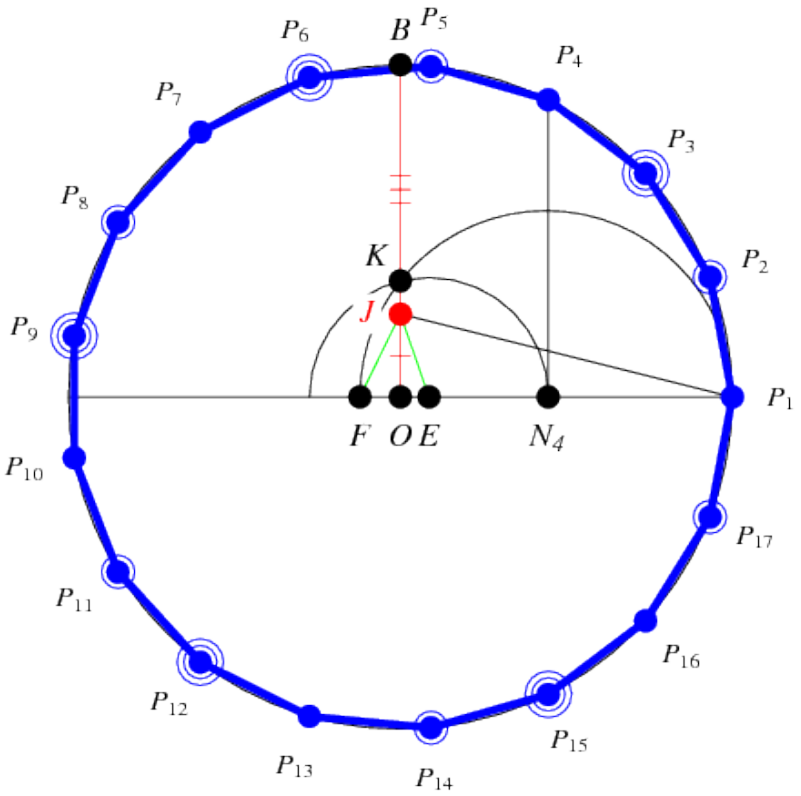
SOBRE LA IMPOSSIBILITAT D'ALGUNES CONSTRUCCIONS AMB REGLE I COMPÀS

Nikolai Lobachevsky



Agraïments

Vull donar les gràcies al meu tutor, José Manuel Mora, per haver acceptat dirigir el meu treball tot i que no em coneixia, pel seguiment que n'ha fet durant aquest temps i, sobretot, pels seus comentaris sempre positius sobre el treball que m'animaven a seguir amb més ganes. D'altra banda, també vull agrair als meus pares tot l'ajut que m'han proporcionat des del primer dia. A la meva mare, pel munt d'hores que l'estiu passat ens vam passar juntes aprenent sobre espais vectorials, cossos, subcossos, extensions, polinomis mínims, grau d'una extensió, etc. I al meu pare, per haver-me proposat el tema d'aquest treball i per haver-me guiat en tot moment al llarg de la seva realització, ajudant-me a entendre coses que no entenia i fent-me veure detalls que a mi constantment se m'escapaven.

Índex

Capítol 1

Introducció

1.1 Les construccions amb regle i compàs

El tema d'aquest treball es remunta a més de 2000 anys enrere i té els seus orígens a la Grècia clàssica. La matemàtica grega d'aquella època es reduïa bàsicament a la geometria i l'aritmètica. A més, la geometria es feia sempre de manera "sintètica", és a dir, sense coordenades, ja que encara no s'havia introduït la idea de les coordenades. [1] És per això que per resoldre molts dels problemes geomètrics que es plantejaven era habitual utilitzar un regle (no graduat) i un compàs, i fer-ho traçant rectes i circumferències. De fet, molts dels problemes que es plantejaven en aquella època ja se'ls plantejaven directament com a construccions d'algun objecte geomètric utilitzant només el regle i el compàs. Per exemple, les primeres tres proposicions del primer dels 13 llibres de què consten els famosos *Elements* d'Euclides (s. III a.C.) plantegen construccions amb regle i compàs, com ara la construcció d'un triangle equilàter que tingui per costat un segment donat (veure Exemple 2.3). En realitat, de les 172 proposicions que apareixen en els primers sis llibres, 48 es refereixen a construccions amb regle i compàs.

Però què s'entén exactament per una construcció amb regle i compàs? En general, en una construcció amb regle i compàs es parteix d'un o més objectes geomètrics, com ara punts, rectes, segments, triangles, circumferències, etc, i es demana de construir algun altre objecte geomètric *utilitzant únicament un regle no graduat i un compàs*. En realitat, encara que les figures de partida no siguin simples punts, sempre vindran determinades per una certa família de punts, i el mateix passarà amb la nova figura que es vol construir, així que en el fons el problema consisteix en partir d'un cert conjunt de punts i construir-ne uns altres. El fet que la construcció s'hagi de fer amb regle i compàs vol dir que els nous punts s'han de construir únicament traçant rectes i circumferències que estiguin determinades per punts que ja es tenen prèviament. Així, les rectes han de passar per dos punts ja donats, i les circumferències han de tenir el centre en un punt ja donat i passar per un altre punt també donat o tenir per radi la distància entre dos punt ja donats. Els nous punts seran aleshores les interseccions entre aquestes rectes i circumferències, i s'afegiran als punts que ja es tenien. Al capítol següent donarem una definició més precisa, però per tal que s'entengui millor què volem dir, donem a continuació alguns exemples senzills. Tots ells apareixen ja als *Elements* d'Euclides.

[1] El mètode de coordenades l'introdueix molt més tard, al s.XVII, el matemàtic i filòsof francès René Descartes.

Exemple 1 (*Punt mig d'un segment*) Es tenen donats dos punts A i B qualssevol i es tracta de construir el punt mig del segment que els té per extrems. El procediment és el següent (veure Fig. 1.1).

1. Tracem la recta que passa per A i B.
2. Dibuixem la circumferència de centre A i radi AB.
3. Dibuixem la circumferència de centre B i radi AB. Siguin P i Q els dos punts d'intersecció d'aquesta circumferència amb l'anterior.
4. Tracem la recta que passa per P i Q. Sigui M el punt d'intersecció d'aquesta recta amb la recta AB.

El punt M és aleshores el punt mig buscat.

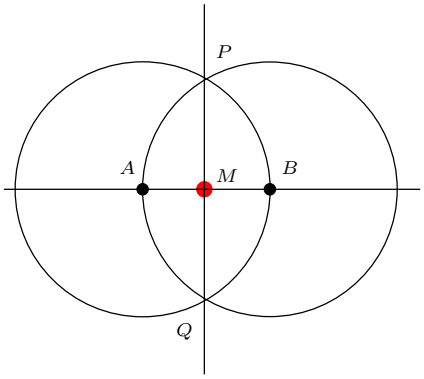


Figura 1.1: Punt mig d'un segment

Exemple 2 (*Perpendicular a una recta per un punt donat*) Es té donada una recta r i un punt P (pot ser de r o no) i es tracta de construir la perpendicular a r per P. Donar la recta r equival a donar-ne dos punts A i B, de manera que en aquest cas el conjunt de punts partida consta de dos o tres punts, segons que P sigui un dels punts A, B o no. Es tracta aleshores de construir almenys un altre punt de la perpendicular buscada. El procediment és el següent (veure Fig. 1.2).

1. Dibuixem la circumferència de centre P i de radi AP (si $P \neq A$; si $P = A$, agafem radi BP). Tallarà la recta r en un altre punt A'.
2. Tracem la circumferència de centre A i radi AA'.
3. Tracem la circumferència de centre A' i radi AA'. Siguin Q i Q' els dos punts de tall d'aquesta circumferència amb l'anterior.

La recta que passa per Q i Q' és aleshores la perpendicular buscada. La primera demostració d'aquesta construcció s'atribueix a un matemàtic grec, Oenipides de Chios, del s. V a.C.

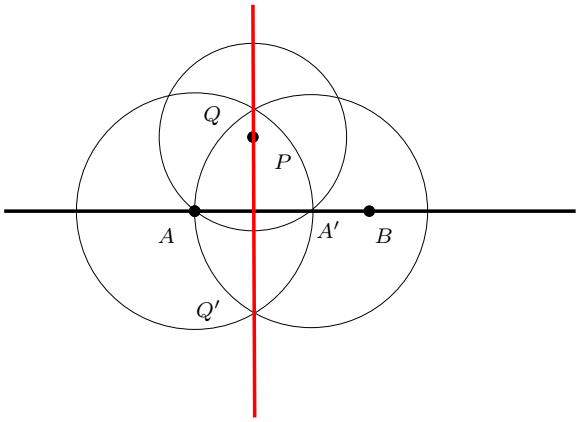


Figura 1.2: Perpendicular a una recta per un punt (exterior) donat

Exemple 3 (*Paral·lela a una recta per un punt donat*) Es té donada una recta r i un punt P (pot ser de r o no) i es tracta de construir la paral·lela a r que passa per P. Com abans, donar r equival a donar-ne dos punts A i B, de manera que el conjunt de punts de partida consta també de dos o tres punts, segons que P sigui un dels punts A, B o no. Es tracta aleshores de construir almenys un altre punt de la paral·lela buscada. El procediment és el següent.

1. Tracem la perpendicular a r per P pel mètode anterior. Sigui s aquesta perpendicular.
2. Tracem la perpendicular a s per P pel mètode anterior. Sigui t aquesta perpendicular.

La recta t és aleshores la paral·lela buscada.

Exemple 4 (*Bisectriu d'un angle*) Es té donat un angle qualsevol i es tracta de construir-ne la seva bisectriu. Donar l'angle equival a donar-ne el vèrtex O i un punt de cada costat A i B, de manera que el conjunt de punts de partida ara consta de tres punts. Es tracta aleshores de construir almenys un altre punt de la bisectriu buscada. El procediment és el següent (veure Fig. 1.3).

1. Dibuixem la circumferència amb radi OB i centre O. Sigui D el punt d'intersecció de la circumferència anterior amb la recta OA.
2. Tracem la circumferència de centre B i radi BD.
3. Tracem la circumferència de centre D i radi BD. Sigui C qualsevol dels punts d'intersecció d'aquesta circumferència amb l'anterior.

La recta OC és aleshores la bisectriu buscada.

