante e se llama la *excentricidad* de la cónica, y ya sabemos que las de excentricidad 1 son las parábolas.

Consideremos la elipse canónica \mathcal{E} con centro en el origen y semiejes $a > b$ en los ejes coordenados. Sabemos que sus focos están en $\mathbf{p} = (c, 0)$ y $\mathbf{q} = (-c, 0)$, donde $c > 0$ es tal que $b^2 + c^2 = a^2$. Sabemos además que \mathcal{E} intersecta al eje x , que es su eje focal, en los puntos $(a, 0)$ y $(-a, 0)$.

Resulta que la directriz de \mathcal{E} correspondiente al foco \mathbf{p} es ortogonal al eje x y lo intersecta en el conjugado armónico de \mathbf{p} respecto a los puntos de intersección; es decir, es la línea polar a \mathbf{p} respecto al círculo $\mathbf{x} \cdot \mathbf{x} = a^2$ (que, notémoslo, es doblemente tangente a \mathcal{E}). Sea entonces ℓ definida por la ecuación

$$x = \frac{a^2}{c}.$$



De los datos que tenemos podemos deducir su excentricidad (pues del punto $(a, 0) \in \mathcal{E}$ conocemos sus distancias a \mathbf{p} y ℓ), que resulta $e = c/a < 1$. Desarrollemos entonces a la ecuación (2.21) en coordenadas para ver si son ciertas las afirmaciones que nos hemos sacado de la manga de la historia.

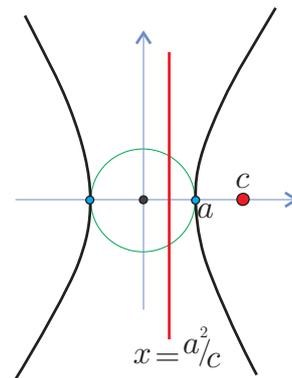
Puesto que ambos lados de (2.21) son positivos, esta ecuación es equivalente a su cuadrado que en coordenadas (y nuestra definición de \mathbf{p} , ℓ y e) toma la forma

$$\begin{aligned} (x - c)^2 + y^2 &= \left(\frac{c}{a}\right)^2 \left(x - \frac{a^2}{c}\right)^2 \\ x^2 - 2cx + c^2 + y^2 &= \left(\frac{c^2}{a^2}\right) \left(x^2 - 2\frac{a^2}{c}x + \frac{a^4}{c^2}\right) \\ \left(\frac{a^2 - c^2}{a^2}\right) x^2 + y^2 &= a^2 - c^2. \end{aligned}$$

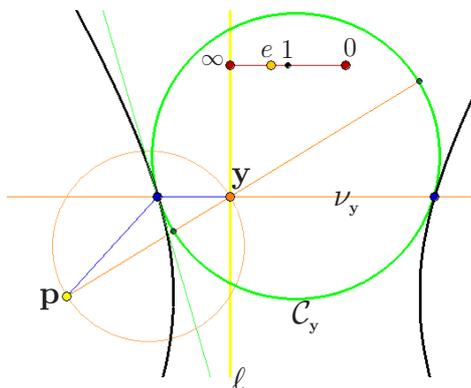
De aquí, al dividir entre $b^2 = a^2 - c^2$ se obtiene la ecuación canónica de la elipse \mathcal{E} y quedan demostradas todas nuestras aseveraciones.

Para una hipérbola \mathcal{H} , de nuevo la directriz correspondiente a un foco es perpendicular al eje focal y lo intersecta en su conjugado armónico respecto a los puntos de intersección con \mathcal{H} . De tal manera que podemos tomar exactamente los mismos datos del caso anterior ($\mathbf{p} = (c, 0)$, $\ell : cx = a^2$ y $e = c/a$), y la ecuación (2.21) se desarrolla igual. Pero ahora tenemos que $c > a$ (la excentricidad es mayor que 1, $e = c/a > 1$) y entonces definimos $b > 0$ tal que $a^2 + b^2 = c^2$; de tal manera que en el último paso hay que dividir entre $-b^2$ para obtener

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



La excentricidad de una cónica es un invariante de semejanza, es decir, si magnificamos (hacemos *zoom*) o disminuimos una cónica, su excentricidad se mantiene. Pues esta magnificación (o contracción) se logra alejando (o acercando) al foco \mathbf{p} y a la directriz ℓ . Pero si fijamos a estos últimos y variamos la excentricidad e , se cubre por cónicas todo el plano (exceptuando a \mathbf{p} y ℓ que aparecen como los límites $e \rightarrow 0$ y $e \rightarrow \infty$ respectivamente) pues cualquier punto da una razón específica entre las distancias y por tanto una excentricidad que lo incluye.



Recreemos la construcción de la parábola a partir de las mediatrices, pero ahora con círculos de Apolonio. Si \mathcal{P} es la cónica con foco \mathbf{p} , directriz ℓ y excentricidad e (puede ser elipse, parábola o hipérbola), y para cada punto \mathbf{y} en ℓ tomamos el círculo de Apolonio \mathcal{C}_y entre \mathbf{p} y \mathbf{y} con razón de armonía e , entonces está claro que $\mathcal{C}_y \cap \nu_y \subset \mathcal{P}$ (donde ν_y es la normal a ℓ en \mathbf{y}). Pues la distancia de un punto en ν_y a ℓ corresponde a su distancia a \mathbf{y} ; y si está en \mathcal{C}_y la razón de las distancias a \mathbf{p} y \mathbf{y} es la adecuada. Pero ahora $\mathcal{C}_y \cap \nu_y$ no es necesariamente un punto pues, salvo el

caso $e = 1$ donde \mathcal{C}_y es la mediatriz, un círculo intersecta a una recta en 2, 1 o 0 puntos.

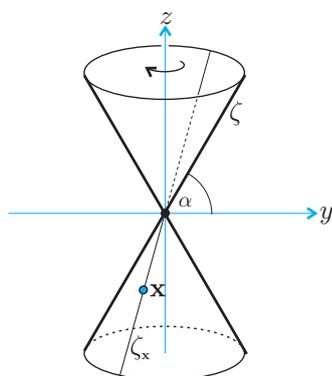
Para el caso $e < 1$, los círculos de Apolonio \mathcal{C}_y contienen a \mathbf{p} en su interior; crecen conforme \mathbf{y} se aleja en ℓ , pero llega un momento (que corresponde al semieje menor de la elipse) en el que ν_y ya no los corta. Por el contrario, para $e > 1$ los círculos de Apolonio contienen a \mathbf{y} , y entonces ν_y siempre los corta en dos puntos que dibujan las dos ramas de la hipérbola.

En ambos casos, si \mathcal{C}_y corta a ν_y en \mathbf{x} , digamos, entonces por un argumento análogo al de la parábola se puede ver que está fuera de \mathcal{P} , es decir, que \mathcal{C}_y es *tangente* a \mathcal{P} en \mathbf{x} , pero en este caso, y salvo por el último toque a una elipse en el semieje menor, también es tangente a \mathcal{P} en otro punto. Para la hipérbola es fácil

verlos, simplemente dejando caer un círculo hacia el centro y si es más grande que la cinturita: ahí donde se detiene es un \mathcal{C}_y , con \mathbf{y} correspondiendo a la altura de los dos puntos de \mathcal{P} donde se atoró.

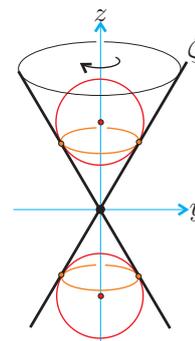
2.7 *Esferas de Dandelin

Hemos desarrollado todo este capítulo a partir de las definiciones planas de las cónicas por propiedades de distancias. Sin embargo, en su introducción empezamos como los griegos: definiéndolas en el espacio como intersección de un cono con planos. Veremos brevemente que tal definición corresponde a la que usamos, basándonos en una demostración que data de cerca del año de 1800 debida a un tal Dandelin.



Para definir un cono \mathcal{C} en \mathbb{R}^3 , consideremos una línea ζ por el origen en el plano yz con un ángulo α respecto al eje y : y hagámosla girar alrededor del eje z . De tal manera que la intersección de \mathcal{C} con un plano horizontal ($z = c$, digamos) es un círculo centrado en “el origen” $((0, 0, c)$, y de radio $c \cos \alpha / \sin \alpha$). Por la definición, es claro que para cualquier punto $\mathbf{x} \in \mathcal{C}$, su recta por el origen, que denotaremos con ζ_x , está contenida en el cono, y la llamaremos *la recta del cono por \mathbf{x}* . Queremos hacer ver que cualquier plano Π que no contenga al origen intersecta el cono \mathcal{C} en una cónica (definida por distancias).

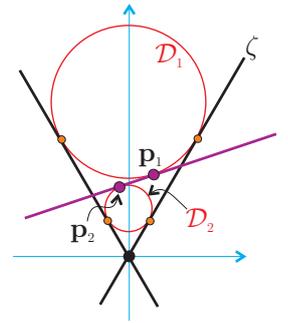
Las *esferas tangentes* al cono son esferas con centro en el eje z y que tocan al cono pero que no lo cortan, es decir, que tienen a todas las líneas del cono como líneas tangentes. Si dejamos caer una esfera dentro del cono se asienta en una esfera tangente. Es claro que las esferas tangentes se obtienen al hacer girar en el eje z un círculo centrado en él y tangente a la línea generadora ζ ; que para cada radio hay exactamente dos esferas tangentes (una arriba y una abajo), y que para cada punto del eje z hay una única esfera tangente centrada en él. Además, cada esfera tangente al cono lo intersecta en un círculo horizontal llamado su *círculo de tangencia* o de *contacto*.



Tomemos ahora un plano Π casi horizontal. Este plano Π tiene exactamente dos esferas tangentes que también son tangentes al cono \mathcal{C} , que llamaremos sus *esferas de Dandelin*, \mathcal{D}_1 y \mathcal{D}_2 . Veamos y, para fijar ideas, pensemos que el plano Π intersecta el cono arriba del origen, pero no muy lejos. Si pensamos en una esfera muy grande tangente al cono, ésta está muy arriba; al desinflarla (manteniéndola siempre tangente al cono) va bajando. Hay un primer momento en el que toca a Π , ésa es \mathcal{D}_1 . Al seguir desinflándola, intersecta a Π en círculos, y hay un último momento en que lo toca —esa es \mathcal{D}_2 — para seguir decreciendo, insignificante, hacia el origen.

Más formalmente, y sin perder generalidad, podemos suponer que el plano Π contiene la dirección del eje x ; entonces las esferas de Dandelin de Π corresponden a dos de los cuatro círculos tangentes a las tres rectas que se obtienen de intersectar el plano yz con el plano Π y el cono \mathcal{C} (precisamente, los dos que caen dentro del cono).

Sean \mathbf{p}_1 y \mathbf{p}_2 los puntos donde \mathcal{D}_1 y \mathcal{D}_2 tocan, respectivamente, el plano Π . Estos serán los focos de la elipse $\Pi \cap \mathcal{C}$. Para verlo necesitamos un lema que es más difícil enunciar que probar.

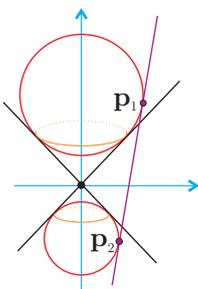
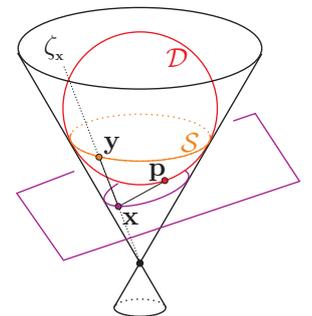


Lema 2.2 *Sea \mathcal{D} una esfera de Dandelin de un plano Π . Sea \mathbf{p} su punto de contacto con Π ; y sea \mathcal{S} el círculo de tangencia de \mathcal{D} en \mathcal{C} . Para cualquier $\mathbf{x} \in \Pi \cap \mathcal{C}$, sea \mathbf{y} la intersección de la recta del cono por \mathbf{x} y el círculo de tangencia, i.e. $\mathbf{y} = \mathcal{S} \cap \zeta_{\mathbf{x}}$; se tiene entonces que*

$$d(\mathbf{x}, \mathbf{p}) = d(\mathbf{x}, \mathbf{y}).$$

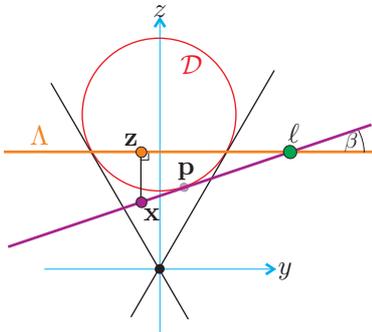
Basta observar que los segmentos $\overline{\mathbf{x}\mathbf{p}}$ y $\overline{\mathbf{x}\mathbf{y}}$ son ambos tangentes a la esfera \mathcal{D} y que, en general, dos segmentos que salen del mismo punto y llegan tangentes a una esfera miden lo mismo. \square

Ahora sí, podemos demostrar que $\Pi \cap \mathcal{C}$ es una elipse tal como la definimos. Sean \mathcal{S}_1 y \mathcal{S}_2 los círculos de tangencia de las esferas de Dandelin \mathcal{D}_1 y \mathcal{D}_2 , respectivamente. Dado $\mathbf{x} \in \Pi \cap \mathcal{C}$, según el lema, sus distancias a los respectivos focos se pueden medir en la línea del cono como las distancias de \mathbf{x} a los círculos de tangencia \mathcal{S}_1 y \mathcal{S}_2 . Como \mathbf{x} está entre estos dos círculos, la suma de las distancias es constante (a saber, es la distancia, medida en líneas del cono, entre \mathcal{S}_1 y \mathcal{S}_2).



Cuando el plano Π es cercano a un plano vertical (incluyendo este caso), se tiene que Π intersecta ambos lados del cono \mathcal{C} . Pero también tiene dos esferas de Dandelin tangentes a él y al cono. De nuevo, sus puntos de tangencia con el plano son los focos de la hipérbola $\Pi \cap \mathcal{C}$, y la demostración es análoga (usando el lema). Dejamos los detalles al lector (siguiente ejercicio), pues para el caso de la parábola necesitamos encontrar la directriz, y su análisis incluirá los tres casos.

Supongamos que el plano Π tiene un ángulo $\beta > 0$ respecto al plano horizontal xy , y que \mathcal{D} es su esfera de Dandelin, es decir, que es tangente a Π y al cono \mathcal{C} con punto y círculo de contacto \mathbf{p} y \mathcal{S} respectivamente. Sea Λ el plano horizontal que contiene al círculo de tangencia \mathcal{S} . Sea $\ell = \Pi \cap \Lambda$, línea que podemos pensar como paralela al eje x y que llamaremos la *directriz*.



Consideremos un punto $\mathbf{x} \in \Pi \cap \mathcal{C}$. De la proyección al plano yz , es fácil deducir que

$$d(\mathbf{x}, \ell) \operatorname{sen} \beta = d(\mathbf{x}, \mathbf{z}),$$

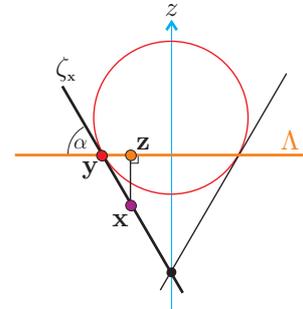
donde \mathbf{z} es la proyección ortogonal de \mathbf{x} al plano horizontal Λ . Tomemos ahora el plano vertical que pasa por \mathbf{x} y el eje

z . Claramente contiene a \mathbf{z} y a la línea del cono $\zeta_{\mathbf{x}}$ que pasa por \mathbf{x} ; y es fácil ver en él que

$$d(\mathbf{x}, \mathbf{z}) = d(\mathbf{x}, \mathbf{y}) \operatorname{sen} \alpha,$$

donde $\mathbf{y} = \zeta_{\mathbf{x}} \cap \Lambda$ (justo la \mathbf{y} que aparece en el lema). De estas dos ecuaciones y el lema se sigue que

$$d(\mathbf{x}, \mathbf{p}) = \frac{\operatorname{sen} \beta}{\operatorname{sen} \alpha} d(\mathbf{x}, \ell).$$



Esto demuestra que los planos que no pasan por el origen (los que tienen esferas de Dandelin) intersectan el cono \mathcal{C} en cónicas cuya excentricidad depende de los ángulos (tal como queríamos demostrar).

Obsérvese que no eliminamos los planos horizontales ($\beta = 0$) de la afirmación anterior. Cuando el plano Π tiende a uno horizontal, la cónica se hace un círculo, el foco \mathbf{p} tiende a su centro y la directriz ℓ se aleja hacia el infinito.

EJERCICIO 2.23 Concluye la demostración de que $\Pi \cap \mathcal{C}$ es una hipérbola cuando las dos esferas de Dandelin de Π están en los dos lados del cono usando a éstas y a los dos focos.

EJERCICIO 2.24 Demuestra que si $\beta = \alpha$, entonces el plano Π tiene una única esfera de Dandelin (el caso de la parábola).

Capítulo 3

Transformaciones

Sin duda, el concepto de función tiene un papel fundamental en todas las ramas de la matemática de hoy en día: en la geometría elemental (aunque sea antigua), se dice que dos figuras son *congruentes* si existe una *transformación rígida* que lleve una sobre la otra, y que son *semejantes* si existe una de estas transformaciones que seguida de una *homotecia* (un cambio de escala) lleve una a la otra; en el análisis o el cálculo se estudian las funciones *integrables* o *diferenciables*; en la topología, las funciones *continuas*; en la teoría de los grupos se estudian los *homomorfismos* (funciones entre grupos que preservan la operación ahí definida); en el álgebra lineal se estudian las funciones *lineales* y las *afines*, etc.; siempre que hay algún tipo interesante de “objetos matemáticos” parece haber una correspondiente noción de “funciones” que los relacionan.

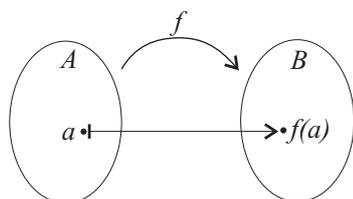
Sin haberlo hecho explícito, los griegos manejaban intuitivamente el concepto de transformación rígida y de semejanza; por ejemplo, en el axioma “todos los ángulos rectos son iguales”, en la palabra “iguales” se incluye la idea de que se puede *mover* uno hasta trasladarse sobre el otro, y en sus teoremas de semejanza la noción ya se hace explícita y habla en el fondo de un cierto tipo de funciones. Más adelante, en el surgimiento del cálculo (Newton y Leibniz) así como en el de la geometría analítica (Descartes) o en el algebra de los arabes, las funciones tenían un papel importante pero debajo del agua o en casos muy concretos (funciones reales de variable real dadas por fórmulas en el cálculo, por ejemplo). Sin embargo, el aislamiento del concepto de función, la generalidad de la noción —que englobaba cosas al parecer distantes—, es muy reciente: termina de afinarse con la teoría de los conjuntos que arranca Cantor en la segunda mitad del siglo XIX. Pero es tan inmediata su aceptación (o bien, ya estaba tan madura su concepción) que a finales del siglo XIX, Felix Klein, en una famosísima conferencia (conocida como “El Programa de Erlangen”), se atreve a afirmar, basado en el trabajo de su gran amigo Sophus Lie, que la geometría es el estudio de un espacio (un conjunto de puntos, piénsese en el plano) junto con un *grupo de transformaciones* (un conjunto específico de funciones del espacio en sí mismo) y de las estructuras que permanecen *invariantes* bajo el grupo. En fin, todavía el estudiante no tiene ejemplos

claros de estas nociones y es casi imposible que lo aprecie. Pero el punto es que, de alguna manera, Klein dijo: para hacer geometría es importantísimo estudiar las *transformaciones* (funciones) del plano en sí mismo, conocerlas de arriba a abajo; y en este capítulo eso intentamos pues el siglo XX, al transcurrir, le fue dando más y más razón al visionario.

Damos muy rápidamente las nociones generales de función en la primera sección. Pues, aunque parece que se nos cae el nivel por un ratito, vale la pena establecer cierta terminología muy general que se usa en su manejo y ciertos ejemplos muy particulares que serán de gran interés en el estudio de las transformaciones geométricas y que, a su vez, servirán para familiarizarnos con los conceptos básicos. Después entramos a las transformaciones geométricas. El orden en que lo hacemos no es el que técnicamente facilita las cosas (empezar por transformaciones lineales), sino el que intuitivamente parece más natural (empezar por transformaciones rígidas); y de ahí, deducir la noción de transformación lineal para regresar de nuevo y desarrollar las fórmulas analíticas.

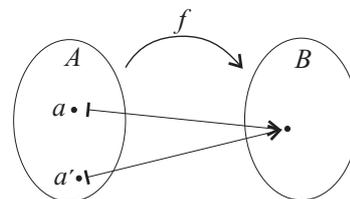
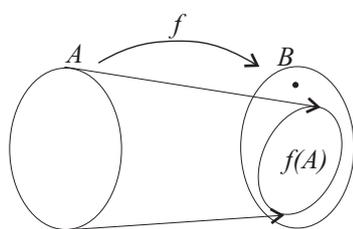
3.1 Funciones y transformaciones

En los siguientes párrafos se usa el tipo de letra *este* para las nociones que se definen formalmente y *este otro* para terminología cómoda y coloquial que facilita mucho el manejo de las funciones.



Dados dos conjuntos A y B, una *función* f de A a B, denotado $f : A \rightarrow B$, es una manera de asociar a cada elemento $a \in A$ un elemento de B, denotado $f(a)$. Por ejemplo, las funciones de \mathbb{R} en \mathbb{R} se describen comúnmente mediante fórmulas como $f(x) = x^2$ o $g(x) = 2x + 3$, que dan la regla para asociar a cada número otro número (por ejemplo, $f(2) = 4$ o $g(-1) = 1$). Otro ejemplo: en el capítulo 1 vimos las funciones de \mathbb{R} en \mathbb{R}^2 que describen rectas parametrizadas. Pero en general, y ésa es la maravillosa idea generalizadora (valga el pleonasma), una función no tiene por qué estar dada por una fórmula; a veces es algo dado por “Dios”, una “caja negra”, que “sabe” como asociar a los elementos del *dominio* (el conjunto A, en la notación con que empezamos) elementos del *contradominio* B.

Se dice que una función $f : A \rightarrow B$ es *inyectiva* si $f(a) = f(a')$ implica que $a = a'$. Es decir, si elementos diferentes de A van bajo f a elementos diferentes de B ($a \neq a' \Rightarrow f(a) \neq f(a')$). Se dice que una función $f : A \rightarrow B$ es *suprayectiva* o *sobre* si para cada elemento de B hay uno en A que le “pega”, es decir, si para cada $b \in B$ existe $a \in A$ tal que $f(a) = b$. Y se dice que es *biyectiva*



o una *biyección*,

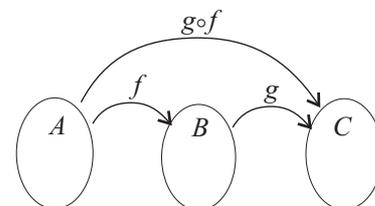
si es inyectiva y suprayectiva; también se le llama *correspondencia biunívoca*. (Diagramas de funciones *no* inyectiva y *no* sobre en las figuras adjuntas).

Las funciones tienen una noción natural de composición, al aplicarse una tras otra para dar una nueva función. Dadas dos funciones $f : A \rightarrow B$ y $g : B \rightarrow C$, su *composición*, denotada $g \circ f$ y a veces llamada *f seguida de g* o bien “*g bolita f*” es la función

$$g \circ f : A \rightarrow C$$

definida por la fórmula

$$(g \circ f)(a) = g(f(a)).$$



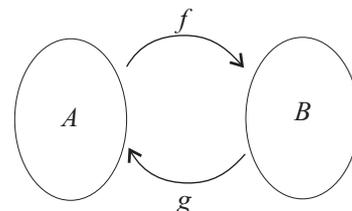
Hay que resaltar que la composición de funciones está definida sólo cuando el *dominio* de g (es decir, de donde “sale” o en donde está definida, el conjunto B en nuestro caso) es igual al *contradominio* de f (es decir, a donde “llega”, el final de la flecha, donde “caen”, otra vez B en nuestro caso); si no fuera así, “ g no sabría que hacer con algunos $f(a)$ ”. También hay que resaltar que la dirección de la escritura “*g bolita f*” es la contraria a la de la acción o lectura (“ f seguida de g ” o bien “*f compuesta con g*”); y esto se ha convenido así por la costumbre y puesto que la fórmula, que a final de cuentas es la que manda, queda mucho más natural.

Por ejemplo, con las funciones $f, g : \mathbb{R} \rightarrow \mathbb{R}$ que definimos con fórmulas unos párrafos antes se tiene que $(g \circ f)(x) = 2x^2 + 3$ mientras que $(f \circ g)(x) = 4x^2 + 12x + 9$, así que *sí* importa el orden; la composición está lejos de ser conmutativa; en general, aunque $g \circ f$ esté definida, $f \circ g$ ni siquiera tiene sentido.

Cada conjunto A , trae consigo una función llamada su *identidad* definida por

$$\begin{aligned} \text{id}_A : A &\rightarrow A \\ \text{id}_A(a) &= a, \end{aligned}$$

que es la función que no *hace nada*, que deja a todos en su lugar. Y, aunque parezca inocua, es fundamental darle un nombre, pues entonces podemos escribir mucho, por ejemplo:



Lema 3.1 Dada una función $f : A \rightarrow B$, se tiene

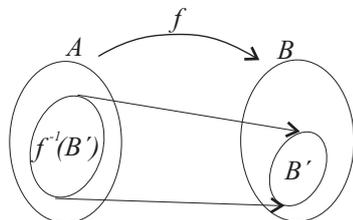
- i) f es inyectiva $\Leftrightarrow \exists g : B \rightarrow A$ tal que $g \circ f = \text{id}_A$,
- ii) f es suprayectiva $\Leftrightarrow \exists g : B \rightarrow A$ tal que $f \circ g = \text{id}_B$.

EJERCICIO 3.1 Da ejemplos de funciones inyectivas que no son sobre y a la inversa.

EJERCICIO 3.2 Demuestra el lema anterior y el siguiente corolario de él.

Corolario 3.2 Dada una función $f : A \rightarrow B$, entonces f es biyectiva si y sólo si existe una función $g : B \rightarrow A$ tal que $g \circ f = \text{id}_A$ y $f \circ g = \text{id}_B$.

En este caso, a g se le da el nombre de *inversa* de f y se le denota f^{-1} . Aunque aquí hay que hacer notar que el símbolo f^{-1} se usa también de una manera más general para denotar conjuntos. Pues para cualquier subconjunto $B' \subset B$ podemos definir su *imagen inversa o pre-imagen* $f^{-1}(B') \subset A$ como



$$f^{-1}(B') := \{a \in A \mid f(a) \in B'\} .$$

Tenemos entonces que f es inyectiva si y sólo si para todo $b \in B$ se tiene que $\#f^{-1}(b) \leq 1$ (donde $\#$ denota cardinalidad y hemos identificado $f^{-1}(b)$ con $f^{-1}(\{b\})$), es decir, “a cada elemento de B le pega a lo más uno de A ”; y que f es sobre si y sólo si para todo $b \in B$ se tiene que $\#f^{-1}(b) \geq 1$ (es decir, que $f^{-1}(b)$ no es el conjunto vacío). Por lo tanto, f es biyectiva si y sólo si $\#f^{-1}(b) = 1$ para todo $b \in B$, es decir, si y sólo si f^{-1} es una función bien definida al aplicarla a *singuletes* de B .

También se puede definir la *imagen directa* de subconjuntos $A' \subset A$, o simplemente su *imagen*, como

$$f(A') := \{f(a) \mid a \in A'\} \subset B;$$

y la *imagen* de la función es $f(A)$, la imagen del dominio.

Finalmente llegamos a la definición que más nos interesa.

Definición 3.1.1 Una *transformación* de A es una función biyectiva de A en A .

Hay que hacer notar que el término “transformación” se usa de diferentes maneras en otros textos y en otros contextos. Pero aquí estaremos tan enfocados a funciones biyectivas de un conjunto en sí mismo que lo asignaremos a ellas.

EJERCICIO 3.3 Demuestra que si $f : A \rightarrow B$ y $g : B \rightarrow C$ son biyectivas entonces $g \circ f$ también es biyectiva. (Demuestra que $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.)

EJERCICIO 3.4 Demuestra que si f y g son transformaciones de un conjunto A entonces $(f \circ g)$ también es una transformación de A .

3.1.1 Grupos de transformaciones

Veremos ahora ciertos ejemplos pequeños de conjuntos de transformaciones, que por su importancia reciben un nombre especial, el de *grupo*.

Consideremos un conjunto con dos elementos, $\{0, 1\}$; llamémoslo Δ_2 , que se lee “delta-dos”. Las funciones de Δ_2 en sí mismo son cuatro :

$$\begin{array}{cccc} \text{id} & & c_0 & & c_1 & & \rho \\ 0 \mapsto 0 & & 0 \mapsto 0 & & 0 \mapsto 1 & & 0 \mapsto 1 \\ 1 \mapsto 1 & & 1 \mapsto 0 & & 1 \mapsto 1 & & 1 \mapsto 0, \end{array}$$

donde hemos usado la notación $x \mapsto y$ para especificar que el elemento x va a dar al elemento y bajo la función en cuestión (no hay que confundir la flechita con “raya

de salida” \mapsto con la flecha \rightarrow que denota función; inclusive puede pensarse que esta última (\rightarrow) es el conjunto de todas las flechitas (\mapsto) entre los elementos). En nuestro ejemplo, las dos funciones de enmedio son *funciones constantes*; y las únicas transformaciones son las de los extremos, id y ρ , donde hay que observar que ρ es su propio inverso, es decir, $\rho \circ \rho = \text{id}$.

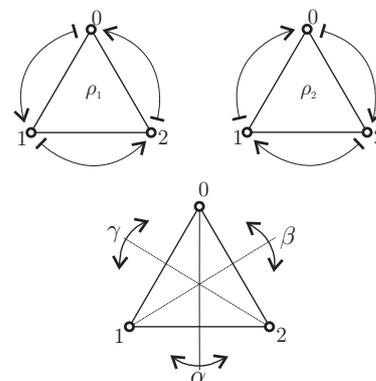
Consideremos ahora a $\Delta_3 := \{0, 1, 2\}$, un conjunto con tres elementos. Una función de Δ_3 en sí mismo puede especificarse en una tablita:

$$\begin{array}{l} 0 \mapsto x \\ 1 \mapsto y \\ 2 \mapsto z, \end{array}$$

donde $x, y, z \in \Delta_3$. Como las imágenes ($x, y, z \in \Delta_3$) son arbitrarias, tenemos que hay $3^3 = 27$ funciones en total; pero de éstas solamente $3 \times 2 \times 1 = 6$ son transformaciones. Pues si queremos que sea biyectiva, una vez que el 0 escoge su imagen, al 1 sólo le quedan dos opciones para escoger y cuando lo hace, al 2 ya no le queda más que una opción obligada.

Estas seis transformaciones son:

	id		ρ_1		ρ_2
0	\mapsto 0		\mapsto 1		\mapsto 2
1	\mapsto 1		\mapsto 2		\mapsto 0
2	\mapsto 2		\mapsto 0		\mapsto 1
	α		β		γ
0	\mapsto 0		\mapsto 2		\mapsto 1
1	\mapsto 2		\mapsto 1		\mapsto 0
2	\mapsto 1		\mapsto 0		\mapsto 2



que podemos visualizar como las “simetrías” de un triángulo equilátero. Vistas así, las tres de arriba corresponden a rotaciones (la identidad rota 0 grados), y cumplen con que ρ_1 y ρ_2 son inversas, es decir $\rho_1 \circ \rho_2 = \rho_2 \circ \rho_1 = \text{id}$, lo que corresponde a que una rota 120° en una dirección y la otra 120° en la dirección contraria. Pero también cumplen que $\rho_1 \circ \rho_1 = \rho_2$ (que podríamos escribir $\rho_1^2 = \rho_2$ si convenimos en que

$$f^n = \overbrace{f \circ f \circ \dots \circ f},$$

donde f es cualquier transformación; es decir, f^n es f compuesta consigo misma n veces, lo cual tiene sentido sólo cuando f sale de y llega al mismo conjunto). Por su parte, las tres transformaciones siguientes se llaman *transposiciones* y geométricamente se ven como *reflexiones*. Cumplen que $\alpha^2 = \beta^2 = \gamma^2 = \text{id}$, y además cumplen otras relaciones como que $\alpha \circ \beta = \rho_1$, lo cual se ve persiguiendo elementos:

	β		α
0	\mapsto 2		\mapsto 1
1	\mapsto 1		\mapsto 2
2	\mapsto 0		\mapsto 0;

o bien, que $\alpha \circ \beta \circ \alpha = \beta \circ \alpha \circ \beta = \gamma$ (¡compruébalo persiguiendo elementos!). Para poder concluir con elegancia, conviene introducir las siguientes nociones generales. Hay ciertos conjuntos de transformaciones que son tan importantes que conviene darles un nombre específico, el de **grupo**:

Definición 3.1.2 A un conjunto G de transformaciones de un conjunto A se le llama un *grupo de transformaciones* de A si cumple

- i) $\text{id}_A \in G$
- ii) $f, g \in G \Rightarrow g \circ f \in G$
- iii) $f \in G \Rightarrow f^{-1} \in G$.

Ejemplos triviales de grupos serían el conjunto de *todas* las transformaciones de A , o bien, sólo su identidad, $\{\text{id}_A\}$.

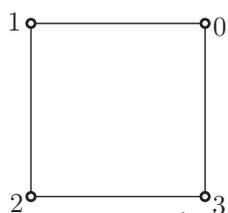
En el caso de $A = \Delta_3$, el grupo de todas sus transformaciones tiene seis elementos; pero también hay otro grupo de transformaciones que consiste en las tres del primer renglón, las rotaciones; pues contiene a la identidad (es un conjunto no vacío), es *cerrado bajo composición* (cumple **(ii)**) y es *cerrado bajo inversas* (cumple **(iii)**). Y también cada una de las transposiciones (o reflexiones) junto con la identidad forman un grupo (con dos elementos) de transformaciones de Δ_3 .

Dado un conjunto cualquiera de transformaciones de A , el grupo que *genera* es el grupo de transformaciones que se obtiene de todas las posibles composiciones con elementos de él o sus inversos.

Por ejemplo, α y β generan todas las transformaciones de Δ_3 (pues ya las hemos descrito como composiciones de α y β); mientras que ρ_1 genera el grupo de rotaciones de Δ_3 (que consiste de id , ρ_1 y $\rho_1^2 = \rho_2$). Las *relaciones* que cumplen α y β son

$$\alpha^2 = \beta^2 = (\beta \circ \alpha)^3 = \text{id}.$$

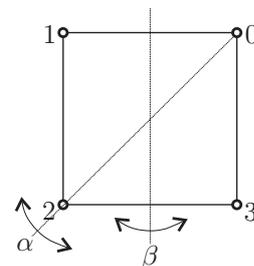
Si consideramos ahora a $\Delta_4 := \{0, 1, 2, 3\}$, tendríamos que el conjunto de sus funciones tiene $4^4 = 256$ elementos, mientras que el grupo de todas sus transformaciones tiene $4! = 4 \times 3 \times 2 \times 1 = 24$ elementos (demasiados para escribirlos todos). Pero dentro de éste, podemos encontrar otros grupos, que podemos llamar *subgrupos*.



Uno importante es el de aquellas transformaciones que mantienen intacta (que *preservan*) la estructura del cuadrado regular cuyos vértices se etiquetan 0, 1, 2, 3 en orden cíclico; por ejemplo, la que mantiene fijos al 0 y al 1 pero que transpone al 2 y al 3 no preserva al cuadrado pues, por ejemplo, la arista del 1 al 2 va a una diagonal, la 1-3. No es difícil ver que éstas (las *simetrías* del cuadrado) son ocho: las cuatro rotaciones (incluyendo a la identidad), y cuatro reflexiones.

Pero en vez de escribirlas todas podemos describirlas mediante generadores. Sean ahora

$$\begin{array}{cc}
 & \alpha & & \beta \\
 0 & \mapsto & 0 & 0 \mapsto 1 \\
 1 & \mapsto & 3 & 1 \mapsto 0 \\
 2 & \mapsto & 2 & 2 \mapsto 3 \\
 3 & \mapsto & 1 & 3 \mapsto 2,
 \end{array} \tag{3.1}$$



entonces $\beta \circ \alpha$ es la rotación $(0 \mapsto 1 \mapsto 2 \mapsto 3 \mapsto 0)$ y las otras dos rotaciones del cuadrado son $(\beta \circ \alpha)^2$ y $(\beta \circ \alpha)^3$; mientras que las dos reflexiones que faltan describir son $(\alpha \circ \beta \circ \alpha)$ y $(\beta \circ \alpha \circ \beta)$. Podemos entonces concluir que el grupo de simetrías del cuadrado está generado por α y β que cumplen las relaciones

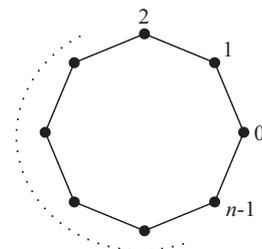
$$\alpha^2 = \beta^2 = (\beta \circ \alpha)^4 = \text{id},$$

mientras que el grupo de rotaciones tiene un solo generador $\rho = (\beta \circ \alpha)$ que cumple $\rho^4 = \text{id}$.

Para referencia posterior, y a reserva de que se estudien con más detenimiento en el caso general en la Sección 3.3.2, conviene poner nombre a los grupos de transformaciones que hemos descrito.

Al conjunto de todas las transformaciones de un conjunto con n elementos $\Delta_n := \{0, 1, \dots, n-1\}$ se le llama el *grupo simétrico de orden n* ; se le denota \mathbf{S}_n y consta de $n! := n \times (n-1) \times (n-2) \times \dots \times 2 \times 1$ (n factorial) elementos también llamados *permutaciones*.

Dentro de este grupo hemos considerado dos subgrupos (para $n = 3, 4$): el que preserva la estructura del polígono regular con n lados cuyos vértices se etiquetan $0, 1, 2, \dots, (n-1)$ en orden cíclico, al cual se le llama el *grupo diédrico de orden n* , y se le denota \mathbf{D}_n , que tiene $2n$ elementos; y el subgrupo de éste generado por la *rotación* $(0 \mapsto 1 \mapsto 2 \mapsto \dots \mapsto (n-1) \mapsto 0)$ que tiene n elementos, se le llama *grupo cíclico de orden n* y se le denota \mathbf{C}_n . Hay que observar que para $n = 3$ el diédrico y el simétrico coinciden ($\mathbf{S}_3 = \mathbf{D}_3$) pero esto ya no sucede para $n = 4$.



EJERCICIO 3.5 Para el caso $n = 4$, escribe explícitamente (en forma de tabla de asignaciones) las ocho transformaciones del grupo diédrico \mathbf{D}_4 ; su expresión más económica (en el número de símbolos usados) como composición de α y β (definidas por las tablas de asignaciones (3.1)), y también da su representación geométrica.

EJERCICIO 3.6 ¿Puedes encontrar un subgrupo de \mathbf{S}_4 “esencialmente igual” a \mathbf{S}_3 ?

EJERCICIO 3.7 ¿Puedes encontrar un subgrupo de \mathbf{S}_4 de orden 12? (Piensa en un tetraedro regular en \mathbb{R}^3 y describe el grupo geoméricamente.)

EJERCICIO 3.8 ¿Puedes encontrar una permutación $\gamma : \Delta_4 \rightarrow \Delta_4$ tal que α, β y γ generen todas las permutaciones \mathbf{S}_4 (por supuesto que $\gamma \notin \mathbf{D}_4$)?

EJERCICIO 3.9 Da explícitamente (con la tabla de asignación $\Delta_n \rightarrow \Delta_n$) dos generadores α y β para los grupos diédricos de orden $n = 5$ y $n = 6$ (es decir, para las simetrías del pentágono y el hexágono). ¿Qué relaciones cumplen? ¿Puedes intuir y describir lo equivalente para el caso general?

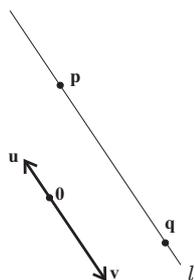
***EJERCICIO 3.10** Considera un cubo Q en \mathbb{R}^3 . Etiqueta sus vértices con Δ_8 . Dentro de S_8 hay un subgrupo que es el que preserva la estructura geométrica del cubo. ¿Cuántos elementos tiene? ¿Puedes describirlo con generadores y relaciones?

3.2 Las transformaciones afines de \mathbb{R}

El primer grupo de transformaciones geométrico-analíticas que estudiaremos es el que surge de la ambigüedad en la parametrización de rectas. Recuérdese que nuestra definición original de una recta fue con un parámetro real, pero para una sola recta hay muchas parametrizaciones (dependen de escoger un punto base y un vector direccional); la forma de relacionarse de estos parámetros serán las transformaciones afines.

Supongamos que tenemos una misma recta l parametrizada de dos maneras distintas. Es decir, que tenemos

$$\begin{aligned} l &= \{\mathbf{p} + t\mathbf{v} : t \in \mathbb{R}\} \\ l &= \{\mathbf{q} + s\mathbf{u} : s \in \mathbb{R}\}, \end{aligned}$$

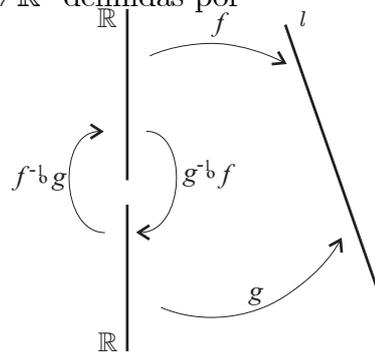


donde $\mathbf{p}, \mathbf{q}, \mathbf{v}, \mathbf{u} \in \mathbb{R}^2$ (aunque también funciona el razonamiento que sigue en cualquier espacio vectorial). Sabemos además que los vectores direccionales, \mathbf{u} y \mathbf{v} son distintos de $\mathbf{0}$ para que efectivamente describan una línea recta. La pregunta es ¿cómo se relacionan los parámetros t y s ?