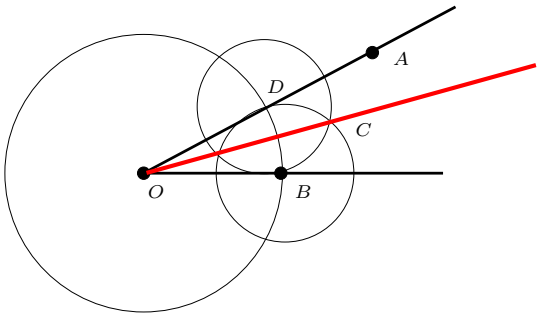


Figura 1.3: Bisectriu d'un angle

Exemple 5 (*Duplicació del quadrat*) Es té donat un quadrat i es tracta de construir un altre quadrat d'àrea doble amb un dels vèrtexs igual a un dels del quadrat original. Donar el quadrat equival a donar-ne els vèrtexs A, B, C, D de manera que el conjunt de punts de partida consta ara de quatre punts. L'objectiu és construir els altres tres vèrtexs del nou quadrat. El procediment és el següent (veure Fig. 1.4).

1. Dibuixem la circumferència de centre A i radi AC (diagonal del quadrat). Siguin P, P' les interseccions d'aquesta circumferència amb la recta AB i R, R' les interseccions amb la recta AD.
3. Tracem la perpendicular a la recta AB per P seguint els passos de l'Exemple 2 anterior.
4. Tracem la perpendicular a la recta AD per R pel mateix mètode. Sigui Q la intersecció d'aquesta perpendicular amb la perpendicular anterior.

El quadrilàter $APQR$ és aleshores el quadrat buscat.

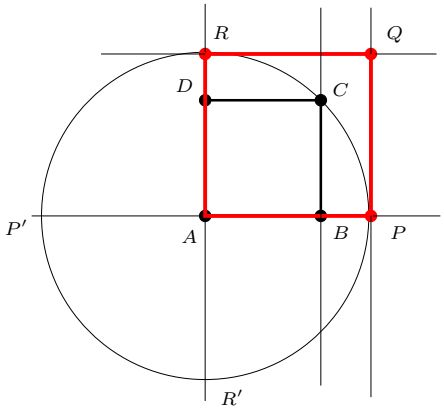


Figura 1.4: Duplicació d'un quadrat

Evidentment, en cada cas és necessari justificar que l'objecte construït és realment el que es demana. Per exemple, en el cas de la construcció del punt mig d'un segment, els punts P i Q

equidisten de A i de B i la seva distància a qualsevol dels dos punts és igual a la distància entre A i B. Per tant, ABP i ABQ són triangles equilàters i $AQBP$ és un rombe, i sabem que les diagonals d'un rombe es tallen en el punt mig.

1.2 Els tres problemes clàssics

Ja al s.V a.C. els matemàtics grecs es van plantejar construccions amb regle i compàs aparentment simples però que no eren capaços de fer. Amb el pas del temps, alguns d'aquests problemes van esdevenir famosos perquè continuaven sense saber-se resoldre. Especialment famosos es van fer els que es coneixen com els *tres problemes clàssics* de construccions amb regle i compàs, que són:

(1) la duplicació del cub,

(2) la trisecció de l'angle, i

(3) la quadratura del cercle.

Van continuar sense resoldre durant més de 2000 anys i no va ser fins al s. XIX que finalment es va acabar demostrant que en realitat cap dels tres tenia solució. Consistien en el següent.

(1) La duplicació del cub. Es tractava de partir d'un cub de volum donat i construir-ne un altre de volum doble. En realitat, tot i que el problema es plantejava com un problema a l'espai, en el fons es pensava com un problema en el pla, en el sentit que el que de fet es volia era construir amb regle i compàs una cara del nou cub. Ara bé, donat un segment qualsevol, ja se sabia construir amb regle i compàs un quadrat que tingués aquest segment per costat. Com que el volum d'un cub no és més que el cub de la longitud del seu costat, del que en el fons es tractava, doncs, era de construir amb regle i compàs un segment de longitud $a\sqrt[3]{2}$ (corresponent a un dels costats del cub de volum doble) a partir d'un segment donat de longitud a (corresponent a un dels costats del cub original), on a era una longitud qualsevol. Per exemple, si el cub original era de costat 1, s'havia de construir un segment de longitud $\sqrt[3]{2}$.

(2) La trisecció de l'angle. Es tractava de partir d'un angle qualsevol i construir amb regle i compàs les dues semirectes amb origen al vèrtex de l'angle que el dividien en tres angles iguals. En realitat, a diferència dels altres dos problemes, aquest si que té solució per alguns angles, però no en general. Més endavant (veure Teorema 42 del Capítol 4) donarem un criteri per saber quan un angle es pot trisecar i quan no amb regle i compàs.

(3) La quadratura del cercle. Es tractava de partir d'un cercle qualsevol i construir amb regle i compàs un quadrat de la mateixa àrea. [2] Tenint en compte que l'àrea d'un cercle de radi r és πr^2, del què es tractava en el fons era de, donat un segment de longitud r (el radi del cercle), construir amb regle i compàs un segment de longitud $r\sqrt{\pi}$ (el costat del quadrat de la mateixa àrea). Per exemple, si el segment era de longitud 1, s'havia de construir un segment de longitud $\sqrt{\pi}$.

Dels tres, aquest últim és el que més famós es va fer perquè en certa manera és més impossible encara que els altres dos. Més endavant en veurem el motiu.

[2] Existeix un altre problema totalment diferent que també es coneix com la quadratura del cercle. El va plantejar el matemàtic polonès Alfred Tarski l'any 1925. Es tracta de saber si és possible tallar un cercle en infinits trossos i tornar-los a unir després per formar un quadrat. Veure Ian Stewart, *De aquí al infinito*, Cap. 12, Ed. Crítica (1998).

1.3 Objectiu del treball i metodologia de presentació

L'objectiu d'aquest treball no és descriure construccions més o menys sofisticades amb regle i compàs, sinó explicar per què aquestes tres construccions són impossibles de fer amb regle i compàs. En realitat, la motivació principal del treball ha estat entendre com s'ho van poder fer els matemàtics per *demostrar* que les tres construccions eren impossibles. D'entrada, em semblava increïble que es pogués demostrar una cosa així. Però en aquest treball veurem que realment es pot fer i explicaré els detalls de com es fa. Per poder-ho demostrar, caldrà entendre més a fons els nombres reals i, en particular, estudiar els anomenats *subcossos* dels reals i les seves *extensions*. Pel camí, també entendrem per què es va trigar tant a demostrar que les construccions eren impossibles. Veurem que, en part, va ser precisament perquè faltava que algú introduís la idea de coordenades en el pla.

Pel que fa a la presentació, tot i que no sempre serà així, he intentat anar explicant el per què de cada afirmació que faig. Això farà que en alguns moments, especialment al Capítol 3, l'exposició resulti més complicada de seguir. Per aquest motiu, he mirat d'anar posant exemples de tot el que vaig explicant (en color blau). Malgrat que al principi pensava evitar ser massa formal en la manera de presentar les coses, al final he acabat fent, en part, el que és bastant habitual en els textos matemàtics: enunciar les definicions i alguns dels resultats en un format diferent per destacar-los. En el cas dels resultats, que apareixen com a Proposicions o com a Teoremes, la prova apareix a continuació.

1.4 Una mica d'història

Abans d'entrar pròpiament en matèria, repassem una mica la història dels tres problemes, des que es van plantejar i per què, fins que es va provar que no tenien solució.

És molt probable que l'origen del problema de la duplicació del cub es trobi en l'intent de generalitzar a tres dimensions el problema anàleg que es podia plantejar en el cas d'un quadrat, un problema que se sabia resoldre i que en realitat és molt senzill de resoldre (veure Exemple 5). Ja en el diàleg de Plató *Menon* apareix Sòcrates dialogant amb un jove esclau per fer-li entendre com es pot duplicar un quadrat. Existeix, però, una llegenda que atribueix l'origen del problema a la petició del déu Apol·lo que se li fes un altar el doble de gran que l'altar cúbic que se li havia construït al temple de Delos, a canvi que ell fes desaparèixer la pesta que s'havia escampat pel poble. Segons la llegenda, es va construir un altar de costat el doble i la pesta va continuar... Per aquest motiu, el problema de la duplicació del cub també s'anomena *problema de Delos*.

Pel que fa al problema de la trisecció de l'angle, és probable que també es plantegés com un anàleg del problema de la bisecció d'un angle o bé del de la trisecció d'un segment, ja que tots dos es podien resoldre fàcilment amb regle i compàs (veure Exemple 4).

Finalment, el problema de la quadratura del cercle es creu que es devia plantejar com a conseqüència del fet que no es coneixia amb exactitud el valor del nombre π i, per tant, que no es podia calcular amb exactitud l'àrea d'un cercle. La idea era intentar reduir el càlcul de l'àrea d'un cercle al de l'àrea d'un quadrat. De fet, ja en un papir egipci que data del 1650 a.C., conegut com a *papir de Rhind*, i que conté un total de 85 problemes resolts, se'n troba un on es pot llegir el següent: "*Construir un quadrat equivalent a un cercle. Resposta. Retirar 1/9 al diàmetre i construir el quadrat amb el que queda*". Es pot comprovar que això equival a agafar

com a valor de π

$$\pi = \frac{256}{81} = 3 + \frac{13}{81} \simeq 3.16$$

En realitat, els grecs es van plantejar el problema de la quadratura per qualsevol figura plana. Quadrar una figura plana volia dir trobar un quadrat que tingués àrea igual a la de la figura. En el cas de figures poligonals els mateixos grecs ja van veure que era possible de fer i com fer-ho. Només calia dividir-la en triangles i quadrar els diferents triangles, així que el problema es reduïa a saber quadrar triangles, i això era fàcil de fer un cop se sabia com quadrar un rectangle qualsevol. [3]

Com que no eren capaços de resoldre cap d'aquests problemes utilitzant només rectes i circumferències, els matemàtics grecs van dedicar-se a buscar solucions utilitzant altres tipus de corbes, com ara còniques que no fossin circumferències o corbes especials que definien "mecànicament".

No és fins al Renaixement que es comença a sospitar que els problemes no tenen solució. Molt resumida, la història a partir d'aquí és la següent.

El matemàtic francès François Viéte (1540-1603), estudiant l'equació de tercer grau, se n'adona de la relació que existeix entre el problema de la trisecció de l'angle i el de la resolució d'una equació d'aquest tipus. A la mateixa època, el filòsof i matemàtic també francès René Descartes (1596-1650) també se n'adona de la relació entre les construccions amb regle i compàs i la resolució d'equacions de primer i segon grau, i és capaç de demostrar (1637) que, donats segments de longituds a i b, es poden construir amb regle i compàs segments de longituds $a + b$, $|a - b|$, ab, a/b i $\sqrt{a}$. Més endavant veurem com es pot fer. En el seu treball, hi ha la llavor de la demostració de la impossibilitat de la trisecció de l'angle i de la duplicació del cub.