Otra manera de pensar estas rectas es como la imagen de funciones cuyo dominio son los reales, es decir, tenemos dos funciones $f, g : \mathbb{R} \rightarrow \mathbb{R}^2$ definidas por

$$\begin{aligned} f(t) &= \mathbf{p} + t\mathbf{v} \\ g(s) &= \mathbf{q} + s\mathbf{u}, \end{aligned}$$

cuya imagen es la misma recta l . Puesto que ambas son biyecciones sobre l , podemos restringir el codominio y pensarlas a ambas como funciones de \mathbb{R} en l . Y entonces tiene sentido hablar de sus inversas $f^{-1} : l \rightarrow \mathbb{R}$ y $g^{-1} : l \rightarrow \mathbb{R}$. La pregunta es entonces ¿cómo se expresan las funciones $(g^{-1} \circ f)$ y $(f^{-1} \circ g)$ que van de \mathbb{R} en \mathbb{R} ?



Puesto que \mathbf{u} y \mathbf{v} son paralelos (definen la misma recta), existe un número real $\mathbf{a} \in \mathbb{R}$ tal que $\mathbf{u} = \mathbf{a}\mathbf{v}$; de hecho $\mathbf{a} = (\mathbf{u} \cdot \mathbf{v}) / (\mathbf{v} \cdot \mathbf{v})$, y obsérvese que $\mathbf{a} \neq 0$. Y además como $\mathbf{q} \in \ell$, se puede expresar en términos de la primera parametrización; es decir, existe $\mathbf{b} \in \mathbb{R}$ tal que $\mathbf{q} = f(\mathbf{b}) = \mathbf{p} + \mathbf{b}\mathbf{v}$. Tenemos entonces que para cualquier $s \in \mathbb{R}$:

$$\begin{aligned} g(s) &= \mathbf{q} + s\mathbf{u} \\ &= \mathbf{p} + \mathbf{b}\mathbf{v} + s(\mathbf{a}\mathbf{v}) \\ &= \mathbf{p} + (\mathbf{as} + \mathbf{b})\mathbf{v} \\ &= f(\mathbf{as} + \mathbf{b}); \end{aligned}$$

al aplicar la función f^{-1} a ambos lados de esta ecuación se obtiene

$$(f^{-1} \circ g)(s) = \mathbf{as} + \mathbf{b}.$$

Para obtener la otra composición, podemos proceder más directamente, considerando esta última expresión como la que da al parametro \mathbf{t} . Es decir

$$\mathbf{t} = \mathbf{as} + \mathbf{b},$$

de donde, como $\mathbf{a} \neq 0$, podemos despejar

$$s = \mathbf{a}^{-1}\mathbf{t} - \mathbf{ba}^{-1}.$$

Definición 3.2.1 Una función $f : \mathbb{R} \rightarrow \mathbb{R}$ se llama *afín* si se escribe como

$$f(x) = \mathbf{ax} + \mathbf{b}, \tag{3.2}$$

donde $\mathbf{a}, \mathbf{b} \in \mathbb{R}$; y cuando $\mathbf{a} \neq 0$ la llamaremos *transformación afín*.

Nuestro uso de el término transformación se justifica por el siguiente ejercicio.

EJERCICIO 3.11 Demuestra que la función afín (3.2) es biyectiva si y sólo si $\mathbf{a} \neq 0$.

EJERCICIO 3.12 Sean $f, g : \mathbb{R} \rightarrow \mathbb{R}$ definidas por $f(x) = 2x + 1$ y $g(x) = -x + 2$. Encuentra las fórmulas para f^{-1} , g^{-1} , f^2 , $f \circ g$ y $g \circ f$.

EJERCICIO 3.13 Demuestra que las transformaciones afines son cerradas bajo inversas y composición.

Obsérvese que las gráficas de las funciones afines son las rectas no verticales, pues éstas están dadas por la ecuación $\mathbf{y} = \mathbf{ax} + \mathbf{b}$. Y que de éstas las rectas no horizontales corresponden a las transformaciones afines.

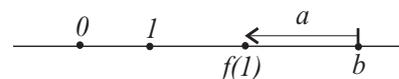
Al conjunto de todas las transformaciones afines de \mathbb{R} , que por el ejercicio anterior forman un grupo, lo denotaremos $\mathbf{Af}(1)$.

Podemos ahora resumir lo que hicimos en los párrafos anteriores.

Lema 3.3 *Dos parametrizaciones de una misma recta se relacionan por una transformación afín entre los parámetros.* □

Así como las parametrizaciones de rectas dependen de la elección de dos puntos en ellas, las transformaciones afines dependen únicamente de dos valores. Pues las constantes que la definen, a y b en nuestro caso (3.2), se obtienen como

$$\begin{aligned} b &= f(0) \\ a &= f(1) - f(0) \end{aligned}$$



y representan un cambio de escala (multiplicar por a) y luego una traslación (sumarle b). De tal manera que si nos dicen que f es una función afín que manda al 0 en b ($f(0) = b$) y al 1 en c ($f(1) = c$), recuperamos toda la función por la fórmula

$$f(x) = (c - b)x + b,$$

y obsérvese que es transformación cuando $f(0) \neq f(1)$; si no: es constante. Si el 0 y el 1 pueden ir a cualquier pareja de números por una transformación afín, entonces cualquier pareja de números (distintos) puede ser enviada al $0, 1$ (por la inversa) y de ahí a cualquier otra pareja. Hemos demostrado así el:

Teorema 3.4 (Dos en Dos) *Dados dos pares de puntos x_0, x_1 y y_0, y_1 en \mathbb{R} , (donde por par se entiende que son distintos, i.e., $x_0 \neq x_1$ y $y_0 \neq y_1$), existe una única transformación afín $f \in \mathbf{Af}(1)$ que manda una en la otra, i.e., tal que $f(x_0) = y_0$ y $f(x_1) = y_1$. \square*

EJERCICIO 3.14 Encuentra la transformación afín $f: \mathbb{R} \rightarrow \mathbb{R}$ que cumple:

- a) $f(2) = 4$ y $f(5) = 1$
- b) $f(1) = -1$ y $f(2) = 3$
- c) $f(-1) = 0$ y $f(3) = 2$

EJERCICIO 3.15 Encuentra la fórmula explícita (en términos de x_0, x_1, y_0, y_1) de la función f del Teorema 3.4.

EJERCICIO 3.16 Discute qué pasa en el Teorema 3.4 (y en la fórmula del ejercicio anterior) si se permite que en los pares se dé la igualdad.

3.2.1 Isometrías de \mathbb{R}

Hemos dicho que las transformaciones afines de la recta consisten de un “cambio de escala” (determinado por la constante a) y luego una traslación (determinada por b). Pero podemos ser más precisos. Como la distancia en \mathbb{R} se mide por la fórmula $d(x, y) = |x - y|$, podemos demostrar que todas las distancias cambian por el mismo factor bajo la transformación afín $f(x) = ax + b$:

$$\begin{aligned} d(f(x), f(y)) &= |ax + b - (ay + b)| = |ax - ay| \\ &= |a(x - y)| = |a||x - y| = |a|d(x, y). \end{aligned}$$

Tenemos entonces una familia distinguida de transformaciones afines que son las que **no** cambian la escala, las que mantienen rígida la recta real y que llamaremos *isometrías* de \mathbb{R} porque preservan la métrica (la distancia). Denotemos

$$\mathbf{Iso}(1) := \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f(x) = ax + b \quad \text{con} \quad |a| = 1\}$$

(dejamos como ejercicio demostrar que es un grupo de transformaciones).

Tenemos que $\mathbf{Iso}(1)$ se divide naturalmente en dos clases de transformaciones: las *traslaciones*, cuando $a = 1$ (y la función es entonces $f(x) = x + b$), que consisten en deslizar la recta rígidamente hasta que el 0 caiga en b ; o bien las *reflexiones*, cuando $a = -1$, y que entonces se escriben $g(x) = -x + b$. Veámos que estas últimas tienen un *punto fijo*, es decir, un punto que se queda en su lugar bajo la transformación. Éste debe satisfacer la ecuación $g(x) = x$ que es

$$x = -x + b$$

y que implica $x = b/2$. Entonces, lo que hacen las reflexiones es intercambiar rígidamente los dos lados de un punto que podemos llamar su *espejo* (pues por ejemplo, $g(b/2 + 1) = b/2 - 1$ y en general se tiene que $g(b/2 + x) = b/2 - x$), y de ahí el nombre de “reflexión”.

EJERCICIO 3.17 Demuestra que $\mathbf{Iso}(1)$ es un grupo de transformaciones de \mathbb{R} .

EJERCICIO 3.18 Demuestra que las traslaciones de \mathbb{R} , que denotaremos $\mathbf{Tra}(1)$, forman un grupo de transformaciones.

EJERCICIO 3.19 Demuestra que si $f \in \mathbf{Af}(1)$ no tiene puntos fijos (es decir, que $f(x) \neq x$ para toda $x \in \mathbb{R}$), entonces es una traslación no trivial. (La trivial es trasladar por 0, que es la identidad y tiene a todo \mathbb{R} como puntos fijos).

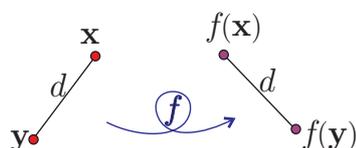
EJERCICIO 3.20 Demuestra, usando la fórmula, que la inversa de cualquier reflexión es ella misma.

EJERCICIO 3.21 Demuestra que la composición de dos reflexiones es una traslación del doble de la distancia (dirigida) entre sus espejos.

3.3 Isometrías y transformaciones ortogonales

En esta sección se estudian las transformaciones más importantes de la geometría euclidiana en su sentido estricto, pues son las que preservan la métrica, la noción de distancia y, por lo tanto, la estructura rígida de las figuras geométricas. Puesto que la definición general es intuitivamente nítida, partiremos de ella y deduciremos las propiedades básicas de las isometrías que nos llevarán, en las secciones siguientes, a obtener las fórmulas explícitas que las definen y que entonces nos facilitarán la obtención de nuevos resultados y su comprensión cabal.

Es importante señalar que muchos de los resultados de esta sección, así como las definiciones, sólo dependen de las nociones de distancia y producto interior en un espacio vectorial. Así que procederemos en general (para $n = 1, 2, 3$, puede pensarse) y hasta el final de la sección, salvo por los ejemplos, concretaremos los resultados al plano.



Definición 3.3.1 Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una *isometría* si preserva distancia. Es decir, si para todo par $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ se cumple que

$$d(\mathbf{x}, \mathbf{y}) = d(f(\mathbf{x}), f(\mathbf{y})).$$

También se les llama *transformaciones rígidas* (aunque esta terminología no cobrará sentido hasta que demos demos que son biyectivas). Se denota por $\mathbf{Iso}(n)$ al conjunto de todas las isometrías de \mathbb{R}^n .

Nótese que la definición tiene sentido en cualquier espacio métrico, y que coincide para $n = 1$ con las de la sección anterior (3.2.1). Las isometrías mandan al espacio en sí mismo de manera tal que cualquier estructura rígida se mantiene. Estas funciones o transformaciones las vemos a diario. Por ejemplo, mover una silla de un lugar a otro induce una isometría si pensamos que el espacio euclidiano se puede generar y definir respecto de ella; el efecto de esta isometría en la silla es ponerla en su destino.

Antes de ver más ejemplos en detalle, demostraremos tres lemas generales que usan solamente la definición de isometría y tienden hacia la demostración de que las isometrías forman un grupo de transformaciones.

Lema 3.5 Una isometría $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es *inyectiva*.

Demostración. Esto se debe a que la distancia entre puntos diferentes es estrictamente positiva. Formalmente, supongamos que $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ son tales que $f(\mathbf{x}) = f(\mathbf{y})$. Esto implica que $d(f(\mathbf{x}), f(\mathbf{y})) = 0$. Como f es isometría, entonces $d(\mathbf{x}, \mathbf{y}) = d(f(\mathbf{x}), f(\mathbf{y})) = 0$ y por lo tanto $\mathbf{x} = \mathbf{y}$. \square

Lema 3.6 Si $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ son isometrías entonces $g \circ f$ también lo es.

Demostración. Usando la regla de composición y la definición de isometría dos veces (primero para g y luego para f), se obtiene que para cualquier $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$:

$$\begin{aligned} d((g \circ f)(\mathbf{x}), (g \circ f)(\mathbf{y})) &= d(g(f(\mathbf{x})), g(f(\mathbf{y}))) \\ &= d(f(\mathbf{x}), f(\mathbf{y})) \\ &= d(\mathbf{x}, \mathbf{y}) \end{aligned}$$

y por tanto $g \circ f \in \mathbf{Iso}(n)$. \square

Lema 3.7 Si $f \in \mathbf{Iso}(\mathbf{n})$ y tiene inversa f^{-1} , entonces $f^{-1} \in \mathbf{Iso}(\mathbf{n})$.

Demostración. Dados $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, se tiene que

$$d(f^{-1}(\mathbf{x}), f^{-1}(\mathbf{y})) = d(f(f^{-1}(\mathbf{x})), f(f^{-1}(\mathbf{y})))$$

pues f es isometría. Pero la última expresión es $d(\mathbf{x}, \mathbf{y})$, lo cual demuestra que $f^{-1} \in \mathbf{Iso}(\mathbf{n})$. \square

Nos falta entonces demostrar que las isometrías son suprayectivas para concluir que $\mathbf{Iso}(\mathbf{n})$ es un grupo de transformaciones. Aunque esto sea intuitivamente claro (“una transformación rígida del plano no puede dejar partes descubiertas”), la demostración sería de momento complicada; con un poco más de técnica será muy sencilla. Así que vale suponer que es cierto por un rato, desarrollar los ejemplos, la intuición y la técnica y luego volver a preocuparnos.

3.3.1 Ejemplos

Ya hemos visto los ejemplos en la recta (cuando $\mathbf{n} = 1$), que son traslaciones y reflexiones. Ejemplos en el plano serían rotar alrededor de un punto fijo, reflejar en una recta o bien trasladar por un vector fijo todo el plano. Estas últimas son muy fáciles de definir:

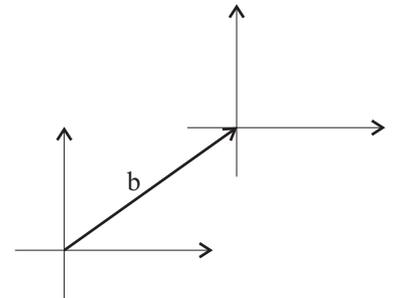
Traslaciones

Dado un vector $\mathbf{b} \in \mathbb{R}^n$, la *traslación por \mathbf{b}* es la función

$$\begin{aligned}\tau_{\mathbf{b}} : \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ \tau_{\mathbf{b}}(\mathbf{x}) &= \mathbf{x} + \mathbf{b},\end{aligned}$$

que claramente es una transformación (inyectiva, pues $\mathbf{x} + \mathbf{b} = \mathbf{y} + \mathbf{b} \Rightarrow \mathbf{x} = \mathbf{y}$; y sobre, pues $\tau_{\mathbf{b}}^{-1} = \tau_{-\mathbf{b}}$). Y además es una isometría ($\tau_{\mathbf{b}} \in \mathbf{Iso}(\mathbf{n})$) ya que

$$\begin{aligned}d(\tau_{\mathbf{b}}(\mathbf{x}), \tau_{\mathbf{b}}(\mathbf{y})) &= |\tau_{\mathbf{b}}(\mathbf{x}) - \tau_{\mathbf{b}}(\mathbf{y})| \\ &= |(\mathbf{x} + \mathbf{b}) - (\mathbf{y} + \mathbf{b})| = |\mathbf{x} - \mathbf{y}| = d(\mathbf{x}, \mathbf{y}).\end{aligned}$$

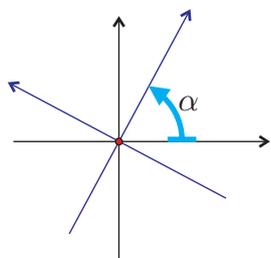


De hecho, las traslaciones forman un grupo de transformaciones de \mathbb{R}^n , al que denotaremos $\mathbf{Tra}(\mathbf{n})$, que puede identificarse con el grupo aditivo \mathbb{R}^n ; en lenguaje de grupos “son *isomorfos*”, pues hay una traslación por cada elemento de \mathbb{R}^n , y su regla de composición es claramente

$$\tau_{\mathbf{b}} \circ \tau_{\mathbf{a}} = \tau_{(\mathbf{a} + \mathbf{b})}.$$

Rotaciones

Definir explícitamente las rotaciones es más difícil, aunque intuitivamente sea claro a qué nos referimos: “clavar una tachuela en algún punto y luego rotar el plano alrededor de ella un cierto ángulo”. Usando coordenadas polares sí es fácil definir la rotación de un ángulo α alrededor del origen, pues al ángulo de cualquier vector simplemente le sumamos el ángulo de rotación. Así que podemos definir la *rotación de un ángulo α* alrededor del origen como



$$\rho_\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

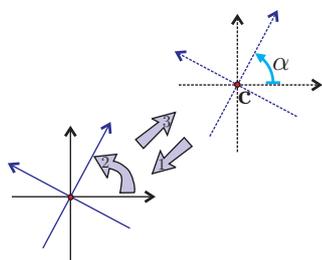
$$\rho_\alpha(\theta, r) = (\theta + \alpha, r) \quad (\text{en coordenadas polares}).$$

Nótese que entonces $\rho_0 = \rho_{2\pi} = \text{id}_{\mathbb{R}^2}$ y se cumple, como en las traslaciones, que

$$\rho_\beta \circ \rho_\alpha = \rho_{(\alpha+\beta)},$$

pero ahora sumando ángulos. Además son transformaciones (biyectivas, insistimos en ello) pues tienen inversa $\rho_\alpha^{-1} = \rho_{-\alpha}$ de tal manera que las rotaciones alrededor del origen forman un grupo de transformaciones de \mathbb{R}^2 (isomorfo al de los ángulos con la suma: los elementos de este grupo —transformaciones por definición, que más adelante denotaremos con $\text{SO}(2)$ — están en correspondencia uno a uno con los puntos del círculo unitario \mathbb{S}^1 de tal manera que la composición corresponde a la suma de ángulos).

Ahora, usando las traslaciones podemos definir las rotaciones con centro en cualquier otro lado con un truco llamado *conjugación*. Para obtener la rotación de un ángulo α con centro en el punto \mathbf{c} , denotémosla con $\rho_{\alpha, \mathbf{c}}$; podemos llevar el centro \mathbf{c} al origen por medio de la traslación $\tau_{-\mathbf{c}}$, rotamos ahí y luego regresamos a \mathbf{c} a su lugar; es decir, podemos definir



$$\rho_{\alpha, \mathbf{c}} = \tau_{\mathbf{c}} \circ \rho_\alpha \circ \tau_{-\mathbf{c}}.$$

Aunque sea intuitivamente cristalino que las rotaciones son isometrías (se puede rotar un vidrio), no tenemos aún una demostración formal de ello; y con coordenadas polares se ve “en chino” pues expresar distancias (o traslaciones) en términos de ellas está igual de complicado. Mejor nos esperamos a tener buenas expresiones cartesianas de las rotaciones para demostrar que preservan distancias y que al componer dos rotaciones cualesquiera se obtiene otra (el problema por ahora es encontrar su centro). Y entonces veremos que junto con las traslaciones forman el grupo de *movimientos rígidos del plano*, llamados así pues se puede llegar a cualquiera de estas transformaciones *moviendo* continuamente (poco a poco) el plano.

EJERCICIO 3.22 Demuestra formalmente (con las definiciones anteriores) que el conjunto de rotaciones alrededor de un punto dado \mathbf{c} es un grupo.

EJERCICIO 3.23 ¿Puedes encontrar el centro de la rotación de 90° que se obtiene rotando 45° en el origen y luego rotando otros 45° en el punto $(2, 0)$? (Haz dibujos —o juega con tachuelas y un acetato sobre un papel cuadriculado— y busca un punto que se quede en su lugar después de las dos rotaciones.) ¿Cuál es el centro de rotación si se invierte el orden de la composición?

EJERCICIO 3.24 A partir de tu experiencia con el ejercicio anterior, y sin preocuparte por las formalidades, ¿puedes dar una receta geométrica para encontrar el centro de rotación de la composición de dos rotaciones (con centros distintos, por supuesto)?

EJERCICIO 3.25 ¿Puedes conjeturar qué transformación es $\rho_{-\alpha, \mathbf{b}} \circ \rho_{\alpha, \mathbf{c}}$? ¿Puedes demostrarlo? Si no, ¿qué hace falta?

Reflexiones

Las reflexiones (que intercambian rígidamente los dos lados de una “línea espejo”) las podemos definir usando la proyección ortogonal a una recta. Dada una recta $\ell \subset \mathbb{R}^2$, se le puede definir por la ecuación normal $\mathbf{u} \cdot \mathbf{x} = c$ con \mathbf{u} un vector unitario y $c \in \mathbb{R}$ una constante (lo que habíamos llamado ecuación unitaria). Sabemos que para cualquier $\mathbf{x} \in \mathbb{R}^2$ el vector que lleva a \mathbf{x} ortogonalmente a la recta ℓ es

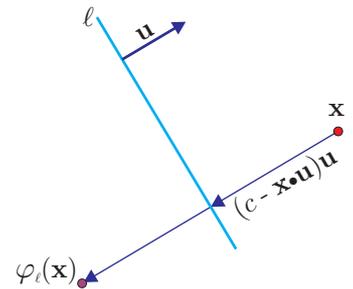
$$(c - \mathbf{u} \cdot \mathbf{x}) \mathbf{u}$$

pues

$$\mathbf{u} \cdot (\mathbf{x} + (c - \mathbf{u} \cdot \mathbf{x}) \mathbf{u}) = \mathbf{u} \cdot \mathbf{x} + (c - \mathbf{u} \cdot \mathbf{x}) (\mathbf{u} \cdot \mathbf{u}) = c.$$

Entonces, el reflejado de \mathbf{x} en el espejo ℓ será *empujarlo* otro tanto del otro lado de ℓ . Así que definimos la *reflexión* de \mathbb{R}^2 a lo largo de $\ell : \mathbf{u} \cdot \mathbf{x} = c$ (con $|\mathbf{u}| = 1$) como

$$\begin{aligned} \varphi_\ell : \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ \varphi_\ell(\mathbf{x}) &= \mathbf{x} + 2(c - \mathbf{u} \cdot \mathbf{x}) \mathbf{u}. \end{aligned} \tag{3.3}$$



Obsérvese que no nos preocupamos por las propiedades de grupo respecto a las reflexiones porque **no** lo son. Ni siquiera contienen la identidad.

EJERCICIO 3.26 Usando coordenadas (x, y) , da fórmulas más sencillas para las reflexiones en los ejes coordenados y verifica que coinciden con la fórmula anterior.

EJERCICIO 3.27 Usando coordenadas da fórmulas sencillas para las reflexiones en las rectas $x = y$ y $x = -y$.

EJERCICIO 3.28 Usando coordenadas da fórmulas sencillas para las reflexiones en las rectas $x = c$ y $y = c$ donde c es cualquier constante.

EJERCICIO 3.29 Define reflexiones en planos de \mathbb{R}^3 .

EJERCICIO 3.30 Demuestra que las reflexiones son isometrías.

EJERCICIO 3.31 Demuestra que si φ es una reflexión, entonces $\varphi^{-1} = \varphi$.

EJERCICIO 3.32 Demuestra que si φ_ℓ es la reflexión en la recta ℓ , entonces $\mathbf{x} \in \mathbb{R}^2$ es un punto fijo de φ_ℓ , i.e. $\varphi_\ell(\mathbf{x}) = \mathbf{x}$, si y sólo si $\mathbf{x} \in \ell$.

3.3.2 Grupos de simetrías

Supongamos por un momento (luego lo demostraremos) que las isometrías de \mathbb{R}^n ($n = 1, 2, 3$) forman un grupo (algo que no es difícil de creer). Entonces podemos definir los “grupos de simetrías” de las figuras, en el plano o de los cuerpos en el espacio, correspondiendo a nuestra intuición milenaria de lo que queremos decir cuando decimos que algo “tiene simetría”, por ejemplo al hablar de una flor, de un dibujo o de un edificio.

Para fijar ideas, pensemos en una figura en un papel, y al papel como el plano \mathbb{R}^2 . La abstracción más elemental es que la figura es un subconjunto F de \mathbb{R}^2 : los puntos del plano están en F (pintados de negro) o no (de blanco como el papel). Si nos ponemos quisquillosos, podríamos distinguir colores en los puntos (algo que hacen las computadoras hoy día) y tener entonces una función $\gamma : F \rightarrow \{\text{Colores}\}$ para ser más precisos. Pero quedémonos con el bulto, una *figura* F es un subconjunto de \mathbb{R}^n (de una vez englobamos los sólidos en \mathbb{R}^3). Una *simetría* de F (y hasta el término “sí-metría” lo indica) es una transformación que preserva la métrica y la figura, es decir una $f \in \mathbf{Iso}(n)$ tal que $f(F) = F$ (donde estamos aplicando la transformación a un conjunto para obtener un nuevo conjunto). Podemos definir entonces el *grupo de simetrías* de la figura $F \subset \mathbb{R}^n$ como el conjunto

$$\mathbf{Sim}(F) := \{f \in \mathbf{Iso}(n) \mid f(F) = F\}.$$

Debemos demostrar entonces que además de ser un conjunto de transformaciones de \mathbb{R}^n cumple con las propiedades que lo hacen un grupo; que estamos usando los términos con propiedad:

i) Claramente $\text{id}_{\mathbb{R}^n} \in \mathbf{Sim}(F)$ pues $\text{id}(F) = F$. De tal manera que cualquier figura tiene al menos la simetría trivial; y si la figura F no tiene otra simetría diríamos coloquialmente que **no** tiene simetrías, que es asimétrica.

ii) Supongamos que $f \in \mathbf{Sim}(F)$; entonces al sustituir $F = f(F)$ obtenemos que

$$f^{-1}(F) = f^{-1}(f(F)) = (f^{-1} \circ f)(F) = \text{id}(F) = F,$$

donde estamos usando nuestra suposición de que $\mathbf{Iso}(n)$ es un grupo y entonces que $f^{-1} \in \mathbf{Iso}(n)$ ya existe.

iii) Si $f, g \in \mathbf{Sim}(F)$ entonces, como ninguna de las dos cambia la figura, se obtiene que

$$(f \circ g)(F) = f(g(F)) = f(F) = F.$$

Que dice lo obvio: si un movimiento deja una figura en su lugar le podemos aplicar otro que la deje en su lugar y sigue en su lugar.

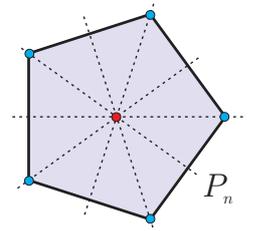
De tal manera que cualquier figura o cuerpo tiene asociado un grupo de isometrías; qué tan grande es ese grupo es nuestra medida intuitiva de “su simetría”. Por ejemplo, el cuerpo humano (abstracto, por supuesto) tiene como simetría no trivial una reflexión, que al ser su propia inversa forma, junto con la identidad, un grupo. Y ahora sí, podemos definir formalmente los grupos que esbozamos (para $n \geq 5$, y definimos para $n = 3, 4$) en la primera sección de este capítulo.

Grupos diédricos

Consideremos el polígono regular de n lados P_n , con $n \geq 2$, cuyos vértices son el conjunto

$$\{\mathbf{v}_{k,n} := (\cos(2\pi k/n), \sin(2\pi k/n)) \mid k = 0, 1, \dots, n-1\}$$

y que consiste en los segmentos entre puntos sucesivos $\overline{\mathbf{v}_{k,n}\mathbf{v}_{k+1,n}}$ llamados sus *caras, lados o aristas*; y si se quiere se puede pensar que también lo de adentro es parte de P_n . Definimos entonces el *grupo diédrico de orden n* como su grupo de simetrías:



$$\mathbf{D}_n := \mathbf{Sim}(P_n).$$

Es claro que \mathbf{D}_n contiene las n rotaciones en el origen $\rho_{2\pi k/n}$ que incluyen la identidad ($k = 0$ o $k = n$). Pero además tiene reflexiones cuyos espejos son las líneas por el origen y los vértices y las líneas por el origen y el punto medio de sus lados (que para n impar ya estaban contados). En cualquier caso, son exactamente n reflexiones pues en nuestro conteo hubo repeticiones. De tal manera que \mathbf{D}_n tiene $2n$ elementos.

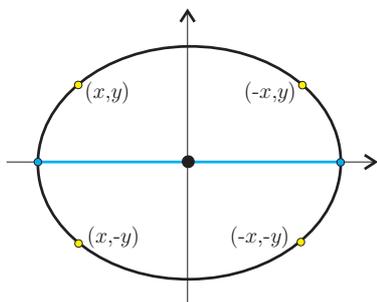
Esta clase de simetría es la que tienen los rosetones de las iglesias (por lo que también se les conoce como “grupos de rosetas o bien de rosetones”); muchas de las figuras de papel picado que se obtienen al doblar radialmente y luego cortar con tijeras; algunas plazas y edificios famosos y múltiples flores así como las estrellas de mar.

Más cercano al “corazoncito” de este libro, podemos considerar las cónicas. Sea \mathcal{E} la elipse dada por la ecuación canónica

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

con $a > b$ (aunque lo que importa es que $a \neq b$). Es claro por el dibujo que las simetrías de \mathcal{E} son las reflexiones en los ejes. Pero lo podemos demostrar formalmente pues, como se debió haber respondido al ejercicio correspondiente, la reflexión en el eje x está dada por la fórmula

$$\varphi_x(x, y) = (x, -y).$$



Y análogamente, la reflexión en el eje y es $\varphi_y(x, y) = (-x, y)$, de tal manera que su composición (en cualquier orden) es la rotación de π en el origen $((x, y) \mapsto (-x, -y))$. Si el punto (x, y) está en la elipse, satisface la ecuación y entonces los otros tres puntos que se obtienen al cambiar el signo a una o a las dos coordenadas también la satisfacen pues cambiar el signo no afecta los cuadrados. Esto implica que \mathcal{E} va en sí misma al reflejar en cualquiera de los ejes, que en adelante llamaremos sus *ejes de simetría*. Falta ver que no tiene ninguna otra simetría y esto se sigue de que su *diámetro* (el segmento máximo entre sus puntos, que mide $2a$) es único. Y puesto que \mathbf{D}_2 es, por definición, el grupo de simetrías de un segmento, entonces

$$\mathbf{Sim}(\mathcal{E}) = \mathbf{D}_2.$$

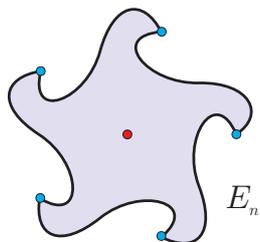
Análogamente, el grupo de simetrías de una hipérbola es el diédrico de orden 2. Las parábolas tienen menos simetrías, sólo una reflexión, que junto con la identidad forma un grupo que podemos llamar \mathbf{D}_1 para completar la definición de los diédricos (finitos). Nótese que este tipo de simetría (el de una sola reflexión, junto con la identidad) es muy común, el cuerpo humano (en abstracto, por supuesto), casi todos los animales, una silla, etcétera.

Grupos cíclicos

Son el subgrupo del diédrico correspondiente que consiste en tomar únicamente las rotaciones. Pero lo importante es que también son grupos de simetría de figuras.

Para obtener una, podemos reemplazar las aristas del polígono P_n por curvas que distingan entre salida y llegada, por ejemplo, un símbolo de “integral” para obtener a la “estrella ninja de n picos” E_n . Y entonces el *grupo cíclico de orden n* es

$$\mathbf{C}_n := \mathbf{Sim}(E_n).$$



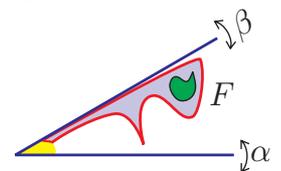
El grupo cíclico de orden n , \mathbf{C}_n , tiene n elementos. Se puede identificar con el grupo de residuos módulo n , denotado con \mathbb{Z}_n , donde la composición corresponde a la suma.

Parece que el gran Leonardo da Vinci notó que cualquier figura plana que se pueda dibujar en una hoja de papel (traducida a una figura $F \subset \mathbb{R}^2$ acotada) y que no tenga todas las simetrías de un círculo (que no sea un conjunto de anillos concéntricos) tiene como grupo de simetrías a alguno de estos dos tipos; es decir, o es diédrico o es cíclico. Una versión de este resultado la demostraremos hacia el final del capítulo, aunque desde ahora se puede entender su enunciado. Dice que si G es un grupo de isometrías que, como conjunto, es finito, entonces es cíclico o es diédrico; que con estos dos ejemplos ya acabamos.

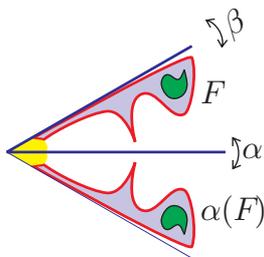
Sin embargo, hay otros grupos de simetría interesantes. Si pensamos en figuras no acotadas; por ejemplo, una cuadrícula o un mosaico (un cubrimiento del plano con copias de unas cuantas piezas), M digamos, entonces $\mathbf{Sim}(M)$ es lo que se llama un grupo cristalográfico. De estos solamente hay 17, pero no lo demostraremos en este libro.

Diédrico y cíclico infinitos

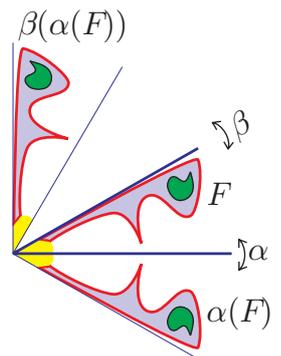
Para entender el porqué del nombre en sus versiones infinitas, vale la pena retomar la idea de generadores de un grupo usada en la primera sección. Para generar una figura con simetría diédrica (que quizá, como veremos, debía llamarse *caleidoscópica*) basta poner dos espejos en un ángulo π/n y colocarlos sobre cualquier dibujo; la porción de éste que queda entre ellos se “reproduce” caleidoscopicamente y forma alrededor de la arista donde se juntan los espejos una figura con simetría \mathbf{D}_n .



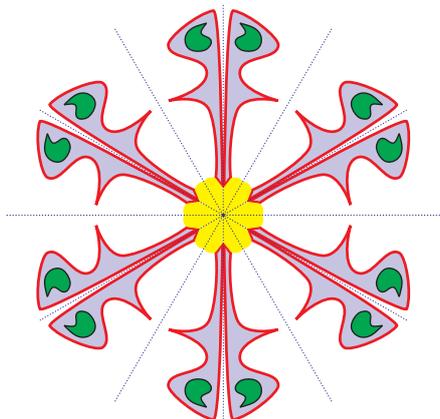
Consideremos esta situación en el plano del dibujo. Sean α y β las reflexiones en los dos espejos (que ahora son líneas en el plano y los espejos verdaderos son los planos que emanan perpendicularmente de ellas) y sea F la figura que queda en el ángulo relevante. La figura $\alpha(F)$ tiene la orientación contraria a F , es un reflejado de ella.



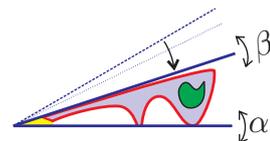
Al volverla a reflejar en β ($\alpha(F)$) esta nueva copia de F ya tiene la orientación original y se obtiene rotando a F el doble del ángulo entre los espejos, a saber, $2\pi/n$ alrededor de su punto de intersección. Pero esto no se cumple sólo para la figura F : es una propiedad de las



transformaciones, es decir, $\beta \circ \alpha$ es dicha rotación. La composición en el otro orden, $\alpha \circ \beta$, resulta ser la rotación inversa (de $-2\pi/n$ en el punto de intersección). Así, a cada copia o imagen de la figura F que se ve en el ángulo de espejos corresponde un elemento del grupo diédrico que es justo la transformación que lleva a F ahí y que se obtiene componiendo sucesivamente las dos reflexiones originales α y β . La figura con simetría diédrica o caleidoscópica que se genera es la unión de todas estas copias (justo $2n$) de lo que se podría llamar la *figura fundamental* que habíamos denotado F .

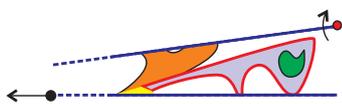


Hemos visto que, como en los casos $n = 3, 4$, los grupos diédricos están generados por dos reflexiones; a saber, D_n está generado por dos reflexiones cuyos espejos se encuentran en un ángulo π/n (en las figuras tomamos $n = 6$). ¿Qué pasa si hacemos crecer n ? Es claro que la figura fundamental F tenderá a hacerse más flaca y astillada, mientras que la figura caleidoscópica que generan (F y las reflexiones α y β), que podemos denotar $D_n(F)$, pues resulta natural definir

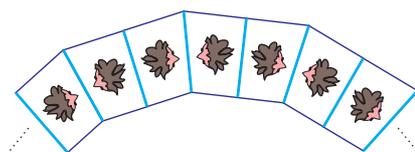


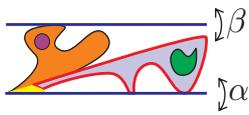
$$D_n(F) := \bigcup_{g \in D_n} g(F),$$

se hará más pesada y confusa, pues el ángulo π/n disminuye. Sin embargo, podemos evitar este tumulto si al mismo tiempo que n crece y π/n disminuye, alejamos el punto de intersección de los espejos; es decir, podemos girar uno de los espejos (el correspondiente a β , digamos) alrededor de algún punto fijo y esto produce que el punto de intersección se aleje.



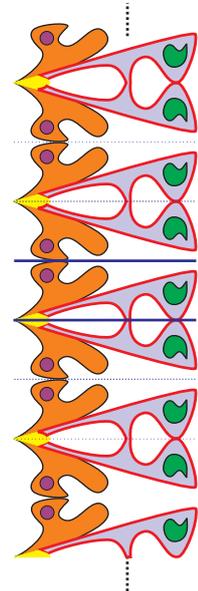
La figura $D_n(F)$ que se genera entonces es un anillo cuyo radio es grande: es como esos pasillos circulares que se generan en un baño o closet con espejos encontrados pero no perfectamente paralelos; aunque los espejos se acaben, los planos que definen se intersectan en una línea (imaginaria) que es el eje, el centro, del pasillo circular donde están nuestras imágenes viendo alternadamente en ambos sentidos y girando poco a poco.





Desde luego, este proceso tiene un limite: cuando n tiende a infinito el espejo de β llega sin problemas a una paralela al espejo de α ; y en ese momento generan el *grupo diédrico infinito* \mathbf{D}_∞ . Es el grupo que aparece en las peluquerías donde ponen espejos en paredes opuestas y, en teoría, paralelas. Cada elemento de este grupo es una composición (*palabra*) del estilo $\alpha\beta\alpha\beta \cdots \beta\alpha$, aunque puede empezar y terminar con cualesquiera de α o β . El subgrupo que consiste de palabras pares (incluyendo a la vacía que representa la identidad), que son composición de un número par de reflexiones, es el *cíclico infinito*, \mathbf{C}_∞ : consta de puras traslaciones y está generado por la más chica de ellas: $\alpha\beta$ o $\beta\alpha$, da lo mismo. De tal manera que naturalmente se identifica con los enteros (\mathbb{Z}) y la suma corresponde a la composición.

Pero el grupo diédrico infinito, como la figura con su simetría lo indica, realmente vive en la recta. Cualquier par de reflexiones distintas en $\mathbf{Iso}(1)$ lo generan. Y si lo definimos como subgrupo de isometrías de \mathbb{R}^2 , fue porque gráficamente es más fácil de ver y para justificar la naturalidad de su pomposo nombre.



Nótese, por último, que aunque al escoger distintas reflexiones para generar el diédrico infinito (o cualquiera de los finitos, para este caso) se pueden obtener subconjuntos (estrictamente) distintos de isometrías, pero la estructura algebraica de ellos (cómo se comportan bajo composición) es esencialmente la misma. De tal manera que los grupos diédricos y cíclicos que hemos definido deben pensarse como algo abstracto que tiene muchas instancias de realización geométrica. Así que, estrictamente hablando, cuando decimos que el grupo generado por tal par de reflexiones es el diédrico fulano, realmente deberíamos decir “es **un** diédrico de zutanito orden”.

EJERCICIO 3.33 Sean α y β las dos reflexiones de \mathbb{R} con espejos en el 0 y en el 1 respectivamente, y sea (para este y los siguientes ejercicios) $\mathbf{D}_\infty \subset \mathbf{Iso}(1)$ el grupo de transformaciones que generan α y β . Da las fórmulas explícitas de α , β , $\alpha \circ \beta$, $\beta \circ \alpha$ y $\alpha \circ \beta \circ \alpha$.

EJERCICIO 3.34 ¿Cuál es el grupo $\mathbf{D}_\infty \cap \mathbf{Tra}(1)$? Da la fórmula explícita de todos sus elementos y su correspondiente expresión como generado por α y β . (*El próximo ejercicio te puede ayudar.*)

EJERCICIO 3.35 Demuestra que las reflexiones en \mathbf{D}_∞ son precisamente las reflexiones con espejo entero. Es decir, cuyo espejo está en $\mathbb{Z} = \{\dots -2, -1, 0, 1, 2, \dots\} \subset \mathbb{R}$.

EJERCICIO 3.36 ¿Cuáles son los generadores del grupo diédrico infinito cuyo subgrupo de traslaciones corresponde precisamente a las traslaciones por números enteros \mathbb{Z} .

3.3.3 Transformaciones ortogonales

Hemos definido isometrías, visto los ejemplos básicos y desarrollado la importante noción de simetría —y no sólo importante en las matemáticas sino en la naturaleza donde parece ser muy abundante—, pero tenemos aún muy pocas herramientas técnicas para demostrar cosas tan elementales como que $\mathbf{Iso}(\mathbb{R}^n)$ es un grupo. Para esto serán fundamentales las transformaciones ortogonales. Pero su motivación más natural viene de recapitular cómo introdujimos la noción de distancia: como una consecuencia natural del producto interior que era más elemental y fácil de trabajar; y a éste lo hemos tenido olvidado. Si las isometrías se definieron como las funciones que preservan distancia, algo similar se debe poder hacer con el producto interior.

Definición 3.3.2 Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una *transformación ortogonal* si preserva el producto interior. Es decir, si para todo par $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ se cumple que

$$\mathbf{x} \cdot \mathbf{y} = f(\mathbf{x}) \cdot f(\mathbf{y}).$$

Se denota por $\mathbf{O}(n)$ el grupo de *transformaciones ortogonales* de \mathbb{R}^n .

El lector observador debió haber notado que nos estamos adelantando al usar los términos “transformación” y “grupo” en la definición anterior. Estrictamente hablando, deberíamos usar los términos más modestos de “función” y “conjunto”, pero como sé (autor) que sí va a tener sentido, vale la pena irse acostumbrando; además, en este caso me suenan horrible los términos modestos, no me hallo usándolos. Por lo pronto, para limpiar un poco la conciencia, dejamos al lector quisquilloso que cumpla los primeros trámites hacía la “liberación total” de los términos verdaderos.

EJERCICIO 3.37 Demuestra que la composición de funciones ortogonales es ortogonal.

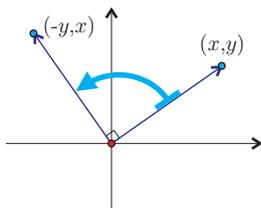
EJERCICIO 3.38 Demuestra que si una función ortogonal tiene inversa entonces ésta también es ortogonal.

EJERCICIO 3.39 Demuestra que $\mathbf{O}(1) = \{\text{id}_{\mathbb{R}}, -\text{id}_{\mathbb{R}}\}$, donde $(-\text{id}_{\mathbb{R}})(x) = -x$. Recuerda que el producto interior en \mathbb{R} es el simple producto. Observa que entonces las transformaciones ortogonales de \mathbb{R} son sus isometrías que dejan fijo el origen.

Antes de entrar a las demostraciones, veamos un ejemplo que nos ha sido muy útil, el compadre ortogonal en \mathbb{R}^2 . Hemos usado al compadre ortogonal de un vector

dado, pero, como lo indicamos en el momento de su presentación, si pensamos en todas las asignaciones al mismo tiempo obtenemos una función de \mathbb{R}^2 en \mathbb{R}^2 , que podemos designar como $\rho : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ y que está dada por la fórmula explícita

$$\rho(x, y) = (-y, x).$$



Para demostrar que el compadre ortogonal es una transformación ortogonal (valga y explíquese la redundancia), hay que escribir la fórmula para $\rho(\mathbf{x}) \cdot \rho(\mathbf{y})$ con (conviene cambiar a subíndices) $\mathbf{x} = (x_1, x_2)$ y $\mathbf{y} = (y_1, y_2)$ cualquier par de vectores:

$$\begin{aligned}\rho(\mathbf{x}) \cdot \rho(\mathbf{y}) &= (-x_2, x_1) \cdot (-y_2, y_1) \\ &= (-x_2)(-y_2) + x_1 y_1 \\ &= x_1 y_1 + x_2 y_2 = \mathbf{x} \cdot \mathbf{y}\end{aligned}$$

(de hecho esto fue el inciso (iii) del Lema 1.12 que presentaba sus monerías en sociedad). Si el compadre ortogonal que, como función, consiste en girar 90° alrededor del origen es ortogonal, es de esperarse que todas las rotaciones en el origen también lo sean, y así será. Pero antes, relacionemos las transformaciones ortogonales con las isometrías.

Puesto que la distancia se define en términos del producto interior, debe cumplirse que una función ortogonal también es isometría.

Lema 3.8 $O(n) \subset Iso(n)$.

Demostración. Sea $f \in O(n)$. Debemos demostrar que preserva distancias sabiendo que preserva producto interior. Pero esto es fácil pues la distancia se escribe en términos del producto interior. Dados \mathbf{x} y \mathbf{y} en \mathbb{R}^n , la definición de distancia es

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})}.$$

Al elevar al cuadrado y usar las propiedades del producto interior se tiene entonces

$$d(\mathbf{x}, \mathbf{y})^2 = \mathbf{x} \cdot \mathbf{x} + \mathbf{y} \cdot \mathbf{y} - 2\mathbf{x} \cdot \mathbf{y}. \quad (3.4)$$

Como esto se cumple para cualquier par de puntos, en particular para $f(\mathbf{x})$ y $f(\mathbf{y})$, tenemos

$$\begin{aligned}d(f(\mathbf{x}), f(\mathbf{y}))^2 &= f(\mathbf{x}) \cdot f(\mathbf{x}) + f(\mathbf{y}) \cdot f(\mathbf{y}) - 2f(\mathbf{x}) \cdot f(\mathbf{y}) \\ &= \mathbf{x} \cdot \mathbf{x} + \mathbf{y} \cdot \mathbf{y} - 2\mathbf{x} \cdot \mathbf{y} \\ &= d(\mathbf{x}, \mathbf{y})^2,\end{aligned}$$

donde usamos que $f \in O(n)$ en la segunda igualdad. Puesto que las distancias son positivas, de aquí se sigue que

$$d(f(\mathbf{x}), f(\mathbf{y})) = d(\mathbf{x}, \mathbf{y}),$$

y por tanto que $f \in Iso(n)$. □

EJERCICIO 3.40 Demuestra que las funciones ortogonales son inyectivas.

Ahora veremos que lo único que le falta a una isometría para ser ortogonal es respetar el origen. Para que tenga sentido, hay que remarcar que una función ortogonal $f \in \mathbf{O}(n)$ preserva el origen pues éste es el único vector cuyo producto interior consigo mismo se anula, es decir, $f(\mathbf{0}) \cdot f(\mathbf{0}) = \mathbf{0} \cdot \mathbf{0} = 0 \Rightarrow f(\mathbf{0}) = \mathbf{0}$.

Lema 3.9 *Sea $f \in \mathbf{Iso}(n)$ tal que $f(\mathbf{0}) = \mathbf{0}$, entonces $f \in \mathbf{O}(n)$.*

Demostración. Con la misma idea del lema anterior, hay que escribir ahora el producto interior en términos de distancias. De la ecuación (3.4), se sigue

$$\mathbf{x} \cdot \mathbf{y} = \frac{1}{2}(\mathbf{x} \cdot \mathbf{x} + \mathbf{y} \cdot \mathbf{y} - d(\mathbf{x}, \mathbf{y})^2).$$

Y como $\mathbf{x} \cdot \mathbf{x} = d(\mathbf{x}, \mathbf{0})^2$, tenemos entonces

$$\mathbf{x} \cdot \mathbf{y} = \frac{1}{2}(d(\mathbf{x}, \mathbf{0})^2 + d(\mathbf{y}, \mathbf{0})^2 - d(\mathbf{x}, \mathbf{y})^2)$$

y ya estamos armados para enfrentar el lema.

Sea $f \in \mathbf{Iso}(n)$ (preserva distancias) tal que $f(\mathbf{0}) = \mathbf{0}$. Por lo anterior, se obtiene

$$\begin{aligned} f(\mathbf{x}) \cdot f(\mathbf{y}) &= \frac{1}{2}(d(f(\mathbf{x}), \mathbf{0})^2 + d(f(\mathbf{y}), \mathbf{0})^2 - d(f(\mathbf{x}), f(\mathbf{y}))^2) \\ &= \frac{1}{2}(d(f(\mathbf{x}), f(\mathbf{0}))^2 + d(f(\mathbf{y}), f(\mathbf{0}))^2 - d(f(\mathbf{x}), f(\mathbf{y}))^2) \\ &= \frac{1}{2}(d(\mathbf{x}, \mathbf{0})^2 + d(\mathbf{y}, \mathbf{0})^2 - d(\mathbf{x}, \mathbf{y})^2) \\ &= \mathbf{x} \cdot \mathbf{y}. \end{aligned}$$

□

Estamos ya en posición de expresar todas las isometrías en términos de las ortogonales y las traslaciones.

Proposición 3.10 *Sea $f \in \mathbf{Iso}(n)$. Entonces existen $g \in \mathbf{O}(n)$ y $\mathbf{b} \in \mathbb{R}^n$ tales que para toda $\mathbf{x} \in \mathbb{R}^n$*

$$f(\mathbf{x}) = g(\mathbf{x}) + \mathbf{b},$$

y además son únicas.

Demostración. Veamos primero la unicidad, es decir, que si suponemos que existen $g \in \mathbf{O}(n)$ y $\mathbf{b} \in \mathbb{R}^n$ tales que f se escribe $f(\mathbf{x}) = g(\mathbf{x}) + \mathbf{b}$, entonces g y \mathbf{b} están forzadas. Puesto que las transformaciones ortogonales dejan fijo el origen, obtenemos que

$$f(\mathbf{0}) = g(\mathbf{0}) + \mathbf{b} = \mathbf{0} + \mathbf{b} = \mathbf{b},$$

Dibujo

y entonces la constante \mathbf{b} está forzada a ser $\mathbf{b} := f(\mathbf{0})$. Pero entonces \mathbf{g} ya queda definida por la ecuación

$$\mathbf{g}(\mathbf{x}) = f(\mathbf{x}) - \mathbf{b}$$

para toda $\mathbf{x} \in \mathbb{R}^n$. Nos falta entonces demostrar que con esta última definición de la función \mathbf{g} y el vector constante $\mathbf{b} \in \mathbb{R}^n$ se cumple que \mathbf{g} es ortogonal.

Observéese que otra manera de definir a \mathbf{g} sería $\mathbf{g} = \tau_{-\mathbf{b}} \circ f$, donde hay que recordar que $\tau_{-\mathbf{b}}$ es la traslación por $-\mathbf{b}$. Por el Lema 3.6 (la composición de isometrías es isometría) y puesto que las traslaciones son isometrías, se obtiene que $\mathbf{g} \in \mathbf{Iso}(n)$. Pero además, como

$$\mathbf{g}(\mathbf{0}) = f(\mathbf{0}) - \mathbf{b} = \mathbf{b} - \mathbf{b} = \mathbf{0},$$

entonces $\mathbf{g} \in \mathbf{O}(n)$ por el lema anterior y la existencia queda demostrada. \square

Tenemos entonces que las isometrías se obtienen de las transformaciones ortogonales al permitir que interactúen, que se involucren, que se compongan, con las traslaciones (que **no** son ortogonales, de hecho sólo se intersectan en la identidad, $\mathbf{O}(n) \cap \mathbf{Tra}(n) = \{\text{id}\}$, pues la única traslación que deja fijo el origen es por $\mathbf{0}$). Corresponde esto al hecho de que la distancia se obtiene del producto interior trasladando un punto al origen. Y por tanto, para acabar de entender $\mathbf{Iso}(n)$ debemos estudiar $\mathbf{O}(n)$. Como en un ejercicio ya establecimos el caso de $\mathbf{O}(1) \simeq \{1, -1\}$, concentrémonos en $\mathbf{O}(2)$.

Sea $f \in \mathbf{O}(2)$. Como f preserva el producto interior, también preserva la norma y en particular manda al círculo unitario \mathbb{S}^1 en sí mismo:

$$\mathbf{x} \in \mathbb{S}^1 \Rightarrow \mathbf{x} \cdot \mathbf{x} = 1 \Rightarrow f(\mathbf{x}) \cdot f(\mathbf{x}) = 1 \Rightarrow f(\mathbf{x}) \in \mathbb{S}^1.$$

Pero también preserva ortogonalidad (definida con base en el producto interior). Entonces debe mandar bases ortonormales en bases ortonormales. Veámoslo con el ejemplo más simple.

Dibujo

Sean $\mathbf{e}_1, \mathbf{e}_2$ la base canónica de \mathbb{R}^2 ; es decir $\mathbf{e}_1 = (1, 0)$ y $\mathbf{e}_2 = (0, 1)$. Y sean

$$\begin{aligned} \mathbf{u} &:= f(\mathbf{e}_1) \\ \mathbf{v} &:= f(\mathbf{e}_2). \end{aligned}$$

Afirmamos que \mathbf{u} y \mathbf{v} son una base ortonormal, pues de que $f \in \mathbf{O}(2)$ (preserva producto interior) se obtiene

$$\begin{aligned} \mathbf{u} \cdot \mathbf{u} &= \mathbf{e}_1 \cdot \mathbf{e}_1 = 1 \\ \mathbf{v} \cdot \mathbf{v} &= \mathbf{e}_2 \cdot \mathbf{e}_2 = 1 \\ \mathbf{u} \cdot \mathbf{v} &= \mathbf{e}_1 \cdot \mathbf{e}_2 = 0. \end{aligned}$$

Y entonces podemos encontrar el valor de $f(\mathbf{x})$ para cualquier \mathbf{x} usando el teorema de las bases ortonormales (Teorema 1.23) sobre $f(\mathbf{x})$:

$$\begin{aligned} f(\mathbf{x}) &= (f(\mathbf{x}) \cdot \mathbf{u})\mathbf{u} + (f(\mathbf{x}) \cdot \mathbf{v})\mathbf{v} \\ &= (f(\mathbf{x}) \cdot f(\mathbf{e}_1))\mathbf{u} + (f(\mathbf{x}) \cdot f(\mathbf{e}_2))\mathbf{v} \\ &= (\mathbf{x} \cdot \mathbf{e}_1)\mathbf{u} + (\mathbf{x} \cdot \mathbf{e}_2)\mathbf{v}, \end{aligned}$$

Dibujo

que se puede escribir, tomando como de costumbre $\mathbf{x} = (x, y)$, como

$$f(x, y) = x\mathbf{u} + y\mathbf{v}. \quad (3.5)$$

Hemos llegado entonces a una fórmula explícita para las funciones ortogonales (de \mathbb{R}^2) que depende únicamente de sus valores en la base canónica. Por ejemplo, si f fuera la función “compadre ortogonal”, entonces $\mathbf{u} = f(\mathbf{e}_1) = \mathbf{e}_2$ y $\mathbf{v} = f(\mathbf{e}_2) = -\mathbf{e}_1$ y por lo tanto $f(x, y) = x\mathbf{e}_2 - y\mathbf{e}_1 = (-y, x)$ como ya sabíamos. Podemos resumir con la:

Proposición 3.11 *Si $f \in \mathbf{O}(2)$, existe una base ortonormal \mathbf{u}, \mathbf{v} de \mathbb{R}^2 para la cual f se escribe*

$$f(x, y) = x\mathbf{u} + y\mathbf{v}.$$

De aquí podemos sacar por el momento dos resultados: el primero, que estas funciones ortogonales son suprayectivas (y por lo tanto que ya les podemos decir “transformaciones” y a $\mathbf{O}(2)$ “grupo” con todas las de la ley); y el segundo, que es una descripción geométrica de todas ellas.

Dibujo

Primero, con la notación de arriba, dado cualquier $\mathbf{y} \in \mathbb{R}^2$, es fácil encontrar \mathbf{x} tal que $f(\mathbf{x}) = \mathbf{y}$. Sea $\mathbf{x} := (\mathbf{y} \cdot \mathbf{u})\mathbf{e}_1 + (\mathbf{y} \cdot \mathbf{v})\mathbf{e}_2$, entonces

$$\begin{aligned} f(\mathbf{x}) &= (\mathbf{x} \cdot \mathbf{e}_1)\mathbf{u} + (\mathbf{x} \cdot \mathbf{e}_2)\mathbf{v} \\ &= (\mathbf{y} \cdot \mathbf{u})\mathbf{u} + (\mathbf{y} \cdot \mathbf{v})\mathbf{v} = \mathbf{y} \end{aligned}$$

por el Teorema 1.23, y esto demuestra que f es sobre.

Segundo, cuántas transformaciones (¡ahora sí! sin pruritos) ortogonales de \mathbb{R}^2 hay es equivalente a cuántas bases ortonormales podemos escoger. El primer vector \mathbf{u} (la imagen de \mathbf{e}_1 , insistamos) puede ser cualquiera en el círculo unitario \mathbb{S}^1 ; pero una vez escogido ya nada más nos quedan dos posibilidades para escoger el segundo, pues éste debe ser perpendicular y unitario. Puede ser el compadre ortogonal ($\mathbf{v} = \mathbf{u}^\perp$) o bien su negativo ($\mathbf{v} = -\mathbf{u}^\perp$). En el primer caso, la transformación resulta ser una rotación que manda a la base canónica $\mathbf{e}_1, \mathbf{e}_2$ en la base rotada $\mathbf{u}, \mathbf{u}^\perp$. Y en el segundo caso veremos que es una reflexión. Podríamos decir entonces que $\mathbf{O}(2)$ “consiste en dos copias” de \mathbb{S}^1 parametrizadas por \mathbf{u} : una corresponde a las rotaciones (cuando se escoge $\mathbf{v} = \mathbf{u}^\perp$) y la otra a las reflexiones.

Pero para que todo esto sea un hecho nos falta demostrar que sin importar que base ortonormal tomemos, siempre se obtiene una transformación ortogonal.

Proposición 3.12 Sea \mathbf{u}, \mathbf{v} una base ortonormal de \mathbb{R}^2 . Entonces la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, definida por

$$f(\mathbf{x}, \mathbf{y}) = x\mathbf{u} + y\mathbf{v},$$

es una transformación ortogonal.

Demostración. Sean $\mathbf{x} = (x_1, x_2)$ y $\mathbf{y} = (y_1, y_2)$ cualquier par de puntos en \mathbb{R}^2 . Entonces

$$\begin{aligned} f(\mathbf{x}) \cdot f(\mathbf{y}) &= (x_1\mathbf{u} + x_2\mathbf{v}) \cdot (y_1\mathbf{u} + y_2\mathbf{v}) \\ &= (x_1y_1)\mathbf{u} \cdot \mathbf{u} + (x_2y_2)\mathbf{v} \cdot \mathbf{v} + (x_1y_2 + x_2y_1)\mathbf{u} \cdot \mathbf{v} \\ &= x_1y_1 + x_2y_2 = \mathbf{x} \cdot \mathbf{y}. \end{aligned}$$

□

EJERCICIO 3.41 Demuestra que las isometrías de \mathbb{R}^2 son suprayectivas, para concluir con la demostración de que $\mathbf{Iso}(2)$ es un grupo de transformaciones.

EJERCICIO 3.42 Demuestra que las funciones ortogonales de \mathbb{R}^3 son suprayectivas y concluye las demostraciones de que $\mathbf{O}(3)$ e $\mathbf{Iso}(3)$ son grupos de transformaciones.

EJERCICIO 3.43 Sea $f \in \mathbf{O}(2)$ con base ortonormal asociada \mathbf{u}, \mathbf{v} (es decir, con $\mathbf{u} = f(\mathbf{e}_1)$ y $\mathbf{v} = f(\mathbf{e}_2)$). Con nuestra demostración de la suprayectividad de f se puede encontrar la base ortonormal asociada a la inversa de f , llamémosla \mathbf{u}', \mathbf{v}' . Da sus coordenadas explícitas tomando $\mathbf{u} = (u_1, u_2)$ y $\mathbf{v} = (v_1, v_2)$.

EJERCICIO 3.44 Sean $f \in \mathbf{O}(2)$ y ℓ una recta dada por la ecuación $\mathbf{n} \cdot \mathbf{x} = c$. Demuestra que $f(\ell)$ es la recta dada por la ecuación $f(\mathbf{n}) \cdot \mathbf{x} = c$ (ojo: tienes que demostrar dos contenciones). Concluye que f manda al haz paralelo con dirección \mathbf{d} en el haz paralelo con dirección $f(\mathbf{d})$.

Un ejemplo

Para fijar ideas, concluyamos esta sección con un ejemplo que nos ayude a comprender el poder de los resultados obtenidos.

Dibujo

Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la rotación de 45° (o $\pi/4$) alrededor del punto $\mathbf{p} = (3, 2)$; el problema es escribir una fórmula explícita para f . Es decir, supongamos que queremos dar la función a una computadora para que haga algo con ella, que sepa aplicarla a puntos concretos; no le podemos dar la descripción que acabamos de hacer (y que cualquier ser humano entiende) para definir f , sino que tiene que ser mucho más concreta: una fórmula analítica.

Por la Proposición 3.10, sabemos que f se escribe como

$$f(\mathbf{x}) = g(\mathbf{x}) + \mathbf{b},$$

para alguna $g \in \mathbf{O}(2)$ y alguna $\mathbf{b} \in \mathbb{R}^2$. El problema se divide entonces en dos: cómo se escribe g y quién es \mathbf{b} . Primero, g debe ser la rotación de $\pi/4$ en el origen pues f cambia las líneas horizontales en líneas a 45° y en la expresión anterior sólo g puede hacer esta gracia (pues una vez que apliquemos g la traslación mantendrá la orientación de las líneas). Por la Proposición 3.11, para expresar g basta ver que le hace a la base canónica \mathbf{e}_1 , \mathbf{e}_2 y sabemos que tiene que ir a \mathbf{u} , \mathbf{u}^\perp , donde \mathbf{u} es el vector unitario a 45° . Es decir, sea

$$\mathbf{u} = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right),$$

y entonces por (3.5) g se escribe

$$\begin{aligned} g(x, y) &= x\mathbf{u} + y\mathbf{u}^\perp \\ &= x\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) + y\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) \\ &= \frac{\sqrt{2}}{2}(x - y, x + y) = \frac{1}{\sqrt{2}}(x - y, x + y), \end{aligned}$$

y esto sí que lo entiende una computadora. Para convencernos de que vamos por buen camino basta checar que efectivamente esta fórmula da $g(\mathbf{e}_1) = \mathbf{u}$ y $g(\mathbf{e}_2) = \mathbf{u}^\perp$; e inclusive que $g(\mathbf{u}) = \mathbf{e}_2$ como dice la geometría.

El segundo problema es encontrar \mathbf{b} . Sabemos por la demostración de la Proposición 3.10 que $\mathbf{b} = f(\mathbf{0})$. Y entonces bastará encontrar este valor particular. Esto se puede razonar de varias maneras, pero lo haremos de manera muy general. Recordemos que definimos rotar en el punto \mathbf{c} como trasladar \mathbf{c} al origen, rotar ahí y luego regresar \mathbf{c} a su lugar. En términos de funciones esto se escribe

$$f = \tau_{\mathbf{c}} \circ g \circ \tau_{-\mathbf{c}},$$

donde $\tau_{\mathbf{x}}$ es la traslación por \mathbf{x} . De aquí se podría sacar directamente la fórmula para f pues ya sabemos expresar g y las traslaciones, y falta nada más desarrollar las expresiones; o bien, podemos sólo aplicar esto al $\mathbf{0}$:

$$\begin{aligned} \mathbf{b} &= f(\mathbf{0}) = (\tau_{\mathbf{c}} \circ g \circ \tau_{-\mathbf{c}})(\mathbf{0}) \\ &= (\tau_{\mathbf{c}} \circ g)(\mathbf{0} - \mathbf{c}) = \tau_{\mathbf{c}}(g(-\mathbf{c})) = g(-\mathbf{c}) + \mathbf{c} \\ &= g(-3, -2) + (3, 2) = \frac{1}{\sqrt{2}}(-1, -5) + (3, 2) \\ &= \left(3 - \frac{1}{\sqrt{2}}, 2 - \frac{5}{\sqrt{2}}\right). \end{aligned}$$

De ahí concluimos que la expresión analítica de f es

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= g(\mathbf{x}, \mathbf{y}) + f(\mathbf{0}) \\ &= \frac{1}{\sqrt{2}}(\mathbf{x} - \mathbf{y}, \mathbf{x} + \mathbf{y}) + \left(3 - \frac{1}{\sqrt{2}}, 2 - 5\frac{1}{\sqrt{2}}\right) \\ &= \frac{1}{\sqrt{2}}\left(\mathbf{x} - \mathbf{y} - 1 + 3\sqrt{2}, \mathbf{x} + \mathbf{y} - 5 + 2\sqrt{2}\right). \end{aligned}$$

EJERCICIO 3.45 Encuentra la expresión analítica de las siguientes isometrías:

- La reflexión en la recta $\ell : \mathbf{x} + \mathbf{y} = 1$.
- La rotación de $\pi/2$ en el punto $(-2, 3)$.
- La rotación de π en el punto $(4, 3)$.
- La reflexión en la recta $\ell : \mathbf{y} = 1$ seguida de la traslación por $(2, 0)$.

EJERCICIO 3.46 Describe geoméricamente (con palabras) las siguientes isometrías:

- $f(\mathbf{x}, \mathbf{y}) = (\mathbf{x} + 1, -\mathbf{y})$.
- $f(\mathbf{x}, \mathbf{y}) = (-\mathbf{x} + 2, -\mathbf{y})$.
- $f(\mathbf{x}, \mathbf{y}) = (-\mathbf{y}, -\mathbf{x} + 2)$.

3.4 Las funciones lineales

Hemos visto que las transformaciones ortogonales de \mathbb{R}^2 quedan determinadas por lo que le hacen a la base canónica; es decir, que se escriben

$$f(\mathbf{x}, \mathbf{y}) = \mathbf{x} \mathbf{u} + \mathbf{y} \mathbf{v},$$

donde \mathbf{u} y \mathbf{v} son vectores fijos ($\mathbf{u} = f(\mathbf{e}_1)$ y $\mathbf{v} = f(\mathbf{e}_2)$, bien por definición o bien por la fórmula); y además, vimos que si la función es ortogonal entonces \mathbf{u} y \mathbf{v} forman una base ortonormal. Pero si en esta fórmula no pedimos nada a \mathbf{u} y a \mathbf{v} , simplemente que sean vectores cualesquiera en \mathbb{R}^2 , obtenemos una familia de funciones mucho más grande: las *lineales* de \mathbb{R}^2 en \mathbb{R}^2 (ver el siguiente teorema). Estas funciones son la base del álgebra lineal y serán el fundamento para las transformaciones que más nos interesan, así que vale la pena definir las y estudiarlas un rato en general.

Definición 3.4.1 Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es *lineal* si para todos los vectores \mathbf{x} , $\mathbf{y} \in \mathbb{R}^n$ y todo número $t \in \mathbb{R}$ se cumple que:

- $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$
- $f(t\mathbf{x}) = t f(\mathbf{x})$.

Estas funciones, al preservar la suma y la multiplicación por escalares, conservan las operaciones básicas de un espacio vectorial y por eso son tan importantes. Pero en el caso que nos ocupa, tenemos el siguiente teorema.

Teorema 3.13 Una función $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ es lineal, si y sólo si se escribe

$$f(\mathbf{x}, \mathbf{y}) = x \mathbf{u} + y \mathbf{v},$$

donde \mathbf{u} y \mathbf{v} son vectores fijos en \mathbb{R}^2 .

Demostración. Supongamos que $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ es lineal y sean $\mathbf{u} := f(\mathbf{e}_1)$ y $\mathbf{v} := f(\mathbf{e}_2)$, donde, recuérdese, $\mathbf{e}_1 = (1, 0)$ y $\mathbf{e}_2 = (0, 1)$ son la base canónica de \mathbb{R}^2 . Dado cualquier $\mathbf{x} \in \mathbb{R}^2$, tenemos que si $\mathbf{x} = (x, y)$ entonces $\mathbf{x} = x \mathbf{e}_1 + y \mathbf{e}_2$, de donde, usando las propiedades (i) y (ii) respectivamente, se obtiene

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= f(\mathbf{x}) = f(x \mathbf{e}_1 + y \mathbf{e}_2) \\ &= f(x \mathbf{e}_1) + f(y \mathbf{e}_2) \\ &= x f(\mathbf{e}_1) + y f(\mathbf{e}_2) \\ &= x \mathbf{u} + y \mathbf{v}. \end{aligned}$$

Lo cual demuestra que cualquier función lineal se expresa en la forma deseada.

Por el otro lado, veamos que la función $f(\mathbf{x}, \mathbf{y}) = x \mathbf{u} + y \mathbf{v}$, con \mathbf{u} y \mathbf{v} arbitrarios, es lineal. Cambiando la notación de las coordenadas, sean $\mathbf{x} = (x_1, x_2)$ y $\mathbf{y} = (y_1, y_2)$. De la definición de f y las propiedades elementales de la suma vectorial tenemos

$$\begin{aligned} f(\mathbf{x} + \mathbf{y}) &= f(x_1 + y_1, x_2 + y_2) \\ &= (x_1 + y_1) \mathbf{u} + (x_2 + y_2) \mathbf{v} \\ &= (x_1 \mathbf{u} + x_2 \mathbf{v}) + (y_1 \mathbf{u} + y_2 \mathbf{v}) \\ &= f(\mathbf{x}) + f(\mathbf{y}), \end{aligned}$$

lo cual demuestra que f cumple (i). Análogamente, f cumple (ii), pues para cualquier $t \in \mathbb{R}$

$$\begin{aligned} f(t \mathbf{x}) &= (tx_1) \mathbf{u} + (tx_2) \mathbf{v} \\ &= t(x_1 \mathbf{u} + x_2 \mathbf{v}) = t(f(\mathbf{x})). \end{aligned}$$

Lo cual concluye la demostración del teorema. □

Tenemos entonces que las transformaciones ortogonales son funciones lineales, así que ya hemos visto varios ejemplos. Pero son muchas más las lineales de \mathbb{R}^2 en \mathbb{R}^2 que las ortogonales (véase el Ejercicio 3.50). Si tomamos la *cuadrícula canónica* con verticales y horizontales en los enteros de los ejes, una transformación lineal la manda en la “rombícula” que generan los vectores imágenes de la base canónica y que consiste en tomar las paralelas a uno de ellos que pasan por los múltiplos enteros del otro. Sólo cuando esta rombícula se sigue viendo como cuadrícula, y del mismo tamaño, tenemos una transformación ortogonal. Pero además, nótese que hemos definido funciones lineales entre diferentes espacios vectoriales y que en la demostración del

teorema no tuvimos que explicitar en dónde viven \mathbf{u} y \mathbf{v} . Si vivieran en \mathbb{R}^3 , por ejemplo, la imagen sería la misma, un plano rombiculturalo, que tiene la libertad extra de bambolearse.

EJERCICIO 3.47 Demuestra que una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ lineal preserva el origen, es decir, que cumple $f(\mathbf{0}) = \mathbf{0}$.

EJERCICIO 3.48 Demuestra que una función lineal manda líneas en líneas (de ahí su nombre) o a veces en puntos.

EJERCICIO 3.49 Demuestra que la composición de funciones lineales es lineal.

EJERCICIO 3.50 Demuestra que la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, definida por $f(x, y) = x\mathbf{u} + y\mathbf{v}$ es ortogonal si y sólo si \mathbf{u}, \mathbf{v} forman una base ortonormal.

EJERCICIO 3.51 Demuestra que la función $f : \mathbb{R} \rightarrow \mathbb{R}$ definida como $f(x) = 3x$ es una función lineal.

EJERCICIO 3.52 Demuestra que una función $f : \mathbb{R} \rightarrow \mathbb{R}$ es lineal si y sólo si es la función $f(x) = ax$ para alguna $a \in \mathbb{R}$.

EJERCICIO 3.53 Demuestra que la función $f : \mathbb{R} \rightarrow \mathbb{R}^2$ definida como $f(x) = (3x, 5x)$ es una función lineal.

EJERCICIO 3.54 Demuestra que una función $f : \mathbb{R} \rightarrow \mathbb{R}^2$ es lineal si y sólo si se escribe $f(x) = (ax, bx)$, para algunos $a, b \in \mathbb{R}$.

EJERCICIO 3.55 Demuestra que la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida como $f(x, y) = 2x + 4y$ es una función lineal.

EJERCICIO 3.56 Demuestra que la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida como $f(x, y) = (x + y, y - x)$ es una función lineal.

EJERCICIO 3.57 Exhibe una función lineal $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ que mande el eje x a la recta $\ell : 2x - y = 0$.

EJERCICIO 3.58 Demuestra que una función $f : \mathbb{R}^3 \rightarrow \mathbb{R}^n$ es lineal, si y sólo si se escribe

$$f(x, y, z) = x\mathbf{u} + y\mathbf{v} + z\mathbf{w},$$

donde $\mathbf{u} = f(\mathbf{e}_1)$, $\mathbf{v} = f(\mathbf{e}_2)$ y $\mathbf{w} = f(\mathbf{e}_3)$.

3.4.1 Extensión lineal

De los ejercicios anteriores, en particular el último, debe quedar claro que el Teorema 3.13 es sólo un caso particular de un resultado general que haremos explícito en esta sección. Aunque estemos especialmente interesados en las dimensiones 1, 2, 3, para ocuparnos de todas ellas a la vez, y evitarnos el molesto trabajo de 6, seis, casos particulares, es necesario trabajar con los abstractos \mathbb{R}^n y \mathbb{R}^m .

Como definimos en el capítulo 1, dados los vectores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ en \mathbb{R}^n , una

combinación lineal de ellos es el vector

$$\sum_{i=1}^k \lambda_i \mathbf{v}_i \in \mathbb{R}^n,$$

donde $\lambda_i \in \mathbb{R}$, $i = 1, \dots, k$, son escalares cualesquiera llamados los *coeficientes* de la combinación lineal. Por ejemplo, todas las combinaciones lineales de un solo vector ($k = 1$) forman la recta que genera (si es no nulo) y las de dos vectores forman un plano (si no son paralelos). En una combinación lineal se están utilizando simultáneamente las dos operaciones básicas de un espacio vectorial. Ahora bien, si tenemos una función lineal $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, se cumple que *preserva combinaciones lineales* pues

$$f\left(\sum_{i=1}^k \lambda_i \mathbf{v}_i\right) = \sum_{i=1}^k f(\lambda_i \mathbf{v}_i) = \sum_{i=1}^k \lambda_i f(\mathbf{v}_i),$$

donde se usa la propiedad **(i)** de las funciones lineales en la primera igualdad y la **(ii)** en la segunda. Así que podemos redefinir función lineal como aquella que cumple

$$f\left(\sum_{i=1}^k \lambda_i \mathbf{v}_i\right) = \sum_{i=1}^k \lambda_i f(\mathbf{v}_i), \quad (3.6)$$

para cualesquiera $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in \mathbb{R}^n$ y $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R}$; pues esta condición claramente incluye la **(i)** y la **(ii)** de la definición original. Hay que hacer énfasis en que en esta igualdad la combinación lineal de la derecha se da en \mathbb{R}^m (el codominio), mientras que la que tenemos en el lado izquierdo ocurre en \mathbb{R}^n (el dominio), antes de aplicarle la función, que no necesariamente son el mismo espacio.

Por otro lado, \mathbb{R}^n tiene una *base canónica*: los vectores $\mathbf{e}_1 = (1, 0, \dots, 0)$, $\mathbf{e}_2 = (0, 1, \dots, 0)$, ..., $\mathbf{e}_n = (0, 0, \dots, 1)$; es decir, para $i = 1, \dots, n$ el vector $\mathbf{e}_i \in \mathbb{R}^n$ tiene i -ésima coordenada 1 y el resto 0. Estos se llaman canónicos pues la combinación lineal de ellos que nos da cualquier otro vector $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ se obtiene trivialmente (“canónicamente”) de sus coordenadas:

$$(x_1, x_2, \dots, x_n) = \sum_{i=1}^n x_i \mathbf{e}_i.$$

De tal manera que para cualquier función lineal f que sale de \mathbb{R}^n se tiene, usando (3.6) y esta última expresión, que

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n x_i f(\mathbf{e}_i).$$

Nos falta ver que los valores que puede tomar f en la base canónica son arbitrarios. Es decir, si tomamos n vectores $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n \in \mathbb{R}^m$ arbitrariamente, y queremos

encontrar una función lineal $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ que cumpla $f(\mathbf{e}_i) = \mathbf{u}_i$ para $i = 1, 2, \dots, n$, entonces debemos definirla en todo \mathbb{R}^n como

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n x_i \mathbf{u}_i.$$

A esta función se le llama la *extensión lineal* a \mathbb{R}^n de lo que determinamos para la base canónica. Pero nos falta demostrar que efectivamente es lineal. Sin embargo, esto es muy fácil de hacer usando la definición original de función lineal y las de suma y multiplicación por escalares; hay que copiar la del Teorema 3.13 con más coordenadas, y dejamos este trabajo al lector. Podemos resumir entonces con el siguiente teorema.

Teorema 3.14 Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es lineal si y sólo si se escribe

$$f(x_1, x_2, \dots, x_n) = x_1 \mathbf{u}_1 + x_2 \mathbf{u}_2 + \dots + x_n \mathbf{u}_n = \sum_{i=1}^n x_i \mathbf{u}_i,$$

donde $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n \in \mathbb{R}^m$. Nótese que $\mathbf{u}_i = f(\mathbf{e}_i)$. □

Corolario 3.15 Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ son dos funciones lineales tales que $f(\mathbf{e}_i) = g(\mathbf{e}_i)$ para $i = 1, 2, \dots, n$, entonces $f(\mathbf{x}) = g(\mathbf{x})$ para toda $\mathbf{x} \in \mathbb{R}^n$.

Demostración. Por el teorema, tienen la misma expresión. □

El teorema, o bien su corolario, pueden llamarse “Teorema de extensión única de funciones lineales”. La moraleja es que una función lineal depende únicamente de unos cuantos valores: los que le asigna a la base canónica, y éstos, una vez fijado el codominio, son arbitrarios.

EJERCICIO 3.59 Describe el lugar geométrico definido como $\mathcal{L} = \{f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}\}$, donde $f : \mathbb{R} \rightarrow \mathbb{R}^3$ es una función lineal.

EJERCICIO 3.60 Describe el lugar geométrico definido como $\Pi = \{f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^2\}$, donde $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ es una función lineal.

EJERCICIO 3.61 ¿Qué ejercicios del bloque anterior se siguen inmediatamente del Teorema 3.14? Reesríbelos.

3.4.2 La estructura de las funciones lineales

En el capítulo 1 dejamos en suspenso un ejercicio donde se pregunta al lector si conoce algún otro espacio vectorial que no sea \mathbb{R}^n . Sería desleal dejar pasar ese ejemplo cuando lo tenemos en las narices. Efectivamente, cuando fijamos el dominio y el contradominio, las funciones lineales entre ellos tienen naturalmente la estructura de un espacio vectorial.

Fijemos la notación. Sea $\mathcal{L}(\mathbf{n}, \mathbf{m})$ el conjunto de todas las funciones lineales de \mathbb{R}^n a \mathbb{R}^m , es decir

$$\mathcal{L}(\mathbf{n}, \mathbf{m}) = \{f : \mathbb{R}^n \rightarrow \mathbb{R}^m \mid f \text{ es lineal}\}.$$

Como en el codominio se tienen las operaciones de suma y multiplicación por escalares, estas operaciones se pueden definir también en las funciones; en cierta manera, las funciones las heredan. Dadas $f, g \in \mathcal{L}(\mathbf{n}, \mathbf{m})$ y $t \in \mathbb{R}$, sean

$$\begin{aligned} f + g &: \mathbb{R}^n \rightarrow \mathbb{R}^m \\ (f + g)(\mathbf{x}) &:= f(\mathbf{x}) + g(\mathbf{x}) \end{aligned}$$

y

$$\begin{aligned} t f &: \mathbb{R}^n \rightarrow \mathbb{R}^m \\ (t f)(\mathbf{x}) &:= t(f(\mathbf{x})). \end{aligned}$$

Hay que demostrar que estas nuevas funciones también son lineales. Y esto es muy fácil a partir de la definición original, pues dados $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ y $\lambda, \mu \in \mathbb{R}$,

$$\begin{aligned} (f + g)(\lambda \mathbf{x} + \mu \mathbf{y}) &= f(\lambda \mathbf{x} + \mu \mathbf{y}) + g(\lambda \mathbf{x} + \mu \mathbf{y}) \\ &= \lambda f(\mathbf{x}) + \mu f(\mathbf{y}) + \lambda g(\mathbf{x}) + \mu g(\mathbf{y}) \\ &= \lambda(f(\mathbf{x}) + g(\mathbf{x})) + \mu(f(\mathbf{y}) + g(\mathbf{y})) \\ &= \lambda(f + g)(\mathbf{x}) + \mu(f + g)(\mathbf{y}). \end{aligned}$$

Formalmente nos falta demostrar que estas operaciones cumplen todas las propiedades que se les exigen a los espacios vectoriales (el Teorema ??). Esto depende esencialmente de que \mathbb{R}^m es espacio vectorial (no de que sean funciones lineales) y dejamos los detalles como ejercicio lateral (no esencial) a la línea del texto. Los nuevos ejemplos de espacio vectorial son entonces los conjuntos de funciones que salen de algún lugar (con alguna propiedad —el caso que desarrollamos fue que son lineales—) pero que **caen** en un espacio vectorial (en nuestro caso, \mathbb{R}^m).

EJERCICIO 3.62 Sea V un espacio vectorial y sea X cualquier conjunto. Demuestra que el conjunto de todas las funciones de X en V , $\mathcal{F}(X, V)$, tiene una estructura natural de espacio vectorial. (En particular $\mathcal{L}(\mathbf{n}, \mathbf{m})$ es un subespacio vectorial de $\mathcal{F}(\mathbb{R}^n, \mathbb{R}^m)$ que es inmensamente más grande.)

EJERCICIO 3.63 Sea $\Delta_n = \{1, 2, \dots, n\}$ el conjunto finito “canónico” con n elementos. Como corolario del Teorema 3.14, demuestra que los espacios vectoriales $\mathcal{L}(\mathbf{n}, \mathbf{m})$ y $\mathcal{F}(\Delta_n, \mathbb{R}^m)$ se pueden identificar naturalmente.

3.5 Matrices

Las matrices son arreglos rectangulares de números, tablas podría decirse, donde los renglones y las columnas no tienen ningún significado extra, como podrían tener, por ejemplo, en la tabla de calificaciones parciales de alumnos en un curso. Así de simple, las matrices son tablas limpias, abstractas. En nuestro contexto actual, habrán de ser los “paquetes” que cargan toda la información de las funciones lineales y nos darán la herramienta para hacer cálculos respecto a ellas, además de simplificar la notación.

En matemáticas, la notación es importante. Una notación clara y sencilla permite ver la esencia de las cosas, de las ideas y de los problemas. Por el contrario, con una notación complicada o confusa es fácil perderse en la labor de desentrañarla y nunca llegar a las ideas profundas; piénsese, por ejemplo, en diseñar un algoritmo para multiplicar con números romanos. Viene esto al caso, pues esta sección trata básicamente de cómo simplificar la notación para lograr manejar las transformaciones geométricas muy en concreto. Encontraremos la mecánica y la técnica para hacer cálculos con facilidad. Para empezar, cambiaremos nuestra notación de vectores.

3.5.1 Vectores columna

Como crítica a la notación que traemos, escribamos una función lineal $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ explícitamente. Digamos que $f(\mathbf{e}_1) = (1, 2, 3)$, $f(\mathbf{e}_2) = (6, 5, 4)$ y $f(\mathbf{e}_3) = (2, 3, 1)$; entonces sabemos que f se escribe

$$\begin{aligned} f(x, y, z) &= x(1, 2, 3) + y(6, 5, 4) + z(2, 3, 1) \\ &= (x + 6y + 2z, 2x + 5y + z, 3x + 4y + z). \end{aligned}$$

Para hacer este cálculo, la vista tuvo que andar brincoteando en una línea, buscando y contando comas; es muy fácil equivocarse. Es más, ¿dónde está el error? Sin embargo, sabemos que el cálculo es sencillísimo, mecánico. Si en lugar de considerar los vectores como renglones los tomamos como columnas, el mismo cálculo se hace evidente a la vista

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + y \begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix} + z \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} x + 6y + 2z \\ 2x + 5y + z \\ 3x + 4y + z \end{pmatrix} \quad (3.7)$$

y eliminamos las comas (y corregimos el error). Ahora cada renglón se refiere a una coordenada, y se ve todo de golpe; mucho mejor notación que haremos oficial.