Uns anys més tard, el 1668, el matemàtic escocès James Gregory (1638-1675) publica un treball en el que pretén haver demostrat la impossibilitat de la quadratura del cercle, però ell mateix se n'adona poc després que la seva demostració és incorrecta.

Al llarg del s. XVIII, no hi ha cap progrés important en el tema, i és al segle següent quan ja es troben per fi demostracions de la impossibilitat de cadascuna de les tres construccions. La primera demostració satisfactòria tant de la impossibilitat de la duplicació del cub com de la trisecció de l'angle la troba un altre matemàtic francès, Pierre Wantzel (1814-1848), l'any 1837. El seu treball apareix publicat al *Journal de Mathématiques* amb el títol de "Recherche sur les moyens de reconnaître si un problème de géometrie peut se résoudre par la règle et le compas" (veure Fig. 1.5). Aquest treball també va permetre entendre millor el problema de la quadratura del cercle perquè va fer veure que el que importava no era el valor de π sinó la seva naturalesa. La qüestió era si el nombre π era algebraic o transcendent (més endavant explicarem el significat d'això). Però no és fins a l'any 1882 que el matemàtic alemany Ferdinand von Lindemann (1852-1939) aconsegueix provar que π és un nombre transcendent. És amb aquesta demostració que queda definitivament provada també la impossibilitat de quadrar el cercle amb regle i compàs.

[3] Una manera de fer-ho es pot trobar en el llibre de William Dunham, *Viaje a través de los genios. Biografías i teoremas de los grandes matemáticos*, Ed. Pirámide (2004).

Recherches sur les moyens de reconnaître si un Problème de Géométrie peut se résoudre avec la règle et le compas ;

PAR M. L. WANTZEL,
Élève-Ingénieur des Ponts-et-Chaussées.

I.

Supposons qu'un problème de Géométrie puisse être résolu par des intersections de lignes droites et de circonférences de cercle : si l'on joint les points ainsi obtenus avec les centres des cercles et avec les points qui déterminent les droites on formera un enchaînement de triangles rectilignes dont les éléments pourront être calculés par les formules de la Trigonométrie ; d'ailleurs ces formules sont des équations algébriques qui ne renferment les côtés et les lignes trigonométriques des angles qu'au premier et au second degré ; ainsi l'inconnue principale du problème s'obtiendra par la résolution d'une série d'équations du second degré dont les coefficients seront fonctions rationnelles des données de la question et des racines des équations précédentes. D'après cela, pour reconnaître si la construction d'un problème de Géométrie peut s'effectuer avec la règle et le compas, il faut chercher s'il est possible de faire dépendre les racines de l'équation à laquelle il conduit de celles d'un système d'équations du second degré composées comme on vient de l'indiquer. Nous traiterons seulement ici le cas où l'équation du problème est algébrique.

II.

Considérons la suite d'équations :

$$(A)\begin{cases} x_1^2+Ax_1+B=0,\ x_2^2+A_1x_2+B_1=0\ldots x_{n-1}^2+A_{n-2}x_{n-1}+B_{n-2}=0, \\ x_n^2+A_{n-1}x_n+B_{n-1}=0, \end{cases}$$

dans lesquelles A et B représentent des fonctions rationnelles des quantités données p, q, $r\ldots$; A_1 et B_1 des fonctions rationnelles de x_1, p, q, $\ldots$; et, en général, A_m et B_m des fonctions rationnelles de x_m, $x_{m-1}\ldots x_1$, p, $q\ldots$.

Toute fonction rationnelle de x_m telle que A_m ou B_m, prend la forme $\frac{C_{m-1}x_m+D_{m-1}}{E_{m-1}x_m+F_{m-1}}$ si l'on élimine les puissances de x_m supérieures à la pre-

Figura 1.5: Primera pàgina de l'article original de P. Wantzel del 1837

Capítol 2

Formalització de les construccions amb regle i compàs

En aquest capítol formalitzem les construccions amb regle i compàs. Per fer-ho, introduirem el concepte de *punt constructible* i, més en general, el de *figura constructible* (pot ser una recta, una circumferència, un polígon, etc). Reduirem el fet que una construcció sigui possible o no a què siguin constructibles certs punts. Després explicarem breument com el problema geomètric d'identificar quins punts són constructibles i quins no es pot traduir a un problema algebraic utilitzant coordenades i introduint el concepte de *nombre (real) constructible*. Els detalls d'aquesta traducció els abordarem al capítol següent.

2.1 Punts i figures constructibles

Com ja hem dit abans, en una construcció amb regle i compàs es parteix d'un conjunt finit d'objectes geomètrics, que poden o no ser simples punts, i es tracta de construir-ne algun altre, que pot també ser un simple punt o alguna altra figura. També hem remarcat a la introducció que sempre és possible reduir-se al cas que el conjunt d'objectes geomètrics de partida és una família de punts i que el que cal construir és una nova família de punts que determinin totalment la figura que es demana.

Per exemple, en la construcció de la bisectriu d'un angle, les dades de partida són tres punts (el vèrtex de l'angle i un punt qualsevol de cada costat de l'angle diferent del vèrtex) i cal construir un altre punt de la bisectriu diferent del vèrtex de l'angle. Quan es demana duplicar un quadrat, les dades de partida són els vèrtexs del quadrat i cal construir els quatre vèrtexs del nou quadrat.

Per tant, qualsevol construcció amb regle i compàs en el fons consistirà en construir una sèrie de punts a partir d'una família de punts donats. Anomenarem **punts base** els punts de partida, i denotarem per $\mathcal{B}$ el conjunt de punts base.

El que caracteritza les construccions amb regle i compàs és com es construeixen nous punts a partir dels punts base. Només es poden fer les dues operacions següents:

(1) traçar amb el regle una recta que passi per dos punts base o, en general, per dos punts prèviament construïts a partir dels punts base, i

(2) traçar amb el compàs una circumferència centrada en un punt base o, en general, en un

punt prèviament construït a partir dels punts base, i que tingui per radi la distància entre dos punts base o, en general, dos punts prèviament construïts a partir dels punts base.

Els punts d'intersecció d'aquestes rectes i circumferències entre elles donaran nous punts que s'afegiran als punts base i als ja construïts i que es podran fer servir per traçar noves rectes i circumferències. Per la seva banda, les interseccions d'aquestes noves rectes i circumferències entre elles portaran a una nova família de punts que s'afegiran als punts base i a tots els construïts prèviament, i així successivament.

A partir d'un conjunt finit $\mathcal{B}$ de punts base, i per moltes vegades que repetim el procés, no s'arribarà a construir qualsevol altre punt del pla per aquest mètode. Els que sí que es poden construir s'anomenen *punts constructibles* a partir de $\mathcal{B}$. La definició precisa és la següent:

Definició 6 *Siguin $\mathcal{B} = \{P_1, ..., P_k\}$ una família de punts base del pla, amb $k \geq 2$. Aleshores:*

(1) S'anomena **punt constructible en un pas a partir de $\mathcal{B}$** *tot punt que sigui la intersecció de dues rectes, dues circumferències o una recta i una circumferència definides cadascuna a partir de punts de $\mathcal{B}$.*

(2) S'anomena **punt constructible a partir de $\mathcal{B}$** *tot punt Q per al qual existeix una successió de punts $Q_1, \ldots, Q_n$ tal que $Q = Q_n$ i cada Q_i, per $1 \leq i \leq n$, és constructible en un pas a partir del conjunt $\mathcal{B} \cup \{Q_1, \ldots, Q_{i-1}\}$.*

Observar que els punts base sempre són constructibles en un pas perquè qualsevol punt base P és la intersecció de la recta que passa per P i per un altre punt base qualsevol P' amb la circumferència de centre P' i radi la distància de P a P'.

Exemple 7 Si $\mathcal{B}$ consta només de dos punts P, P' (punts negres), a partir d'ells només es poden traçar la recta PP' i les dues circumferències de centres P i P' i radi la distància de P a P' (veure Fig. 2.1). Per tant, els únics punts constructibles en un pas a partir de $\mathcal{B}$ (punts vermells) són els punts base P, P' i els punts Q_1, Q_2, Q_3, Q_4. Un punt no constructible en un pas però sí constructible és el punt Q (en verd). És constructible a partir de $\{P, P', Q_3\}$.

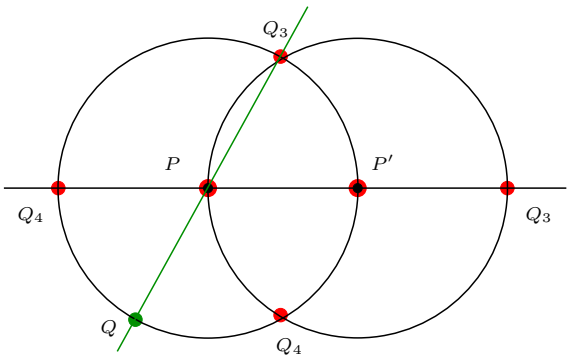


Figura 2.1: Punts constructibles en un pas a partir de $\mathcal{B} = \{P, P'\}$

Exemple 8 Si es parteix dels punts base $\mathcal{B} = \{P_1, P_2, P_3, P_4, P_5\}$ de la Figura 2.2, els punts Q_1, Q_2, Q_3 (en vermell) són constructibles en un pas a partir de $\mathcal{B}$. Per exemple, Q_2 és la intersecció de la recta que passa per P_1 i P_2 amb la circumferència de centre P_3 i radi la distància entre P_4 i P_5. Però a diferència d'abans hi haurà molts més punts constructibles en un pas perquè a partir dels P_i's es poden construir moltes més rectes i circumferències de les que hem dibuixat. Per altra banda, el punt Q_4 (en verd) potser no és constructible en un pas però si que és constructible, perquè es pot construir en un pas a partir de $\mathcal{B} \cup \{Q_2\}$ (és la intersecció de la recta per P_3 i Q_2 amb la recta per P_4 i P_5).