Notación 3.1 De ahora en adelante, los vectores en \mathbb{R}^n se escribirán como columnas y no como renglones. Es decir, si antes escribíamos $\mathbf{x} = (x_1, x_2, \dots, x_n)$, ahora escribiremos

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

El inconveniente de que entonces se complica escribir un vector dentro del texto (aquí, por ejemplo), lo subsanamos al llamar a los vectores renglón, *transpuestos* de los vectores columna, y los denotaremos con $\mathbf{x}^\top = (x_1, x_2, \dots, x_n)$; o bien, $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top$ conviniendo que los *transpuestos* de vectores renglón son los vectores. Aunque a veces se nos va a olvidar poner el exponente \top de “transpuesto” y el lector nos lo perdonará; y con el tiempo nos lo agradecerá.

Así, por ejemplo, tenemos que la base canónica de \mathbb{R}^2 es $\mathbf{e}_1 = (1, 0)^\top$ y $\mathbf{e}_2 = (0, 1)^\top$, que quiere decir

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{y} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Vale la pena insistir en que no hay ningún cambio esencial. Sólo cambiamos de convención, una pareja ordenada de números se puede pensar horizontal (de izquierda a derecha) o vertical (de arriba a abajo); estamos conviniendo en que lo haremos y escribiremos en sentido vertical.

3.5.2 La matriz de una función lineal

Una matriz de $m \times n$ es un arreglo rectangular (o tabla) de números reales con m renglones y n columnas. Si, usando dos subíndices, denotamos con a_{ij} al número que está en el renglón i y la columna j , llamado *entrada*, tenemos que una matriz de $m \times n$ es

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Podríamos también definirla como un conjunto ordenado de n vectores en \mathbb{R}^m : sus columnas. Es decir, la matriz A anterior se puede escribir como

$$A = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n), \quad \text{donde} \quad \mathbf{u}_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix} \in \mathbb{R}^m, \quad i = 1, 2, \dots, n.$$

De tal manera que, de acuerdo con el Teorema 3.14, una matriz $m \times n$ tiene justo la información de una función lineal de \mathbb{R}^n en \mathbb{R}^m (ojo: se invierte el orden por la convención de que las funciones se componen “hacia atrás”). Explícitamente, a la matriz A se le asocia la función lineal $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ que manda al vector canónico $\mathbf{e}_i \in \mathbb{R}^n$ en su columna i -ésima, es decir, tal que $f(\mathbf{e}_i) = \mathbf{u}_i$ para $i = 1, 2, \dots, n$. O

bien, a la inversa, a cada función lineal $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ se le asocia la matriz $m \times n$ que tiene como columnas a sus valores en la base canónica, es decir, se le asocia la matriz

$$A = (f(\mathbf{e}_1), f(\mathbf{e}_2), \dots, f(\mathbf{e}_n)).$$

Por ejemplo, a la función lineal de \mathbb{R}^3 en \mathbb{R}^3 definida en (3.7) se le asocia la matriz

$$\begin{pmatrix} 1 & 6 & 2 \\ 2 & 5 & 3 \\ 3 & 4 & 1 \end{pmatrix}$$

que es ya una compactación considerable en la notación; a la transformación “compadre ortogonal” de \mathbb{R}^2 en \mathbb{R}^2 se le asocia la matriz

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

y a la función lineal que a cada vector en \mathbb{R}^3 lo manda a su segunda coordenada (en \mathbb{R}) se le asocia la matriz 1×3 : $\begin{pmatrix} 0 & 1 & 0 \end{pmatrix} = (0, 1, 0)$.

Ya logramos nuestro primer objetivo, empaquetar en una matriz toda la información de una función lineal. Ahora usémosla como herramienta. Primero, definiremos el producto de una matriz por un vector para que el resultado sea lo que su función lineal asociada hace al vector. Es decir, si A es una matriz $m \times n$, podrá multiplicar sólo a vectores $\mathbf{x} \in \mathbb{R}^n$ y el resultado será un vector en \mathbb{R}^m . Como ya vimos, A se puede escribir $A = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$ donde $\mathbf{u}_i \in \mathbb{R}^m$, para $i = 1, 2, \dots, n$; y por su parte, sea $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top$. Definimos entonces el *producto* de A por \mathbf{x} como

$$A\mathbf{x} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \dots + x_n\mathbf{u}_n, \quad (3.8)$$

de tal manera que el Teorema 3.14 junto con su corolario y lo que hemos visto de matrices se puede resumir en el siguiente teorema.

Teorema 3.16 *Las funciones lineales de \mathbb{R}^n en \mathbb{R}^m están en correspondencia natural y biunívoca con las matrices de $m \times n$, de tal manera que a la función f le corresponde la matriz A que cumple*

$$f(\mathbf{x}) = A\mathbf{x}$$

para todo $\mathbf{x} \in \mathbb{R}^n$.

Obsérvese que cuando $m = 1$ el producto que acabamos de definir corresponde a nuestro viejo conocido, el producto interior. Es decir, si $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ entonces

$$\mathbf{y}^\top \mathbf{x} = \mathbf{y} \cdot \mathbf{x}.$$

De tal manera que si escribimos A como una columna de renglones, en vez de un renglón de columnas, entonces hay vectores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m \in \mathbb{R}^n$ (los renglones) tales que

$$A\mathbf{x} = \begin{pmatrix} \mathbf{v}_1^\top \\ \mathbf{v}_2^\top \\ \vdots \\ \mathbf{v}_m^\top \end{pmatrix} \mathbf{x} = \begin{pmatrix} \mathbf{v}_1 \cdot \mathbf{x} \\ \mathbf{v}_2 \cdot \mathbf{x} \\ \vdots \\ \mathbf{v}_m \cdot \mathbf{x} \end{pmatrix}.$$

Esta definición del producto de una matriz por un vector es, aunque equivalente, más “desagradable” que la anterior, pues los vectores $\mathbf{v}_i \in \mathbb{R}^n$ no tienen ningún significado geométrico; son simplemente la colección ordenada de las i -ésimas coordenadas de los vectores $\mathbf{u}_j \in \mathbb{R}^m$. Aunque somos injustos, sí tienen un significado, los renglones \mathbf{v}_i^\top son las matrices asociadas a las m funciones lineales $\mathbb{R}^n \rightarrow \mathbb{R}$ que dan las coordenadas de la función original.

Para fijar ideas, y no perdernos en las abstracciones, terminamos esta sección con la expresión explícita del caso que más nos interesa en este libro: la multiplicación de una matriz de 2×2 por un vector en \mathbb{R}^2 , que tiene la fórmula general

$$\begin{pmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{ax} + \mathbf{by} \\ \mathbf{cx} + \mathbf{dy} \end{pmatrix}$$

donde todos los protagonistas son simples números.

EJERCICIO 3.64 Encuentra las matrices asociadas a las funciones lineales de algunos de los ejercicios anteriores.

EJERCICIO 3.65 Un vector $\mathbf{v} \in \mathbb{R}^n$ se puede pensar como matriz $n \times 1$. Como tal, ¿a qué función lineal representa según el Teorema 3.16?

EJERCICIO 3.66 El conjunto de todas las funciones lineales de \mathbb{R}^n en \mathbb{R}^m , que denotamos con $\mathcal{L}(n, m)$ en la Sección 3.4.2, tienen la estructura de espacio vectorial y el Teorema 3.16 dice que está en biyección natural con las matrices $m \times n$. Si definimos la suma de matrices y la multiplicación de escalares por matrices de modo que correspondan a las operaciones análogas en $\mathcal{L}(n, m)$, demuestra que $A + B$ (que tiene sentido sólo cuando A y B son del mismo tipo $m \times n$) es sumar entrada por entrada, y que tA es multiplicar todas las entradas por t . ¿Puedes dar una demostración fácil de que $\mathcal{L}(n, m)$ es un espacio vectorial? (Observa, del ejercicio anterior, que $\mathcal{L}(1, n)$ se identifica naturalmente con \mathbb{R}^n .)

3.5.3 Multiplicación de matrices

Vamos ahora a definir la multiplicación de matrices correspondiendo a la composición de funciones lineales. Tiene sentido componer dos funciones sólo cuando una “acaba” donde la otra “empieza”; más formalmente, cuando el codominio de una coincide con el dominio de la otra. De igual forma, sólo tendrá sentido multiplicar las matrices cuyas dimensiones se “acoplen”. Veámoslo.

Sean $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $g: \mathbb{R}^m \rightarrow \mathbb{R}^k$ dos funciones lineales. Por el Ejercicio 3.49, $g \circ f: \mathbb{R}^n \rightarrow \mathbb{R}^k$ también es lineal. Sean A la matriz $m \times n$ y B la matriz $k \times m$ que corresponden a f y a g respectivamente. Definimos el producto BA como la matriz $k \times n$ que corresponde a la función lineal $g \circ f$. Es decir, BA es la única matriz $k \times n$ que cumple

$$(g \circ f)(\mathbf{x}) = (BA)\mathbf{x} \quad \text{para todo } \mathbf{x} \in \mathbb{R}^n.$$

Aunque esta definición ya es precisa según el Teorema 3.16, todavía no nos dice cómo multiplicar. Esto habrá que deducirlo. Recordemos que las columnas de una matriz son las imágenes de la base canónica bajo la función asociada. Así que si $A = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$ donde $\mathbf{u}_i = f(\mathbf{e}_i) \in \mathbb{R}^m$, entonces $(g \circ f)(\mathbf{e}_i) = g(f(\mathbf{e}_i)) = g(\mathbf{u}_i) = B\mathbf{u}_i$. Y por lo tanto

$$BA = B(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) = (B\mathbf{u}_1, B\mathbf{u}_2, \dots, B\mathbf{u}_n).$$

Lo cual es muy natural: para obtener las columnas de la nueva matriz, se usa la multiplicación de B por los vectores columna de A , multiplicación que ya definimos. Para expresar cada una de las entradas de la matriz BA , habrá que expresar B como una columna de vectores renglon, y se obtiene

$$BA = \begin{pmatrix} \mathbf{w}_1^\top \\ \mathbf{w}_2^\top \\ \vdots \\ \mathbf{w}_k^\top \end{pmatrix} (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) = \begin{pmatrix} \mathbf{w}_1 \cdot \mathbf{u}_1 & \mathbf{w}_1 \cdot \mathbf{u}_2 & \cdots & \mathbf{w}_1 \cdot \mathbf{u}_n \\ \mathbf{w}_2 \cdot \mathbf{u}_1 & \mathbf{w}_2 \cdot \mathbf{u}_2 & \cdots & \mathbf{w}_2 \cdot \mathbf{u}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{w}_k \cdot \mathbf{u}_1 & \mathbf{w}_k \cdot \mathbf{u}_2 & \cdots & \mathbf{w}_k \cdot \mathbf{u}_n \end{pmatrix}, \quad (3.9)$$

En esta fórmula queda claro por qué los renglones de B (los transpuestos de los vectores \mathbf{w}_i) y las columnas de A (los vectores \mathbf{u}_j) tienen que estar en el mismo espacio (\mathbb{R}^m); y cuál es la mecánica¹ para obtener las entradas de una matriz $k \times n$ (BA) a partir de una $k \times m$ (B) y una $m \times n$ (A). Esta fórmula da, en el caso de matrices 2×2 , lo siguiente

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} a\alpha + b\gamma & a\beta + b\delta \\ c\alpha + d\gamma & c\beta + d\delta \end{pmatrix}.$$

EJERCICIO 3.67 Sean

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 3 \\ -2 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} -2 & 0 \\ 1 & 2 \end{pmatrix}.$$

Calcula AB , $A(B + C)$, $(C - B)A$.

¹La mecánica de la multiplicación de matrices corresponde a la *ley del karatazo*: un karateca en guardia pone su antebrazo derecho vertical y el izquierdo horizontal —¡(hhuouui)!— y se apresta a multiplicar dos matrices: recoge con la mano derecha una columna de la matriz derecha y —¡(aah₁, aah₂, ..., ah_k)[⊤]!— le va pegando karatazos a los renglones de la izquierda para obtener la columna correspondiente en el producto; y si es zurdo puede recoger con su buena renglones de la izquierda y —¡(uuuf₁, uuuf₂, ..., uf_n)!— golpear a las columnas derechas (codo arriba) para ir sacando, de una en una, las entradas del producto.

EJERCICIO 3.68 Exhibe matrices A y B de 2×2 tales que $AB \neq BA$. De tal manera que el producto de matrices no es conmutativo.

EJERCICIO 3.69 Exhibe matrices A y B de 2×2 tales que $AB = \mathbf{0}$, pero $A \neq \mathbf{0}$ y $B \neq \mathbf{0}$; donde $\mathbf{0}$ es la matriz cero (con todas sus entradas 0). (*Piensa en términos de funciones.*)

EJERCICIO 3.70 Demuestra que si A, B, C son matrices 2×2 , entonces $A(BC) = (AB)C$, es decir, que el producto de matrices es asociativo, y por lo tanto tiene sentido escribir ABC . (*No necesariamente tienes que escribir todo el producto y tal vez tu demostración valga en general.*)

EJERCICIO 3.71 Demuestra que si A, B, C son matrices 2×2 , entonces $A(B + C) = AB + AC$ y $(A + B)C = AC + BC$.

EJERCICIO 3.72 Demuestra que si A, B son matrices 2×2 entonces $A(kB) = (kA)B = k(AB)$ para cualquier $k \in \mathbb{R}$.

EJERCICIO 3.73 Usando que la composición de funciones es asociativa demuestra que si A, B, C son matrices $n \times n$ entonces $A(BC) = (AB)C$.

3.5.4 Algunas familias distinguidas de matrices

La matriz identidad

Aunque propiamente dicho no sea una familia se merece su párrafo aparte y, además, ser el primero. La *matriz identidad*, o simplemente “la identidad”, es la matriz asociada a la función identidad, que se denota por I , y es la matriz que tiene 1 en la diagonal y 0 fuera de ella, es decir

$$I := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Cumple además que $AI = A$ e $IA = A$ para cualquier matriz A que se deje multiplicar por ella: es la identidad multiplicativa. Nótese que en realidad hay una identidad para cada dimensión, empezando por el 1; así que deberíamos escribir algo así como I_n , pero el contexto en que se usa siempre trae la información de la dimensión.

Homotecias

Las homotecias, como funciones, son simples “cambios de escala”. Si tomamos un número $k \in \mathbb{R}$, $k \neq 0$, la función $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por $f(\mathbf{x}) = k\mathbf{x}$ es claramente lineal y se llama una *homotecia*. Entonces tiene una matriz asociada que es kI , la matriz que tiene sólo k en la diagonal y 0 fuera de ella. Como función, lo que hace es

expandir uniformemente desde el origen (si $k > 1$), o bien contraer (cuando $k < 1$); son un *zoom*, pues.

Las homotecias tienen la propiedad de conmutar con cualquier otra matriz, pues $A(kI) = k(AI) = kA = k(IA) = (kI)A$ (y aquí, pudo haber sido que las dos I sean diferentes, lo que depende de A). En términos de funciones, el cambio de escala se puede hacer antes o después de una función lineal y da lo mismo. Y además son las únicas matrices (pensando en matrices cuadradas) que tienen esta propiedad; pero esto lo dejamos como ejercicio.

EJERCICIO 3.74 Demuestra que si una matriz A de 2×2 es tal que $AB = BA$ para todas las matrices B de 2×2 , entonces $A = kI$ para alguna $k \in \mathbb{R}$. (*Tienes que encontrar matrices apropiadas B que te vayan dando información; quizás conviene ver el siguiente párrafo y regresar.*)

Matrices de permutaciones

Recordemos de la Sección 3.1 que una permutación de n elementos es una función biyectiva $\rho : \Delta_n \rightarrow \Delta_n$, donde podemos tomar $\Delta_n = \{1, 2, \dots, n\}$; y que todas estas permutaciones forman el grupo simétrico de orden n , \mathbf{S}_n , con $n!$ elementos. Ahora bien, a cada permutación $\rho : \Delta_n \rightarrow \Delta_n$ se le puede asociar la función lineal $f_\rho : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por $f_\rho(\mathbf{e}_i) = \mathbf{e}_{\rho(i)}$; es decir, f_ρ permuta a los elementos de la base canónica según ρ y se extiende linealmente a \mathbb{R}^n . A la matriz asociada a esta función lineal se le llama la matriz *de* la permutación ρ ; denotémosla A_ρ . Es fácil ver que la multiplicación de matrices de permutaciones es de nuevo una matriz de permutación y corresponde a la composición de las permutaciones correspondientes. Es decir, si $\rho, \sigma \in \mathbf{S}_n$, entonces $A_\rho A_\sigma = A_{\rho \circ \sigma}$. Se tiene entonces que el grupo simétrico se puede ver como un “grupo de matrices”.

Otra manera equivalente de definir las es que son las matrices cuadradas de ceros y unos tales que tienen exactamente un 1 en cada renglón y en cada columna. Por ejemplo, la matriz

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

corresponde a la permutación $\{1 \mapsto 3, 2 \mapsto 2, 3 \mapsto 4, 4 \mapsto 1\}$.

EJERCICIO 3.75 Escribe todas las matrices de permutaciones de orden 2.

EJERCICIO 3.76 Escribe todas las matrices de permutaciones de orden 3. Identifica sus inversas (es decir la matriz asociada a la permutación inversa), y describe en cada caso la transformación lineal de \mathbb{R}^3 correspondiente.

EJERCICIO 3.77 Sea A_ρ la matriz de la permutación $\rho: \Delta_n \rightarrow \Delta_n$, y sea B cualquier otra matriz de $n \times n$. Describe a las matrices BA_ρ y $A_\rho B$ en términos de la permutación ρ .

***EJERCICIO 3.78** Sea C el cubo en \mathbb{R}^3 centrado en el origen, de lado 2 y con lados paralelos a los ejes coordenados; es decir, sus vértices son los ocho puntos cuyas tres coordenadas son 1 o -1 . Describe el conjunto de matrices cuyas funciones lineales asociadas mandan al cubo en sí mismo. En particular, ¿cuántas son? (*Quizás valga la pena empezar con el problema análogo en dimensión 2.*)

Rotaciones

Regresamos, después de un provechoso paréntesis, a la motivación original que nos llevó a adentrarnos en las funciones lineales.

Sabemos que una rotación es una transformación lineal que manda al vector canónico \mathbf{e}_1 en un vector unitario $\mathbf{u} \in S^1$ y al otro vector canónico \mathbf{e}_2 en su compadre ortogonal \mathbf{u}^\perp . Si escribimos a \mathbf{u} en términos de su ángulo respecto del eje x , es decir, respecto a \mathbf{e}_1 , entonces obtenemos que la matriz asociada a la rotación por un ángulo θ es

$$R_\theta := \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Primero veamos que la rotación de un ángulo $-\theta$ nos regresa a la identidad. Como

$$\cos(-\theta) = \cos \theta \quad \text{y} \quad \sin(-\theta) = -\sin \theta \quad (3.10)$$

entonces

$$\begin{aligned} R_{-\theta} R_\theta &= \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \\ &= \begin{pmatrix} \cos^2 \theta + \sin^2 \theta & \cos \theta(-\sin \theta) + \sin \theta \cos \theta \\ -\sin \theta \cos \theta + \cos \theta \sin \theta & \sin^2 \theta + \cos^2 \theta \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I. \end{aligned}$$

Podemos obtener las fórmulas trigonométricas clásicas para el coseno y el seno de la suma de ángulos como consecuencia de la composición de funciones y la multiplicación de matrices. Está claro que si rotamos un ángulo β y después un ángulo α , habremos rotado un ángulo $\alpha + \beta$. Así que

$$R_\alpha R_\beta = R_{\alpha+\beta}.$$

Pero por otro lado, al multiplicar las matrices tenemos

$$\begin{aligned} R_\alpha R_\beta &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix} = \\ &= \begin{pmatrix} \cos \alpha \cos \beta - \sin \alpha \sin \beta & -\cos \alpha \sin \beta - \sin \alpha \cos \beta \\ \sin \alpha \cos \beta + \cos \alpha \sin \beta & -\sin \alpha \sin \beta + \cos \alpha \cos \beta \end{pmatrix}. \end{aligned}$$

Y por lo tanto

$$\begin{aligned}\cos(\alpha + \beta) &= \cos \alpha \cos \beta - \sin \alpha \sin \beta \\ \sin(\alpha + \beta) &= \sin \alpha \cos \beta + \cos \alpha \sin \beta,\end{aligned}\tag{3.11}$$

pues si dos matrices coinciden ($R_{\alpha+\beta}$ y $R_\alpha R_\beta$) lo hacen entrada por entrada. Junto con las igualdades trigonométricas de inversos de ángulos, estas últimas dan

$$\begin{aligned}\cos(\alpha - \beta) &= \cos \alpha \cos \beta + \sin \alpha \sin \beta \\ \sin(\alpha - \beta) &= \sin \alpha \cos \beta - \cos \alpha \sin \beta.\end{aligned}\tag{3.12}$$

EJERCICIO 3.79 Da una expresión numérica bonita de las matrices asociadas a las rotaciones de $\pi/6, \pi/4, \pi/3, 2\pi/3, \pi, -\pi/2$.

EJERCICIO 3.80 Demuestra que

$$\begin{aligned}\cos 2\alpha &= \cos^2 \alpha - \sin^2 \alpha \\ \sin 2\alpha &= 2 \sin \alpha \cos \alpha.\end{aligned}$$

Reflexiones

Animacion

Conviene parametrizar las reflexiones no por el ángulo de la imagen de \mathbf{e}_1 , sino por su mitad, que es el ángulo de la recta-espejo de la reflexión. Pues si, como es natural, a una recta por el origen le asociamos su ángulo (entre 0 y π) con el rayo positivo del eje x , entonces la reflexión, E_θ , en la recta con ángulo θ manda a \mathbf{e}_1 en el vector unitario de ángulo 2θ , y por lo tanto su matriz asociada es

$$E_\theta := \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix},$$

pues el segundo vector no es el compadre ortogonal sino su negativo. Es muy fácil ver que $E_\theta E_\theta = I$, que corresponde a que una reflexión es su propia inversa.

Podemos ahora ver cómo se comporta la composición de reflexiones:

$$\begin{aligned}E_\alpha E_\beta &= \\ &= \begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix} \begin{pmatrix} \cos 2\beta & \sin 2\beta \\ \sin 2\beta & -\cos 2\beta \end{pmatrix} \\ &= \begin{pmatrix} \cos 2\alpha \cos 2\beta + \sin 2\alpha \sin 2\beta & \cos 2\alpha \sin 2\beta - \sin 2\alpha \cos 2\beta \\ \sin 2\alpha \cos 2\beta - \cos 2\alpha \sin 2\beta & \sin 2\alpha \sin 2\beta + \cos 2\alpha \cos 2\beta \end{pmatrix} \\ &= \begin{pmatrix} \cos 2(\alpha - \beta) & -\sin 2(\alpha - \beta) \\ \sin 2(\alpha - \beta) & \cos 2(\alpha - \beta) \end{pmatrix} = R_{2(\alpha-\beta)}\end{aligned}$$

donde hemos usado (3.12). Así que la composición de dos reflexiones es la rotación del doble del ángulo entre sus espejos. Tal y como lo habíamos visto al hablar de generadores de grupos diédricos.

EJERCICIO 3.81 Da una expresión numérica bonita de las matrices asociadas a las reflexiones en espejos a $0, \pi/6, \pi/4, \pi/3, \pi/2, 2\pi/3, 3\pi/4$.

EJERCICIO 3.82 ¿Quiénes son $E_\alpha R_\beta$ y $R_\beta E_\alpha$?

EJERCICIO 3.83 Demuestra con matrices que $R_\beta E_0 R_{-\beta} = E_\beta$.

EJERCICIO 3.84 Demuestra que la fórmula (3.3) obtenida en la Sección 3.3.1 para la reflexión φ_ℓ en la línea ℓ , coincide con la que nos da una matriz E_θ cuando la línea ℓ pasa por el origen.

Matrices ortogonales (y transpuestas)

Decimos que una matriz es *ortogonal* si su función lineal asociada es ortogonal. Según nuestro estudio de las transformaciones ortogonales en \mathbb{R}^2 , las matrices ortogonales de 2×2 son las que acabamos de describir: las rotaciones y las reflexiones. Podemos considerar entonces al grupo de transformaciones $\mathbf{O}(2)$ como un grupo de matrices con la operación de multiplicación.

Nótese que en las matrices de rotaciones y reflexiones, la matriz *inversa*, esto es, la matriz asociada a la transformación inversa, resultó ser muy parecida a la original; en las reflexiones es la misma y en las rotaciones se intercambian los elementos fuera de la diagonal (que difieren sólo en el signo). Esto no es coincidencia sino parte de algo mucho más general, que haremos explícito en esta sección y que será importante en capítulos posteriores.

Llamemos a una matriz *ortogonal* si es cuadrada y su función lineal asociada es una transformación ortogonal (que preserva producto interior). Entonces, si A es ortogonal, todas sus columnas son vectores unitarios (pues son la imagen de vectores unitarios, los canónicos) y además por parejas son ortogonales (pues así es la base canónica y esto se detecta por el producto interior). Dicho más explícitamente, si $A = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$ entonces $\mathbf{u}_i = f(\mathbf{e}_i)$ donde f es su función asociada ($f(\mathbf{x}) = A\mathbf{x}$), y si $f \in \mathbf{O}(2)$ se cumple que

$$\mathbf{u}_i \cdot \mathbf{u}_j = \mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} := \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j, \end{cases}$$

donde δ_{ij} es conocida como “la delta de Kroenecker”. En este caso, al conjunto ordenado de vectores $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ se le llama *base ortonormal* (pues son vectores unitarios o normalizados y mutuamente ortogonales). Pero nótese que estos n^2 pro-

ductos interiores $(\mathbf{u}_i \cdot \mathbf{u}_j)$ son las entradas de la matriz

$$\begin{pmatrix} \mathbf{u}_1^\top \\ \mathbf{u}_2^\top \\ \vdots \\ \mathbf{u}_n^\top \end{pmatrix} (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) = \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & \mathbf{u}_1 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_1 \cdot \mathbf{u}_n \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \mathbf{u}_2 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_2 \cdot \mathbf{u}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_n \cdot \mathbf{u}_1 & \mathbf{u}_n \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_n \cdot \mathbf{u}_n \end{pmatrix}$$

y que el que sean iguales a la delta de Kroenecker quiere decir que ésta es la matriz identidad. Así que si llamamos *transpuesta* de una matriz a la que se obtiene cambiando columnas por renglones, hemos demostrado que la “inversa” de una matriz ortogonal es su transpuesta.

En general, llamemos la *transpuesta* de una matriz $A = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ a la matriz

$$A^\top = \begin{pmatrix} \mathbf{v}_1^\top \\ \mathbf{v}_2^\top \\ \vdots \\ \mathbf{v}_n^\top \end{pmatrix}$$

que se obtiene al reflejar las entradas en la diagonal; si A es $m \times n$ entonces A^\top es $n \times m$ y se generaliza nuestra noción previa de transpuestos de vectores. Hemos demostrado la mitad del siguiente teorema. Su otra mitad, para el caso que más nos interesa, $n = 2$, es consecuencia inmediata de la Proposición 3.12. Sin embargo damos una nueva demostración, muy general, para ejemplificar el poder de la notación matricial.

Teorema 3.17 *Una matriz A de $n \times n$ es ortogonal (la matriz de una transformación ortogonal) si y sólo si*

$$A^\top A = I.$$

Demostración. Hemos demostrado que si la función $f(\mathbf{x}) = A\mathbf{x}$ preserva producto interior entonces su matriz cumple que $A^\top A = I$. Supongamos ahora que $A^\top A = I$ y demostremos que su transformación asociada preserva producto interior. Para esto, expresaremos al producto interior como un caso más del producto de matrices. Para cualquier $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, se tiene

$$f(\mathbf{x}) \cdot f(\mathbf{y}) = A\mathbf{x} \cdot A\mathbf{y} = (A\mathbf{x})^\top A\mathbf{y} = (\mathbf{x}^\top A^\top) A\mathbf{y} = \mathbf{x}^\top (A^\top A) \mathbf{y} = \mathbf{x}^\top (I) \mathbf{y} = \mathbf{x} \cdot \mathbf{y},$$

donde hemos usado que $(A\mathbf{x})^\top = \mathbf{x}^\top A^\top$ (lo cual es un caso particular de la fórmula $(AB)^\top = B^\top A^\top$ que se deja como ejercicio). Por lo tanto, f es ortogonal (y también A lo es). \square

Así que aunque las entradas de las matrices ortogonales sean complicadas, encontrar sus inversas es muy fácil. Y se explica por qué las inversas de las ortogonales 2×2 fueron tan parecidas a sí mismas.

EJERCICIO 3.85 Sea $\mathbf{u}, \mathbf{v}, \mathbf{w}$ una base ortonormal de \mathbb{R}^3 , y sea $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definida por

$$f(x, y, z) = x\mathbf{u} + y\mathbf{v} + z\mathbf{w}.$$

Demuestra directamente que f es ortogonal (preserva producto interior). (Compara con la Proposición 3.12.)

EJERCICIO 3.86 Demuestra que $(A\mathbf{x})^\top = \mathbf{x}^\top A^\top$ cuando A es una matriz $m \times n$ y \mathbf{x} un vector ($n \times 1$).

EJERCICIO 3.87 Demuestra que $(AB)^\top = B^\top A^\top$ cuando A es una matriz $m \times n$ y B una matriz ($n \times k$).

EJERCICIO 3.88 Demuestra que la “matriz de Pitágoras”

$$\frac{1}{5} \begin{pmatrix} 4 & 3 \\ -3 & 4 \end{pmatrix}$$

es ortogonal.

3.6 El Grupo General Lineal ($\mathbf{GL}(2)$)

En esta sección estudiamos las matrices cuyas funciones lineales asociadas son transformaciones, es decir, biyectivas. La definición abstracta será entonces muy fácil, aunque después habrá que hacerla más concreta.

Definición 3.6.1 Una matriz A de $n \times n$ es *invertible* si existe otra matriz de $n \times n$, llamada su *inversa* y denotada A^{-1} tal que $AA^{-1} = A^{-1}A = I$. Al conjunto de todas las matrices invertibles de $n \times n$ se le llama el *grupo general lineal de orden n* y se le denota $\mathbf{GL}(n)$. Así que $A \in \mathbf{GL}(n)$ quiere decir que A es invertible y de $n \times n$.

Para demostrar que las transformaciones lineales $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ tienen matrices asociadas que son invertibles sólo habría que ver que su función inversa también es lineal (ver el ejercicio siguiente), y entonces la matriz asociada a ésta será la inversa. Y al revés, si una matriz tiene inversa entonces las funciones lineales asociadas también son inversas. Así que a $\mathbf{GL}(2)$ se le puede pensar como grupo de transformaciones, tal y como lo hicimos en las primeras secciones de este capítulo, o bien como grupo de matrices (donde, hay que enfatizar, la operación del grupo es la multiplicación de matrices).

Ejemplos de matrices invertibles son las ortogonales que ya hemos visto. Así, tenemos que $\mathbf{O}(n)$ es un subgrupo de $\mathbf{GL}(n)$.

EJERCICIO 3.89 Sea $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ una transformación lineal; es decir, una función lineal con inversa f^{-1} . Demuestra que su inversa también es lineal.

3.6.1 El determinante

Veamos ahora el problema de detectar de manera sencilla las matrices invertibles de 2×2 . Sea

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

una matriz cualquiera de 2×2 . Consideremos el problema general de encontrar su inversa, y de paso ver cuándo ésta existe. Tomemos entonces una matriz X cuyas entradas son incógnitas, digamos x, y, z, w , que cumplan

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & z \\ y & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Esto nos da un sistema de cuatro ecuaciones con cuatro incógnitas, que en realidad se parte en los dos sistemas de dos ecuaciones con dos incógnitas siguientes

$$\begin{array}{l} ax + by = 1 \\ cx + dy = 0 \end{array} \quad y \quad \begin{array}{l} az + bw = 0 \\ cz + dw = 1, \end{array}$$

que provienen de cada una de las columnas de la identidad, o bien de las columnas de incógnitas. Recordando la Sección 1.6, o bien volviéndola a resolver, se obtiene que si $ad - bc \neq 0$, entonces los dos sistemas tienen soluciones únicas que son precisamente

$$\begin{array}{l} x = \frac{d}{ad - bc} \\ y = \frac{-c}{ad - bc} \end{array} \quad \begin{array}{l} z = \frac{-b}{ad - bc} \\ w = \frac{a}{ad - bc}. \end{array} \quad (3.13)$$

Para enunciar esto de manera elegante conviene definir el *determinante* de la matriz A como el número

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} := ad - bc.$$

Nótese que si $A = (\mathbf{u}, \mathbf{v})$, entonces $\det(A) = \det(\mathbf{u}, \mathbf{v})$ según lo definimos en la Sección 1.7, de tal manera que el determinante de una matriz es el determinante de sus columnas o el determinante de un sistema de ecuaciones asociado y sólo estamos ampliando naturalmente el espectro del término.

Teorema 3.18 Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, entonces A es invertible si y sólo si $\det(A) = ad - bc \neq 0$; y en este caso

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Hemos demostrado que si $\det(A) \neq 0$ entonces A es invertible y su inversa tiene la expresión indicada por (3.13). Nos falta ver que si A tiene inversa entonces su determinante no es cero. Pero esto será consecuencia inmediata del siguiente lema que nos da el determinante de un producto.

Lema 3.19 Sean A y B matrices 2×2 , entonces $\det(AB) = \det(A) \det(B)$.

Demostración. Sea A como arriba y consideremos B con letras griegas. Se tiene entonces que

$$\begin{aligned} \det(AB) &= \det \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \right) \\ &= \det \begin{pmatrix} a\alpha + b\gamma & a\beta + b\delta \\ c\alpha + d\gamma & c\beta + d\delta \end{pmatrix} \\ &= (a\alpha + b\gamma)(c\beta + d\delta) - (a\beta + b\delta)(c\alpha + d\gamma) \\ &= a\alpha c\beta + a\alpha d\delta + b\gamma c\beta + b\gamma d\delta \\ &\quad - a\beta c\alpha - a\beta d\gamma - b\delta c\alpha - b\delta d\gamma \\ &= ad(\alpha\delta - \beta\gamma) + bc(\beta\gamma - \alpha\delta) \\ &= (ad - bc)(\alpha\delta - \beta\gamma) = \det(A) \det(B). \end{aligned}$$

□

Demostración. (del Teorema 3.18). Si A es una matriz invertible entonces existe A^{-1} tal que $AA^{-1} = I$. Por el lema anterior se tiene entonces que $\det(A) \det(A^{-1}) = \det(I) = 1$, y por lo tanto $\det(A)$ no puede ser 0. □

Veamos por último que el determinante de una matriz tiene un significado geométrico. Por el Teorema 1.30, corresponde al área dirigida del paralelogramo generado por sus vectores columna.

Teorema 3.20 Sea A una matriz 2×2 , entonces $\det(A)$ es el área del paralelogramo definido por sus vectores columna. □

Se dice que la matriz A , o que la función $f(\mathbf{x}) = A\mathbf{x}$, *invierte orientación* si $\det(A) < 0$, o bien, que la *preserva* si $\det(A) > 0$. Esto corresponde a que cualquier figura que no tenga simetrías de espejo (por ejemplo, un “manquito” de la mano izquierda) se puede orientar, y bajo transformaciones lineales que preservan orientación, la figura puede deformarse pero conserva su orientación (el manquito seguirá siendo manquito de la mano izquierda); pero bajo transformaciones que invierten orientación se verá como manquito de la otra mano.



Pero además, el determinante de A representa lo que la función f “distorsiona” las áreas. Es decir, si una figura cualquiera F tiene área α , entonces $f(F)$ tendrá área $\alpha|\det(A)|$. Sin pretender dar una demostración formal de esto, en general se puede argumentar como sigue. Puesto que la imagen del cuadrado unitario, C , bajo la función f asociada a la matriz $A = (\mathbf{u}, \mathbf{v})$, es el paralelogramo generado por \mathbf{u} y \mathbf{v} , primero hay que ver que cuadraditos con lados paralelos a los ejes y de lado ε (y área ε^2) van a dar a paralelogramos de área $\varepsilon^2 \det(A)$; y luego observar que las áreas se pueden aproximar por cuadraditos disjuntos de este tipo. Con esta interpretación geométrica del determinante es natural que el determinante de un producto sea el producto de los determinantes, pues el factor de distorsión de áreas de una composición debía de ser el producto de factores de distorsión.

Observemos por último que el determinante distingue nuestras dos clases de matrices ortogonales (en $\mathbf{O}(2)$); es 1 para las rotaciones, y -1 para las reflexiones. Es decir, las rotaciones preservan orientación y las reflexiones la invierten. Además su multiplicación corresponde a cómo se multiplican 1 y -1 .

EJERCICIO 3.90 Demuestra que el determinante de las rotaciones es 1 y el de las reflexiones -1 .

EJERCICIO 3.91 Sea $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la transformación lineal asociada a la matriz $A \in \mathbf{GL}(2)$, y sea T cualquier triángulo en el plano. Demuestra que el área de $f(T)$ es $|\det(A)|$ por el área de T .

EJERCICIO 3.92 Sea $\mathbb{C} = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$; y sea $\mathbb{C}^* = \mathbb{C} - \{\mathbf{0}\}$, es decir, todo \mathbb{C} menos la matriz $\mathbf{0}$. Describe las funciones lineales asociadas. Demuestra que \mathbb{C}^* es un subgrupo de $\mathbf{GL}(2)$. Compara estas matrices con los números complejos.

EJERCICIO 3.93 Encuentra (en cada caso) la matriz de la transformación lineal $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ que cumple:

- a) $f(2, 1) = (1, 0)$ y $f(3, 2) = (0, 1)$
 b) $f(2, 5) = (1, 0)$ y $f(1, 2) = (0, 1)$
 c) $f(2, 1) = (2, 5)$ y $f(3, 2) = (1, 2)$,

donde hemos usado vectores renglón en vez de vectores columna por simplicidad.

EJERCICIO 3.94 Demuestra que una transformación lineal de \mathbb{R}^2 manda rectas en rectas.

EJERCICIO 3.95 Considera las rectas $(\ell_1 : x + y = 2)$; $(\ell_2 : 2x - y = 1)$; $(\ell_3 : 2x + y = 6)$ y $(\ell_4 : x - 4y = 3)$. Encuentra la transformación lineal $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ que cumple $f(\ell_1) = \ell_3$ y $f(\ell_2) = \ell_4$.

3.6.2 Sistemas de ecuaciones lineales II

Vale la pena resumir y reinterpretar explícitamente lo que hemos hecho sobre sistemas de dos ecuaciones lineales en dos incógnitas con nuestro nuevo lenguaje de matrices y funciones. Un sistema tal es una ecuación

$$A \mathbf{x} = \mathbf{b},$$

donde A es una matriz 2×2 dada y \mathbf{b} es un vector constante en \mathbb{R}^2 . Si f es la función lineal asociada a la matriz, resolver el sistema es encontrar un vector \mathbf{x} que bajo f vaya a \mathbf{b} . Si A tiene inversa, resolver el sistema es fácil pues se multiplican ambos lados por A^{-1} para obtener

$$\mathbf{x} = A^{-1} \mathbf{b},$$

que es justo la primera fórmula que obtuvimos en la Sección 1.6.1. Y podemos resumir lo que hemos hecho en diferentes lados con el siguiente teorema que esencialmente ya hemos demostrado.

Teorema 3.21 Sea A una matriz de 2×2 y $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ su función lineal asociada. Entonces son equivalentes:

- i) $\det(A) \neq 0$.
- ii) La matriz A tiene inversa.
- iii) El sistema $A \mathbf{x} = \mathbf{b}$ tiene solución para toda $\mathbf{b} \in \mathbb{R}^2$.
- iv) El sistema $A \mathbf{x} = \mathbf{b}$ tiene solución única para alguna $\mathbf{b} \in \mathbb{R}^2$.
- v) La función f es inyectiva.
- vi) La función f es sobre.
- vii) La función f es biyectiva. □

Además sabemos que si $\det(A) = 0$ entonces la imagen de f es una recta por el origen (o el origen cuando A es la matriz $\mathbf{0}$) y entonces para los puntos \mathbf{b} fuera de esa recta, el sistema $A \mathbf{x} = \mathbf{b}$ no tiene solución, mientras que para los puntos en ella hay una recta de soluciones; además estas rectas de soluciones son paralelas.

3.7 Transformaciones afines

El papel que tuvieron las transformaciones ortogonales respecto de las isometrías lo tendrán ahora las transformaciones lineales respecto de las que llamaremos afines. Es decir, si componemos una transformación lineal con una traslación no trivial, ya no

obtenemos una lineal (pues el origen se mueve) sino algo que llamaremos transformación afín, generalizando para $n > 1$ a las transformaciones afines de \mathbb{R} , que ya estudiamos.

Definición 3.7.1 Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una *transformación afín* si se escribe Dibujo

$$f(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$$

para toda $\mathbf{x} \in \mathbb{R}^n$, donde $A \in \mathbf{GL}(n)$ y \mathbf{b} es un vector fijo en \mathbb{R}^n . Dicho de otra manera, f es una *transformación afín* si existe una transformación lineal $f_o : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ($f_o(\mathbf{x}) = A\mathbf{x}$) y una translación $\tau_{\mathbf{b}}$ tales que

$$f = \tau_{\mathbf{b}} \circ f_o.$$

El conjunto de todas las transformaciones afines de \mathbb{R}^n forman un grupo de transformaciones, llamado el *grupo afín de \mathbb{R}^n* y denotado con $\mathbf{Af}(n)$.

Lo primero que hay que observar es que una transformación afín es efectivamente una transformación; y esto es claro pues es la composición de dos funciones biyectivas. Para ver que $\mathbf{Af}(n)$ es efectivamente un grupo, hay que ver que es cerrado bajo inversas y composición, pues la identidad (y de hecho cualquier transformación lineal) está en $\mathbf{Af}(n)$. Si $f \in \mathbf{Af}(n)$ (dada por la fórmula de la definición) para obtener la fórmula explícita de su inversa, despejaremos \mathbf{x} en la ecuación

$$\mathbf{y} = A\mathbf{x} + \mathbf{b},$$

donde \mathbf{y} desempeña el papel de $f(\mathbf{x})$. Pasando \mathbf{b} del otro lado y luego multiplicando por A^{-1} (que existe pues $A \in \mathbf{GL}(n)$) se obtiene

$$\mathbf{x} = (A^{-1})\mathbf{y} - (A^{-1})\mathbf{b}.$$

Sea entonces $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por la ecuación $g(\mathbf{x}) = (A^{-1})\mathbf{x} - (A^{-1})\mathbf{b}$ (nótese que $g \in \mathbf{Af}(n)$ por definición) y se tiene que

$$\begin{aligned} (g \circ f)(\mathbf{x}) &= g(A\mathbf{x} + \mathbf{b}) \\ &= (A^{-1})(A\mathbf{x} + \mathbf{b}) - (A^{-1})\mathbf{b} \\ &= (A^{-1}A)\mathbf{x} + A^{-1}\mathbf{b} - A^{-1}\mathbf{b} = \mathbf{x}, \end{aligned}$$

y análogamente se tiene que $(f \circ g)(\mathbf{x}) = \mathbf{x}$, así que g es la inversa de f .

Si tomamos ahora $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ como cualquier otra transformación afín, entonces se escribe $g(\mathbf{x}) = B\mathbf{x} + \mathbf{c}$ para algunos $B \in \mathbf{GL}(n)$ y $\mathbf{c} \in \mathbb{R}^n$ de tal forma que $(g \circ f)$ también es una transformación afín pues se escribe

$$(g \circ f)(\mathbf{x}) = B(A\mathbf{x} + \mathbf{b}) + \mathbf{c} = (BA)\mathbf{x} + (B\mathbf{b} + \mathbf{c}).$$

Por lo tanto, $\mathbf{Af}(n)$ sí es un grupo de transformaciones.

Dibujo

Las transformaciones afines preservan rectas, es decir, mandan rectas en rectas. Pues si una recta ℓ está definida como

$$\ell = \{\mathbf{p} + t\mathbf{v} \mid t \in \mathbb{R}\}$$

y $f \in \mathbf{Af}(n)$, es como arriba, entonces

$$\begin{aligned} f(\mathbf{p} + t\mathbf{v}) &= A(\mathbf{p} + t\mathbf{v}) + \mathbf{b} \\ &= (A\mathbf{p} + \mathbf{b}) + t(A\mathbf{v}) \\ &= f(\mathbf{p}) + t(A\mathbf{v}), \end{aligned}$$

de tal manera que

$$f(\ell) \subseteq \{f(\mathbf{p}) + t(A\mathbf{v}) \mid t \in \mathbb{R}\}.$$

Y este último conjunto es la recta que pasa por $f(\mathbf{p})$ con dirección $A\mathbf{v}$ (nótese que $A\mathbf{v} \neq \mathbf{0}$, pues $\mathbf{v} \neq \mathbf{0}$ y A es invertible); así que más que una contención, la última es justo una igualdad (ambos conjuntos están parametrizados por \mathbb{R}).

Resulta que las transformaciones afines son precisamente las que preservan rectas. Pero la demostración de que si una transformación preserva rectas es afín es complicada y requiere argumentos de continuidad que no vienen al caso en este momento. Así que mejor nos enfocamos en otra cosa que sí preservan.

3.7.1 Combinaciones afines (el Teorema de 3 en 3)

Otra manera de ver que las transformaciones afines preservan rectas es usando coordenadas baricéntricas. Sea ℓ la recta que pasa por los puntos \mathbf{p} y \mathbf{q} . Entonces ℓ se puede escribir como

$$\ell = \{s\mathbf{p} + t\mathbf{q} \mid s, t \in \mathbb{R}, \quad s + t = 1\}.$$

Veamos qué hace la transformación afín f al punto $s\mathbf{p} + t\mathbf{q}$ cuando $s + t = 1$:

$$\begin{aligned} f(s\mathbf{p} + t\mathbf{q}) &= A(s\mathbf{p} + t\mathbf{q}) + \mathbf{b} \\ &= s(A\mathbf{p}) + t(A\mathbf{q}) + (s + t)\mathbf{b} \\ &= s(A\mathbf{p} + \mathbf{b}) + t(A\mathbf{q} + \mathbf{b}) \\ &= s f(\mathbf{p}) + t f(\mathbf{q}), \end{aligned}$$

donde hemos usado que $\mathbf{b} = s\mathbf{b} + t\mathbf{b}$ pues $s + t = 1$. Así que la recta por \mathbf{p} y \mathbf{q} va, bajo f , a la recta por $f(\mathbf{p})$ y $f(\mathbf{q})$ y preserva las coordenadas baricéntricas. Y esto se puede generalizar.

Una combinación lineal $\sum_{i=0}^k \lambda_i \mathbf{v}_i$ se llama *combinación afín* si los coeficientes cumplen que $\sum_{i=0}^k \lambda_i = 1$. Le hemos llamado “combinación baricéntrica” cuando intervienen dos puntos distintos o tres no colineales, pero en el caso general no nos interesa qué propiedades tengan los puntos involucrados $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_k$. Su importancia, entre otras cosas, es que las transformaciones afines las tratan como las lineales tratan a las combinaciones lineales:

Teorema 3.22 Una transformación $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es afín si y sólo si preserva combinaciones afines; es decir, si y sólo si

$$f\left(\sum_{i=0}^k \lambda_i \mathbf{v}_i\right) = \sum_{i=0}^k \lambda_i f(\mathbf{v}_i) \quad \text{cuando} \quad \sum_{i=0}^k \lambda_i = 1.$$

Demostración. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es afín, entonces se escribe $f(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ para alguna matriz A de $n \times n$ y un vector $\mathbf{b} \in \mathbb{R}^n$. Veamos que preserva combinaciones afines. Sean $\lambda_i \in \mathbb{R}$, $i = 0, 1, \dots, k$, tales que $\sum_{i=0}^k \lambda_i = 1$; y sean $\mathbf{v}_i \in \mathbb{R}^n$, $i = 0, 1, \dots, k$, vectores cualesquiera. Entonces

$$\begin{aligned} f\left(\sum_{i=0}^k \lambda_i \mathbf{v}_i\right) &= A\left(\sum_{i=0}^k \lambda_i \mathbf{v}_i\right) + \mathbf{b} = \sum_{i=0}^k \lambda_i (A\mathbf{v}_i) + \mathbf{b} \\ &= \sum_{i=0}^k \lambda_i (A\mathbf{v}_i) + \sum_{i=0}^k \lambda_i \mathbf{b} = \sum_{i=0}^k \lambda_i (A\mathbf{v}_i + \mathbf{b}) = \sum_{i=0}^k \lambda_i f(\mathbf{v}_i). \end{aligned}$$

Donde se usó la linealidad en la segunda igualdad, y luego fue crucial que $\sum_{i=0}^k \lambda_i = 1$ para poder “distribuir la \mathbf{b} ” en la sumatoria.

Por el otro lado, si nos dan f que preserva combinaciones afines, debemos encontrar la matriz A (o su función lineal asociada) y el vector \mathbf{b} que la definan. Encontrar el vector \mathbf{b} es fácil pues es dónde cualquier transformación afín manda el origen; definamos entonces

$$\mathbf{b} := f(\mathbf{0}).$$

Y, como sucedió con las isometrías, debemos definir la función $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ con

$$g(\mathbf{x}) := f(\mathbf{x}) - \mathbf{b}$$

y demostrar que es lineal para concluir la demostración; pues entonces $f(\mathbf{x}) = g(\mathbf{x}) + \mathbf{b}$. Veamos entonces que g preserva combinaciones lineales.

Sean $\mathbf{v}_1, \dots, \mathbf{v}_k$ vectores en \mathbb{R}^n y $\lambda_1, \dots, \lambda_k$ coeficientes (en \mathbb{R}) arbitrarios. Debemos demostrar que $g\left(\sum_{i=1}^k \lambda_i \mathbf{v}_i\right) = \sum_{i=1}^k \lambda_i g(\mathbf{v}_i)$, pero nuestra hipótesis sólo nos dice cómo se comporta f respecto a combinaciones afines. Podemos convertir esta combinación lineal en una afín usando el vector $\mathbf{0}$ como comodín. Sean

$$\mathbf{v}_0 := \mathbf{0} \quad \text{y} \quad \lambda_0 := 1 - \sum_{i=1}^k \lambda_i$$

de tal manera que $\sum_{i=0}^k \lambda_i = 1$ (ojo, empezamos la sumatoria en 0) y $\sum_{i=0}^k \lambda_i \mathbf{v}_i =$

$\sum_{i=1}^k \lambda_i \mathbf{v}_i$; y ahora sí podemos usar nuestra hipótesis para obtener

$$\begin{aligned} g\left(\sum_{i=1}^k \lambda_i \mathbf{v}_i\right) &= f\left(\sum_{i=1}^k \lambda_i \mathbf{v}_i\right) - \mathbf{b} = f\left(\sum_{i=0}^k \lambda_i \mathbf{v}_i\right) - \left(\sum_{i=0}^k \lambda_i\right) \mathbf{b} \\ &= \sum_{i=0}^k \lambda_i f(\mathbf{v}_i) - \sum_{i=0}^k \lambda_i \mathbf{b} = \sum_{i=0}^k \lambda_i (f(\mathbf{v}_i) - \mathbf{b}) \\ &= \sum_{i=0}^k \lambda_i g(\mathbf{v}_i) = \lambda_0 g(\mathbf{v}_0) + \sum_{i=1}^k \lambda_i g(\mathbf{v}_i) = \sum_{i=1}^k \lambda_i g(\mathbf{v}_i), \end{aligned}$$

pues $g(\mathbf{v}_0) = g(\mathbf{0}) = \mathbf{0}$. Lo cual concluye la demostración del teorema. \square

Vale la pena notar que el concepto de transformación afín se puede generalizar al de función afín cuando la matriz \mathbf{A} no es necesariamente invertible; una *función afín* será entonces una función lineal seguida de una traslación; ni siquiera necesitamos pedir que el dominio y el codominio coincidan. Si se revisa con cuidado el teorema anterior, se notará que vale en esta generalidad. Pero no lo hicimos así pues nuestro interés principal está en las transformaciones.

Dibujo

En el caso que más nos interesa, $n = 2$, este teorema y su demostración tienen como corolario que una transformación afín de \mathbb{R}^2 queda determinada por sus valores en el *triángulo canónico* con vértices $\mathbf{e}_0 := \mathbf{0}, \mathbf{e}_1, \mathbf{e}_2$, y que estos pueden ir a cualquier otro triángulo. Explícitamente, dado un triángulo T con vértices $\mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2$, la transformación afín que manda al triángulo canónico en T es

$$f(\mathbf{x}) = (\mathbf{a}_1 - \mathbf{a}_0, \mathbf{a}_2 - \mathbf{a}_0)\mathbf{x} + \mathbf{a}_0,$$

que cumple claramente que $f(\mathbf{e}_i) = \mathbf{a}_i$. Hay que observar que esta función está bien definida independientemente de quiénes sean $\mathbf{a}_0, \mathbf{a}_1$ y \mathbf{a}_2 , pues tanto la matriz 2×2 como el vector por el que se traslada lo están. Pero que es una transformación sólo cuando $\mathbf{a}_0, \mathbf{a}_1$ y \mathbf{a}_2 no son colineales (que es lo que entendemos como *triángulo*), pues entonces la matriz es invertible.

Y como corolario de este hecho obtenemos el siguiente teorema.

Dibujo

Teorema 3.23 (3 en 3) *Dados dos triángulos con vértices $\mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2$ y $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2$ respectivamente, existe una única transformación afín $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ que cumple $f(\mathbf{a}_i) = \mathbf{b}_i$ para $i = 0, 1, 2$.*

Demostración. Basta componer la inversa de la transformación que manda al triángulo canónico en el $\mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2$ con la que lo manda al $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2$. La unicidad se sigue de la unicidad de las dos transformaciones usadas. \square

Un ejemplo

Para fijar ideas veamos un ejemplo numérico. Sean

$$\begin{aligned}\mathbf{a}_0 &= \begin{pmatrix} 0 \\ 2 \end{pmatrix}, & \mathbf{a}_1 &= \begin{pmatrix} 3 \\ -1 \end{pmatrix}, & \mathbf{a}_2 &= \begin{pmatrix} -1 \\ 1 \end{pmatrix}; \\ \mathbf{b}_0 &= \begin{pmatrix} -5 \\ 6 \end{pmatrix}, & \mathbf{b}_1 &= \begin{pmatrix} 4 \\ 3 \end{pmatrix}, & \mathbf{b}_2 &= \begin{pmatrix} -4 \\ 1 \end{pmatrix}.\end{aligned}$$

Queremos encontrar la transformación afín $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ que cumple $f(\mathbf{a}_i) = \mathbf{b}_i$ para $i = 0, 1, 2$.

Dibujos

Un método infalible (si no cometemos un error en el camino) es seguir la demostración del último teorema (el 6.4). Es decir, la transformación que manda al canónico en el triángulo de las \mathbf{a} 's es

$$\begin{pmatrix} 3 & -1 \\ -3 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \end{pmatrix},$$

luego habrá que encontrar su inversa y componerla con la que manda al canónico en el triángulo de las \mathbf{b} 's. Pero dejaremos este método para que el lector lo siga paso a paso y compruebe el resultado con una solución diferente que es la que aquí seguiremos.

Ataquemos el problema directamente. Estamos buscando una matriz

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \quad \text{y un vector} \quad \mathbf{b} = \begin{pmatrix} \lambda \\ \mu \end{pmatrix}$$

para definir la función $f(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ que cumpla $f(\mathbf{a}_i) = \mathbf{b}_i$ para $i = 0, 1, 2$. Como las \mathbf{a} 's y las \mathbf{b} 's tienen valores numéricos explícitos, estas condiciones nos darán ecuaciones explícitas con las seis incógnitas $\alpha, \beta, \gamma, \delta, \lambda, \mu$ que podemos resolver directamente. La condición $f(\mathbf{a}_0) = \mathbf{b}_0$ nos da

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} -5 \\ 6 \end{pmatrix},$$

que se traduce en una ecuación para cada coordenada, a saber:

$$\begin{aligned}0\alpha + 2\beta + \lambda &= -5 \\ 0\gamma + 2\delta + \mu &= 6.\end{aligned}$$

Análogamente, las condiciones en las otras dos parejas de puntos nos dan las siguientes ecuaciones:

$$\begin{aligned}3\alpha - \beta + \lambda &= 4 \\ 3\gamma - \delta + \mu &= 3 \\ -\alpha + \beta + \lambda &= -4 \\ -\gamma + \delta + \mu &= 1\end{aligned}$$

Resulta que en vez de seis ecuaciones con seis incógnitas, tenemos dos sistemas de tres ecuaciones con tres incógnitas, y además los dos sistemas tienen los mismos coeficientes y sólo se diferencian por las constantes. De hecho los podemos reescribir como

$$A \begin{pmatrix} \alpha \\ \beta \\ \lambda \end{pmatrix} = \begin{pmatrix} -5 \\ 4 \\ -4 \end{pmatrix} \quad \text{y} \quad A \begin{pmatrix} \gamma \\ \delta \\ \mu \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \\ 1 \end{pmatrix},$$

donde $A = \begin{pmatrix} 0 & 2 & 1 \\ 3 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$

Observemos que al resolver los dos sistemas de ecuaciones con el método de eliminar variables sumando múltiplos de una ecuación a otra, estaremos realizando las mismas operaciones, y éstas dependen nada más de los coeficientes, no del nombre de las incógnitas. Así que la manera de resolver los dos sistemas al mismo tiempo será añadiendo dos columnas a la matriz A que correspondan a las constantes de los dos sistemas, y luego sumando múltiplos de renglones a otros para llevar la parte 3×3 de la matriz a una de permutaciones que nos diga el valor explícito de todas las incógnitas. Hagámoslo. La matriz con que empezamos es

$$\left(\begin{array}{ccccc} 0 & 2 & 1 & -5 & 6 \\ 3 & -1 & 1 & 4 & 3 \\ -1 & 1 & 1 & -4 & 1 \end{array} \right).$$

Sumándole 3 por el tercer renglón al segundo (para hacer 0 su primera entrada), operación que podemos denotar con $3\mathbf{r}_3 \vec{+} \mathbf{r}_2$ y que equivale a sumar un múltiplo de una ecuación a otra, obtenemos

$$\begin{aligned} & \xrightarrow{3\mathbf{r}_3 \vec{+} \mathbf{r}_2} \left(\begin{array}{ccccc} 0 & 2 & 1 & -5 & 6 \\ 0 & 2 & 4 & -8 & 6 \\ -1 & 1 & 1 & -4 & 1 \end{array} \right) \xrightarrow{-\mathbf{r}_2 \vec{+} \mathbf{r}_1} \left(\begin{array}{ccccc} 0 & 0 & -3 & 3 & 0 \\ 0 & 2 & 4 & -8 & 6 \\ 1 & -1 & -1 & 4 & -1 \end{array} \right) \\ & \xrightarrow{(1/2)\mathbf{r}_2; -(1/3)\mathbf{r}_1} \left(\begin{array}{ccccc} 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 2 & -4 & 3 \\ 1 & -1 & -1 & 4 & -1 \end{array} \right) \xrightarrow{\mathbf{r}_2 \vec{+} \mathbf{r}_3} \left(\begin{array}{ccccc} 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 2 & -4 & 3 \\ 1 & 0 & 1 & 0 & 2 \end{array} \right) \\ & \xrightarrow{-2\mathbf{r}_1 \vec{+} \mathbf{r}_2; -\mathbf{r}_1 \vec{+} \mathbf{r}_3} \left(\begin{array}{ccccc} 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -2 & 3 \\ 1 & 0 & 0 & 1 & 2 \end{array} \right). \end{aligned}$$

Ahora, recordando nuestras incógnitas, el primer sistema nos dice que $\lambda = -1$, $\beta = -2$ y $\alpha = 1$ y el segundo sistema que $\mu = 0$, $\delta = 3$ y $\gamma = 2$, lo cual lo resuelve por completo. En resumen, la función que buscábamos es

$$f(\mathbf{x}) = \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix} \mathbf{x} + \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

y es fácil comprobarlo; ya sea revisando los valores que se pedían o con el otro método.

EJERCICIO 3.96 Para cada caso, encuentra la transformación afín que manda \mathbf{a}_i en \mathbf{b}_i

- a) $\mathbf{a}_0 = (1, 2)^\top, \mathbf{a}_1 = (1, -1)^\top, \mathbf{a}_2 = (0, -1)^\top$ y $\mathbf{b}_0 = (1, 4)^\top, \mathbf{b}_1 = (4, 4)^\top, \mathbf{b}_2 = (3, 3)^\top$
 b) $\mathbf{a}_0 = (1, 2)^\top, \mathbf{a}_1 = (0, 1)^\top, \mathbf{a}_2 = (3, -1)^\top$ y $\mathbf{b}_0 = (5, 3)^\top, \mathbf{b}_1 = (2, 0)^\top, \mathbf{b}_2 = (6, -1)^\top$
 c) $\mathbf{a}_0 = (1, 2)^\top, \mathbf{a}_1 = (0, 1)^\top, \mathbf{a}_2 = (3, -1)^\top$ y $\mathbf{b}_0 = (10, 5)^\top, \mathbf{b}_1 = (3, 3)^\top, \mathbf{b}_2 = (-4, 4)^\top$

(¿Verdad que conviene relajar la notación sobre vectores?)

3.8 Isometrías II

Ya con la herramienta técnica que nos dan las matrices y en el contexto general que nos dan las transformaciones afines, regresamos brevemente a estudiar isometrías del plano y ver algunas cosas que se nos quedaron pendientes. Ahora sabemos que una isometría de \mathbb{R}^2 es una transformación que se escribe

$$f(\mathbf{x}) = A\mathbf{x} + \mathbf{b}, \quad (3.14)$$

donde la matriz A es ortogonal ($A^\top = A^{-1}$) y es llamada la *parte lineal* de la isometría. Hay de dos tipos. Las que preservan orientación (con $\det(A) = 1$) y las que invierten orientación ($\det(A) = -1$). Para las primeras, la matriz A es una rotación en el origen y para las segundas, una reflexión en una recta por el origen. Puesto que al componer isometrías (o, en general, transformaciones afines) la parte lineal se obtiene componiendo las partes lineales, entonces las isometrías (o las afines) que preservan orientación forman un grupo, denotado con $\mathbf{Iso}^+(2)$ (respectivamente, $\mathbf{Af}^+(2)$). Ojo, las que invierten orientación **no** forman un grupo, ni siquiera contienen a la identidad y la composición de dos de ellas preserva orientación.

Lo primero que queremos ver es que las isometrías que preservan orientación son rotaciones en algún punto o traslaciones.

3.8.1 Rotaciones y traslaciones

Recordemos que definimos la rotación $\rho_{\theta, \mathbf{c}}$ de un ángulo θ con centro \mathbf{c} , conjugando con la traslación de \mathbf{c} al origen, es decir, como $\tau_{\mathbf{c}} \circ \rho_{\theta} \circ \tau_{-\mathbf{c}}$, donde $\tau_{\mathbf{y}}$ es la traslación por \mathbf{y} y ρ_{θ} es rotar un ángulo θ en el origen. Así que, usando matrices, la fórmula para $\rho_{\theta, \mathbf{c}}$ es

$$\begin{aligned} \rho_{\theta, \mathbf{c}}(\mathbf{x}) &= R_{\theta}(\mathbf{x} - \mathbf{c}) + \mathbf{c} \\ &= R_{\theta}\mathbf{x} + (\mathbf{c} - R_{\theta}\mathbf{c}). \end{aligned} \quad (3.15)$$

Está claro entonces que $\rho_{\theta, \mathbf{c}} \in \mathbf{Iso}^+(2)$, pues se escribe como en (3.14) con una \mathbf{b} complicada, pero constante al fin. Y al revés, aunque cierto, no está tan claro. Para ver que $f \in \mathbf{Iso}^+(2)$ dada como en (3.14) es la rotación en algún centro, hay que encontrar un punto fijo, es decir, un punto \mathbf{c} para el cual $f(\mathbf{c}) = \mathbf{c}$ (pues a su centro no lo mueve la rotación); que nos da la ecuación en \mathbf{c}

$$\begin{aligned}\mathbf{c} &= A\mathbf{c} + \mathbf{b} \\ \mathbf{c} - A\mathbf{c} &= \mathbf{b}.\end{aligned}$$

Esto coincide con la expresión (3.15), donde se usa el centro de rotación y nos da la traslación \mathbf{b} en términos de él. Entonces, el problema de encontrar un punto fijo para la transformación (3.14) equivale a encontrar una solución a la ecuación

$$\mathbf{x} - A\mathbf{x} = \mathbf{b}$$

que en realidad es un sistema de ecuaciones. Y como tal, se puede reescribir

$$\begin{aligned}I\mathbf{x} - A\mathbf{x} &= \mathbf{b} \\ (I - A)\mathbf{x} &= \mathbf{b},\end{aligned}\tag{3.16}$$

donde nos ha sido muy útil la suma de matrices. Sabemos que este sistema tiene solución única si y sólo si su determinante es distinto de cero. Para el caso que nos ocupa, cuando A es una matriz de rotación, por un ángulo θ digamos, se tiene

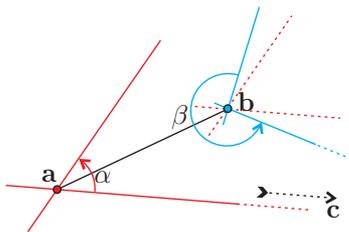
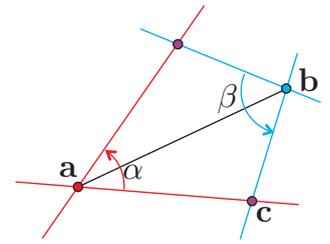
$$\begin{aligned}\det(I - R_\theta) &= \det\begin{pmatrix} 1 - \cos\theta & \text{sen}\theta \\ -\text{sen}\theta & 1 - \cos\theta \end{pmatrix} \\ &= (1 - \cos\theta)^2 + \text{sen}^2\theta \\ &= 1 - 2\cos\theta + \cos^2\theta + \text{sen}^2\theta = 2(1 - \cos\theta).\end{aligned}$$

Podemos concluir que si $\theta \neq 0$ entonces $\det(I - R_\theta) \neq 0$ y tenemos una solución única del sistema (3.16) con $A = R_\theta$, que es el centro de rotación. Así que si $f \in \mathbf{Iso}^+(2)$, entonces es una rotación (alrededor de un punto) o bien una traslación (cuando el ángulo θ es cero en la discusión anterior). Además, hemos demostrado que una rotación (por un ángulo no cero, se sobreentiende) no tiene puntos fijos además de su centro.

Puesto que una rotación o una traslación se puede obtener *moviendo* continuamente (poco a poco) el plano, a veces se le llama a $\mathbf{Iso}^+(2)$ el grupo de *movimientos rígidos del plano*.

Como corolario se obtiene que la composición de dos rotaciones es una nueva rotación. El nuevo centro de rotación se puede obtener analíticamente resolviendo el correspondiente sistema de ecuaciones, pero lo veremos geoméricamente.

Sean $\rho_{\alpha, \mathbf{a}}$ y $\rho_{\beta, \mathbf{b}}$ las rotaciones de ángulos α y β y centros \mathbf{a} y \mathbf{b} respectivamente. La composición de estas dos rotaciones será una rotación de un ángulo $\alpha + \beta$ (pues las partes lineales se componen) pero su centro depende del orden en que se componen. Para obtener el centro de la rotación $\rho_{\beta, \mathbf{b}} \circ \rho_{\alpha, \mathbf{a}}$, se traza el segmento de \mathbf{a} a \mathbf{b} , la línea por \mathbf{a} con ángulo $-\alpha/2$ respecto de este y la línea por \mathbf{b} con ángulo $\beta/2$; la intersección \mathbf{c} de estas dos líneas es el centro de rotación. Para verlo, persígase este punto según las rotaciones. Nótese que su reflejado respecto de la línea por \mathbf{a} y \mathbf{b} es el centro de rotación de la composición en el otro orden, $\rho_{\alpha, \mathbf{a}} \circ \rho_{\beta, \mathbf{b}}$.



Para que nuestra construcción de \mathbf{c} tuviera sentido se necesita que las dos rectas cuya intersección lo definen sean no paralelas y éste es justo el caso cuando $\alpha + \beta \neq 0$. Pero nos podemos aproximar al valor $\beta = -\alpha$ poco a poco, y en este proceso el punto de intersección \mathbf{c} se “va al infinito”, así que las traslaciones son, en cierta manera, “rotaciones con centro al infinito”.

EJERCICIO 3.97 Las *homotecias* son transformaciones de la forma $f(\mathbf{x}) = k\mathbf{x} + \mathbf{b}$, donde a $k \neq 0$ se le llama el *factor de expansión*. Demuestra que las homotecias con factor de expansión no trivial (distinto de 1) tienen un *centro de expansión*, es decir, un punto fijo.

3.8.2 Reflexiones y “pasos”

Consideremos ahora las isometrías que invierten orientación. Se escriben como

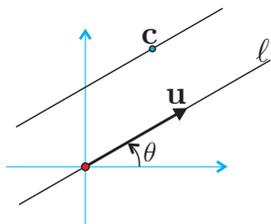
$$f(\mathbf{x}) = E_{\theta}\mathbf{x} + \mathbf{b},$$

donde E_{θ} es una matriz de reflexión. Veamos primero si tiene puntos fijos, para lo cual hay que resolver el correspondiente sistema (3.16). En este caso el determinante es

$$\begin{aligned} \det(I - E_{\theta}) &= \det \begin{pmatrix} 1 - \cos 2\theta & -\sin 2\theta \\ -\sin 2\theta & 1 + \cos 2\theta \end{pmatrix} \\ &= (1 - \cos 2\theta)(1 + \cos 2\theta) - \sin^2 2\theta \\ &= 1 - \cos^2 2\theta - \sin^2 2\theta = 1 - 1 = 0. \end{aligned}$$

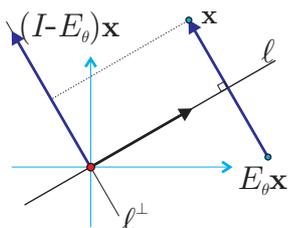
Y por lo tanto no tiene solución única: o no tiene solución o bien tiene muchas; y veremos que ambos casos son posibles.

Para $\mathbf{b} = \mathbf{0}$ tenemos que f es una reflexión y entonces las soluciones debieran ser los puntos de la recta espejo (a un ángulo θ con el eje x positivo), que llamaremos ℓ . Veamos: sea \mathbf{u} el vector unitario que genera a ℓ , es decir, $\mathbf{u} = (\cos \theta, \sin \theta)$. Entonces

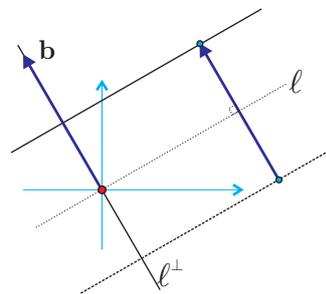


$$\begin{aligned} E_\theta \mathbf{u} &= \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \\ &= \begin{pmatrix} \cos 2\theta \cos \theta + \sin 2\theta \sin \theta \\ \sin 2\theta \cos \theta - \cos 2\theta \sin \theta \end{pmatrix} = \begin{pmatrix} \cos (2\theta - \theta) \\ \sin (2\theta - \theta) \end{pmatrix} = \mathbf{u}, \end{aligned}$$

donde hemos usado las fórmulas trigonométricas para la diferencia de ángulos (3.12). Y por lo tanto $E_\theta(\mathbf{t}\mathbf{u}) = \mathbf{t}(E_\theta\mathbf{u}) = \mathbf{t}\mathbf{u}$; que nos dice que todos los puntos de la recta ℓ satisfacen la ecuación homogénea $(I - E_\theta)\mathbf{x} = \mathbf{0}$. Ahora bien, si para cierta \mathbf{b} , el sistema $(I - E_\theta)\mathbf{x} = \mathbf{b}$ tiene una solución particular, \mathbf{c} digamos, entonces toda la recta $\ell + \mathbf{c}$, paralela a ℓ que pasa por \mathbf{c} , consta de soluciones para el sistema; es una reflexión con espejo $\ell + \mathbf{c}$. Pero nos falta determinar para cuáles \mathbf{b} hay solución.



Para esto, conviene pensar geoméricamente. Para cualquier $\mathbf{x} \in \mathbb{R}^2$, $(I - E_\theta)\mathbf{x} = \mathbf{x} - E_\theta\mathbf{x}$ es el vector que va de $E_\theta\mathbf{x}$ a \mathbf{x} , y como E_θ es una reflexión debe ser perpendicular al espejo; es más, debe ser el doble del vector que va de ℓ a \mathbf{x} y es perpendicular a ℓ (recuérdese que así definimos reflexiones por primera vez). De tal manera que $(I - E_\theta)$ es, como función, la proyección ortogonal a $\ell^\perp : \mathbf{u} \cdot \mathbf{x} = 0$ seguida de la homotecia “multiplicar por 2”. La imagen de la función $(I - E_\theta)$ es entonces ℓ^\perp , y por lo tanto, sólo cuando $\mathbf{b} \in \ell^\perp$ la isometría $f(\mathbf{x}) = E_\theta\mathbf{x} + \mathbf{b}$ tiene puntos fijos. Más concretamente, si $\mathbf{u} \cdot \mathbf{b} = 0$, entonces reflejar en ℓ y luego trasladar por \mathbf{b} , regresa de la recta $\ell + (1/2)\mathbf{b}$ a su lugar (y punto a punto).

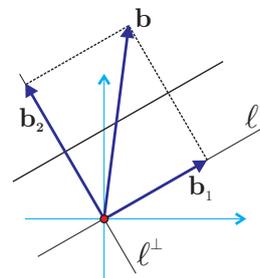


En el caso general, podemos expresar a cualquier $\mathbf{b} \in \mathbb{R}^2$ como una suma de sus *componentes* respecto a la base ortonormal $\mathbf{u}, \mathbf{u}^\perp$, es decir como

$$\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2,$$

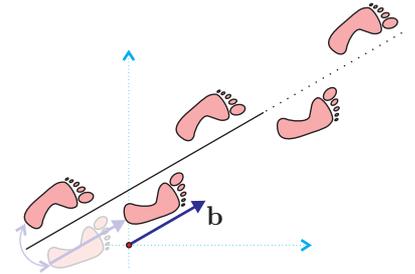
con $\mathbf{b}_1 \in \ell$ y $\mathbf{b}_2 \in \ell^\perp$ (de hecho, por el teorema de bases ortonormales: $\mathbf{b}_1 := (\mathbf{b} \cdot \mathbf{u})\mathbf{u}$ y $\mathbf{b}_2 := (\mathbf{b} \cdot \mathbf{u}^\perp)\mathbf{u}^\perp$). Y entonces la transformación $f(\mathbf{x}) = E_\theta\mathbf{x} + \mathbf{b}$ se puede escribir como

$$f(\mathbf{x}) = (E_\theta\mathbf{x} + \mathbf{b}_2) + \mathbf{b}_1,$$



que es una reflexión en la recta $\ell + (1/2) \mathbf{b}_2$ (por el caso anterior) y que en la fórmula es $(E_\theta \mathbf{x} + \mathbf{b}_2)$, seguida de una traslación $(\dots + \mathbf{b}_1)$ en la dirección del espejo. Nótese que si $\mathbf{b}_1 \neq \mathbf{0}$, entonces f no tiene puntos fijos, los puntos de un lado del espejo van al otro lado y los que están en el espejo se trasladan en él, aunque el espejo $\ell + (1/2) \mathbf{b}_2$ como línea sí se queda en su lugar (todos sus puntos se mueven dentro de ella) y éste es entonces el caso en que nuestro sistema de ecuaciones que detecta puntos fijos no tiene solución.

A una isometría tal, una reflexión seguida de una traslación no trivial ($\mathbf{b} \neq \mathbf{0}$) en la dirección de la línea espejo, la llamaremos un *paso* pues genera la simetría que tiene un caminito de huellas (infinito) y corresponde al acto de dar un paso: cambiar de pie (reflejar) y adelantarlo un poco (trasladar). Si la aplicamos dos veces, nos da una traslación (por $2\mathbf{b}$), y si la seguimos aplicando (a ella y a su inversa) nos da un subgrupo de isometrías que son las simetrías de un camino infinito. Pero en fin, hemos demostrado:



Teorema 3.24 Una isometría que invierte orientación es un paso o una reflexión (paso con traslación trivial). \square

EJERCICIO 3.98 Con la notación de los párrafos anteriores, demuestra con coordenadas que $(I - E_\theta) \mathbf{u}^\perp = 2\mathbf{u}^\perp$. (Tienes que usar las identidades trigonométricas para el doble del ángulo.)

EJERCICIO 3.99 Demuestra que si f es una isometría que invierte orientación, entonces $f^2 = f \circ f$ es una traslación.

EJERCICIO 3.100 Sea $f(\mathbf{x}) = E_\theta \mathbf{x} + \mathbf{b}$. Encuentra la expresión para f^{-1} y argumentala geoméricamente.

EJERCICIO 3.101 Demuestra analíticamente que la composición de dos reflexiones en líneas paralelas es una traslación en la dirección ortogonal a sus espejos.

3.8.3 Homotecias y semejanzas

El clásico *zoom* que aumenta (o disminuye) de tamaño una figura es lo que llamaremos una *homotecia*. Siempre tiene un centro de expansión, que cuando es el origen resulta ser una transformación lineal con matriz asociada

$$kI = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix},$$

donde $k > 0$; para $k > 1$ magnifica (es un *zoom* estrictamente hablando) y para $k < 1$ disminuye. Si la componemos con una traslación, por $\mathbf{b} \in \mathbb{R}^2$ digamos, se obtiene

una *homotecia* de factor k . Su centro de expansión \mathbf{c} es el punto que se queda fijo y que se obtiene entonces resolviendo la ecuación

$$k\mathbf{x} + \mathbf{b} = \mathbf{x}.$$

Una *semejanza* es una transformación afín que preserva ángulos. Claramente las homotecias y las isometrías son semejanzas. Veremos que estas últimas se obtienen simplemente al dejar interactuar a las dos primeras.

Teorema 3.25 Si $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ es una semejanza, entonces existen $k > 0$, $A \in \mathbf{O}(2)$ y $\mathbf{b} \in \mathbb{R}^2$ tales que

$$f(\mathbf{x}) = kA\mathbf{x} + \mathbf{b}.$$

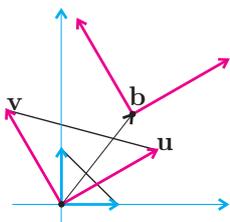
Demostración. Consideremos la transformación lineal $\mathbf{g}(\mathbf{x}) = f(\mathbf{x}) - \mathbf{b}$, donde $\mathbf{b} := f(\mathbf{0})$, y notemos que preserva ángulos pues las traslaciones lo hacen. Sea $B = (\mathbf{u}, \mathbf{v})$ la matriz asociada a \mathbf{g} . Como los vectores canónicos son ortogonales entonces \mathbf{u} y \mathbf{v} (sus imágenes bajo \mathbf{g}) también lo son. Además tienen la misma norma pues si no fuera así, los ángulos chiquitos del triángulo $\mathbf{0}, \mathbf{u}, \mathbf{v}$ no medirían $\pi/4$ que son los del triángulo canónico. Sean $k = |\mathbf{u}| = |\mathbf{v}|$ y

$$A = \frac{1}{k}B.$$

Se tiene que $A \in \mathbf{O}(2)$, pues sus columnas son ortonormales, y además

$$f(\mathbf{x}) = \mathbf{g}(\mathbf{x}) + \mathbf{b} = B\mathbf{x} + \mathbf{b} = kA\mathbf{x} + \mathbf{b}.$$

□



EJERCICIO 3.102 Encuentra la expresión de la homotecia de factor k y centro \mathbf{c} .

EJERCICIO 3.103 Demuestra que una transformación $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ es una semejanza si y sólo si existe $k > 0$ tal que $d(f(\mathbf{x}), f(\mathbf{y})) = kd(\mathbf{x}, \mathbf{y})$ para todo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$.

3.9 *Simetría plana

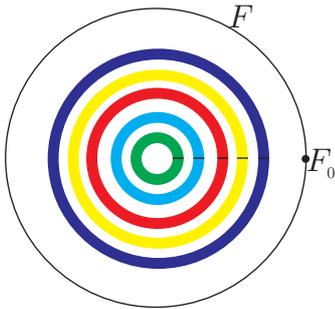
En la Sección 3.3.2 se definió el grupo de simetrías de una figura plana; en esta última sección estudiamos diferentes instancias de cuales pueden ser estos grupos y cómo son sus figuras asociadas.

3.9.1 El Teorema de Leonardo

Versión topológica

Por una *figura acotada* F entendemos un subconjunto del plano que a su vez está contenido en el interior de algún círculo; y por medio de una homotecia, podemos suponer que está dibujada en una hoja de papel. El problema general que nos guía es qué podemos decir sobre su *grupo de simetrías*

$$\mathbf{Sim}(F) = \{f \in \mathbf{Iso}(2) \mid f(F) = F\}.$$



O bien, ¿qué grupos son de este tipo? Además de los grupos cíclicos y diédricos que vimos en la Sección 3.3.2 tenemos otros. Por ejemplo, si F es un círculo centrado en el origen, entonces $\mathbf{Sim}(F) = \mathbf{O}(2)$. Nótese que también un conjunto de anillos concéntricos —un tiro al blanco— tiene este mismo grupo de simetrías. La figura F se genera por un pedacito F_0 —un simple punto para el círculo y para el tiro al blanco un intervalo por cada anillo— juntando todas las copias $f(F_0)$ al correr $f \in \mathbf{O}(2)$; es decir, dejando que el grupo “barra” o “actúe” en una “figura fundamental” F_0 . Y como en este caso el grupo es “continuo”

(los ángulos de las rotaciones son un continuo) la figura que se genera también es “continua”. Podemos pensar que la figura F se obtiene “poniendo tinta” (de diferentes colores para el tiro al blanco) en los puntos de F_0 y luego dejando que el grupo mueva a F_0 y vaya dejando huella. Todas las figuras que tienen simetría se deben obtener así, una porción de ella y el grupo la definen: “simetría, se dice coloquialmente, es una relación armoniosa entre las partes y el todo”.

Otro ejemplo es tomar el grupo generado por una *rotación irracional*. Sea θ un ángulo para el cual $n\theta$ (con $n \in \mathbb{Z}$) nunca es múltiplo de 2π , es decir para el que no existen $p, q \in \mathbb{Z}$ tales que $\theta = (p/q)2\pi$ (por ejemplo, $\theta = \sqrt{2}\pi$, o $\theta = 1$ pues π como número ya es irracional)². Entonces el conjunto $\langle R_\theta \rangle := \{R_{n\theta} \mid n \in \mathbb{Z}\}$ es un grupo de rotaciones; para $n = 0$ tenemos la identidad y la composición de dos claramente está en él. En este caso, si tomamos un punto \mathbf{x} distinto del origen (en nuestra notación anterior, $F_0 = \{\mathbf{x}\}$) las copias $R_{n\theta}(\mathbf{x})$ van poblando poco a poco (al crecer n en ambos sentidos) el círculo \mathcal{C} con centro en \mathbf{O} y que contiene a \mathbf{x} , pues nunca caen sobre un punto ya “pintado”, pero nunca llegan a llenar todo el círculo; lo “pueblan”, lo “atascan”, son “densos en él” pero no lo llenan. Si definimos $F = \{R_{n\theta}(\mathbf{x}) \mid n \in \mathbb{Z}\}$ entonces $\langle R_{n\theta} \rangle \subset \mathbf{Sim}(F)$ (no pusimos la igualdad pues aparecen reflexiones como simetrías). Sin embargo, esta “figura” no se le aparecería nunca a un artista (o vil ser humano) pintando en un papel con lápiz o plumón de cierto grosor. Para eliminar de nuestras consideraciones este tipo de figuras raras o abstractas, que sólo se le aparecen a un matemático quisquilloso, tendríamos que usar conceptos de “topología de conjuntos” que expresen nuestras nociones intuitivas de “continuidad”.

La versión topológica del Teorema de Leonardo consiste en poner restricciones topológicas a las figuras (con “cerradas y acotadas” basta) para concluir que su grupo

²Un número es *irracional* si no se puede expresar como una fracción p/q con p y q enteros. Por ejemplo $\sqrt{2}$ o π ; aunque sea difícil demostrarlo.

de simetrías es $\mathbf{O}(2)$ o bien finito. Intuitivamente, lo de “cerrado” implica que si sus puntos se acumulan cerca de otro, ese también está, de tal manera que los “huecos” de nuestro último ejemplo se llenarían y las simetrías crecerían a todas las rotaciones y las reflexiones ($\mathbf{O}(2)$). Pero dejemos ya de lado esta motivación para el estudio de la topología de conjuntos (cómo controlar particularidades sobre la “continuidad” de subconjuntos de \mathbb{R}^2) y entremos de lleno en lo que sí estamos en posición de hacer: poner restricciones sobre el grupo en vez de sobre la figura; usar técnicas de geometría discreta. Demostraremos entonces la versión “grupera” del Teorema de Leonardo.