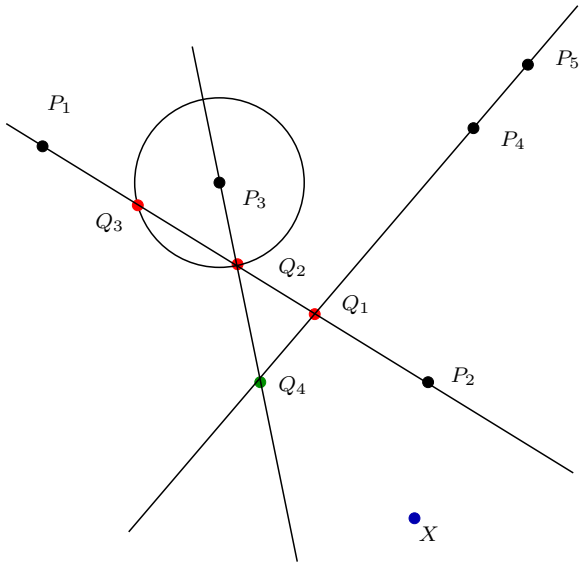


Figura 2.2: Alguns punts constructibles a partir de $\{P_1, P_2, P_3, P_4, P_5\}$

Un cop definit el concepte de punt constructible a partir de $\mathcal{B}$ ja és fàcil definir què s'entén en general per *figura constructible* a partir de $\mathcal{B}$.

Definició 9 *Siguin $\mathcal{B} = \{P_1, ..., P_k\}$ una família de punts base del pla amb $k \geq 2$. S'anomena* **figura constructible a partir de $\mathcal{B}$** *qualsevol figura que quedi totalment determinada per una família finita de punts constructibles a partir de $\mathcal{B}$. En particular:*

(1) Una **recta constructible a partir de $\mathcal{B}$** *és qualsevol recta que conté almenys dos punts constructibles a partir de $\mathcal{B}$ (no cal que ho siguin tots).*

(2) Una **circumferència constructible a partir de $\mathcal{B}$** *és qualsevol circumferència tal que el centre i radi són constructibles a partir de $\mathcal{B}$, entenent que el radi és constructible si ho són dos punts que es troben a una distància igual al radi.*

(3) Un **polígon constructible a partir de $\mathcal{B}$** *és qualsevol polígon tal que tots els seus vèrtexs són constructibles a partir de $\mathcal{B}$.*

Exemple 10 Si $\mathcal{B} = \{P, P'\}$, els dos triangles equilàters que tenen el segment PP' com a costat són constructibles a partir de $\mathcal{B}$. Concretament, el tercer vèrtex Q_1 o Q_2 (veure Fig. 2.3) és constructible en un pas a partir de $\mathcal{B}$ perquè és la intersecció de les circumferències centrades en cadascun dels punts P i P' de radi la distància entre ells. Aquesta construcció és la que apareix com a primera proposició del llibre dels *Elements* d'Euclides.

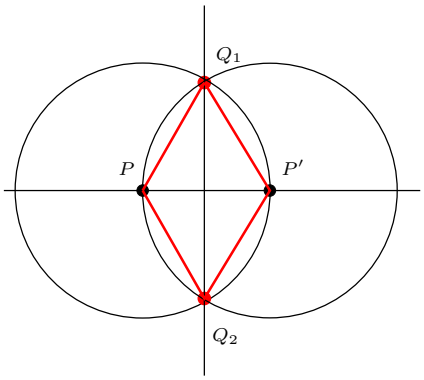


Figura 2.3: Construcció del triangle equilàter amb un segment donat com a costat

Exemple 11 Si $\mathcal{B} = \{P, P'\}$, els dos hexàgons regulars que tenen el segment PP' com a costat són constructibles a partir de $\mathcal{B}$. Concretament, el tercer vèrtex Q_1 (veure Fig. 2.4; només dibuixem una de les dues solucions) és constructible a partir de $\mathcal{B}$ perquè es té la successió X, Q_1, on X és constructible en un pas a partir de P, P' (és la intersecció de les circumferències centrades a P i P' de radi PP') i Q_1 ho és a partir de X, P' (és la intersecció de les circumferències centrades a X i P' de radi XP'). La resta de vèrtexs Q_2, Q_3, Q_4 de l'hexàgon es construeixen de manera anàloga: Q_2 en un pas a partir de X i Q_1, Q_3 en 1 pas a partir de X i Q_2 i Q_4 en un pas a partir de X i Q_3. Per tant, tots són constructibles a partir de $\mathcal{B}$.

En general, entendrem que una construcció es pot fer amb regle i compàs només quan tots els punts i figures que calgui construir siguin constructibles a partir dels punts base donats.

2.2 Nombres constructibles

A partir d'un conjunt $\mathcal{B}$ de punts base ja hem remarcat que no es podrà construir qualsevol altre punt del pla. En realitat, encara que continuéssim construint punts indefinidament, es pot demostrar que el conjunt de punts que es poden construir a partir d'un $\mathcal{B}$ finit donat és molt petit comparat amb el conjunt de tots els punts del pla. [1]

[1] Més exactament, és el que s'anomena un *conjunt numerable*, és a dir, un conjunt tal que existeix una bijecció del conjunt dels nombres naturals cap a ell (la bijecció *numerarà* els elements del conjunt). Es pot demostrar que el conjunt de tots els punts del pla no és numerable.

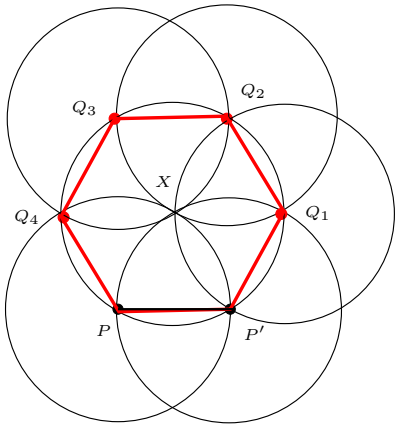


Figura 2.4: Construcció de l'hexàgon regular amb un segment donat com a costat

El nostre objectiu és saber identificar quins punts són constructibles a partir de $\mathcal{B}$ i quins no. No és un problema evident. Per ex., si tornem a la Figura 2.2, ens podem preguntar si un punt, com ara el punt X, és constructible a partir del $\mathcal{B}$ donat. Veure si ho és en principi vol dir veure si existeix una successió finita de punts $Q_1, Q_2, \ldots, Q_r$ tal que l'últim sigui el punt X buscat i que cadascun d'ells sigui constructible en un pas a partir dels anteriors i dels punts base. Però això no és gens evident si el punt es tria a l'atzar.

Per identificar "fàcilment" quins punts són constructibles i quins no, serà necessari convertir el problema geomètric en un problema algebraic. Aquest és un pas clau i és precisament el pas que els matemàtics grecs no van arribar a fer. El motiu és que la traducció es fa introduint un sistema de coordenades en el pla i identificant cada punt amb les seves coordenades, i els matemàtics grecs desconeixien la idea de les coordenades.

A partir d'ara, doncs, suposarem fixat un sistema de coordenades al pla, de manera que cada punt tindrà unes coordenades concretes (x, y). Sempre suposarem que el sistema de coordenades l'agafem amb origen a un dels punts base i amb el punt $(1, 0)$ en un altre punt base. D'aquesta manera, els eixos de coordenades són rectes constructibles a partir de $\mathcal{B}$. L'eix d'abscisses ho és perquè ho són dos dels seus punts i l'eix d'ordenades gràcies a l'Exemple 2. Definim aleshores el següent:

Definició 12 *Donat un conjunt $\mathcal{B}$ de punts base, anomenarem* **nombre (real) constructible** *a partir de $\mathcal{B}$ qualsevol nombre real $t \in \mathbb{R}$ que sigui l'abscissa o l'ordenada d'un punt constructible a partir de $\mathcal{B}$.*

Exemple 13 Si els punts P, P' de l'Exemple 11 són els de coordenades $P = (0, 0)$ i $P' = (1, 0)$, aleshores la resta de vèrtexs de l'hexagon regular construit són els punts de coordenades $Q_1 = (3/2, \sqrt{3}/2)$, $Q_2 = (1, \sqrt{3})$, $Q_3 = (0, \sqrt{3})$ i $Q_4 = (-1/2, \sqrt{3}/2)$. Per tant, els nombres reals

$$0,\ 1,\ -\frac{1}{2},\ \frac{3}{2},\ \sqrt{3},\ \frac{\sqrt{3}}{2}$$

són constructibles a partir de $\mathcal{B} = \{(0, 0), (1, 0)\}$.

És important observar que la constructibilitat d'un real t és equivalent a què sigui constructible el punt $(t, 0)$, ja que ho farem anar de seguida. La raó és la següent. D'acord amb la definició anterior, si t és constructible és que existeix algun punt constructible Q de coordenades (t, y) o bé (x, t) (o bé l'abscissa o bé l'ordenada val t). Discutim cada cas per separat.

- Si és $Q = (t, y)$, $(t, 0)$ serà constructible perquè és la intersecció de l'eix OX, una recta constructible, amb la perpendicular a OX que passa per P, que també és constructible (veure Exemple 2 més amunt).
- Si és $Q = (x, t)$, $(t, 0)$ és igualment constructible, però ara pel procediment següent. Tracem la perpendicular a l'eix OY per Q (recta constructible). La seva intersecció amb OY dóna el punt constructible $(0, t)$. A continuació, tracem la circumferència centrada a l'origen i que passa per $(0, t)$. Talla l'eix OX en el punt buscat $(t, 0)$.

Anàlogament, es pot veure que la constructibilitat d'un real t' equival a la del punt $(0, t')$. Així doncs, tenim el següent.

Lema 14 *Si $t \in \mathbb{R}$, les condicions següents són equivalents:*

(1) t és constructible a partir de $\mathcal{B}$.

(2) El punt $(t, 0)$ és constructible a partir de $\mathcal{B}$.

(3) El punt $(0, t)$ és constructible a partir de $\mathcal{B}$.

La importància del concepte de nombre constructible està en el resultat següent, que mostra que si coneixem el conjunt de tots els nombres constructibles podem deduir quins són els punts constructibles. Per tant, redueix el problema d'identificar quins punts són constructibles a partir d'un $\mathcal{B}$ donat al d'identificar el conjunt de tots els nombres constructibles a partir de $\mathcal{B}$.