Versión grupera

Teorema 3.26 (Leonardo) *Si G es un grupo finito de isometrías del plano entonces G es diédrico o cíclico.*

La demostración consta de cuatro pasos muy bien diferenciados y con ideas de distinta índole en cada uno.

Paso 1 (El centro de simetría). Como G es un grupo finito, sea $n \geq 1$ su cardinalidad. Entonces podemos enumerar sus elementos como

$$G = \{g_0, g_1, \dots, g_{n-1}\},$$

donde no importa el orden, salvo que podemos distinguir a g_0 como la identidad, que sabemos que está en G ; y digamos de una vez que a $g_0 = \text{id}_{\mathbb{R}^2}$ también lo llamamos g_n para facilitar la notación.

Vamos a demostrar que existe un punto $\mathbf{c} \in \mathbb{R}^2$ *invariante bajo G* , es decir, tal que $g_i(\mathbf{c}) = \mathbf{c}$ para toda $i = 1, \dots, n$, que es naturalmente el *centro de simetría* del grupo. Encontrarlo es fácil. Tomamos cualquier punto \mathbf{x} de \mathbb{R}^2 , y \mathbf{c} es el baricentro de su *órbita* $G\mathbf{x} := \{g_1(\mathbf{x}), \dots, g_n(\mathbf{x})\}$. Es decir, sea

$$\mathbf{c} := \sum_{i=1}^n \left(\frac{1}{n}\right) g_i(\mathbf{x}).$$

Dibujo

(Estamos generalizando la noción de baricentro de un triángulo, 3 puntos, a conjuntos arbitrarios, pero finitos, de puntos.) Para ver que \mathbf{c} es invariante bajo G , usamos el hecho de que las isometrías son funciones afines y que éstas preservan combinaciones afines (Teorema 3.22), pues definimos \mathbf{c} como una combinación tal ($\sum_{i=1}^n (1/n) = 1$). Se tiene entonces que para cualquier $g \in G$ (podemos eliminar su subíndice para aliviar notación):

$$g(\mathbf{c}) = \sum_{i=1}^n \left(\frac{1}{n}\right) g(g_i(\mathbf{x})) = \sum_{i=1}^n \left(\frac{1}{n}\right) (g \circ g_i)(\mathbf{x}). \quad (3.17)$$

Afirmamos que esta última sumatoria es justo la original (que definió a \mathbf{c}) salvo que el orden de los sumandos está permutado. Esto se sigue de que “componer con g ” es una biyección del grupo en sí mismo (una permutación de los elementos del grupo).

Más formalmente, definamos la función *multiplicación* por g como

$$\begin{aligned}\mu_g : G &\rightarrow G \\ \mu_g(g_i) &= g \circ g_i.\end{aligned}$$

Que está bien definida (que siempre caemos en elementos de G) se sigue de que G es cerrado bajo composición; y para ver que es biyección basta exhibir su inversa: $\mu_{g^{-1}}$, multiplicar por g^{-1} (de nuevo bien definida pues $g^{-1} \in G$). Claramente se tiene que $\mu_{g^{-1}} \circ \mu_g = \text{id}_G$ pues

$$(\mu_{g^{-1}} \circ \mu_g)(g_i) = \mu_{g^{-1}}(\mu_g(g_i)) = \mu_{g^{-1}}(g \circ g_i) = g^{-1} \circ g \circ g_i = g_i$$

y análogamente, $\mu_g \circ \mu_{g^{-1}} = \text{id}_G$.

Así que la combinación afín que da $g(\mathbf{c})$ (3.17) es un simple reordenamiento de la que da \mathbf{c} y por lo tanto $g(\mathbf{c}) = \mathbf{c}$ para toda $g \in G$. Hemos demostrado que G tiene centro de simetría y, en particular, que no contiene ni traslaciones ni pasos pues estas transformaciones no tienen puntos fijos.

Paso 2 (Vámonos al origen). Ahora veremos que podemos suponer que el centro de simetría de G es el origen, y simplificar así nuestras consideraciones subsecuentes, usando nuestro viejo truco de conjugar con la traslación de \mathbf{c} al origen. Más formalmente, sea

$$G_0 = \tau_{-\mathbf{c}} \circ G \circ \tau_{\mathbf{c}} := \{\tau_{-\mathbf{c}} \circ g \circ \tau_{\mathbf{c}} \mid g \in G\},$$

(el conjunto de los conjugados de todos los elementos de G). Nótese que si $g \in G$ entonces $(\tau_{-\mathbf{c}} \circ g \circ \tau_{\mathbf{c}})(\mathbf{0}) = \tau_{-\mathbf{c}}(g(\tau_{\mathbf{c}}(\mathbf{0}))) = \tau_{-\mathbf{c}}(g(\mathbf{c})) = \tau_{-\mathbf{c}}(\mathbf{c}) = \mathbf{0}$ así que $\mathbf{0}$ es el centro de simetría de G_0 . Es muy fácil ver que G_0 también es un grupo y que es *isomorfo* al grupo original G (si G fuera el grupo de simetrías de una figura F , G_0 es el grupo de simetrías de $\tau_{-\mathbf{c}}(F)$ — F trasladada al origen— y ambas figuras claramente tienen la “misma” simetría), pues la conjugación respeta composición

$$\tau_{-\mathbf{c}} \circ (h \circ g) \circ \tau_{\mathbf{c}} = \tau_{-\mathbf{c}} \circ h \circ (\tau_{\mathbf{c}} \circ \tau_{-\mathbf{c}}) \circ g \circ \tau_{\mathbf{c}} = (\tau_{-\mathbf{c}} \circ h \circ \tau_{\mathbf{c}}) \circ (\tau_{-\mathbf{c}} \circ g \circ \tau_{\mathbf{c}})$$

e inversos $((\tau_{-\mathbf{c}} \circ g \circ \tau_{\mathbf{c}})^{-1} = \tau_{\mathbf{c}}^{-1} \circ g^{-1} \circ \tau_{-\mathbf{c}}^{-1} = \tau_{-\mathbf{c}} \circ g^{-1} \circ \tau_{\mathbf{c}})$. De hecho, la función inversa de la conjugación es la conjugación por la inversa, $G = \tau_{\mathbf{c}} \circ G_0 \circ \tau_{-\mathbf{c}}$. Así que si entendemos G_0 (con centro de simetrías $\mathbf{0}$) entenderemos el grupo original G . Simplificando nuestra notación, suponemos en adelante que G tiene como centro de simetría el origen, que todos sus elementos dejan fijo el $\mathbf{0}$ y, por lo tanto, que es subgrupo de las transformaciones ortogonales. Es decir, en adelante $G \subset \mathbf{O}(2)$.

Paso 3 (Cuando G preserva orientación). Supongamos que todos los elementos de $G \subset \mathbf{O}(2)$ preservan orientación. Entonces todos los elementos de G son rotaciones con un cierto ángulo θ y tenemos una nueva enumeración natural (no necesariamente la que habíamos usado, salvo la primera) usando sus ángulos:

$$\begin{aligned}G &= \{R_{\theta_0}, R_{\theta_1}, R_{\theta_2}, \dots, R_{\theta_{n-1}}\}, \\ \text{donde } 0 &= \theta_0 < \theta_1 < \theta_2 < \dots < \theta_{n-1} < 2\pi.\end{aligned}$$

Queremos demostrar que $\theta_1 = 2\pi/n$ y que $\theta_i = i\theta_1$, lo cual es justo decir que $G = C_n$, que G es el grupo cíclico de orden n generado por la rotación $R_{2\pi/n}$. Puesto que θ_1 es el más chico de estos ángulos (positivos) y todos sus múltiplos $2\theta_1, 3\theta_1, \dots$ están en el conjunto de ángulos de G (pues se obtienen al componer R_{θ_1} consigo misma), entonces la primera vez que $k\theta_1$ exceda a 2π tiene que ser justo por θ_1 y después de haber pasado por todos los θ_i ; es decir, $\theta_1 = 2\pi/n$.

Paso 4 (Jaque mate). Supongamos, por último, que algún elemento de G , digamos que f_0 , no preserva orientación. Consideremos el subgrupo de G que sí preserva orientación, y llamémosle G^+ ; es decir, sea

$$G^+ := \{g \in G \mid \det(g) = 1\}.$$

Claramente G^+ es un grupo (contiene a la identidad y tanto el producto como los inversos de transformaciones que preservan orientación preservan orientación) que, por el paso anterior, es un grupo cíclico C_m para alguna $m < n$ (como hasta ahora, n es el *orden* del grupo G , es decir, $n = \#G$).

Primero veamos que $2m = n$. En el Paso 1 definimos la biyección de G “multiplicar por $_$ ”; veamos qué hace μ_{f_0} (multiplicar por nuestro elemento dado que no preserva orientación). Al componer con una transformación que invierte orientación (f_0), a una que la preserva ($g \in G^+$) nos da una que la invierte ($\mu_{f_0}(g) \notin G^+$) y a una que la invierte ($g \notin G^+$), resulta que preserva ($\mu_{f_0}(g) \in G^+$); podemos resumir este trabalenguas como

$$g \in G^+ \iff \mu_{f_0}(g) \notin G^+.$$

Esto quiere decir que μ_{f_0} es una biyección entre G^+ y su complemento, G^- digamos —y vale la notación aunque no sea grupo—, de tal manera que ambos tienen el mismo número de elementos, m habíamos dicho, y entonces $n = 2m$. Por el paso anterior, G^+ es cíclico de orden m y está generado por la rotación $R_{2\pi/m}$.

Por otro lado, como $G \subset O(2)$, f_0 tiene que ser una reflexión en una línea ℓ_0 que pasa por el origen. En el Ejercicio 3.82 se pidió demostrar analíticamente que $f_0 \circ R_\theta$ es la reflexión en la línea $R_{-\theta/2}(\ell_0)$. Geométricamente, si aplicamos $f_0 \circ R_\theta$ a está última línea, R_θ la gira a la que está a un ángulo $\theta/2$ más allá de ℓ_0 y luego f_0 la regresa a su lugar, así que es cierto lo que se pidió (aunque no hayan hecho su tarea). De tal manera que

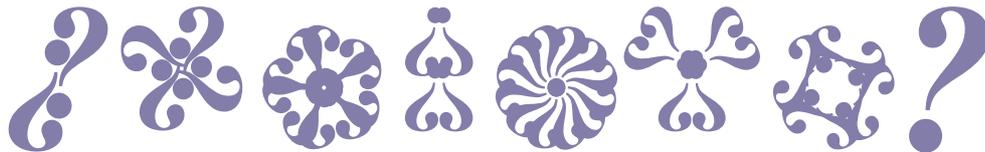
$$f_1 := \mu_{f_0}(R_{2\pi/m}) = f_0 \circ R_{2\pi/m} \in G^-$$

es la reflexión en la recta ℓ_1 cuyo ángulo hacia ℓ_0 es π/m . Por nuestra definición de grupo diédrico, las reflexiones f_0 y f_1 generan a D_m ; pero también generan a G y al verlo concluiremos. A saber, $R_{2\pi/m}$ se obtiene como $f_0 \circ f_1$, pues $f_0 = f_1^{-1}$, así que generan a G^+ (generado por $R_{2\pi/m}$) y luego cualquier elemento de G^- se obtiene como $\mu_{f_0}(g) = f_0 \circ g$ para algún $g \in G^+$. Así que $G = D_m$ y “colorín colorado, el Teorema de Leonardo ha quedado demostrado”. \square

Vale la pena hacer notar de la demostración anterior que, aunque el grupo cíclico quede totalmente definido al fijar su centro de simetría (la \mathbf{c} que encontramos en el Paso 1), o bien que esté conjuntistamente determinado como subgrupo de $\mathbf{O}(2)$, no sucede lo mismo para el diédrico, o **los** diédricos para ser más claros. Nótese que nunca usamos el ángulo de la recta ℓ (el espejo básico) en la que se basó (valga la redundancia) el final de la demostración. Para cualquier recta ℓ que pasa por el origen funciona la demostración y tenemos una instancia, conjuntistamente distinta, del diédrico. Si queremos **el diédrico** con ℓ igual al eje de las x , tendremos que conjugar, como en el Paso 2, pero ahora con la rotación que lleva ℓ a la recta canónica.

Quizá quede más claro a qué nos referimos con eso de “instancias” de un cierto grupo si enunciamos el Teorema de Leonardo como “*un grupo finito de isometrías es conjugado del cíclico $\langle \mathbf{R}_{2\pi/n} \rangle$ o del diédrico $\langle \mathbf{E}_0, \mathbf{E}_{\pi/n} \rangle$* ” que fue, junto con la observación precedente, lo que realmente demostramos.

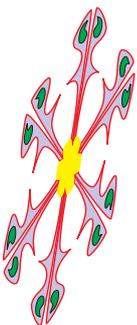
EJERCICIO 3.104 Identifica el grupo de simetrías de las siguientes figuras y, a ojo de buen cubero, píntales su centro de simetría. ¿Qué algunas no tienen tal? ¿Dónde está un error formal (no esencial) de la demostración anterior? Corrígelo.



Leonardo afín

El ojo humano, y la naturaleza en general, identifica simetrías aunque, estrictamente hablando, no las haya; vemos el bulto y no nos preocupamos por los milímetros; el ejercicio anterior tiene sentido aunque es seguro que en el dibujo, o en la impresión, se cometieron errores; el cuerpo humano tiene simetría \mathbf{D}_1 . E igualmente existe la simetría más allá del rígido grupo de isometrías.

Supongamos que una figura F tiene grupo de simetrías \mathbf{G} (será \mathbf{D}_6 en la figura que viene). Si f es una transformación afín que no es isometría, entonces $f(F)$ (ahora sí: la figura) ya no tiene, en general y estrictamente hablando, las mismas simetrías. Sin embargo, es evidente que sí las tiene: las vemos, las detectamos: tiene simetría afín. Podemos definir el *grupo de simetría afín* de (cualquier figura) F (luego regresamos a nuestro ejemplo, $f(F)$) como



$$\mathbf{Sim}_{\mathbf{Af}}(F) = \{g \in \mathbf{Af}(2) \mid g(F) = F\}$$



y la demostración de que es un subgrupo de $\mathbf{Af}(2)$ es, *verbatim*,³ la de que $\mathbf{Sim}(F)$ es subgrupo de $\mathbf{Iso}(2)$.

En nuestro ejemplo ($f(F)$, el dibujo de al lado), $\mathbf{Sim}_{\mathbf{Af}}(f(F))$ será justo el conjugado por f^{-1} del grupo $G = \mathbf{Sim}(F)$; que, puesto que f no es isometría, simplemente es afín, sólo es un subgrupo de las transformaciones afines. Veamos. Si $g \in G$ entonces

$$\begin{aligned} (f \circ g \circ f^{-1})(f(F)) &= (f \circ g)(f^{-1}(f(F))) \\ &= (f \circ g)(F) = f(g(F)) = f(F) \end{aligned}$$

implica que $(f \circ g \circ f^{-1}) \in \mathbf{Sim}_{\mathbf{Af}}(f(F))$. Hemos demostrado que

$$f \circ \mathbf{Sim}(F) \circ f^{-1} \subset \mathbf{Sim}_{\mathbf{Af}}(f(F)),$$

donde estamos extendiendo naturalmente nuestra definición de grupo conjugado (dada en el Paso 2 del Teorema de Leonardo) al contexto más amplio de cualquier grupo (en nuestro caso a $\mathbf{Af}(2)$). Intentemos demostrar la otra contención, pues habíamos aventurado justo la igualdad.

dibujo arriba

Si $g \in \mathbf{Sim}_{\mathbf{Af}}(f(F))$, entonces un cálculo equivalente, aunque más fácil, implica que $(f^{-1} \circ g \circ f)(F) = F$. Pero de esto no podemos concluir que $(f^{-1} \circ g \circ f) \in \mathbf{Sim}(F)$ sino sólo que $(f^{-1} \circ g \circ f) \in \mathbf{Sim}_{\mathbf{Af}}(F)$, pues no sabemos si $(f^{-1} \circ g \circ f)$ es isometría o no: en principio, es simplemente afín. Habíamos aventurado la igualdad justa en el caso concreto del dibujo $f(F)$ que estamos estudiando, porque a simple vista es evidente: el “chuequito” $f(F)$ tiene únicamente las 12 simetrías afines que corresponden a las 12 simetrías métricas del “redondito” F . Y aunque no lo hayamos podido demostrar en el primer intento, surgen las preguntas naturales de si los grupos de simetrías afines de las figuras acotadas (versión topológica) o si los subgrupos finitos (versión grupera) se comportan como los rígidos (teoremas de Leonardo). Más en concreto, y en base a lo que hemos logrado con isometrías, olvidándonos de las figuras y concentrándonos en los grupos, nos debemos hacer dos preguntas:

Pregunta 1. Si G es un subgrupo finito de $\mathbf{Af}(2)$, ¿será cierto que es isomorfo a un cíclico o un diédrico?

Pregunta 2. Si G es un subgrupo finito de $\mathbf{Af}(2)$, ¿será cierto que es conjugado del cíclico $\langle \mathbf{R}_{2\pi/n} \rangle$ o del diédrico $\langle \mathbf{E}_0, \mathbf{E}_{\pi/n} \rangle$ para alguna n ?

Dejamos estas preguntas al lector, pues gran parte del trabajo ya se hizo. Por ejemplo, el Paso 1 en la demostración de Leonardo funciona *verbatim* en el caso afín (de hecho sólo en esa cualidad se basó); y ya que lo mencionamos, ahí está el error que dejamos buscar como ejercicio: para que el centro \mathbf{c} esté bien definido, que no dependa de \mathbf{x} , se necesita que el grupo tenga más de dos elementos. Y el Paso 2, irse al origen, es igual, pero acabamos obteniendo que G es subgrupo de $\mathbf{GL}(2)$ en vez de $\mathbf{O}(2)$. Los problemas aparecen en el Paso 3 que aparentemente se basó en

³Dícese de cuando es literalmente idéntica, salvo los cambios obvios.

sumas, restas y desigualdades de ángulos como números reales, pero en el fondo son una forma más mundana de trabajar con la estructura de grupo, y el orden cíclico que dan se puede obtener al tomar la órbita ($G\mathbf{x} = \{\mathbf{g}(\mathbf{x}) \mid \mathbf{g} \in G\}$) de algún punto. Algunos de los ejercicios siguientes pueden ayudar.

EJERCICIO 3.105 Sea $\mathbf{g} \in \mathbf{GL}^+(2)$ (donde $\mathbf{GL}^+(2) := \{\mathbf{g} \in \mathbf{GL}(2) \mid \det(\mathbf{g}) > 0\}$), tal que para algún vector $\mathbf{v} \in \mathbb{R}^2$, $\mathbf{v} \neq \mathbf{0}$, se tiene que $\mathbf{g}(\mathbf{v}) = t\mathbf{v}$. Demuestra que si \mathbf{g} tiene orden finito, es decir, que si $\mathbf{g}^n = \text{id}$ para alguna n , entonces $\mathbf{g} = \text{id}$ o bien $\mathbf{g} = -\text{id}$. (Primero demuestra que $t = \pm 1$ pensando en la acción de \mathbf{g} sobre la recta generada por \mathbf{v} , $\langle \mathbf{v} \rangle$; y luego usa la hipótesis de que preserva orientación: toma algún vector fuera de $\langle \mathbf{v} \rangle$ y expresa lo que le hace \mathbf{g}).

EJERCICIO 3.106 Demuestra que si $G \subset \mathbf{GL}^+(2)$ es finito, entonces $G\mathbf{v} = \{\mathbf{g}(\mathbf{v}) \mid \mathbf{g} \in G\}$ tiene la cardinalidad de G para cualquier $\mathbf{v} \neq \mathbf{0}$. Y que además se tiene que si $\mathbf{g}(\mathbf{v})$ es paralelo a $\mathbf{h}(\mathbf{v})$ para algunos \mathbf{g} y \mathbf{h} en G , entonces $\mathbf{g} = \mathbf{h}$ o bien $\mathbf{g} = -\mathbf{h}$ (ojo: en este último caso, sus ángulos son opuestos).

EJERCICIO 3.107 Sea $\mathbf{g} \in \mathbf{GL}(2)$, $\mathbf{g} \neq \text{id}$; demuestra que su conjunto de puntos fijos, $\{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{g}(\mathbf{x}) = \mathbf{x}\}$, es una recta que pasa por el origen o sólo el origen (revisa la Sección 8.1). Concluye que los puntos fijos de una transformación afín distinta de la identidad son el vacío, o un punto o una recta.

EJERCICIO 3.108 Demuestra que $\mathbf{Sim}_{\text{Af}}(\mathbb{S}^1) = \mathbf{O}(2) = \mathbf{Sim}(\mathbb{S}^1)$; es decir, que toda simetría afín del círculo unitario es una isometría.

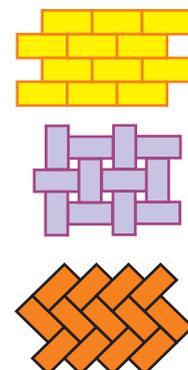
EJERCICIO 3.109 Sea \mathcal{E} la elipse dada por la ecuación

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

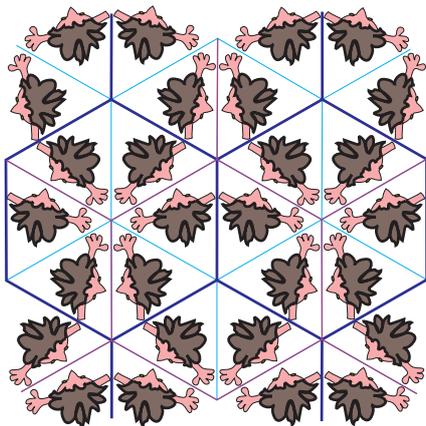
Demuestra que $\mathbf{Sim}_{\text{Af}}(\mathcal{E})$ es conjugado de $\mathbf{O}(2)$ (para esto, encuentra una matriz que mande a \mathbb{S}^1 en \mathcal{E}).

3.9.2 Grupos discretos y caleidoscópicos

Los teoremas tipo Leonardo hablan de las posibles simetrías de las figuras acotadas. Pero además hay figuras infinitas que tienen simetría, y de ellas, más bien de sus grupos, queremos hablar en esta sección. Las más comunes son los pisos por los que caminamos a diario y las paredes que nos rodean. Aunque no sean infinitos, una pequeña porción sugiere claramente cómo debía continuarse de tal manera que cada región sea “igualita” a cualquier otra. Tienen, o dan la idea, de una “regla” para ir pegando piezas que llamaremos *losetas* y que juntas todas ellas llenan el plano. A tal descomposición del plano en losetas la llamaremos un *mosaico*; cada loseta es copia isométrica de una o varias que forman el *muestrario*; por ejemplo, en el mosaico de en medio un “ladrillo” rectangular y un “huequito” cuadrado forman el muestrario y en los otros dos el muestrario consta de una sola pieza. Con una sola pieza se pueden armar varios mosaicos (ver otros ejemplos, además de los dos al margen, en el segundo ejercicio que viene, el 3.111) usando diferentes “reglas”. Y esto de las “reglas” ya lo conocemos: consisten en el fondo de dar un subgrupo de isometrías; que llamaremos el *grupo de simetrías* del mosaico: las isometrías que lo mandan en sí mismo (que dejan a la figura completa —aunque infinita— en su lugar).



Un último y sugestivo ejemplo, antes de entrar a la teoría, surge de los caleidoscopios clásicos que en su versión física, real, consisten de tres espejos rectangulares



del mismo tamaño que forman un prisma triangular equilátero, de tal manera que al observar desde un extremo de él, la figura plana al otro lado (que puede ser cambiante) se reproduce para dar la impresión de un plano infinito. Este es otro ejemplo de un mosaico plano. Podemos pensar en una figura plana (*la figura fundamental*) dentro de un triángulo equilátero de líneas espejo: ésta se reproduce por las tres reflexiones y sus composiciones para tapizar todo el plano y generar una *figura caleidoscópica*; como mosaico, el muestrario consiste de una loseta que es el triángulo equilátero con un dibujo.

El grupo de simetrías de este mosaico es el conjunto de todas las posibles composiciones de las tres reflexiones básicas o *generadoras*; y un elemento del grupo lleva a la figura (o loseta) fundamental a una cierta loseta, que no sólo es un triángulo de la descomposición del plano en triángulos equiláteros, sino que la figura (en el dibujo, el “manquito”) le da una orientación y una posición específicas. En su versión física, la composición que produce una cierta loseta tiene que ver con los espejos y el orden en que rebota el rayo de luz que emana de la figura real para producir la imagen dada.

Basta de ejemplos (ojéense los ejercicios) y hagamos algo de teoría. Por nuestra experiencia reciente, será más fácil definir con precisión a los grupos involucrados en los mosaicos que a los mosaicos mismos (se necesitarían rudimentos de topología). Pero la intuición y los ejemplos bastan. Aunque los grupos de simetría de los mo-

saicos sean infinitos, tienen la propiedad de que localmente son finitos; cada punto del plano está en un número finito de losetas. Esto se reflejará en la primera propiedad de nuestra definición. Y otra cosa importante es que las losetas sean acotadas, que expresaremos como la propiedad dos. Llamaremos *disco* a un círculo (de radio positivo) junto con su interior, es decir a los conjuntos de la forma $D = \{\mathbf{x} \in \mathbb{R}^2 \mid d(\mathbf{x}, \mathbf{c}) \leq r\}$ para algún punto $\mathbf{c} \in \mathbb{R}^2$ y algún $r > 0$.

Definición 3.9.1 Un subgrupo G de las isometrías **Iso** (2) es un *grupo cristalográfico* si cumple:

- i) Para cualquier disco D se tiene que el conjunto $\{g \in G \mid g(D) \cap D \neq \emptyset\}$ es finito.
- ii) Existe un disco D_0 tal que $\bigcup_{g \in G} g(D_0) = \mathbb{R}^2$, es decir, tal que para todo $\mathbf{x} \in \mathbb{R}^2$ se tiene que $\mathbf{x} \in g(D_0)$ para alguna $g \in G$.

La primera propiedad dice que al estudiar un grupo cristalográfico dentro de una región acotada del plano sólo intervendrá un número finito de elementos del grupo; a esto también se le llama que el grupo sea *discreto*. Y la segunda, que al entender una región suficientemente grande, pero acotada, de alguna manera ya acabamos; el grupo se encarga del resto. Un ejemplo importante son los subgrupos generados por dos traslaciones en direcciones independientes, por ejemplo en los vectores canónicos \mathbf{e}_1 y \mathbf{e}_2 , que da lugar a un subgrupo isomorfo a $\mathbb{Z} \times \mathbb{Z}$ (en el ejemplo son precisamente las traslaciones por vectores en $\mathbb{Z}^2 \subset \mathbb{R}^2$).

El nombre “cristalográfico” viene de que los cristales en la naturaleza se reproducen respetando ciertas simetrías; los químicos, interesados en el asunto, se pusieron a estudiar tales simetrías y demostraron, hacia el final del siglo XIX, que en el plano solamente había 17 posibilidades teóricas para estas reglas de crecimiento. Los matemáticos, ya en el siglo XX, pulieron sus observaciones y la demostración, pero preservaron el nombre dado por los químicos.

La demostración de este teorema involucra dos grandes pasos: la existencia (exhibir los 17 grupos cristalográficos) y la unicidad (demostrar que no hay más). Pero resulta que los artistas y artesanos árabes —cuya expresión gráfica se centró mucho en los mosaicos, la simetría y la ornamentación pues en cierta interpretación del Corán se prohibía representar la figura humana— ya habían hecho (¡y de qué manera!) la mitad del trabajo: en la Alhambra, hay ejemplos de mosaicos con los 17 grupos. Así que la demostración de la existencia está en la Alhambra, concluida en el siglo XVI; y si hay un Teorema de Leonardo, éste —digo yo— debía ser el de la Alhambra.

No tenemos tiempo para demostrarlo aquí. Pero sí podemos ejemplificar su demostración con uno mucho más modesto. Diremos que un grupo cristalográfico G es *caleidoscópico* si además cumple con que tiene suficientes reflexiones como para generarlo, es decir, si cada elemento de G puede escribirse como composición de reflexiones en G . Un ejemplo es el que describimos arriba, que estaba generado por tres reflexiones y que denotaremos como $K_{3,3,3}$ pues las líneas de sus espejos se encuentran en los ángulos $\pi/3, \pi/3, \pi/3$.

Dibujo

Otros dos grupos caleidoscópicos surgen del juego de escuadras. Con tres líneas espejo en los lados de un triángulo de ángulos $\pi/6, \pi/3, \pi/2$ (la escuadra “30, 60, 90”) se genera el grupo $K_{6,3,2}$; y con tres espejos en ángulos $\pi/4, \pi/4, \pi/2$ (la escuadra “45”) el $K_{4,4,2}$. El último, que llamamos $K_{4 \times 4}$, es el generado por espejos en los cuatro lados de un rectángulo. Y, salvo ver que efectivamente son cristalográficos en cuyos detalles no entraremos (construyáanse modelos con espejos reales), esto concluye la parte de existencia del siguiente teorema

Teorema 3.27 *Hay cuatro grupos caleidoscópicos, que son $K_{6,3,2}$, $K_{4,4,2}$, $K_{3,3,3}$ y $K_{4 \times 4}$.*

La demostración de la unicidad se basa en un lema general sobre los centros de simetría. Dado un grupo cristalográfico G (aunque de hecho sólo usaremos que es discreto, la propiedad (i)) tenemos que para cualquier punto $\mathbf{x} \in \mathbb{R}^2$ hay un subgrupo de G , llamado el *estabilizador* de \mathbf{x} y denotado con $\mathbf{St}_G(\mathbf{x})$, que consta de los elementos de G que no mueven a \mathbf{x} , que lo “estabilizan”, es decir,

$$\mathbf{St}_G(\mathbf{x}) := \{g \in G \mid g(\mathbf{x}) = \mathbf{x}\}.$$

(Es fácil demostrar que efectivamente es un subgrupo de G .) Si tomamos cualquier disco D centrado en \mathbf{x} , todos los elementos de su estabilizador mandan a D precisamente en D , pues fijan su centro y son isometrías. Por la propiedad (i) los que mandan a D cerca (lo suficiente para intersectarlo) ya son un número finito, entonces $\mathbf{St}_G(\mathbf{x})$ es finito para cualquier $\mathbf{x} \in \mathbb{R}^2$. El Teorema de Leonardo nos da entonces el siguiente lema.

Lema 3.28 *Si G es un grupo discreto, entonces sus subgrupos estabilizadores no triviales son cíclicos o diédricos.* □

Ahora sí, enfoquemos nuestras baterías contra el teorema de los grupos caleidoscópicos. Como el grupo G tiene suficientes reflexiones para generarlo, nos podemos fijar en el conjunto de sus líneas espejo. Es decir, para cada reflexión en G , pintamos de negro su línea espejo; llamemos \mathcal{L} a este conjunto de líneas. Primero vamos a demostrar que podemos encontrar un triángulo o un rectángulo formado por líneas negras (en \mathcal{L}) tal que ninguna línea de \mathcal{L} pasa por su interior y lo llamaremos la *región fundamental*.

Si todas las líneas de \mathcal{L} son paralelas, los elementos de G mueven cualquier disco en la dirección perpendicular y entonces cubren a lo más una franja del tamaño de su diámetro en esa dirección y no existiría el disco D_0 de la condición (ii). Hay, por tanto, al menos dos pendientes en \mathcal{L} .

Si son exactamente dos las pendientes en \mathcal{L} , éstas tienen que ser perpendiculares, pues con reflexiones en dos rectas no ortogonales se generan reflexiones en líneas con nuevas direcciones. En este caso, hay al menos dos rectas de cada tipo, pues si hay sólo una de un tipo, con cualquier disco se cubre, de nuevo, a lo más una franja.

Dibujo

Dibujo

Existe entonces un rectángulo en \mathcal{L} .

Si hay rectas en tres direcciones tenemos un triángulo. Llamemos F' al rectángulo o triángulo que hemos encontrado. Puesto que las reflexiones en G que mandan a F' cerca son finitas (la condición (i) aplicada a un disco que lo contenga), hay, a lo más, un número finito de rectas en \mathcal{L} que pasan por el interior de F' . Recortando con tijeras en todas ellas, un pedacito (cualquiera) de lo que queda es la región fundamental F que buscábamos, y ya no pasan líneas negras por su interior.

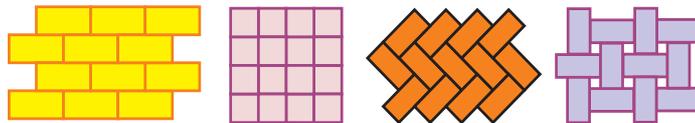
Veamos ahora cómo puede ser la región fundamental F . En un vértice, \mathbf{v} digamos, se intersectan dos líneas de \mathcal{L} de tal manera que en su ángulo definido por F (como ángulo interior) no hay ninguna otra línea de \mathcal{L} . Como las reflexiones correspondientes a estas dos líneas son elementos del grupo estabilizador de \mathbf{v} , $\text{St}_G(\mathbf{v})$, y éste es diédrico o cíclico según el Lema 3.28, entonces es diédrico (pues tiene reflexiones) y el ángulo interior de F en \mathbf{v} es de la forma π/n . Por lo tanto, F tiene que ser un rectángulo (con cuatro ángulos $\pi/2$), o bien un triángulo con ángulos $\pi/p, \pi/q, \pi/r$ donde podemos suponer que $p \leq q \leq r$. En el segundo caso se tiene entonces que p, q, r cumplen

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1,$$

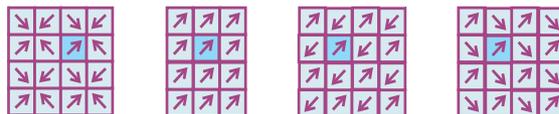
pues los ángulos de un triángulo suman π . Para $p = 2$, tenemos las soluciones $q = 3, r = 6$ y $q = r = 4$; para $p = 3$ tenemos la solución $q = r = 3$ y para $p \geq 4$ ya no hay soluciones.

Entonces hemos demostrado que los lados de F generan uno de los grupos $K_{4 \times 4}, K_{6,3,2}, K_{4,4,2}$ o $K_{3,3,3}$ (respectivamente a su orden de aparición) y por lo tanto que uno de estos grupos, el correspondiente a F , digamos que G' , está contenido en G . Dejamos demostrar que $G' = G$ al lector intrigado; tendrá que usar sustancialmente que las reflexiones generan a G .

EJERCICIO 3.110 Identifica las *líneas de simetría* (espejos de reflexiones y líneas de pasos en el grupo de simetrías) y los *centros de simetría* (centros de subgrupos finitos), de los siguientes mosaicos comunes, que, por supuesto, suponemos infinitos.



EJERCICIO 3.111 Supón que en los siguientes mosaicos la loseta oscura está justo en el cuadrado unitario. Da generadores para sus grupos de simetría.



³Para leer más sobre simetría y mosaicos ver [15], [8] y [11].

Rearmar figura con mosaquito asimétrico

3.9.3 Fractales afinmente autosimilares

Veremos por último algunas figuras que tienen cierta “simetría” aunque no dentro del grupo de isometrías.

Espirales

Consideremos una espiral infinita. Claramente tiene autosemejanzas: siente uno que “gira y crece proporcionalmente”. En efecto, el grupo “espiral”

$$E_c = \{c^t R_{t\pi} \mid t \in \mathbb{R}\},$$

donde $c > 0$, es un subgrupo de las semejanzas y es isomorfo al grupo de los reales con la suma para cualquier $c \neq 1$, pues

$$(c^t R_{t\pi}) \circ (c^r R_{r\pi}) = c^{(t+r)} R_{(t+r)\pi}.$$

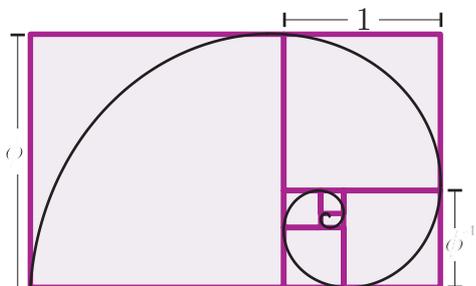
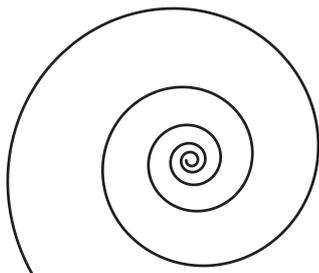
Para cualquier punto $\mathbf{x} \neq \mathbf{0}$, su órbita $E_c(\mathbf{x}) = \{f(\mathbf{x}) \mid f \in E_c\}$ es una espiral *invariante* bajo el grupo, es decir, su imagen bajo cualquier elemento del grupo es ella misma. Los distintos valores de c dan qué tan rápido crece la espiral y en qué orientación.

Especialmente bonita, y que además se da en la naturaleza, es la *espiral áurea* que crece la proporción áurea ϕ en un cuarto de vuelta y entonces en nuestra fórmula anterior corresponde a $c = \phi^2$. La *proporción áurea*, descubierta y ampliamente usada por los griegos que también la llamaban “proporción divina”, es la proporción de los lados de un rectángulo tal que si se le extrae un cuadrado sobre el lado chico nos queda

un rectángulo semejante al original. Nótese que en la figura clásica de al lado, considerando que en ella está también la sucesión infinita de cuadrados a la que alude, tiene como grupo de simetrías afines o de semejanzas al generado por la transformación líneal

$$\begin{pmatrix} 0 & -\phi \\ \phi & 0 \end{pmatrix}$$

(poniendo al origen en el centro de la espiral), que es un subgrupo isomorfo a \mathbb{Z} dentro de E_{ϕ^2} .



EJERCICIO 3.112 Demuestra, usando la definición que acabamos de dar, que la proporción áurea cumple con la ecuación

$$x^2 = x + 1$$

y por tanto que $\phi = (1 + \sqrt{5})/2 \simeq 1.61803$.

EJERCICIO 3.113 Si en la figura anterior el cuadrado superior derecho es el cuadrado unitario, ¿dónde está el centro de la espiral?

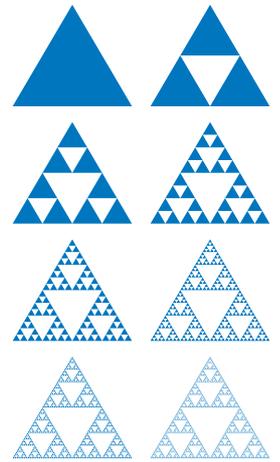
Fractales autosimilares

Si tomamos una espiral desde su centro hasta algún punto dado, entonces al hacer un *zoom* hacia el centro, sin importar de qué tanto aumento, volvemos a encontrar una copia de la figura original. Veremos ahora cómo definir figuras que cumplen esto pero alrededor de cualquiera de sus puntos.

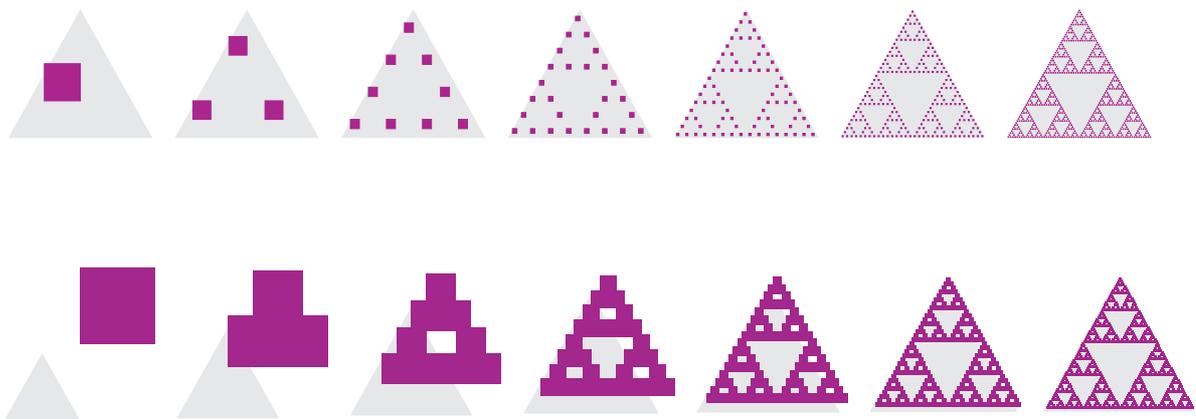
Consideremos primero el ejemplo clásico conocido como el Triángulo de Sierpinsky. Sea Z_0 un triángulo equilátero y sean f_1, f_2 y f_3 las contracciones de factor $1/2$ centradas en sus tres vértices. Tomamos Z_1 como la unión de las tres imágenes de Z_0 bajo estas tres homotecias; y en general, definimos Z_n inductivamente como

$$Z_{n+1} = f_1(Z_n) \cup f_2(Z_n) \cup f_3(Z_n).$$

En la figura vemos hasta Z_7 . El triángulo de Sierpinsky Z es el límite de los Z_n , o bien, como $Z_{n+1} \subset Z_n$, también es la intersección de todos ellos, así que, por ejemplo los tres vértices originales están en Z así como todos los vértices de todos los triángulitos (3^n) en cada paso. La propiedad importante que cumple Z es que si magnificamos pequeñas porciones en cualquier lugar vemos copias del original, y se dice que es un *fractal autosimilar*; en este caso, podemos añadir *autosemejante* pues con *zooms* basta.



La clave es que las tres transformaciones afines que consideramos contraen en todas las direcciones y cumplen entonces que independientemente de con qué conjunto acotado empezamos (en vez de Z_0) el límite es el mismo (aunque ya no se obtenga por intersección). No lo demostraremos aquí, pero ponemos dos dibujos del principio de la sucesión correspondiente a otros dos conjuntos, uno un cuadradito dentro del triángulo Z_0 (que se mantiene en gris al fondo) y el otro un cuadrado grande fuera de él. Después de unos cuantos pasos, las tres figuras ya son casi indistinguibles. Así que Z está realmente asociado a las tres semejanzas (reductoras) f_1, f_2 y f_3 , y no depende del conjunto acotado con el que empezamos.