Proposició 15 *Un punt és constructible a partir de $\mathcal{B}$ si i només si les seves dues coordenades són nombres constructibles a partir de $\mathcal{B}$.*

Prova. Per definició de nombre constructible, si un punt és constructible és clar que ho són les seves dues coordenades. Per tant, el punt només pot ser constructible quan ho són les seves dues coordenades. A l'inrevés, d'acord amb l'observació anterior, si a i b són dos reals constructibles, és que ho són els punts $(a, 0)$ i $(0, b)$. Per tant, també ho és el punt (a, b), ja que és la intersecció de les perpendiculars als eixos per $(a, 0)$ i $(0, b)$. □

2.3 Cap a la traducció algebraica de les construccions amb regle i compàs

Fins ara, el que hem fet és: (1) reduir el problema de determinar si una construcció es pot fer amb regle i compàs al problema de determinar quins punts són constructibles a partir d'uns de donats, i (2) reduir el problema d'identificar quins punts són constructibles i quins no al d'identificar quins nombres reals són constructibles i quins no (a partir dels punts $\mathcal{B}$ donats). Però com es pot saber si un nombre real donat és o no constructible a partir de $\mathcal{B}$?

Bàsicament, la idea és la següent (deixem els detalls pel pròxim capítol). El conjunt $\mathbb{R}$ de tots els nombres reals és el que en matemàtiques s'anomena un *cos*. Això vol dir un conjunt amb una suma i un producte que compleixen certes propietats. Per altra banda, els punts de $\mathcal{B}$ tindran unes coordenades, i totes aquestes coordenades "generaran" el que s'anomena un *subcòs* de $\mathbb{R}$. Això vol dir un subconjunt K_0 de $\mathbb{R}$ que per si sol serà un cos amb la suma i producte de $\mathbb{R}$ restringides a K_0. Aquest subcòs serà l'anàleg algebraic del conjunt $\mathcal{B}$ de punts base i l'anomenarem *cos base*. Normalment, serà el conjunt $\mathbb{Q}$ de tots els racionals, però en general pot ser més gran, depenent de $\mathcal{B}$.

Doncs bé, algebraicament, la construcció geomètrica d'un nou punt a partir dels punts base correspon a construir el que s'anomena una *extensió* de K_0, és a dir, un altre subcòs K_1 de $\mathbb{R}$ que contindrà K_0 i les coordenades del nou punt, que en general no pertanyeran a K_0. Més exactament, K_1 serà el *subcòs més petit* de $\mathbb{R}$ que conté K_0 i les dues coordenades del nou punt. Quan a continuació construïm un nou punt a partir dels punts base i del punt ja construït, el que algebraicament estarem fent és construir una nova extensió K_2 de l'extensió K_1. Com abans, K_2 serà el subcòs més petit de $\mathbb{R}$ que conté K_1 i les dues coordenades del nou punt. I això serà així per cada nou punt que construim.

Per tant, una construcció amb regle i compàs algebraicament correspondrà a partir d'un cos base K_0 i anar-ne fent successives extensions $K_1, K_2, K_3, ...$, cadascuna obtinguda a partir de l'anterior afegint les coordenades del nou punt construït. Per tal d'identificar quins nombres reals seran constructibles i quins no a partir de $\mathcal{B}$ el que d'alguna manera farem és identificar quines extensions de K_0 es poden obtenir quan els nous punts es construeixen utilitzant únicament un regle i un compàs. Una construcció determinada serà aleshores possible només quan les coordenades de tots els punts que calgui construir per fer-la siguin d'alguna d'aquestes extensions.

Capítol 3

El criteri de constructibilitat d'un nombre

L'objectiu d'aquest capítol és desenvolupar amb detall la idea exposada al final del capítol anterior. És el nucli d'aquest treball i també la part més complicada. Aquí explicarem tots els conceptes i resultats que es necessiten per poder demostrar que els tres problemes clàssics no tenen solució. Al final del capítol enuciarem i demostrarem (només parcialment) el criteri per saber quan un nombre real és constructible i quan no (Teorema 39), i al capítol següent utilitzarem aquest criteri per veure que les tres construccions clàssiques són impossibles.

3.1 Cossos, subcossos i extensions de cossos

Comencem veient què s'entén per un cos. L'exemple que ens interessa és el dels nombres reals, però la definició general és senzilla, encara que llarga.

Definició 16 *S'anomena* **cos** *qualsevol conjunt K on hi ha definides dues operacions binàries $+$ i $\cdot$, anomenades suma i producte, que compleixen les següents propietats:*

(1) la suma és commutativa: $a+b=b+a$ per tot $a,b\in K$;

(2) la suma és associativa: $a+(b+c)=(a+b)+c$ per tot $a,b,c\in K$;

(3) la suma té un element neutre 0: $a+0=0+a=a$ per tot $a\in K$;

(4) tot element $a\in K$ té un oposat $-a$ per la suma: $a+(-a)=(-a)+a=0$;

(5) el producte és commutatiu: $a\cdot b=b\cdot a$ per tot $a,b\in K$;

(6) el producte és associatiu: $a\cdot(b\cdot c)=(a\cdot b)\cdot c$ per tot $a,b,c\in K$;

(7) el producte té un element neutre 1: $a\cdot 1=1\cdot a=a$ per tot $a\in K$;

(8) tot element $a\in K$ diferent del 0 té un invers a^{-1} pel producte: $a\cdot a^{-1}=a^{-1}\cdot a=1$;

(2) el producte distribueix respecte la suma: $a\cdot(b+c)=a\cdot b+a\cdot c$ per tot $a,b,c\in K$.

Per exemple, el conjunt $\mathbb{N}$ de tots els naturals amb la suma i producte usuals no és un cos, ja que no hi existeixen ni oposats respecte la suma ni inversos respecte el producte. Tampoc ho és el conjunt $\mathbb{Z}$ de tots els enters, ja que té oposats però no inversos. En canvi, el conjunt $\mathbb{Q}$ de tots els racionals sí que ho és amb la suma i producte habituals, i també ho són els conjunts $\mathbb{R}$ i $\mathbb{C}$ de tots els reals i tots els complexos, respectivament. Aquests tres cossos no són de fet independents, sinó que es tenen les inclusions

$$\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

A més, en cada cas les operacions suma i producte del conjunt petit són les mateixes que les del gran. Quan això passa es diu que el subconjunt petit és un *subcòs* del gran i que el gran és una *extensió* del petit. Per ex., $\mathbb{Q}$ és un subcòs tant de $\mathbb{R}$ com de $\mathbb{C}$, $\mathbb{R}$ és una extensió de $\mathbb{Q}$ i $\mathbb{C}$ és una extensió de $\mathbb{R}$.

En general, per **subcòs** d'un cos K s'entén un subconjunt S de K que per si sol és un cos amb la suma i producte de K restringides a S. Com que la suma i el producte de K ja tenen totes les propietats que es demanen, també les tindran la suma i el producte de S. Per tant, a la pràctica, per veure que un subconjunt de K n'és un subcòs només caldrà comprovar que compleix dues condicions:

(1) conté el 0 i l'1 de K, i

(2) és "estable" respecte les operacions suma, producte, oposat i invers, és a dir, haurà de ser tal que si a, b són de S també ho hauran de ser $a \cdot b$, $a + b$, $-a$ i a^{-1}.

Evidentment, no tot subconjunt d'un cos és un subcòs. Per exemple, ni el conjunt dels naturals ni el dels enters són subcossos de $\mathbb{Q}$.

Per altra banda, per **extensió** d'un cos K en general s'entén un cos L que conté a K com a subcòs. Quan això passa, escriurem simplement $K \subset L$.

Exemple 17 Considerem el subconjunt següent de $\mathbb{R}$:

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2},\ \text{amb } a, b \in \mathbb{Q} \text{ qualssevol}\}.$$

Clarament, conté el conjunt $\mathbb{Q}$ de tots els racionals (només cal fer $b = 0$). En particular, el 0 i l'1 són de $\mathbb{Q}(\sqrt{2})$. Per altra banda, $\mathbb{Q}(\sqrt{2})$ és estable per sumes i productes, ja que si $a + b\sqrt{2},\ a' + b'\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ tenim que

$$(a + b\sqrt{2}) + (a' + b'\sqrt{2}) = (a + a') + (b + b')\sqrt{2} \ \in \ \mathbb{Q}(\sqrt{2})$$

i

$$(a + b\sqrt{2})(a' + b'\sqrt{2}) = (aa' + 2bb') + (ab' + a'b)\sqrt{2} \ \in \ \mathbb{Q}(\sqrt{2}).$$

També és estable per oposats ja que

$$-(a + b\sqrt{2}) = (-a) + (-b)\sqrt{2} \ \in \ \mathbb{Q}(\sqrt{2}).$$

Finalment, és estable per inversos, ja que si a, b no són tots dos zero, racionalitzant s'obté que

$$(a + b\sqrt{2})^{-1} = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2}.$$

El denominador $a^2 - 2b^2$ no pot ser zero perquè $\sqrt{2}$ no és racional ($a^2 - 2b^2 = 0$ implicaria que $\sqrt{2} = a/b$). D'altra banda, com que $a, b \in \mathbb{Q}$, també són de $\mathbb{Q}$ els nombres $a/(a^2 - 2b^2)$ i $b/(a^2 - 2b^2)$ i, per tant, l'invers de $a + b\sqrt{2}$ és de $\mathbb{Q}(\sqrt{2})$. Per tant, $\mathbb{Q}(\sqrt{2})$ és un subcòs de $\mathbb{R}$ i al mateix temps una extensió de $\mathbb{Q}$.

3.2 Subcossos de $\mathbb{R}$

En aquest treball ens interessen únicament els subcossos de $\mathbb{R}$. Un d'ells és $\mathbb{Q}$, tot i que tal i com acabem de veure no és l'únic. De fet, n'hi ha infinits més. Però $\mathbb{Q}$ juga un paper molt especial degut al fet següent:

Proposició 18 *Qualsevol subcòs de $\mathbb{R}$ conté $\mathbb{Q}$. En particular, $\mathbb{Q}$ és el subcòs més petit de $\mathbb{R}$.*

Prova. Qualsevol subcòs de $\mathbb{R}$ ha de contenir el 0 i l'1 per definició. A partir d'aquí, fent servir que ha de ser estable per oposats, sumes i inversos es dedueix que ha de contenir totes les fraccions. Concretament, ha de contenir el -1 perquè és l'oposat de l'1. Però si conté l'1 i el -1 ha de contenir tots els enters perquè s'obtenen sumant 1's o -1's, i si conté tots els enters, ha de contenir totes les fraccions de numerador 1 o -1 perquè són els inversos dels enters no nuls. Per tant, com que qualsevol fracció diferent de zero s'obté sumant fraccions de numerador 1 o -1, ha de contenir totes les fraccions. □

Observar que això significa que considerar subcossos de $\mathbb{R}$ és exactament el mateix que considerar extensions de $\mathbb{Q}$ que estiguin dintre de $\mathbb{R}$.

A part de $\mathbb{Q}$, l'altre exemple de subcòs de $\mathbb{R}$ que hem vist fins ara és $\mathbb{Q}(\sqrt{2})$. Resulta ser el subcòs de $\mathbb{R}$ més petit que conté $\sqrt{2}$. El motiu és que qualsevol altre subcòs que contingui $\sqrt{2}$, a més de contenir $\mathbb{Q}$ (Proposició 18), també haurà de contenir obligatòriament tots els reals de la forma $a + b\sqrt{2}$ amb $a, b \in \mathbb{Q}$ perquè ha de ser estable per sumes i productes. Per tant, ha de contenir tot $\mathbb{Q}(\sqrt{2})$. Com que $\mathbb{Q}(\sqrt{2})$ ja és un subcòs, serà el més petit que conté $\sqrt{2}$.

Es poden obtenir molts exemples més de subcossos de $\mathbb{R}$ utilitzant la mateixa idea. Es tracta de considerar el subcòs més petit de $\mathbb{R}$ que conté un subconjunt donat de reals. Per exemple, un pot considerar el subcòs més petit de $\mathbb{R}$ que conté $\sqrt{3}$, que es representa per $\mathbb{Q}(\sqrt{3})$. No és $\mathbb{Q}$ perquè ha de contenir $\sqrt{3}$, i tampoc és el subcòs $\mathbb{Q}(\sqrt{2})$ perquè $\sqrt{3}$ no pertany a $\mathbb{Q}(\sqrt{2})$. En efecte, si hi pertanyés, seria de la forma $\sqrt{3} = a + b\sqrt{2}$ per algun parell de racionals $a, b \in \mathbb{Q}$. Elevant aleshores al quadrat i aillant el terme $\sqrt{2}$ resultaria que $\sqrt{2}$ seria racional, però sabem que no ho és. Per tant, $\mathbb{Q}(\sqrt{3})$ és un nou subcòs de $\mathbb{R}$. Com en el cas de $\mathbb{Q}(\sqrt{2})$ es pot comprovar que $\mathbb{Q}(\sqrt{3})$ és el conjunt

$$\mathbb{Q}(\sqrt{3}) = \{a + b\sqrt{3}, \text{ amb } a, b \in \mathbb{Q} \text{ qualssevol}\}$$

(més endavant veurem una manera ràpida de veure-ho).

En general, es defineix el següent:

Definició 19 *Donat un subconjunt qualsevol X de $\mathbb{R}$, s'anomena* **subcòs de $\mathbb{R}$ generat per** *X al subcòs més petit de $\mathbb{R}$ que conté X. Se'l representa per $\mathbb{Q}(X)$ o bé, quan $X = \{x_1, \ldots, x_n\}$ és finit, per $\mathbb{Q}(x_1, \ldots, x_n)$.*

Aquesta definició és molt útil perquè en realitat qualsevol subcòs K de $\mathbb{R}$ és igual a un $\mathbb{Q}(X)$ per algun subconjunt X. Només cal agafar com a subconjunt X el propi K. Com que el subcòs més petit que conté el subcòs K és el propi subcòs K, és evident que $K = \mathbb{Q}(K)$.

Però s'ha d'anar molt en compte amb els subcossos $\mathbb{Q}(X)$ perquè subconjunts X diferents no necessàriament generaran subcossos diferents. Pot perfectament passar que $Q(X_1)$ sigui igual a $Q(X_2)$ encara que $X_1 \neq X_2$. Un exemple molt simple d'això és quan X_1 i X_2 són dos subconjunts diferents de racionals. Per ex., $X_1 = \{0, 1\}$ i $X_2 = \{1/2\}$. En tots dos casos el subcòs que es genera és simplement $\mathbb{Q}$. Un altre cas clar on això també passa és, per exemple, quan $X_1 = \{\sqrt{2}\}$ i $X_2 = \mathbb{Q} \cup \{\sqrt{2}\}$. En tots dos casos, el subcòs generat és $\mathbb{Q}(\sqrt{2})$. En realitat, afegint a X tots els racionals, el subcòs que es genera no serà diferent del que es genera només amb X ja que $\mathbb{Q}$ està inclòs en qualsevol subcòs. Per evitar aquest repetició, a partir d'ara sempre que parlem del subcòs $\mathbb{Q}(X)$ suposarem que tots els elements de X són irracionals, i interpretarem que $\mathbb{Q}(\emptyset) = \mathbb{Q}$.

Un exemple no evident de $X_1 \neq X_2$ tal que $\mathbb{Q}(X_1) = \mathbb{Q}(X_2)$ és el següent.

Exemple 20 Els subcossos $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ i $\mathbb{Q}(\sqrt{2}+\sqrt{3})$, corresponents a $X_1 = \{\sqrt{2}, \sqrt{3}\}$ i $X_2 = \{\sqrt{2} + \sqrt{3}\}$, són iguals. Per a veure-ho, provarem que cadascun està inclòs dins de l'altre. Que $\mathbb{Q}(\sqrt{2} + \sqrt{3}) \subseteq \mathbb{Q}(\sqrt{2}, \sqrt{3})$ és molt fàcil. Com que $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ conté $\sqrt{2}$ i $\sqrt{3}$ i és estable per sumes, necessàriament també conté $\sqrt{2}+\sqrt{3}$. I si conté $\sqrt{2} + \sqrt{3}$ contindrà tot el subcòs més petit que conté $\sqrt{2} + \sqrt{3}$, és a dir, contindrà $\mathbb{Q}(\sqrt{2} + \sqrt{3})$. Una mica més difícil és veure que $\mathbb{Q}(\sqrt{2} + \sqrt{3})$ conté $\sqrt{2}$ i $\sqrt{3}$ i per tant, tot $\mathbb{Q}(\sqrt{2}, \sqrt{3})$. Primer de tot, observem que

$$(\sqrt{2} + \sqrt{3})^2 = 5 + 2\sqrt{6}.$$

Això ens diu que $\sqrt{6} \in \mathbb{Q}(\sqrt{2} + \sqrt{3})$. Però per estabilitat per productes obtenim aleshores que

$$\sqrt{6}(\sqrt{2} + \sqrt{3}) = \sqrt{12} + \sqrt{18} = 2\sqrt{3} + 3\sqrt{2} \in \mathbb{Q}(\sqrt{2} + \sqrt{3}).$$

D'aquí deduïm que

$$\sqrt{2} = (2\sqrt{3} + 3\sqrt{2}) - 2(\sqrt{3} + \sqrt{2}) \in \mathbb{Q}(\sqrt{2} + \sqrt{3})$$

i que

$$\sqrt{3} = (2\sqrt{3} + 3\sqrt{2}) - 3(\sqrt{3} + \sqrt{2}) \in \mathbb{Q}(\sqrt{2} + \sqrt{3}).$$

És important remarcar dues coses més sobre els subcossos $\mathbb{Q}(X)$. La primera és que $\mathbb{Q}(X)$ contindrà molts més elements a part dels de $\mathbb{Q}$ i dels de X. Això ja es veu, per exemple, en el cas de $\mathbb{Q}(\sqrt{2})$. El real $1 + \sqrt{2}$ és de $\mathbb{Q}(\sqrt{2})$, però en canvi no pertany a $\mathbb{Q} \cup \{\sqrt{2}\}$. Això passa perquè el fet que un real α pertanyi a $\mathbb{Q}(X)$ implica que hi han de pertànyer molts altres reals degut a la condició d'estabilitat per sumes, productes, oposats i inversos. Per exemple, també hi hauran de pertànyer 2α, α^{-1} (si $\alpha \neq 0$), $-3 + \alpha$, etc.

La segona és que, en el cas que X consti d'un sol element α, no sempre els elements de $\mathbb{Q}(\alpha)$ seran tots de la forma $a + b\alpha$ amb $a, b \in \mathbb{Q}$. És cert que contindrà tots els elements d'aquesta forma, ja que ha de ser estable per sumes i productes, però en general en contindrà molts més perquè aquests sols no formaran un subcòs de $\mathbb{R}$.

Exemple 21 $\mathbb{Q}(\sqrt[3]{2})$ no pot ser el conjunt $S = \{a + b\sqrt[3]{2},\ \text{amb } a, b \in \mathbb{Q} \text{ qualssevol}\}$ perquè, tot i que aquest conjunt conté $\mathbb{Q}$ i $\sqrt[3]{2}$, no és un subcòs de $\mathbb{R}$. Per exemple, si racionalitzem, és fàcil comprovar que l'invers de $\sqrt[3]{2}$ és $\sqrt[3]{4}/2$. Si aquest invers fos de S també ho seria $\sqrt[3]{4}$ (només caldria duplicar els coeficients racionals que expressen $\sqrt[3]{4}/2$ com a combinació lineal de 1 i $\sqrt[3]{2}$). Però $\sqrt[3]{4}$ no pot pertànyer a S. El raonament és el següent. Si existissin $a, b \in \mathbb{Q}$ tals que

$$\sqrt[3]{4} = a + b\sqrt[3]{2},$$

elevant al cub tindríem que

$$4 = a^3 + 3a^2b\sqrt[3]{2} + 3ab^2\sqrt[3]{4} + 2b^3.$$

Substituint ara $\sqrt[3]{4}$ per $a + b\sqrt[3]{2}$ i aillant $\sqrt[3]{2}$ obtindríem que

$$(3a^2b + 3ab^3)\sqrt[3]{2} = 4 - a^3 - 3a^2b^2 - 2b^3.$$

A més, es pot veure que $3a^2b + 3ab^3$ no pot ser mai zero si $a, b \in \mathbb{Q}$ són tals que $\sqrt[3]{4} = a + b\sqrt[3]{2}$. [1] Per tant, tindríem que

$$\sqrt[3]{2} = \frac{4 - a^3 - 3a^2b^2 - 2b^3}{3a^2b + 3ab^3} \in \mathbb{Q},$$

i això no és possible perquè $\sqrt[3]{2}$ no és racional (veure § 5.1 de l'Apèndix). En realitat, el conjunt $\mathbb{Q}(\sqrt[3]{2})$ és

$$\mathbb{Q}(\sqrt[3]{2}) = \{a + b\sqrt[3]{2} + c\sqrt[3]{4},\ \text{amb } a, b, c \in \mathbb{Q} \text{ qualssevol}\}.$$

Més endavant veurem per què és així (veure Exemple 37).