Si en vez de semejanzas tomamos transformaciones afines que contraigan en todas las direcciones, es decir, que cumplan la desigualdad

$$d(f(\mathbf{x}), f(\mathbf{y})) < d(\mathbf{x}, \mathbf{y}) \quad \text{para toda } \mathbf{x}, \mathbf{y} \in \mathbb{R}^2 \text{ con } \mathbf{x} \neq \mathbf{y}, \quad (3.18)$$

se tiene que la figura límite queda compacta y bien definida (teorema que no demostraremos). Así se obtienen fractales afinmente autosimilares.

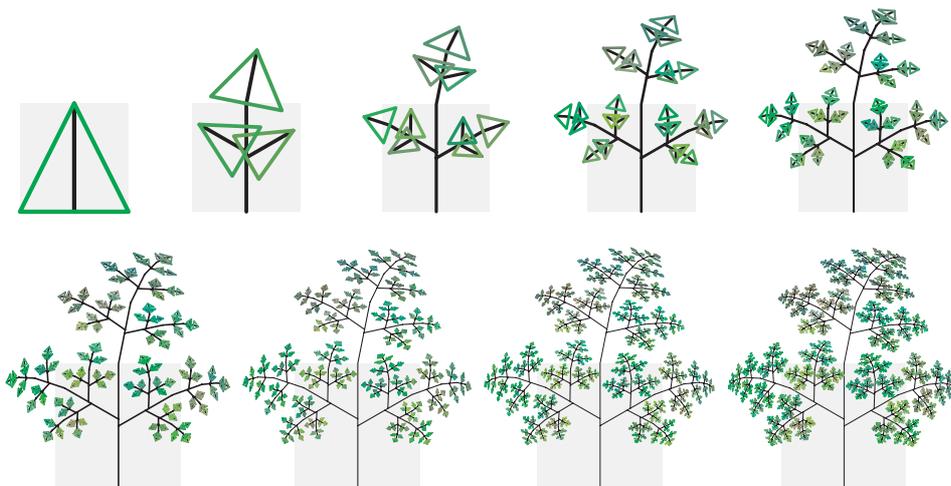
Otro ejemplo interesante es el helecho de la figura siguiente. Está asociado a las tres transformaciones afines g_1 , g_2 y g_3 que mandan al triángulo F_0 del primer cuadro en los tres triángulos del segundo respectivamente (en gris se mantiene el cuadrado unitario para mostrar proporciones y de los segmentos hablaremos en breve). Y de nuevo se define inductiva o recursivamente

$$F_{n+1} = g_1(F_n) \cup g_2(F_n) \cup g_3(F_n)$$

para obtener el “follaje” F como el límite de los conjuntos F_n (en la figura aparece hasta F_8); otra vez no depende del conjunto inicial F_0 sino únicamente de las tres transformaciones. (Puesto que reducen área, las tres transformaciones tienen determinante con valor absoluto menor que 1 y una de ellas lo tiene negativo: ¿puedes decir cuál? se aprecia hasta el tercer cuadro.) Obsérvese que los tallos del helecho se obtienen de una manera sutilmente diferente. Si T_0 es el segmento vertical del primer cuadro, entonces a la siguiente iteración también se le añade lo que ya llevabamos —crece, por decirlo botánicamente, sobre lo ya construido—; esto es, definimos

$$T_{n+1} = T_n \cup g_1(T_n) \cup g_2(T_n) \cup g_3(T_n)$$

y entonces el resultado final T (el límite) sí depende de con qué empezamos (y con ese pequeño intervalo se logra el efecto, de “tallos”, deseado).



Llamemos $H = T \cup F$ al helecho. Las tres funciones generadoras g_1, g_2, g_3 lo mandan en sus tres ramas principales, las composiciones de un par de ellas lo mandan en sus 9 “ramas secundarias”, y así sucesivamente. Si consideramos todas las posibles composiciones finitas de g_1, g_2, g_3 , junto con la identidad se obtiene un *semigrupo* (algunos autores lo llaman *grupoide*) que cumple con dos de los axiomas de grupo. A saber, tiene a la identidad y es cerrado bajo composición, pero no tiene inversos; de alguna manera, es el análogo de los grupos de simetría para este tipo de fractales. Y cualquier elemento g de este semigrupo cumple que $g(H) \subset H$; manda al helecho completo en una ramita por ahí que es afinmente equivalente a H .

Como ya mencionamos, es importante que las tres transformaciones afines que usamos para definir el helecho sean contracciones; en particular, entonces F tampoco depende de con qué figura acotada hayamos empezado. Esbozaremos por último, algunas de las ideas que indican el porqué de este hecho.

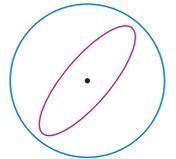
Primero, para una transformación afín $f(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$, que sea una contracción (que cumpla (3.18)), implica algo más fuerte que llamaremos *contracción fuerte* que es que existe una constante $c, 0 < c < 1$ para la cual

$$d(f(\mathbf{x}), f(\mathbf{y})) \leq c d(\mathbf{x}, \mathbf{y}) \quad \text{para toda } \mathbf{x}, \mathbf{y} \in \mathbb{R}^2 \text{ con } \mathbf{x} \neq \mathbf{y}. \quad (3.19)$$

Pues la parte lineal es la que determina las distancias porque

$$|f(\mathbf{x}) - f(\mathbf{y})| = |(A\mathbf{x} + \mathbf{b}) - (A\mathbf{y} + \mathbf{b})| = |A\mathbf{x} - A\mathbf{y}|;$$

y entonces basta ver qué le hace A al círculo unitario. Como quedará claro en el capítulo siguiente, A lo manda en una elipse centrada en el origen y si se cumple (3.18) para $\mathbf{x} = \mathbf{0}$ entonces esta elipse está contenida en el círculo unitario y su semieje mayor $c < 1$ sirve como la constante mínima que cumple (3.19).



Con (3.19) se puede demostrar en general que una contracción fuerte f tiene un punto fijo \mathbf{p} . Para una función afín esto es más fácil (véase el Ejercicio 3.115). Y entonces para cualquier \mathbf{x} se tiene que la sucesión $f^n(\mathbf{x})$ tiende (se acerca indefinidamente) a \mathbf{p} , pues si d es la distancia de \mathbf{x} a \mathbf{p} , es fácil ver que

$$d(f^n(\mathbf{x}), \mathbf{p}) \leq c^n d \rightarrow 0$$

(obsérvese que esto no necesariamente es cierto si sólo se cumple (3.18)). Así que las iteraciones bajo f de cualquier conjunto acotado se van aproximando a \mathbf{p} y lo tienen como límite. Esto implica que el punto fijo de cualquier transformación en los dos semigrupos “contractivos” que hemos considerado va a aparecer en el límite (Z o F), y que también lo harán sus imágenes bajo otros elementos del semigrupo. Pero quizás aparezcan más puntos. Por lo menos, esto medio explica por qué el conjunto inicial no importa para definir el fractal.

Por último, hay que mencionar que esto de los fractales es mucho más amplio. Hay de muchos otros tipos y aún la terminología no está bien establecida pues en las

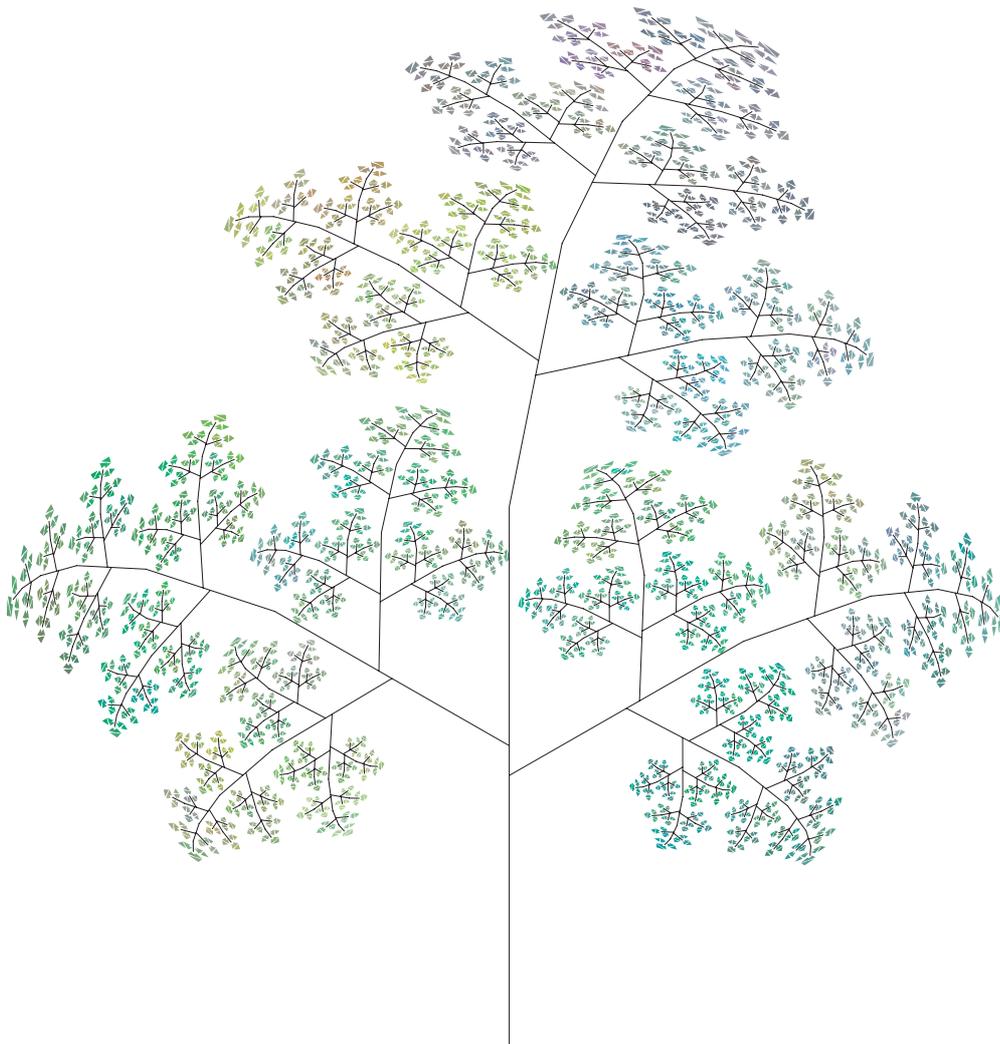
últimas décadas ha habido mucho trabajo sobre ellos y su sorprendente similitud con la naturaleza.

EJERCICIO 3.114 Sea f_1 la transformación afín definitoria del helecho H asociada a su rama principal, es decir, que manda al triángulo F_0 en el más grande y más alto de los tres en F_1 , el central, pues. ¿Cuál es su punto fijo? ¿Puedes determinar los puntos fijos de las otras dos? (Uno es fácil y el otro más escurridizo.) ¿De alguna composición doble?

EJERCICIO 3.115 Sea $f(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ una función afín de \mathbb{R}^2 en \mathbb{R}^2 que no tiene puntos fijos. Demuestra que entonces existe un vector $\mathbf{v} \neq \mathbf{0}$ para el cual $A\mathbf{v} = \mathbf{v}$. (Usa que la ecuación $f(\mathbf{x}) = \mathbf{x}$ es equivalente a $(A - I)\mathbf{x} = -\mathbf{b}$ y en el Teorema 3.21.)

EJERCICIO 3.116 Usando el ejercicio anterior, demuestra que una contracción afín (que cumple (3.18)) tiene un único punto fijo.

³Como bibliografía general sobre fractales, véanse [18] y [19].



Capítulo 4

Cónicas II (clasificación)

En el Capítulo 2 se introdujeron las cónicas (círculos, elipses, hipérbolas y parábolas) como los lugares geométricos de puntos en el plano euclidiano que cumplen cierta propiedad descrita en términos de distancias. De esa descripción se dedujo que son los conjuntos de puntos cuyas coordenadas cumplen cierta ecuación cuadrática. Una ecuación cuadrática de éstas se puede considerar como un polinomio cuadrático (de grado dos) en dos variables (las coordenadas) igualado a 0. Al tomar cónicas en diferentes posiciones y con distintos parámetros se obtendrán diversos polinomios. La pregunta que guía este capítulo es la pregunta inversa:

¿Será cierto que cualquier polinomio cuadrático en dos variables define una cónica? Si es así, ¿cómo saber qué tipo de cónica es y cuál es su descripción geométrica?

Por su parte, en el Capítulo 3 estudiamos diferentes grupos de transformaciones geométricas. Ellos nos darán la herramienta y el lenguaje para atacar y entender el problema que nos acabamos de plantear.

4.1 ¿Qué es clasificar?

Muy en general, “clasificar” es describir o enumerar las clases de equivalencia de un conjunto de objetos de acuerdo con ciertos “criterios” que definen una relación de equivalencia entre ellos. En nuestro caso, de lo que depende una “clasificación” es de determinar un conjunto de objetos geométricos (triángulos, digamos, como ejemplo para fijar ideas) y de establecer una noción de equivalencia, es decir, definir las condiciones en que estamos dispuestos a decir que dos objetos son “iguales”, o, mejor dicho, equivalentes. En esto último, la noción de grupo de transformaciones que hemos desarrollado es básica. Concretemos.

Una *figura plana* es cualquier subconjunto $F \subset \mathbb{R}^2$. Sea G un grupo de transformaciones de \mathbb{R}^2 (nos interesan principalmente los que ya hemos definido, es decir, G puede ser el grupo afín $\mathbf{Af}(2)$; el general lineal $\mathbf{Gl}(2)$; el grupo de semejanzas; el de homotecias; el de isometrías $\mathbf{Iso}(2)$; el de transformaciones ortogonales $\mathbf{O}(2)$, o el de traslaciones). Decimos que dos figuras F_1 y F_2 son G -*equivalentes* (o bien, que son

equivalentes módulo G) y lo denotamos $F_1 \stackrel{G}{\sim} F_2$, si existe $g \in G$ tal que $g(F_1) = F_2$. La terminología se puede hacer a veces más simple; por ejemplo, dos figuras son *semejantes* (o *isométricas*) si son equivalentes módulo semejanzas (o, respectivamente, $\mathbf{Iso}(2)$ -equivalentes).

Veamos primero que ésta es efectivamente una relación de equivalencia, demostración que corresponde puntualmente a que G es un grupo de transformaciones:

- i) $\stackrel{G}{\sim}$ es *reflexiva* ($F \stackrel{G}{\sim} F$); pues $\text{id}(F) = F$ y la función identidad, id , está en G .
- ii) $\stackrel{G}{\sim}$ es *simétrica* ($F_1 \stackrel{G}{\sim} F_2 \Rightarrow F_2 \stackrel{G}{\sim} F_1$); pues si existe $g \in G$ tal que $g(F_1) = F_2$ entonces $g^{-1}(F_2) = F_1$ y $g^{-1} \in G$.
- iii) $\stackrel{G}{\sim}$ es *transitiva* ($F_1 \stackrel{G}{\sim} F_2$ y $F_2 \stackrel{G}{\sim} F_3 \Rightarrow F_1 \stackrel{G}{\sim} F_3$); pues si $g_1(F_1) = F_2$ y $g_2(F_2) = F_3$ con $g_1, g_2 \in G$, entonces $(g_2 \circ g_1)(F_1) = F_3$ y $(g_2 \circ g_1) \in G$.

Además, estas relaciones se “anidan” de acuerdo con la contención de grupos; es decir, si $H \subset G$ entonces $F_1 \stackrel{H}{\sim} F_2 \Rightarrow F_1 \stackrel{G}{\sim} F_2$. Así, si, por decir algo, dos figuras son isométricas, también serán afinmente equivalentes (pero no necesariamente linealmente equivalentes).

4.1.1 Clasificación de triángulos

Para acabar de entender con un ejemplo concreto, describamos la clasificación de triángulos. Primero hay que definirlos: sea \mathcal{T} el conjunto de triángulos planos, es decir, ternas (no ordenadas) no colineales de puntos en \mathbb{R}^2 . El Teorema 6.4 nos dice que “hay un sólo triángulo afín”, es decir que cualquier par de triángulos son afinmente equivalentes.

Esta única clase de equivalencia afín se parte en muchas subclases cuando vemos los triángulos con “ojos de semejanza”. Pues, si alguna de las seis transformaciones afines que mandan un triángulo en otro es una semejanza, entonces éstos tienen los mismos ángulos, e inversamente. Así que el teorema “*dos triángulos son semejantes si y sólo si sus ángulos son iguales*” es un “teorema de clasificación” que dice que hay una clase de semejanza por cada terna (no ordenada) de ángulos α, β, γ que cumple $\alpha > 0, \beta > 0, \gamma > 0$ y $\alpha + \beta + \gamma = \pi$.

De nuevo, las clases de semejanza se parten en clases de isometría de acuerdo con un parámetro real y positivo (el tamaño) que podríamos tomar como el área del triángulo o la longitud de un lado dado. Y podríamos de nuevo partir estas clases de isometría en clases módulo translaciones (y aparece un parámetro de “orientación” y uno de “dirección”), o bien en clases módulo ortogonales (y aparece un parámetro de “distancia al origen”). Llegando al extremo de la clasificación con el grupo trivial ($\{\text{id}\}$) donde dos son equivalentes si son estrictamente el mismo.

4.2 Clasificación de cónicas

Primero que nada, un polinomio de segundo grado (o *cuadrático*) en las variables x, y es una expresión

$$P(x, y) = ax^2 + bxy + cy^2 + dx + ey + f,$$

donde los coeficientes a, b, c, d, e, f son números reales, y además al menos uno de a, b, c es distinto de 0 (si no fuera así estaríamos hablando de un polinomio lineal). Los primeros tres términos constituyen su *parte cuadrática*, los siguientes dos su *parte lineal*, y por último f es la *parte constante*. Nótese que al monomio bxy se le considera de grado 2 pues es la suma de los exponentes de las dos variables. A un polinomio cuadrático se le puede pensar también como una función $P: \mathbb{R}^2 \rightarrow \mathbb{R}$ donde a un punto dado $\mathbf{x} \in \mathbb{R}^2$ se le asigna el número $P(\mathbf{x})$ que resulta de sustituir en la expresión los valores de sus coordenadas $\mathbf{x}^\top = (x, y)$. Así que podemos hablar de los *ceros del polinomio* P , o la *curva asociada al polinomio* P , como un subconjunto de \mathbb{R}^2 , a saber:

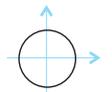
$$\mathcal{C}(P) = \{\mathbf{x} \in \mathbb{R}^2 \mid P(\mathbf{x}) = 0\}.$$

Diremos que un subconjunto $\mathcal{C} \subset \mathbb{R}^2$ es una *curva cuadrática* si consiste de los ceros de un polinomio cuadrático, es decir, si para algún polinomio cuadrático P se tiene que $\mathcal{C} = \mathcal{C}(P)$. Por el momento no queremos usar el nombre más común de “cónica”, para que quede claro nuestro problema, que consiste en clasificar las curvas cuadráticas de acuerdo con los diferentes grupos geométricos, y en particular, ver si son siempre cónicas. La parte fundamental de esta clasificación consistirá en demostrar que cualquier curva cuadrática es afinmente equivalente a alguna de las siguientes, que por el momento sirven de ejemplos.

4.2.1 Las cónicas canónicas (y algo más)

i) **El círculo unitario.** El polinomio

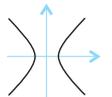
$$x^2 + y^2 - 1 \quad (i)$$



tiene como ceros el círculo unitario.

ii) **La hipérbola unitaria.** El polinomio

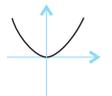
$$x^2 - y^2 - 1 \quad (ii)$$



tiene como ceros la hipérbola que llamaremos unitaria.

iii) **La parábola canónica.** El polinomio

$$x^2 - y \quad (iii)$$



tiene como ceros a una parábola.

Por medio de transformaciones afines estas tres curvas se pueden mandar a muchas otras. Del círculo unitario se obtienen las elipses, y de la hipérbola y la parábola (que hemos escogido por la simpleza de su polinomio) se obtienen todas las demás hipérbolas y parábolas que estudiamos en el Capítulo 2. Pero notemos de una vez que con los polinomios cuadráticos se obtienen nuevos conjuntos:

iv) **El círculo imaginario.** El polinomio

$$x^2 + y^2 + 1 \quad (\text{iv})$$

no tiene ningún cero en los reales pues la suma de dos cuadrados es no negativa, más 1 es estrictamente positiva. Es decir, su curva cuadrática es el vacío $\emptyset \subset \mathbb{R}^2$. La llamamos “círculo imaginario” pues fácilmente le encontramos soluciones complejas (por ejemplo $i^2 + 0^2 + 1 = 0$, donde $i^2 = -1$) pero se nos sale del contexto de este libro; aunque como polinomio tiene plenos derechos y tendremos entonces que incluirlo en nuestro análisis.

v) **Par de rectas.** El polinomio



$$x^2 - y^2 \quad (\text{v})$$

tiene como conjunto de ceros a la unión de las dos rectas $x + y = 0$ y $x - y = 0$; pues al verlo como producto ($x^2 - y^2 = (x + y)(x - y)$), si se hace cero entonces alguno de los factores debe ser cero y viceversa.

vi) **El círculo de radio cero.** Dado por el polinomio



$$x^2 + y^2 \quad (\text{vi})$$

que define al punto $\mathbf{0}$ por una sola ecuación cuadrática que es el caso límite de círculos cuyo radio se hace cero. También podría llamarse **par de rectas imaginarias** pues se factoriza ($x^2 + y^2 = (x + iy)(x - iy)$) y se puede ver entonces como dos “rectas complejas” de las que sólo vemos en el plano real al origen.

vii) **Rectas paralelas.** Por la misma razón que en (v), el polinomio



$$x^2 - 1, \quad (\text{vii})$$

da dos rectas paralelas ($x = 1$ y $x = -1$).

viii) **Rectas paralelas imaginarias.** Dada por

$$x^2 + 1, \quad (\text{viii})$$

que se factoriza $x^2 + 1 = (x + i)(x - i)$.

ix) **Recta doble.** Dada por



$$x^2. \quad (\text{ix})$$

Aunque geoméricamente sólo consiste de una recta ($x = 0$) se le llama “recta doble” pues el polinomio que la define es diferente.

4.2.2 Equivalencia de polinomios

Para que tenga sentido el problema de clasificar curvas cuadráticas módulo transformaciones afines, la primera pregunta que debemos responder es si las imágenes afines de curvas cuadráticas son de nuevo curvas cuadráticas.

Para esto, sean \mathcal{C} una curva cuadrática y $g \in \mathbf{Af}(2)$ una transformación afín. Tenemos entonces que existe un polinomio cuadrático P que define a \mathcal{C} , es decir, tal que $\mathcal{C} = \mathcal{C}(P) = \{\mathbf{x} \in \mathbb{R}^2 \mid P(\mathbf{x}) = 0\}$. Afirmamos que

$$g(\mathcal{C}) = \{\mathbf{y} \in \mathbb{R}^2 \mid (P \circ g^{-1})(\mathbf{y}) = 0\}. \quad (4.1)$$

Esta igualdad de conjuntos se demuestra por dos contenciones. Primera (\subset): cualquier punto en $g(\mathcal{C})$ es de la forma $g(\mathbf{x})$ con $\mathbf{x} \in \mathcal{C}$; se tiene entonces que $P(\mathbf{x}) = 0$ y por lo tanto

$$(P \circ g^{-1})(g(\mathbf{x})) = P(g^{-1}(g(\mathbf{x}))) = P(\mathbf{x}) = 0,$$

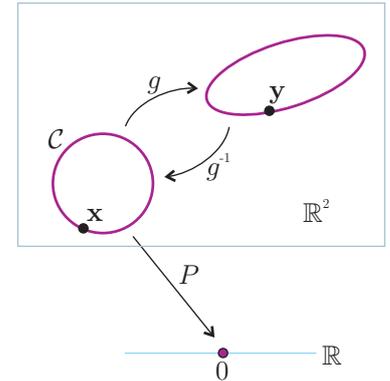
lo cual implica que $g(\mathbf{x})$ está en el conjunto de la derecha.

Y al revés (\supset): si tomamos \mathbf{y} tal que $(P \circ g^{-1})(\mathbf{y}) = 0$; sea $\mathbf{x} := g^{-1}(\mathbf{y})$, se tiene entonces que $P(\mathbf{x}) = P(g^{-1}(\mathbf{y})) = 0$ y por tanto que $\mathbf{x} \in \mathcal{C}$, lo cual implica que $\mathbf{y} = g(\mathbf{x}) \in g(\mathcal{C})$. Queda entonces demostrada la igualdad.

Obsérvese que en la demostración anterior sólo se usó que \mathcal{C} estuviera definida como los ceros de una función y que g fuera invertible (una transformación). Ahora veremos que $g(\mathcal{C})$ también es una curva cuadrática. Para esto bastará ver que si P es un polinomio cuadrático y g una transformación afín entonces la función $(P \circ g) : \mathbb{R}^2 \rightarrow \mathbb{R}$ también es un polinomio cuadrático (para no estar usando g^{-1} , que sólo complica la notación). Puesto que las dos coordenadas de g son polinomios lineales, y se obtiene una expresión de $P \circ g$ al sustituir estos dos polinomios lineales en las variables de P , es claro que (después de simplificar la expresión) obtendremos un polinomio con monomios de grado a lo más dos. Si los coeficientes de los monomios cuadráticos se cancelaran todos, es decir, si el resultado fuera lineal o constante, tendríamos que al volver a precomponer con g^{-1} se vuelve cuadrático y esto es claramente imposible. Tuvimos que usar entonces que g es una transformación. En resumen, hemos demostrado el siguiente lema.

Lema 4.1 *Sea \mathcal{C} una curva cuadrática y $g \in \mathbf{Af}(2)$, entonces $g(\mathcal{C})$ también es una curva cuadrática. Además, si $\mathcal{C} = \mathcal{C}(P)$ entonces $g(\mathcal{C}) = \mathcal{C}(P \circ g^{-1})$. \square*

Hay que observar, antes de definir la equivalencia de polinomios, que al multiplicar un polinomio por una constante $k \neq 0$, sus ceros no cambian. Es decir, que $\mathcal{C}(P) = \mathcal{C}(kP)$ cuando $k \neq 0$; pues entonces $P(\mathbf{x}) = 0 \Leftrightarrow kP(\mathbf{x}) = 0$.



Sea G un subgrupo de $\mathbf{Af}(2)$. Decimos que dos polinomios cuadráticos P_1 y P_2 son G -equivalentes, o equivalentes módulo G , escrito $P_1 \stackrel{G}{\sim} P_2$, si existen $g \in G$ y $k \in \mathbb{R}$, con $k \neq 0$, tales que $kP_1 = P_2 \circ g$. Que ésta es una relación de equivalencia se deja como ejercicio.

De lo anterior se obtiene que polinomios cuadráticos G -equivalentes definen curvas G -equivalentes. El inverso, aunque es cierto, aún no lo podemos demostrar, pues incluye que si dos polinomios cuadráticos definen la misma curva (tienen los mismos ceros) entonces difieren por una constante y para esto (en general) hay que incluir los números complejos. Lo que sí podemos hacer es enunciar el teorema fundamental de este capítulo, cuya demostración nos llevará gran parte del mismo.

Teorema 4.2 *Sea P un polinomio cuadrático en dos variables x, y . Entonces P es afinmente equivalente a uno y sólo uno de los polinomios (i), (ii), ..., (ix).*

Una vez establecido este teorema, la clasificación afín de curvas cuadráticas será inmediata.

Corolario 4.3 *Hay ocho clases de equivalencia afín de curvas cuadráticas que corresponden a los ocho dibujos de la Sección 4.2.1.*

Nótese que redujimos a ocho pues el vacío aparece en dos tipos de polinomios.

EJERCICIO 4.1 Demuestra que para un subgrupo G de $\mathbf{Af}(2)$, la relación de ser G -equivalentes es una relación de equivalencia en los polinomios cuadráticos en dos variables.

4.3 Reducción de polinomios cuadráticos

En esta sección veremos cómo reducir (simplificar) un polinomio cuadrático cualquiera mediante cambios de coordenadas afines. Para empezar, valdrá la pena simplificar la notación de tal polinomio usando el lenguaje matricial y vectorial que hemos desarrollado. A su vez, esto nos dejará ver con nitidez cómo diferentes tipos de transformaciones afines afectan el polinomio.

El polinomio general de segundo grado en dos variables se puede siempre escribir como

$$P(x, y) = ax^2 + 2bxy + cy^2 + dx + ey + f, \quad (4.2)$$

donde la única diferencia con la forma anterior es que hemos puesto un 2 al coeficiente del término mixto (el monomio en xy) por razones que serán evidentes muy pronto. Si consideramos el vector variable $\mathbf{x}^T = (x, y)$ y definimos una matriz A y un vector \mathbf{k} como

$$A := \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \quad \mathbf{k} = \begin{pmatrix} d \\ e \end{pmatrix},$$

el polinomio P se reescribe como

$$P(\mathbf{x}) = \mathbf{x} \cdot A\mathbf{x} + \mathbf{k} \cdot \mathbf{x} + f, \quad (4.3)$$

que llamaremos la *expresión vectorial* de P , pues

$$\begin{aligned} \mathbf{x} \cdot A\mathbf{x} &= \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} ax + by \\ bx + cy \end{pmatrix} \\ &= x(ax + by) + y(bx + cy) = ax^2 + 2bxy + cy^2, \end{aligned}$$

y

$$\mathbf{k} \cdot \mathbf{x} = \begin{pmatrix} d \\ e \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = dx + ey.$$

Debe quedar claro entonces que anteponer el 2 al coeficiente mixto fue para que la expresión de la matriz A quede más simple. Así la matriz A , que debemos enfatizar que es *simétrica* (es decir, $A = A^T$), representa a la parte cuadrática, mientras que el vector \mathbf{k} a la parte lineal (la constante f quedó tal cual). Por ejemplo, el polinomio $x^2 + 4xy + 3y^2 + 5x - y + 2$ queda determinado por su parte constante 2, la matriz

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \quad \text{y el vector} \quad \mathbf{k} = \begin{pmatrix} 5 \\ -1 \end{pmatrix}.$$

Nuestra definición de *polinomio cuadrático* será ahora una expresión (4.3) con

$$A = A^T \neq \mathbf{0}.$$

Nótese que al pedir a la matriz A que sea simétrica obtenemos que la expresión vectorial de P (4.3) es única; se pudo haber tomado cualquier otra matriz con tal de que sus elementos no diagonales sumaran $2b$ pero la más natural, y que resulta única, es la simétrica y esto será importante. Una de las grandes ventajas de usar la expresión vectorial (4.3) en vez de la expresión en variables (4.2) es que la vectorial se escribe igual cuando hay más variables; sólo cambian las dimensiones de los elementos, y da la idea clara de que se pueden generalizar nuestros argumentos. Por ejemplo, para tres variables, tanto el vector variable \mathbf{x} y el constante \mathbf{k} tienen tres entradas, y la matriz A es simétrica de orden 3×3 ; pero en la expresión en variables faltan cuatro monomios a considerar (véase ejercicio abajo). Sin embargo, esta posibilidad no la usaremos por el momento, y seguiremos con dos variables.

Veremos ahora cómo se afecta la expresión vectorial del polinomio cuadrático P al precomponer con diversas transformaciones afines, empezando por las más simples.

EJERCICIO 4.2 Encuentra la matriz simétrica A y el vector constante \mathbf{k} que dan la expresión vectorial de los siguientes polinomios cuadráticos:

$$\begin{aligned} & x^2 + 2y^2 - 6x + 4y + 3 \\ & 9x^2 + 24xy + 16y^2 - 2x - 86y + 39 \\ & 2xy - 6x - 4y - 4. \end{aligned}$$

EJERCICIO 4.3 Da una expresión general para un polinomio cuadrático en tres variables x, y, z y luego define una expresión vectorial para él.

4.3.1 Traslaciones (cómo encontrar el centro)

Consideremos la traslación $g(\mathbf{x}) = \mathbf{x} + \mathbf{h}$, para cualquier vector $\mathbf{h} \in \mathbb{R}^2$, y veamos cómo se escribe el polinomio cuadrático $P \circ g$, donde P está descrito por (4.3) (es decir, donde $P(\mathbf{x}) = \mathbf{x} \cdot A\mathbf{x} + \mathbf{k} \cdot \mathbf{x} + f$).

$$\begin{aligned} (P \circ g)(\mathbf{x}) &= P(\mathbf{x} + \mathbf{h}) = (\mathbf{x} + \mathbf{h}) \cdot A(\mathbf{x} + \mathbf{h}) + \mathbf{k} \cdot (\mathbf{x} + \mathbf{h}) + f \\ &= (\mathbf{x} + \mathbf{h}) \cdot (A\mathbf{x} + A\mathbf{h}) + \mathbf{k} \cdot \mathbf{x} + \mathbf{k} \cdot \mathbf{h} + f \\ &= \mathbf{x} \cdot A\mathbf{x} + \mathbf{x} \cdot A\mathbf{h} + \mathbf{h} \cdot A\mathbf{x} + \mathbf{h} \cdot A\mathbf{h} + \mathbf{k} \cdot \mathbf{x} + \mathbf{k} \cdot \mathbf{h} + f \\ &= \mathbf{x} \cdot A\mathbf{x} + \mathbf{x} \cdot A\mathbf{h} + \mathbf{h} \cdot A\mathbf{x} + \mathbf{k} \cdot \mathbf{x} + (\mathbf{h} \cdot A\mathbf{h} + \mathbf{k} \cdot \mathbf{h} + f) \\ &= \mathbf{x} \cdot A\mathbf{x} + (\mathbf{x} \cdot A\mathbf{h} + \mathbf{h} \cdot A\mathbf{x} + \mathbf{k} \cdot \mathbf{x}) + P(\mathbf{h}). \end{aligned} \tag{4.4}$$

Nótese que los tres sumandos de enmedio corresponden a la parte lineal, que podemos simplificar aún más usando dos lemas generales.

Lema 4.4 Dadas dos matrices A y B que se puedan multiplicar, se tiene $(AB)^T = B^T A^T$.

Demostración. Ver ejercicios 3.86 y 4.4 (abajo). □

Lema 4.5 Sea A una matriz simétrica (es decir, $A = A^T$) entonces para cualquier par de vectores \mathbf{x}, \mathbf{y} se tiene que $\mathbf{x} \cdot A\mathbf{y} = A\mathbf{x} \cdot \mathbf{y}$.

Demostración. Puesto que $\mathbf{x} \cdot A\mathbf{y} = \mathbf{x}^T A\mathbf{y}$ y el transpuesto de un número real (matriz 1×1) es él mismo, se tiene por el lema anterior que

$$\begin{aligned} \mathbf{x} \cdot A\mathbf{y} &= \mathbf{x}^T A\mathbf{y} = (\mathbf{x}^T A\mathbf{y})^T = (A\mathbf{y})^T (\mathbf{x}^T)^T \\ &= \mathbf{y}^T A^T \mathbf{x} = \mathbf{y}^T A\mathbf{x} = \mathbf{y} \cdot A\mathbf{x} = A\mathbf{x} \cdot \mathbf{y}. \end{aligned}$$

□

Por este lema, donde es fundamental que A sea simétrica, podemos simplificar la parte lineal de (4.4) para obtener

$$\mathbf{x} \cdot A\mathbf{h} + \mathbf{h} \cdot A\mathbf{x} + \mathbf{k} \cdot \mathbf{x} = 2(A\mathbf{h} \cdot \mathbf{x}) + \mathbf{k} \cdot \mathbf{x} = (2A\mathbf{h} + \mathbf{k}) \cdot \mathbf{x},$$

de donde concluimos que

$$(P \circ g)(\mathbf{x}) = P(\mathbf{x} + \mathbf{h}) = \mathbf{x} \cdot A\mathbf{x} + (2A\mathbf{h} + \mathbf{k}) \cdot \mathbf{x} + P(\mathbf{h}).$$

De tal manera que si podemos encontrar $\mathbf{h} \in \mathbb{R}^2$ que cumpla

$$2A\mathbf{h} + \mathbf{k} = \mathbf{0}$$

habremos encontrado una traslación que elimine la parte lineal de P . En caso de que tal \mathbf{h} exista, se le llama el *centro* de la curva cuadrática (o de la cónica) asociada a P . En particular, podemos concluir con el siguiente lema.

Lema 4.6 *Sea $P(\mathbf{x}) = \mathbf{x} \cdot A\mathbf{x} + \mathbf{k} \cdot \mathbf{x} + f$ un polinomio cuadrático ($A = A^\top$) tal que $\det(A) \neq 0$. Entonces si definimos $\mathbf{c} := -(1/2)A^{-1}\mathbf{k}$, se tiene que \mathbf{c} es el centro de $\mathcal{C}(P)$ y*

$$P(\mathbf{x} + \mathbf{c}) = \mathbf{x} \cdot A\mathbf{x} + P(\mathbf{c}).$$

□

EJERCICIO 4.4 Demuestra que $(ABC)^\top = C^\top B^\top A^\top$, donde A, B, C son tres matrices que se pueden multiplicar.

EJERCICIO 4.5 Encuentra los centros, si es que tienen, de las curvas asociadas a algunos de los polinomios siguientes:

$$\begin{aligned} & x^2 + 2y^2 - 6x + 4y + 3 \\ & 9x^2 + 24xy + 16y^2 - 2x - 86y + 39 \\ & xy - 3x - 2y - 2 \\ & -6x^2 + 24xy + y^2 - 12x - 26y - 161 \\ & 9x^2 - 4xy + 6y^2 - 58x + 24y + 59 \\ & x^2 + 2xy + y^2 + 2x + 2y - 1. \end{aligned}$$

4.3.2 Rotaciones (cómo encontrar los ejes)

Ya vimos que las traslaciones afectan la parte lineal de los polinomios cuadráticos. Consideremos ahora el caso en que g sea una transformación lineal $g(\mathbf{x}) = B\mathbf{x}$ (donde

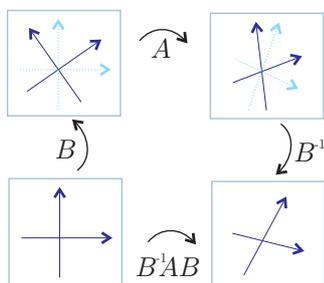
$B \in \mathbf{Gl}(2)$). Si P es el polinomio cuadrático general (4.3) entonces

$$\begin{aligned}
 (P \circ g)(\mathbf{x}) &= P(B\mathbf{x}) = (B\mathbf{x}) \cdot A(B\mathbf{x}) + \mathbf{k} \cdot (B\mathbf{x}) + f \\
 &= (B\mathbf{x})^\top A(B\mathbf{x}) + \mathbf{k}^\top B\mathbf{x} + f \\
 &= (\mathbf{x}^\top B^\top)A(B\mathbf{x}) + (\mathbf{k}^\top B)\mathbf{x} + f \\
 &= \mathbf{x} \cdot (B^\top AB)\mathbf{x} + (B^\top \mathbf{k}) \cdot \mathbf{x} + f.
 \end{aligned} \tag{4.5}$$

Obsérvese primero que la nueva matriz $B^\top AB$ que define la parte cuadrática sigue siendo simétrica pues $(B^\top AB)^\top = B^\top A^\top (B^\top)^\top = B^\top AB$, así que la expresión de $P \circ g$ dada arriba es válida. Simplificar la parte cuadrática de P corresponde entonces a encontrar una matriz B tal que $B^\top AB$ sea más simple; por ejemplo *diagonal*, es decir, con ceros fuera de la diagonal, pues entonces el polinomio $P \circ g$ no tendrá término mixto xy . Veremos, como resultado principal de esta sección, que esto siempre se puede lograr.

Hasta aquí sólo hemos usado que B es una matriz (ni siquiera usamos que fuera no singular, con determinante no nulo), pero si pedimos además que sea ortogonal entonces tenemos que $B^\top = B^{-1}$ y, en este caso, la matriz de la parte cuadrática se transforma en la matriz $B^\top AB = B^{-1}AB$ que tiene un significado geométrico muy claro. Es la matriz que expresa la función A en la base de las columnas de B . Veamos.

Sean \mathbf{u} y \mathbf{v} las columnas de B , es decir $B = (\mathbf{u}, \mathbf{v})$. Si $\mathbf{x}^\top = (x, y)$ entonces $B\mathbf{x} = x\mathbf{u} + y\mathbf{v}$; y podemos pensar a $B\mathbf{x}$ como el vector (con coordenadas respecto a la



base canónica) que tiene coordenadas (x, y) en la base \mathbf{u}, \mathbf{v} . Por el contrario, dado cualquier $\mathbf{y} \in \mathbb{R}^2$, $B^{-1}\mathbf{y}$ nos da los coeficientes para escribir a \mathbf{y} como combinación de la base \mathbf{u}, \mathbf{v} , pues es la solución del sistema $B\mathbf{x} = \mathbf{y}$. Así que $B^{-1}AB$ es la matriz que describe a A , pensada como función, pero en términos de la base \mathbf{u}, \mathbf{v} ; por ejemplo, $(B^{-1}AB)(1, 0)^\top = B^{-1}A(B(1, 0)^\top) = B^{-1}(A\mathbf{u})$ que da, como habíamos quedado, las coordenadas de $A\mathbf{u}$ en la base \mathbf{u}, \mathbf{v} y los datos que le dimos, $(1, 0)$, son las coordenadas de \mathbf{u} en la base \mathbf{u}, \mathbf{v} . Quizá este razonamiento

general no convenza aún, pero no importa, pues en nuestro caso lo podemos concretar aún más.

Supongamos, como es el caso en que estamos interesados, que $B \in \mathbf{O}(2)$, es decir que $BB^\top = I$. Y sean \mathbf{u} y \mathbf{v} las columnas de B que forman una base ortonormal. Supongamos además que A actúa de manera muy sencilla en \mathbf{u} y en \mathbf{v} : que simplemente los alarga por factores λ y μ respectivamente, es decir que $A\mathbf{u} = \lambda\mathbf{u}$ y $A\mathbf{v} = \mu\mathbf{v}$. Entonces, expresar a A (como función lineal) pero en términos de la base B debía ser la matriz

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}.$$

Veamos si corresponde a $B^T A B$:

$$\begin{aligned} B^T A B &= (\mathbf{u}, \mathbf{v})^T A(\mathbf{u}, \mathbf{v}) = \begin{pmatrix} \mathbf{u}^T \\ \mathbf{v}^T \end{pmatrix} (A\mathbf{u}, A\mathbf{v}) = \begin{pmatrix} \mathbf{u}^T A\mathbf{u} & \mathbf{u}^T A\mathbf{v} \\ \mathbf{v}^T A\mathbf{u} & \mathbf{v}^T A\mathbf{v} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{u} \cdot A\mathbf{u} & \mathbf{u} \cdot A\mathbf{v} \\ \mathbf{v} \cdot A\mathbf{u} & \mathbf{v} \cdot A\mathbf{v} \end{pmatrix} = \begin{pmatrix} \mathbf{u} \cdot \lambda\mathbf{u} & \mathbf{u} \cdot \mu\mathbf{v} \\ \mathbf{v} \cdot \lambda\mathbf{u} & \mathbf{v} \cdot \mu\mathbf{v} \end{pmatrix} \\ &= \begin{pmatrix} \lambda(\mathbf{u} \cdot \mathbf{u}) & \mu(\mathbf{u} \cdot \mathbf{v}) \\ \lambda(\mathbf{v} \cdot \mathbf{u}) & \mu(\mathbf{v} \cdot \mathbf{v}) \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}. \end{aligned} \quad (4.6)$$

Efectivamente, y con esto nos basta, pues si encontramos una matriz B que satisfaga lo anterior (que “diagonalice” la matriz A) podremos eliminar el término mixto del polinomio P para acercarnos a los polinomios canónicos.