Finalment, observem també que la notació $\mathbb{Q}(X)$ es pot generalitzar de la manera següent al cas d'un subcòs qualsevol de $\mathbb{R}$, no necessàriament $\mathbb{Q}$. Si $K \subset \mathbb{R}$ és un subcòs qualsevol i X és un subconjunt de reals no contingut a K, $K(X)$ representarà *el subcòs més petit de* $\mathbb{R}$ *que conté* K *i* X. Quan $X = \{x_1, \ldots, x_n\}$, escriurem simplement $K(x_1, \ldots, x_n)$. A diferència d'abans, observem que ara cal exigir explícitament que conté K, i no només X, perquè no tots els subcossos de $\mathbb{R}$ contenen K. De fet, a menys que K sigui $\mathbb{Q}$, el subcòs $K(X)$ serà més gran que $\mathbb{Q}(X)$, i és fàcil d'entendre el per què. D'acord amb la Proposició 18, *qualsevol* subcòs K de $\mathbb{R}$ és una extensió $\mathbb{Q}(X')$ de $\mathbb{Q}$ per algun $X' \subset \mathbb{R}$. Per tant, $K(X)$ és

$$K(X) = \mathbb{Q}(X')(X) = \mathbb{Q}(X' \cup X).$$

En realitat, la notació $K(X)$ només serveix per posar de manifest que el subcòs $K(X)$, a part de ser una extensió de $\mathbb{Q}$, és també una extensió de K. Per exemple, $\mathbb{Q}(\sqrt{2})(\sqrt[3]{3})$ representa el subcòs més petit de $\mathbb{R}$ que conté el subcòs $\mathbb{Q}(\sqrt{2})$ i $\sqrt[3]{3}$, i és el mateix que $\mathbb{Q}(\sqrt{2}, \sqrt[3]{3})$.

[1] Tenim que $3a^2b + 3ab^3 = 3ab(a + b^2)$. Ara, a no pot ser zero perquè tindríem que $\sqrt[3]{4} = b\sqrt[3]{2}$ i, per tant, seria $b = \sqrt[3]{2}$, però $\sqrt[3]{2}$ no és racional (veure § 5.1 de l'Apèndix). Anàlogament, tampoc pot ser $b = 0$ perquè aleshores $\sqrt[3]{4}$ seria racional, que tampoc ho és (veure § 5.1 de l'Apèndix). Finalment, es pot veure que tampoc pot ser $a = -b^2$ perquè també acabaria implicant que $\sqrt[3]{2}$ seria racional. Per tant, $3a^2b + 3ab^3$ no pot ser zero si $a, b \in \mathbb{Q}$ i $\sqrt[3]{4} = a + b\sqrt[3]{2}$.

3.3 El cos base d'una construcció amb regle i compàs

Un exemple de subcòs $\mathbb{Q}(X)$ que és important per aquest treball és el del cos base d'una construcció amb regle i compàs, que representàvem per K_0 al final del capítol anterior. Si denotem per $X_{\mathcal{B}}$ el subconjunt de $\mathbb{R}$ definit per les coordenades no racionals de tots els punts de $\mathcal{B}$, el cos base no és altre que l'extensió

$$K_0 = \mathbb{Q}(X_{\mathcal{B}}).$$

És a dir, és el subcòs més petit de $\mathbb{R}$ que conté $X_{\mathcal{B}}$. En particular, serà $\mathbb{Q}$ quan totes les coordenades dels punts de $\mathcal{B}$ siguin racionals.

Aquí és important observar una cosa. I és que, d'acord amb el que hem dit abans, conjunts X diferents poden generar el mateix subcòs de $\mathbb{R}$. Això es tradueix en el fet que diferents conjunts de punts base poden correspondre en realitat al mateix cos base. Per exemple, els cossos $\mathbb{Q}(X_{\mathcal{B}_1})$ i $\mathbb{Q}(X_{\mathcal{B}_2})$ corresponents a $\mathcal{B}_1 = \{(0,0),(1,0)\}$ i $\mathcal{B}_2 = \{(0,0),(2,0)\}$ són tots dos iguals a $\mathbb{Q}$.

En realitat, els punts que es podran construir a partir d'un $\mathcal{B}$ donat i, en conseqüència, els nombres reals que es podran construir a partir de $\mathcal{B}$ (recordar la Proposició 15), només dependran del cos base $\mathbb{Q}(X_{\mathcal{B}})$, i no de qui és exactament el conjunt $\mathcal{B}$. Això és raonable que sigui així, però més endavant quedarà clar, quan provem el criteri de constructibilitat d'un nombre real a partir d'un $\mathcal{B}$ donat (veure Teorema 39).

Exemple 22 El punt $(1,0)$ és el punt mig entre $(0,0)$ i $(2,0)$ i, per tant, és constructible a partir de $\mathcal{B}_2 = \{(0,0),(2,0)\}$ (veure Exemple 1.1). D'aquí deduïm que tot el que es podrà construir a partir de $\mathcal{B}_1 = \{(0,0),(1,0)\}$ també es pot construir a partir de $\mathcal{B}_2$. Per altra banda, el punt $(2,0)$ és constructible a partir de $\mathcal{B}_1$ perquè és la intersecció de la recta per $(0,0)$ i $(1,0)$ amb la circumferència centrada a $(1,0)$ de radi 1. Per tant, tot el que és constructible a partir de $\mathcal{B}_2$ també ho serà a partir de $\mathcal{B}_1$. A partir de $\mathcal{B}_1 = \{(0,0),(1,0)\}$ es poden construir, doncs, exactament els mateixos punts que a partir de $\mathcal{B}_2 = \{(0,0),(2,0)\}$. I efectivament, com ja hem remarcat abans, tots dos conjunts de punts base corresponen al mateix cos base.

Per aquest motiu, els nombres constructibles a partir de $\mathcal{B}$ també els anomenarem nombres constructibles a partir de $\mathbb{Q}(X_{\mathcal{B}})$.

3.4 El cos dels nombres constructibles

Un resultat bàsic per la resolució del nostre problema d'identificar quins reals són constructibles és que, sigui qui sigui $\mathcal{B}$, el conjunt de tots els nombres reals constructibles a partir de $\mathcal{B}$ és un subcòs de $\mathbb{R}$. Concretament, es té el següent.

Teorema 23 *Sigui $\mathcal{B}$ un subconjunt finit de punts del pla que conté l'origen i el punt $(1,0)$, i sigui $\mathfrak{C}_{\mathcal{B}}$ el conjunt de nombres reals constructibles amb regle i compàs a partir del corresponent subcòs $\mathbb{Q}(X_{\mathcal{B}})$. Aleshores $\mathfrak{C}_{\mathcal{B}}$ és un subcòs de $\mathbb{R}$ estable per arrels quadrades.*

Per conjunt de reals estable per arrels quadrades s'entén un conjunt S de reals tal que si $a \in S$, les seves arrels quadrades (positiva i negativa), quan existeixen, també són de S.

Prova. Com que $(1,0) \in \mathcal{B}$, és clar que el 0 i l'1 són de $\mathfrak{C}_{\mathcal{B}}$, així que només hem de comprovar que és estable per sumes, productes, oposats, inversos i arrels quadrades.

(1) *Estabilitat per oposats.* Considerem a un nombre constructible. Pel Lema 14, això significa que el punt $A = (a,0)$ és constructible. Però aleshores també ho és el punt $A' = (-a,0)$, ja que és l'altre punt d'intersecció de l'eix d'abscisses amb la circumferència de centre O i radi OA. Per tant, $-a$ també és constructible altre cop pel Lema 14.

(2) *Estabilitat per sumes.* Considerem a, b dos nombres constructibles. Pel Lema 14, els punts $A = (a,0)$ i $B = (b,0)$ són constructibles. Els dos punts d'intersecció de l'eix d'abscisses amb la circumferència de centre A i radi OB són aleshores constructibles. Però aquests dos punts són $C = (a + |b|, 0)$ i $C' = (a - |b|, 0)$ (des de A, em desplaço una distància $|b|$ a la dreta o a l'esquerra). Segons que b sigui positiu o negatiu un dels dos és necessàriament el punt $(a + b, 0)$. Per tant, $a + b$ també és constructible altre cop pel Lema 14.

(3) *Estabilitat per productes.* Considerem a, b dos nombres constructibles. Si algun dels dos és nul, el producte també ho és i, per tant, és constructible. Suposem, doncs, que $a, b \neq 0$. Podem suposar que $a, b > 0$ ja que ja sabem que $\mathfrak{C}_{\mathcal{B}}$ és estable per oposats. Pel Lema 14, els punts $A = (a,0)$ i $B = (0,b)$ són constructibles. Per tant, si denotem per I el punt $(1,0) \in \mathcal{B}$, també són constructibles la recta IB i la paral·lela a IB que passa per A (veure Exemple 1.1). Anomenem C la intersecció d'aquesta última amb l'eix d'ordenades (veure Fig. 3.1). És un punt constructible. Ara, si representem per PQ la distància entres dos punts qualssevol P i Q, pel teorema de Tales sabem que $OC/OB = OA/OI$. Però $OI = 1$, de manera que $OC = OA \cdot OB = a \cdot b$ i $C = (0, a \cdot b)$. Per tant, com que C és constructible, $a \cdot b$ és constructible pel Lema 14.

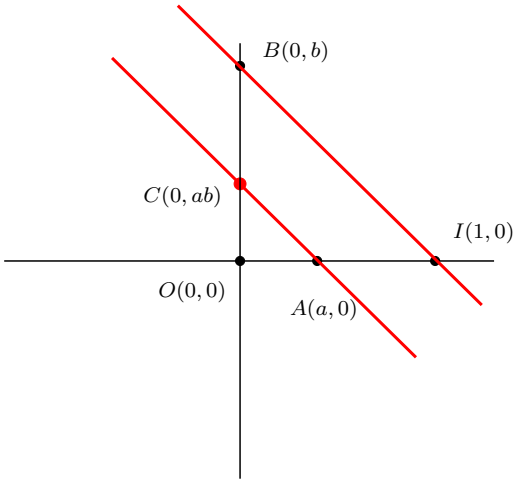


Figura 3.1: Construcció de $a \cdot b$

(4) *Estabilitat per inversos.* Considerem $a \neq 0$ un nombre constructible. Podem suposar-lo positiu per la mateixa raó d'abans. Pel Lema 14, el punt $A = (a,0)$ és constructible.