Valores propios y vectores propios

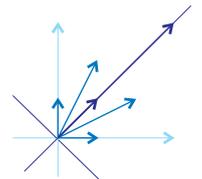
Sea A una matriz cuadrada cualquiera. Se dice que un vector \mathbf{v} es *vector propio* de A con *valor propio* $\lambda \in \mathbb{R}$, si $\mathbf{v} \neq \mathbf{0}$ y se tiene que $A\mathbf{v} = \lambda\mathbf{v}$. Nuestro problema será encontrar *parejas propias* (es decir, λ, \mathbf{v} valor y vector propio correspondientes). A los vectores propios y valores propios también se les conoce como *eigenvectores* y *eigenvalores* por su nombre en alemán.

Un vector propio de una matriz es un vector donde la matriz (por más complicada que ésta sea), como función, actúa de una manera muy sencilla, simplemente lo alarga. Nótese además que eliminamos el $\mathbf{0}$ de los posibles vectores propios pues éste sería tal para cualquier “valor propio”; sin embargo el $0 \in \mathbb{R}$ sí puede ser valor propio si la matriz anula algún vector no nulo.

Antes de abordar el problema de encontrar vectores propios de manera algebraica, veamos un **ejemplo** sencillo donde geoméricamente es fácil encontrarlos. Consideremos la matriz

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Puesto que bajo esta función el eje x gira hacia arriba y el eje y hacia la derecha, el cuadrante positivo se apachurra; las líneas en él cercanas al eje x giran en dirección contraria a las cercanas al eje y y entonces debe haber una línea que se queda fija. Por simetría (pasa esencialmente lo mismo en los dos extremos) esta línea debe ser la de enmedio; efectivamente, el $(1, 1)$ es vector propio de A con valor propio 3. Los otros dos cuadrantes se abren como abanico y de nuevo su línea “media” se queda fija; el vector $(-1, 1)$ es vector propio de A con valor propio 1. Entonces, en la base rotada 45° esta transformación consiste simplemente en expandir por un factor 3 el primer eje y dejar fijo el ortogonal.



Otro ejemplo geoméricamente claro son las reflexiones: tienen vectores propios con valor propio 1 en la línea espejo y vectores propios con valor propio -1 en su ortogonal.

Veamos primero un resultado muy general que agrupa naturalmente a los vectores propios.

Lema 4.7 *Si \mathbf{u} y \mathbf{v} son vectores propios de A con el mismo valor propio $\lambda \in \mathbb{R}$, entonces cualquier combinación lineal no trivial de ellos también es vector propio.*

Demostración. Si $A\mathbf{u} = \lambda\mathbf{u}$ y $A\mathbf{v} = \lambda\mathbf{v}$, entonces para cualquier par de coeficientes $\alpha, \beta \in \mathbb{R}$ se tiene que

$$A(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha(A\mathbf{u}) + \beta(A\mathbf{v}) = \alpha(\lambda\mathbf{u}) + \beta(\lambda\mathbf{v}) = \lambda(\alpha\mathbf{u} + \beta\mathbf{v})$$

y entonces, si $\alpha\mathbf{u} + \beta\mathbf{v} \neq \mathbf{0}$, es vector propio de A con valor propio λ . □

Así que los vectores propios vienen en “paquetes” que son subespacios (líneas por el origen es lo más común) asociados a los valores propios. Los conjuntos de vectores propios asociados a un valor propio se llaman *espacios propios* o *eigenespacios*; en ellos, la matriz actúa como una homotecia. Por ejemplo, la matriz

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

tiene como espacio propio del valor propio 2 al plano xy y como espacio propio del valor propio 1 al eje z .

Supongamos que λ, \mathbf{v} es una pareja propia de la matriz A (en particular $\mathbf{v} \neq \mathbf{0}$), entonces

$$\begin{aligned} A\mathbf{v} &= \lambda\mathbf{v} \\ \Leftrightarrow A\mathbf{v} - \lambda\mathbf{v} &= \mathbf{0} \\ \Leftrightarrow (A - \lambda I)\mathbf{v} &= \mathbf{0}. \end{aligned}$$

Esta última ecuación nos dice que la transformación $(A - \lambda I)$ no es biyectiva pues manda a un vector no cero en el cero. Y como $\mathbf{v} \neq \mathbf{0}$, esto último sucede (regresando formalmente al caso de dimensión 2 que sí hemos demostrado, aunque será cierto en general) si y sólo si

$$\det(A - \lambda I) = 0.$$

Esto demuestra el siguiente lema.

Lema 4.8 *Para cualquier matriz A (de 2×2):*

$$\lambda \text{ es un valor propio de } A \Leftrightarrow \det(A - \lambda I) = 0.$$

□

Lema 4.9 *Si A es simétrica 2×2 entonces tiene dos valores propios λ_1 y λ_2 en \mathbb{R} .*

Demostración. Puesto que A es simétrica, $A = A^\top$, la podemos escribir

$$A = \begin{pmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{b} & \mathbf{c} \end{pmatrix}.$$

Consideremos ahora a λ como una variable o incógnita y entonces tenemos

$$\begin{aligned} \det(A - \lambda I) &= \det \left(\begin{pmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{b} & \mathbf{c} \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right) \\ &= \det \begin{pmatrix} \mathbf{a} - \lambda & \mathbf{b} \\ \mathbf{b} & \mathbf{c} - \lambda \end{pmatrix} = (\mathbf{a} - \lambda)(\mathbf{c} - \lambda) - \mathbf{b}^2 \\ &= \lambda^2 - (\mathbf{a} + \mathbf{c})\lambda + (\mathbf{a}\mathbf{c} - \mathbf{b}^2). \end{aligned}$$

Que resulta ser un polinomio de segundo grado en la variable λ .

Este polinomio es tan importante que tiene nombre: $\det(A - \lambda I)$ es el *polinomio característico* de la matriz A .

Nos interesa, según el lema anterior, determinar para cuáles λ se anula este polinomio, es decir, cuáles son sus raíces. Para esto apliquemos la fórmula clásica “del chicharronero” (página 77), la cual expresa las raíces en términos de los coeficientes y en nuestro caso da:

$$\begin{aligned} \frac{(\mathbf{a} + \mathbf{c}) \pm \sqrt{(\mathbf{a} + \mathbf{c})^2 - 4(\mathbf{a}\mathbf{c} - \mathbf{b}^2)}}{2} &= \frac{(\mathbf{a} + \mathbf{c}) \pm \sqrt{\mathbf{a}^2 + 2\mathbf{a}\mathbf{c} + \mathbf{c}^2 - 4\mathbf{a}\mathbf{c} + 4\mathbf{b}^2}}{2} \\ &= \frac{(\mathbf{a} + \mathbf{c}) \pm \sqrt{\mathbf{a}^2 - 2\mathbf{a}\mathbf{c} + \mathbf{c}^2 + 4\mathbf{b}^2}}{2} = \frac{(\mathbf{a} + \mathbf{c}) \pm \sqrt{(\mathbf{a} - \mathbf{c})^2 + 4\mathbf{b}^2}}{2}. \end{aligned}$$

Puesto que $(\mathbf{a} - \mathbf{c})^2 + 4\mathbf{b}^2 \geq 0$ por ser la suma de dos cuadrados, entonces la raíz cuadrada es real y $\det(A - \lambda I) = 0$ tiene dos soluciones reales, que vienen de tomar $+$ o $-$ en la expresión anterior. Explícitamente, sean

$$\lambda_1 = \frac{(\mathbf{a} + \mathbf{c}) + \sqrt{(\mathbf{a} - \mathbf{c})^2 + 4\mathbf{b}^2}}{2}, \quad \lambda_2 = \frac{(\mathbf{a} + \mathbf{c}) - \sqrt{(\mathbf{a} - \mathbf{c})^2 + 4\mathbf{b}^2}}{2}$$

las dos raíces; de tal manera que podemos concluir, por el Lema 4.8, que λ_1 y λ_2 son valores propios de A y que su polinomio característico se escribe

$$\det(A - \lambda I) = (\lambda - \lambda_1)(\lambda - \lambda_2).$$

□

Obsérvese que en la demostración fue crucial que $A^\top = A$; pues en general este lema no es cierto (véanse los ejercicios): no toda matriz A de 2×2 tiene valores propios, como no todo polinomio cuadrático en una variable tiene raíces reales.

Además, de la demostración podemos concluir (para matrices simétricas) que los dos valores propios coinciden si y solamente si $\mathbf{a} = \mathbf{c}$ y $\mathbf{b} = 0$ (pues se debe tener que $(\mathbf{a} - \mathbf{c})^2 + 4\mathbf{b}^2 = 0$); es decir, los dos valores propios coinciden si y sólo si la matriz original es la de una homotecia. Tenemos entonces:

Corolario 4.10 Sea A simétrica 2×2 y sean λ_1, λ_2 sus dos valores propios. Entonces, sus valores propios coinciden ($\lambda_1 = \lambda_2$) si y sólo si $A = \lambda_1 I$. Y en este caso, cualquier vector $\mathbf{v} \neq \mathbf{0}$ es vector propio. \square

Corolario 4.11 Sea A simétrica 2×2 . Entonces existe una base \mathbf{u}, \mathbf{v} de \mathbb{R}^2 donde \mathbf{u} y \mathbf{v} son vectores propios de A .

Demostración. Por el Lema 4.9 sabemos que A tiene dos valores propios, λ_1 y λ_2 . Si éstos fueran iguales, por el corolario anterior cualquier base funciona. Si fueran diferentes ($\lambda_1 \neq \lambda_2$), entonces por la definición de valor propio existen vectores propios \mathbf{u}, \mathbf{v} distintos de $\mathbf{0}$ correspondientes a los valores propios λ_1 y λ_2 respectivamente. Estos dos vectores no pueden ser paralelos pues por el Lema 4.8 sus valores propios coincidirían. Entonces forman una base de \mathbb{R}^2 . \square

Un ejemplo Vale la pena hacer un ejemplo concreto para fijar ideas. Consideremos la matriz simétrica

$$A = \begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix}.$$

Para encontrar sus valores propios hay que resolver su polinomio característico que es

$$\begin{aligned} \det(A - \lambda I) &= \det \left(\begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right) \\ &= \det \begin{pmatrix} 2 - \lambda & 2 \\ 2 & -1 - \lambda \end{pmatrix} \\ &= (2 - \lambda)(-1 - \lambda) - 4 \\ &= \lambda^2 - \lambda - 6. \end{aligned}$$

La fórmula nos dice que sus raíces son

$$\lambda_1 = (1 + \sqrt{1 + 24})/2 = 3 \quad \text{y} \quad \lambda_2 = (1 - 5)/2 = -2,$$

lo cual implica que

$$\det(A - \lambda I) = (\lambda - 3)(\lambda + 2).$$

Aunque en este caso pudo haber sido resuelto sin usar la fórmula, ya tenemos los dos valores propios. Para encontrar los vectores propios correspondientes, habrá que encontrar una solución no trivial de los sistemas $(A - \lambda_i I)\mathbf{x} = \mathbf{0}$. Para $\lambda_1 = 3$ tenemos

$$(A - \lambda_1 I)\mathbf{x} = \begin{pmatrix} -1 & 2 \\ 2 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x + 2y \\ 2x - 4y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix};$$

y una de sus (múltiples) soluciones no triviales es $\mathbf{u}^\top := (2, 1)$. Compruébese directamente que $A\mathbf{u} = 3\mathbf{u}$ como reza la teoría. Para el otro eigenvalor, $\lambda_2 = -2$, el sistema es

$$(A - \lambda_2 I)\mathbf{x} = \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4x + 2y \\ 2x + y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

y una solución no trivial es $\mathbf{v}^\top = (-1, 2)$.

No sólo obtuvimos que los vectores propios \mathbf{u} y \mathbf{v} forman una base de \mathbb{R}^2 , sino mucho más: que son ortogonales. Pero esto, como veremos, no es coincidencia.

EJERCICIO 4.6 Demuestra que la matriz $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ no tiene vectores propios. Da un argumento algebraico y uno geométrico.

EJERCICIO 4.7 Demuestra que la matriz $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ no tiene vectores propios para $b \neq 0$. ¿Puedes dar un argumento geométrico?

Conclusión (diagonalización ortogonal de matrices simétricas)

En nuestro estudio de los vectores propios de las matrices simétricas, saquémosle jugo ahora a lo que ya sabíamos sobre su buen comportamiento respecto al producto interior.

Lema 4.12 *Sea A una matriz simétrica. Si λ_1, \mathbf{u} y λ_2, \mathbf{v} son pares propios tales que $\lambda_1 \neq \lambda_2$, entonces \mathbf{u} y \mathbf{v} son ortogonales.*

Demostración. Usando el Lema 4.5, tenemos

$$\lambda_1(\mathbf{u} \cdot \mathbf{v}) = (\lambda_1 \mathbf{u}) \cdot \mathbf{v} = A\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot A\mathbf{v} = \mathbf{u} \cdot (\lambda_2 \mathbf{v}) = \lambda_2(\mathbf{u} \cdot \mathbf{v}),$$

de donde

$$(\lambda_1 - \lambda_2)(\mathbf{u} \cdot \mathbf{v}) = 0.$$

Si $\lambda_1 \neq \lambda_2$, esto implica que $\mathbf{u} \cdot \mathbf{v} = 0$. □

Podemos entonces concluir.

Teorema 4.13 *Sea A una matriz simétrica 2×2 . Existe una rotación $B \in O(2)$ tal que $B^\top AB$ es diagonal, es decir, tal que*

$$B^\top AB = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

donde λ_1 y λ_2 son los valores propios de A .

Demostración. Por el Lema 4.9, A tiene valores propios $\lambda_1, \lambda_2 \in \mathbb{R}$. Si $\lambda_1 = \lambda_2$ entonces A ya es diagonal por el Corolario 4.10, y B puede tomarse entonces como la identidad. Si $\lambda_1 \neq \lambda_2$, sean \mathbf{u} y \mathbf{v} vectores propios respectivos. Por el Lema 4.12, \mathbf{u} y \mathbf{v} son ortogonales, entonces \mathbf{v} es paralelo a \mathbf{u}^\perp ; y por el Lema 4.7, \mathbf{u}^\perp también es vector propio correspondiente a λ_2 . Sea

$$B = \frac{1}{|\mathbf{u}|} (\mathbf{u}, \mathbf{u}^\perp);$$

puesto que $\mathbf{u} \neq \mathbf{0}$, claramente B es la matriz de una rotación. Y con un cálculo análogo al que hicimos en (4.6) se concluye el teorema. \square

La teoría que hemos desarrollado se resume en una receta muy simple.

Receta para diagonalizar matrices simétricas. El **ingrediente** es una matriz simétrica $A = A^\top$ de 2×2 .

El **primer paso** consiste en resolver su polinomio característico $\det(A - \lambda I)$.

Una vez que se conocen $\lambda_1, \lambda_2 \in \mathbb{R}$ tales que $\det(A - \lambda I) = (\lambda - \lambda_1)(\lambda - \lambda_2)$, **procédase** a encontrar $\mathbf{u} \neq \mathbf{0}$ tal que $(A - \lambda_1 I)\mathbf{u} = \mathbf{0}$ (esto le resultará muy fácil si recuerda las propiedades del compadre ortogonal, pues basta aplicarlo a un renglón de $(A - \lambda_1 I)$).

Y todo ya está cocinado para **declarar** $B = (1/|\mathbf{u}|)(\mathbf{u}, \mathbf{u}^\perp)$. El resultado será que $B^\top AB$ es la matriz diagonal con entradas λ_1 y λ_2 .

EJERCICIO 4.8 Concluye la demostración del Teorema 4.13.

EJERCICIO 4.9 Encuentra la matriz B de una rotación que diagonalice las siguientes matrices simétricas:

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} -6 & 12 \\ 12 & 1 \end{pmatrix} \\ \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix} \quad \begin{pmatrix} -7 & -6 \\ -6 & 2 \end{pmatrix} \quad \begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix}.$$

Calcula $B^\top AB$, donde A es una de las anteriores y B la rotación correspondiente.

EJERCICIO 4.10 Demuestra que si una matriz A (no necesariamente simétrica) tiene dos valores propios distintos entonces existe una matriz $B \in \mathbf{Gl}(2)$ tal que $B^{-1}AB$ es diagonal.

4.4 Clasificación de curvas cuadráticas

Hemos estudiado cómo afectan las traslaciones y las transformaciones ortogonales a los polinomios cuadráticos; ahora juntaremos nuestros resultados para dar su clasificación

isométrica. Pero empecemos con un ejemplo concreto para fijar las ideas, técnicas y resultados que obtuvimos en la sección anterior.

4.4.1 Ejemplo

Consideremos el polinomio cuadrático

$$P(x, y) = 8x^2 + 12xy + 17y^2 - 16x - 12y - 12;$$

queremos saber qué curva define y dónde está. Para pasarlo a su forma vectorial, $P(\mathbf{x}) = \mathbf{x} \cdot \mathbf{A}\mathbf{x} + \mathbf{k} \cdot \mathbf{x} + f$, definimos $f = -12$,

$$\mathbf{A} = \begin{pmatrix} 8 & 6 \\ 6 & 17 \end{pmatrix} \quad \text{y} \quad \mathbf{k} = \begin{pmatrix} -16 \\ -12 \end{pmatrix}.$$

El determinante de \mathbf{A} es $\det(\mathbf{A}) = 8 \times 17 - 6 \times 6 = 100$ distinto de cero, y por tanto $\mathcal{C}(P)$ tiene centro

$$\mathbf{c} = -\mathbf{A}^{-1} \left(\frac{1}{2}\mathbf{k} \right) = \frac{1}{100} \begin{pmatrix} 17 & -6 \\ -6 & 8 \end{pmatrix} \begin{pmatrix} 8 \\ 6 \end{pmatrix} = \frac{1}{100} \begin{pmatrix} 8 \times 17 - 6 \times 6 \\ -8 \times 6 + 6 \times 8 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

De tal manera que si trasladamos la curva $\mathcal{C}(P)$ por el vector $(-1, 0)$ ésta queda definida por el polinomio

$$P_1(\mathbf{x}) = P(\mathbf{x} + \mathbf{c}) = \mathbf{x} \cdot \mathbf{A}\mathbf{x} + P(\mathbf{c}) = \mathbf{x} \cdot \mathbf{A}\mathbf{x} - 20.$$

Ahora queremos encontrar una rotación \mathbf{B} que elimine el término mixto de P_1 , y para esto, seguimos la receta de diagonalización. El polinomio característico de \mathbf{A} es

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \det \begin{pmatrix} 8 - \lambda & 6 \\ 6 & 17 - \lambda \end{pmatrix} = \lambda^2 - 25\lambda + 100 = (\lambda - 5)(\lambda - 20)$$

y entonces sus valores propios son $\lambda_1 = 5$ y $\lambda_2 = 20$. Para el primero de ellos, se tiene que

$$\mathbf{A} - 5\mathbf{I} = \begin{pmatrix} 3 & 6 \\ 6 & 12 \end{pmatrix}$$

y podemos tomar como solución del sistema homogéneo que define a $\mathbf{u}^\top = (2, -1)$, de tal manera que una rotación que diagonaliza a \mathbf{A} es

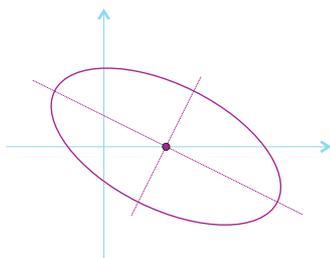
$$\mathbf{B} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}; \quad \mathbf{B}^\top \mathbf{A} \mathbf{B} = \begin{pmatrix} 5 & 0 \\ 0 & 20 \end{pmatrix}.$$

Podemos concluir, por (4.5), que si rotamos el eje x a la línea con dirección $(2, -1)$ el polinomio P_1 toma la forma

$$P_2(\mathbf{x}) = P_1(\mathbf{B}\mathbf{x}) = 5x^2 + 20y^2 - 20$$

que, finalmente, es equivalente a

$$P_0(\mathbf{x}) = \frac{1}{5}P_2(\mathbf{x}) = x^2 + 4y^2 - 4.$$



La curva cuadrática de P_0 es claramente una elipse dada por la ecuación canónica

$$\frac{x^2}{4} + y^2 = 1$$

Entonces podemos concluir que $\mathcal{C}(P)$, la que originalmente queríamos describir, es una elipse con centro en $(1, 0)$, eje mayor de tamaño 2 en la dirección del vector $(2, -1)$ y eje menor 1.

Ya que tenemos los numeritos calculados y frescos, prosigamos con el ejemplo hasta su final afín. Si al círculo unitario lo expandimos en el eje x un factor 2, y luego componemos con la isometría $\mathbf{B}\mathbf{x} + \mathbf{c}$, debemos llegar justo a $\mathcal{C}(P)$. La transformación afín que acabamos de describir tiene la fórmula

$$g \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 2x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{4}{\sqrt{5}}x + \frac{1}{\sqrt{5}}y + 1 \\ -\frac{2}{\sqrt{5}}x + \frac{2}{\sqrt{5}}y \end{pmatrix}.$$

Puesto que dividimos entre 5 y luego entre 4 para obtener la ecuación canónica de la elipse, debe ser cierto, según la teoría que hemos desarrollado, que

$$(P \circ g)(x, y) = 20(x^2 + y^2 - 1).$$

EJERCICIO 4.11 Ensuciate las manos y haz con cuidado todo el trabajo (“talacha” le llamamos en México) que compruebe nuestra última aseveración; simplemente hay que sustituir fórmulas, expandir y simplificar. O bien, prográmalo y revisalo en algún paquete matemático adecuado.

4.4.2 Clasificación isométrica

Consideremos de nuevo el polinomio cuadrático general

$$P(\mathbf{x}) = \mathbf{x} \cdot \mathbf{A}\mathbf{x} + 2\mathbf{b} \cdot \mathbf{x} + c,$$

con $A = A^T \neq \mathbf{0}$; donde el vector constante \mathbf{b} que determina la parte lineal es ahora la mitad del que habíamos tomado antes (\mathbf{k}) para evitar el $\frac{1}{2}$ que nos aparecía; nótese también que hemos cambiado el nombre a la constante. Y sea

$$g(\mathbf{x}) = \mathbf{B}\mathbf{x} + \mathbf{h}$$

una isometría, es decir, $\mathbf{B} \in \mathbf{O}(2)$. Al hacer los cálculos como en las secciones precedentes, se obtiene que

$$\begin{aligned} (\mathbf{P} \circ g)(\mathbf{x}) &= \mathbf{x} \cdot (\mathbf{B}^T \mathbf{A} \mathbf{B}) \mathbf{x} + 2(\mathbf{A}\mathbf{h} + \mathbf{b}) \cdot (\mathbf{B}\mathbf{x}) + \mathbf{P}(\mathbf{h}) \\ &= \mathbf{x} \cdot (\mathbf{B}^T \mathbf{A} \mathbf{B}) \mathbf{x} + 2\mathbf{B}^T(\mathbf{A}\mathbf{h} + \mathbf{b}) \cdot \mathbf{x} + \mathbf{P}(\mathbf{h}). \end{aligned} \quad (4.7)$$

Debemos partir nuestro análisis en dos grandes casos de acuerdo con el determinante de la matriz A .

Caso 1: $\det(A) \neq 0$.

Si escogemos a \mathbf{h} como el centro ($\mathbf{h} = -A^{-1}\mathbf{b}$) y a \mathbf{B} como una rotación que diagonalice A (bien definida salvo el orden y orientación que demos a los vectores propios), entonces \mathbf{P} es isométricamente equivalente a un polinomio de la forma

$$P_1(x, y) = \alpha x^2 + \beta y^2 + \gamma, \quad (4.8)$$

donde α y β son los valores propios de A y γ es el valor del polinomio original evaluado en el centro (nótese que con estos datos bastan).

Ahora bien, tenemos también que α y β son distintos de cero por el siguiente lema que habíamos postergado.

Lema 4.14 *Si A es una matriz simétrica con valores propios α y β , entonces*

$$\det(A) = \alpha\beta$$

Demostración. Por el Teorema de diagonalización, 4.13, sea \mathbf{B} una rotación que diagonaliza a A . Entonces como el determinante es multiplicativo se tiene que

$$\begin{aligned} \alpha\beta &= \det \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} = \det(\mathbf{B}^T \mathbf{A} \mathbf{B}) \\ &= \det(\mathbf{B}^T) \det(A) \det(\mathbf{B}) = 1 \times \det(A) \times 1 \\ &= \det(A). \end{aligned}$$

□

Entonces tenemos naturalmente dos subcasos: cuando $\det(A) > 0$ y los signos de α y β coinciden, o bien cuando son distintos y $\det(A) < 0$.

$\det(A) > 0$: Tenemos tres posibilidades. Primera, si la constante γ es cero (el centro es parte de la “curva” $\mathcal{C}(P)$) entonces la única solución (el único cero de (4.8)) es $(x, y) = (0, 0)$, pues de otra forma es estrictamente del signo de los eigenvalores. Segunda, si la constante γ es del mismo signo que α y β , entonces la curva es vacía pues no hay soluciones (reales, insistamos, pues se le puede llamar “elipse imaginaria”). Y por último, si el signo de γ es el opuesto a los de α y β , los ceros de P_1 (véase (4.8)) claramente coinciden con las soluciones de la ecuación canónica de la elipse:

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

donde $a = \sqrt{-\gamma/\alpha}$ y $b = \sqrt{-\gamma/\beta}$.

$\det(A) < 0$: Sólo tenemos dos posibilidades. Si $\gamma = 0$, P_1 es una diferencia de cuadrados que se puede factorizar en la forma

$$(ax + by)(ax - by)$$

donde hemos supuesto que $\alpha > 0$ (y entonces $a = \sqrt{\alpha}$ y $b = \sqrt{-\beta}$), cuya curva consiste en dos rectas cuya intersección es el centro. O bien, si $\gamma \neq 0$ podemos escoger el primer vector propio (correspondiente a la variable x) de tal manera que su valor propio α tenga signo contrario a γ , y entonces los ceros de P_1 corresponden a las soluciones de la ecuación canónica de la hipérbola:

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

donde $a = \sqrt{-\gamma/\alpha}$ y $b = \sqrt{\gamma/\beta}$.

Caso 2: $\det(A) = 0$.

Puesto que no tenemos la seguridad de eliminar la parte lineal, nos conviene simplificar primero la parte cuadrática (cuidado en los ejercicios, esto modifica la parte lineal). En este caso, por el Lema 4.14, uno de los valores propios es cero, y el otro es distinto de cero pues, si no, por el Corolario 4.10, la matriz A sería la matriz nula que hemos desechado desde el principio. Entonces, podemos diagonalizar (rotar) para que el valor propio correspondiente a x sea el no nulo, y dividiendo entre él obtener que P es isométricamente equivalente a un polinomio de la forma

$$x^2 + 2\alpha x + \beta y + \gamma.$$

Este polinomio siempre se puede reescribir como

$$(x + \alpha)^2 + \beta y + (\gamma - \alpha^2),$$

de tal manera que haciendo el cambio de variable isométrico $x' = x + \alpha$, y olvidando la prima, se simplifica a

$$P_2(x, y) = x'^2 + ay + b. \tag{4.9}$$

donde \mathbf{a} y \mathbf{b} tienen las definiciones obvias. ¿A precomponer con qué traslación corresponde este cambio de variable? Y de nuevo tenemos dos subcasos.

$\mathbf{a} = 0$: Si $\mathbf{b} < 0$ entonces el polinomio (4.9) define (y por tanto $\mathcal{C}(\mathbf{P})$ consiste de) dos rectas paralelas. Si $\mathbf{b} = 0$, es una recta doble; y si $\mathbf{b} > 0$ consiste de “dos rectas imaginarias” (en \mathbb{R}^2 es el conjunto vacío).

$\mathbf{a} \neq 0$: Con una traslación en el eje \mathbf{y} , o equivalentemente, el cambio de variable $\mathbf{y}' = \mathbf{y} + \mathbf{b}/\mathbf{a}$ y olvidando la prima, obtenemos que \mathbf{P} es isométricamente equivalente al polinomio

$$x^2 + \mathbf{a}y,$$

que define una parábola.

En resumen, hemos demostrado que cualquier polinomio cuadrático es isométricamente equivalente a alguna de nueve posibles familias canónicas. Enfatizamos lo de familias, pues nótese que (con isometrías) no podemos deshacernos de ciertos parámetros. Por ejemplo, y para describir los casos importantes, en las elipses nos quedaron los parámetros \mathbf{a} y \mathbf{b} en la ecuación canónica, que podemos suponer positivos (pues como se usan sus cuadrados el signo da lo mismo) y además podemos suponer que $\mathbf{a} \geq \mathbf{b}$ (pues al rotar 90° se intercambian las variables x y y en la ecuación). De tal manera que, como intuitivamente ya sabíamos, hay una clase isométrica de elipses para cada pareja $\mathbf{a} \geq \mathbf{b} > 0$ (los semiejes mayor y menor); los casos extremos $\mathbf{a} = \mathbf{b} > 0$ corresponden a los círculos. En las hipérbolas, no podemos intercambiar las variables (cuyo papel se define respecto a la constante) y tenemos una clase para cada par \mathbf{a}, \mathbf{b} estrictamente positivo. Por último, las parábolas dependen del parámetro $\mathbf{a} > 0$ (pues cambiar \mathbf{y} por $-\mathbf{y}$ es una isometría).

EJERCICIO 4.12 Describe geoméricamente algunas de las curvas cuadráticas definidas por los polinomios siguientes (da el centro, la dirección de los ejes y los parámetros o la ecuación canónica correspondiente). Escoge al menos una elipse, una hipérbola y una parábola.

$$\begin{aligned}
 & -6x^2 + 24xy + y^2 - 12x - 26y - 161 \\
 & 9x^2 - 4xy + 6y^2 - 58x + 24y + 59 \\
 & x^2 + 2xy + y^2 + 2x + 2y - 1 \\
 & 47x^2 + 32xy - 13y^2 - 252x - 12y + 247 \\
 & 5x^2 + 24xy - 5y^2 + 14x - 34y - 37 \\
 & 66x^2 - 24xy + 59y^2 - 108x - 94y + 1 \\
 & 13x^2 + 10xy + 13y^2 + 42x - 6y - 27 \\
 & 9x^2 + 6xy + 17y^2 + 12x - 28y - 52 \\
 & -7x^2 - 12xy + 2y^2 + 40x + 20y - 55 \\
 & 9x^2 - 24xy + 16y^2 + 130x - 90y + 175 \\
 & 18x^2 + 48xy + 32y^2 - 29x + 3y - 22 \\
 & 32x^2 + 48xy + 18y^2 + 31x - 8y - 88 \\
 & -7x^2 + 48xy + 7y^2 + 158x - 6y - 88 \\
 & -24x^2 - 14xy + 24y^2 + 152x - 164y + 151 \\
 & 7x^2 + 48xy - 7y^2 + 116x - 138y - 348.
 \end{aligned}$$

EJERCICIO 4.13 Encuentra un polinomio que defina las siguientes curvas cuadráticas.

- La elipse con semieje mayor 3 en la dirección (3, 4), semieje menor 2 y centro en (-1, 2).
 - La hipérbola con semieje principal 4 en la dirección (2, 1), semieje secundario 1 y centro en (2, 3).
 - La parábola con vértice en (1, 3) directriz en la dirección (1, 4) y distancia focal 1.
-

4.4.3 Clasificación afín y por semejanzas

Para concluir con la clasificación afín que habíamos anunciado al principio de la Sección 4.2, bastará convertir los parámetros que nos quedan (α , β y γ , o bien a y b , en los párrafos anteriores) en 1 o -1 . Y es claro que esto debe lograrse alargando (o encogiéndolo) adecuadamente en los ejes, y saliéndonos entonces del ámbito de las isometrías. Por ejemplo, en el polinomio de la parábola, $P(x, y) = x^2 + ay$ con $a \neq 0$, se obtiene que

$$P\left(x, -\frac{y}{a}\right) = x^2 - y,$$

que equivale a precomponer con la expansión de $-a^{-1}$ en el eje y , y que nos da justo el polinomio de la parábola canónica, (iii) en la Sección 2.1. Éste es el único polinomio con parte lineal de la clasificación isométrica, así que podemos olvidarnos de estos términos.

Si en P_1 (4.8) o en P_2 con $a = 0$, hay un término constante no nulo podemos dividir entre él para hacerlo 1. Y nos queda entonces preocuparnos por los términos cuadráticos puros como αx^2 . Obsérvese que, si en vez de x , sustituimos la expresión ax (una nueva a , cualquier $a \neq 0$), se obtiene $\alpha (ax)^2 = (\alpha a^2) x^2$ de tal manera que el coeficiente puede cambiar, pero nunca cambiará de signo. En términos de matrices, si precomponemos a P_2 (4.9) con una transformación lineal que expanda el eje x por un factor a y el eje y por b , la matriz de la parte cuadrática resultante será:

$$"B^T A B" = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} a^2 \alpha & 0 \\ 0 & b^2 \beta \end{pmatrix}.$$

Así que lo que debemos hacer es tomar $a = (\sqrt{|\alpha|})^{-1}$ y $b = (\sqrt{|\beta|})^{-1}$ para obtener una matriz del tipo

$$\begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}.$$

Con estas observaciones, es fácil ver que los subcasos de la clasificación isométrica dada en la sección anterior corresponden a los incisos y polinomios de la Sección 2.1, con lo que se completa el teorema de clasificación afín.

Si consideramos por último el grupo de semejanzas, entre las isometrías y las afines, en la matriz "B" última que usamos se debe tener que $a = b$. Entonces sólo podemos lograr que uno de los coeficientes se haga ± 1 , y por tanto las elipses (imaginarias o no) y las hipérbolas módulo semejanzas tienen un parámetro que puede ser la razón entre los semiejes o , como habíamos visto en el Capítulo 2, la excentricidad.

EJERCICIO 4.14 ¿Cuál es la matriz de una homotecia que lleve a la parábola dada por $x^2 + ay$, con $a \neq 0$, en la canónica (dada por $x^2 - y$)? Concluye que hay sólo una clase de parábolas módulo semejanzas.

4.5 *Lo que no demostramos

Aunque sí demostramos (detalles más, detalles menos) que cualquier polinomio cuadrático es afinmente equivalente a alguno de los nueve canónicos, no demostramos que estos no son equivalentes entre sí. Es decir, que son justo nueve las clases de equivalencia. Un argumento de tipo topológico va como sigue.

Si un polinomio de una elipse fuera equivalente a uno de una hipérbola, entonces tenemos una transformación afín, *afinidad* debíamos llamarla para resumir,

que manda a la elipse en la hipérbola. Pero esto no es posible porque la hipérbola parte el plano en tres pedazos (que se llaman las “componentes conexas” de su complemento) y la elipse sólo en dos. Para diferenciar a la elipse de la parábola necesitamos usar otra cualidad; la de acotado funciona, pues la elipse lo es mientras que la parábola no. En el fondo estamos usando que las *afinidades* son continuas y que bajo transformaciones continuas estas propiedades se preservan. Para diferenciar a las degeneradas no vacías entre sí, este tipo de argumentos funcionan; y para diferenciarlas con las no degeneradas habrá que usar que estas últimas no contienen rectas y son más que un punto. Sin embargo, con estos argumentos topológicos no podemos diferenciar a los dos polinomios de las curvas vacías, aunque a estos sí de los demás.

Otra cosa que no demostramos es que los polinomios que definen a una misma cónica son únicos salvo constantes diferentes de cero. Como veremos, esto tiene que ver con su grupo de simetrías afines. Consideremos el caso más simple; el del círculo unitario \mathbb{S}^1 cuya matriz asociada es la identidad, pues está definido por el polinomio

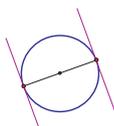
$$P(\mathbf{x}) = \mathbf{x} \cdot \mathbf{x} - 1 = \mathbf{x} \cdot \mathbf{I} \mathbf{x} - 1.$$

Supongamos que otro polinomio cuadrático P_1 también define el círculo unitario, es decir, que $\mathbb{S}^1 = \mathcal{C}(P_1)$. Queremos demostrar que entonces existe $k \neq 0$ tal que $P_1 = kP$. Por el Teorema de Clasificación Afín y el párrafo precedente sabemos que P_1 y P tienen que ser afinmente equivalentes. Entonces existen $g \in \mathbf{Af}(2)$ y $k \neq 0$ tales que

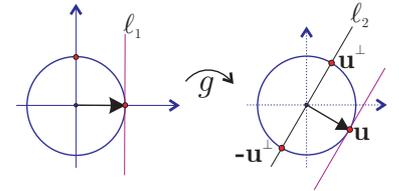
$$P_1 = k(P \circ g).$$

Esto implica que g manda biyectivamente a \mathbb{S}^1 en \mathbb{S}^1 (argumentalo). Es decir, que g es una simetría afín del círculo unitario. Ejemplos de este tipo de las g son las transformaciones ortogonales, y ahora demostraremos con argumentos geométricos que son las únicas.

Para ver que g deja fijo el origen (y que entonces es lineal) hay que determinarlo, como centro del círculo, en términos geométricos: es el punto medio del segmento que une los puntos de contacto de dos tangentes paralelas. Las tangentes son rectas que intersectan el círculo en un único punto; como g manda líneas en líneas, manda entonces tangentes en tangentes. Como además g preserva paralelismo y puntos medios, tiene que dejar fijo el centro y por tanto es lineal.



Sea $\mathbf{u} = g(\mathbf{e}_1) \in \mathbb{S}^1$; nos falta demostrar que $g(\mathbf{e}_2) = \pm \mathbf{u}^\perp$. La tangente a \mathbb{S}^1 en \mathbf{e}_1 , ℓ_1 digamos, es la vertical por \mathbf{e}_1 , que bajo g tiene que ir a la tangente en \mathbf{u} , que es la ortogonal al vector \mathbf{u} por el punto \mathbf{u} . Entonces, el eje y (paralelo a ℓ_1 por el centro) va bajo g a la ortogonal a \mathbf{u} por el origen, sea ésta ℓ_2 . Como $\mathbf{e}_2 \in \mathbb{S}^1 \cap (\text{eje } y)$, entonces $g(\mathbf{e}_2) \in \mathbb{S}^1 \cap \ell_2 = \{\mathbf{u}^\perp, -\mathbf{u}^\perp\}$, y por lo tanto g es ortogonal y se escribe como $g(\mathbf{x}) = \mathbf{B}\mathbf{x}$ con $\mathbf{B} \in \mathbf{O}(2)$. Tenemos entonces la siguiente igualdad de polinomios



$$\begin{aligned} P_1(\mathbf{x}) &= k(P \circ g)(\mathbf{x}) = k(P(\mathbf{B}\mathbf{x})) \\ &= k(\mathbf{x} \cdot (\mathbf{B}^\top \mathbf{I} \mathbf{B}) \mathbf{x} - 1) = k(\mathbf{x} \cdot \mathbf{I} \mathbf{x} - 1) = kP(\mathbf{x}), \end{aligned}$$

pues $\mathbf{B}^\top \mathbf{B} = \mathbf{I}$, que es lo que queríamos demostrar: *los polinomios cuadráticos que definen el círculo unitario, \mathbb{S}^1 , son exactamente los de la forma*

$$k(x^2 + y^2 - 1) \quad \text{con } k \neq 0.$$

Con este caso concreto en la mano, es fácil argumentar que para cualquier elipse los polinomios que la definen difieren por constantes distintas de cero. Pero en general, esta línea de argumentación deja mucho que desear. Habría que argumentar geométricamente en cada caso (llevamos uno de nueve) y obtener el grupo de simetrías afines de una cónica (véase el ejercicio sobre simetrías de la hipérbola). Algo que resulta imposible para las cónicas vacías, de las que tenemos dos, la degenerada y la no degenerada, que no podemos diferenciar geométricamente.

La línea de argumentación clásica va por otro lado; es algebraica. Pero depende de cerrar algebraicamente a los números reales en los complejos, y ahí sí funciona. Si dos polinomios tienen los mismos ceros, entonces, salvo constantes, son potencias de un mismo polinomio, y esto podría considerarse como el principio de la geometría algebraica moderna. Rebasa el contexto de este libro.

EJERCICIO 4.15 Demuestra que el conjunto de afinidades que dejan invariante un polinomio cuadrático P ,

$$\mathbf{Sim}_{\mathbf{Af}}(P) := \{g \in \mathbf{Af}(2) \mid P \circ g = P\},$$

es un grupo.

EJERCICIO 4.16 Para $P(\mathbf{x}) = \mathbf{x} \cdot \mathbf{x} - 1$ demuestra algebraicamente que

$$\mathbf{Sim}_{\mathbf{Af}}(P) = \mathbf{O}(2).$$

EJERCICIO 4.17 Sea \mathcal{H} la hipérbola canónica dada por la ecuación $x^2 - y^2 = 1$. Demuestra que para cualquier $(\mathbf{a}, \mathbf{b}) \in \mathcal{H}$ se tiene que la transformación lineal definida por la matriz

$$\begin{pmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{b} & \mathbf{a} \end{pmatrix}$$

deja invariante \mathcal{H} (es una simetría afín de) \mathcal{H} .

***EJERCICIO 4.18** Demuestra que si $g \in \mathbf{Af}(2)$ es tal que $g(\mathcal{H}) = \mathcal{H}$ (con \mathcal{H} la hipérbola canónica) entonces g es lineal y está dada por una matriz como la anterior, o bien por esa seguida de la reflexión en el eje x .

***EJERCICIO 4.19** Demuestra que dos polinomios que definen la hipérbola canónica difieren por una constante ($P_1 = kP_2$).