Per tant, si $I = (1, 0)$ i $J = (0, 1)$, també són constructibles la recta AJ i la para·lela a AJ que passa per I. Anomenem B la intersecció d'aquesta última amb l'eix d'ordenades (veure Fig. 3.2). És un punt constructible. Pel teorema de Tales un altre cop sabem que $OB/OJ = OI/OA$. Però $OJ = OI = 1$, de manera que $OB = 1/OA = 1/a$ i $B = (0, 1/a)$. Per tant, $1/a$ és constructible pel Lema 14.

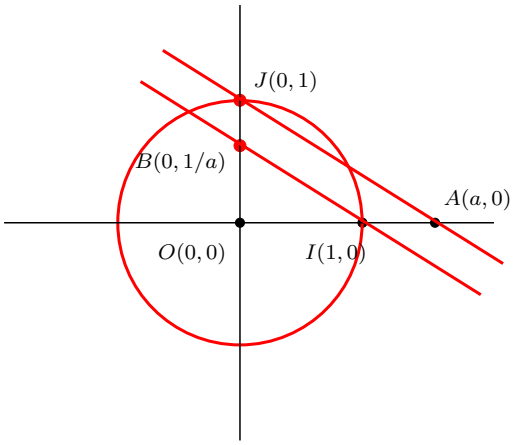


Figura 3.2: Construcció de $1/a$

(5) *Estabilitat per arrels quadrades.* Considerem $a \geq 0$ un nombre constructible. Com que 1 també és constructible i ja sabem que $\mathfrak{C}_{\mathcal{B}}$ és estable per sumes, també és constructible $1+a$. Pel Lema 14 els punts $A = (a, 0)$ i $B = (1 + a, 0)$ són aleshores constructibles. Sigui M el punt mig de OB. És constructible (veure Exemple 1.1). Si $I = (1, 0)$, la perpendicular a l'eix d'abscisses que passa per I i la circumferència de centre M i radi OM també són aleshores constructibles. Sigui C el punt d'intersecció de totes dues d'ordenada positiva (veure Fig. 3.3). És un punt constructible. Es tracta ara de veure que és el punt de coordenades $C = (1, \sqrt{a})$. Per a veure-ho, necessitem dos resultats de geometria clàssica, ja coneguts pels grecs, que demostrem a l'Apèndix (veure § 5.2):

(a) (*Segon teorema de Tales*) Qualsevol triangle inscrit a una circumferència que té per un dels costats un diàmetre de la circumferència és rectangle en el vèrtex oposat al diàmetre.

(b) (*Teorema de l'altura*) Si MNP és un triangle rectangle, amb l'angle recte a P, h és l'altura des de P i X el peu d'aquesta altura, aleshores $h^2 = MX \cdot XN$.

El resultat (a) implica que el triangle OCB és rectangle al vèrtex C. A més, el peu de l'altura des de C d'aquest triangle és justament el punt I. Per tant, aplicant (b) al triangle rectangle OCB obtenim que $IC^2 = OI \cdot IB = a$, ja que $OI = 1$. D'aquí deduïm que $IC = \sqrt{a}$, és a dir, $C = (1, \sqrt{a})$. Com que aquest punt és constructible, un cop més pel Lema 14 ho és el nombre $\sqrt{a}$.

□

D'aquest resultat, no només és important allò que diu, sinó també la demostració. D'una banda, el resultat implica que són constructibles tots els nombres naturals (ja que qualsevol

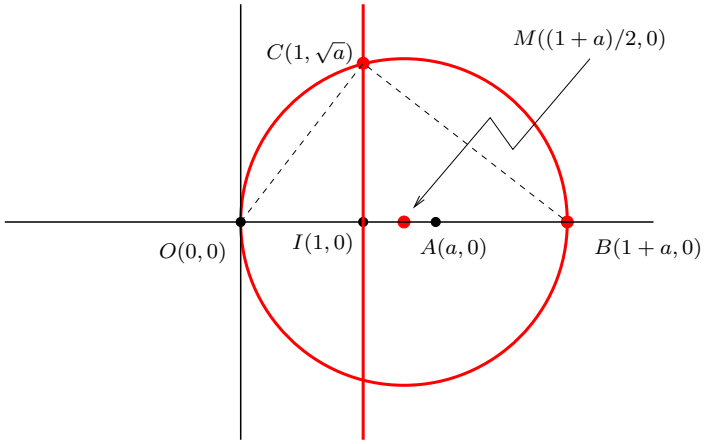


Figura 3.3: Construcció de $\sqrt{a}$

subcòs de $\mathbb{R}$ conté $\mathbb{Q}$ i, per tant, $\mathbb{N}$; veure Proposició 18) i, en conseqüència, qualsevol real que s'obtingui a partir de nombres naturals només fent sumes, oposats, productes, inversos i arrels quadrades. De l'altra, la demostració descriu implícitament un mètode de construir amb regle i compàs qualsevol d'aquests nombres.

Exemple 24 El nombre $a = \sqrt[4]{2}$ és constructible a partir de $\mathcal{B} = \{(0,0),(1,0)\}$. D'acord amb la demostració anterior, una manera de construir-lo és la següent.

1. Construim el punt $A = (2,0)$ com a intersecció de l'eix d'abscisses amb la circumferència centrada a $I = (1,0)$ i radi OI.
2. Apliquem el mètode descrit en demostrar l'estabilitat de $\mathfrak{C}_{\mathcal{B}}$ per arrels quadrades per construir el punt $C = (1,\sqrt{2})$ a partir de $A = (2,0)$.
3. Construim el punt $D = (0,\sqrt{2})$ com la intersecció de l'eix d'ordenades amb la paral·lela a l'eix d'abscisses que passa per C.
4. Construim el punt $A' = (\sqrt{2},0)$ com la intersecció de l'eix d'abscisses amb la circumferència de centre O i radi OD.
5. Apliquem una altra vegada el mètode descrit a la demostració per construir el punt $C' = (1,\sqrt{\sqrt{2}})$ a partir de $A' = (\sqrt{2},0)$.

Com que $\sqrt{\sqrt{2}} = \sqrt[4]{2}$, ja tenim construït el nombre que volíem.

Exemple 25 Tot i que molt més llarg, un procediment similar permetrà construir a partir només de $\mathcal{B} = \{(0,0),(1,0)\}$ un nombre real com ara

$$a = \frac{1}{16}\left(-1+\sqrt{17}+\sqrt{34-2\sqrt{17}}+ \right.$$
$$\left. +\sqrt{68+12\sqrt{17}+2\left(-1+\sqrt{17}\right)\sqrt{34-2\sqrt{17}}-16\sqrt{34+2\sqrt{17}}}\right)$$

L'interès del fet que aquest nombre sigui constructible és que l'any 1796 Gauss va demostrar que és exactament el valor del cosinus de $2\pi/17$, i això implicava que el polígon regular de 17 costats era constructible amb regle i compàs. De fet, si dos dels vèrtexs del polígon són els punts $P_1 = (0,0)$ i $P_2 = (1,0)$, és fàcil veure fent trigonometria que el següent vèrtex és el punt de coordenades

$$P_3 = (1 + \cos(2\pi/17), \sin(2\pi/17)).$$

Ara, com que el nombre $\cos(2\pi/17)$ és constructible, també ho és $\sin(2\pi/17)$, ja que s'obté a partir de l'anterior fent un producte, un oposat, una suma i una arrel quadrada d'acord amb la relació $\sin(2\pi/17) = \sqrt{1 - \cos^2(2\pi/17)}$. Per tant, el vèrtex P_3 és constructible amb regle i compàs. A partir d'aquí, repetint el procés però partint de P_2, P_3 es construeix el següent vèrtex P_4, etc. La fórmula anterior proporciona, doncs, un procediment de construcció de l'*heptadecàgon*. Però és un procediment molt llarg i pesat. Al llarg del s. XIX diversos matemàtics van trobar mètodes alternatius molt més elegants de fer-ho. Un dels més elegants és el descrit l'any 1893 pel matemàtic anglès H.W. Richmond. [2]

Normalment, $\mathcal{B}$ és tal que el cos base corresponent és $\mathbb{Q}$, i en aquest cas es parla simplement del cos dels nombres constructibles (amb regle i compàs), sense precisar qui és $\mathcal{B}$. El representarem aleshores per $\mathfrak{C}$.

D'alguna manera, el cos $\mathfrak{C}$ o, en general, $\mathfrak{C}_\mathcal{B}$ és el "protagonista" d'aquesta història. Tal i com dèiem al final del capítol anterior, el nostre objectiu immediat és identificar qui és $\mathfrak{C}_\mathcal{B}$, perquè un cop tinguem un criteri que ens digui si un nombre real donat és o no de $\mathfrak{C}_\mathcal{B}$, ja sabrem quins punts són constructibles i provar la impossibilitat de les tres construccions clàssiques serà senzill (excepte en el cas de la quadratura del cercle).

3.5 Extensions simples

Abans de poder enunciar i entendre aquest criteri, encara ens calen alguns conceptes més. El primer d'ells és el d'*extensió simple* d'un cos.

Definició 26 *Anomenarem* **extensió simple de** $\mathbb{Q}$ *qualsevol subcòs de* $\mathbb{R}$ *de la forma* $\mathbb{Q}(\alpha)$ *per algun* $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. [3] *En general, si* K *és un subcòs qualsevol de* $\mathbb{R}$*, anomenarem* **extensió simple de** K *qualsevol subcòs de* $\mathbb{R}$ *de la forma* $K(\alpha)$ *per algun* $\alpha \in \mathbb{R} \setminus K$.

S'ha d'anar molt en compte amb el concepte d'extensió simple. Una extensió que aparentment no és simple en realitat pot ser-ho. És el cas de l'Exemple 20 anterior, on hem vist que $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ és en realitat l'extensió simple $\mathbb{Q}(\sqrt{2} + \sqrt{3})$.

La importància de les extensions simples és que qualsevol extensió $\mathbb{Q}(X)$, quan X és un conjunt finit, sempre es pot pensar com el resultat de fer un nombre finit d'extensions simples.

[2] Més detalls tant de la construcció de Richmond com de la prova de Gauss de l'expressió anterior del cosinus de $2\pi/17$ es poden trobar als llibres de J.C. Carrega (Cap. IV) i de I. Stewart (Cap. 17) que s'indiquen a la Bibliografia.

[3] Si A, B són dos conjunts qualssevol, es representa per $A \setminus B$ el conjunt que té per elements tots els elements de A que no són de B.

Per exemple, l'extensió $\mathbb{Q}(\sqrt{2}, \sqrt[3]{5}, \pi)$ de $\mathbb{Q}$ es pot pensar com l'extensió simple $K_1(\pi)$ del cos $K_1 = \mathbb{Q}(\sqrt{2}, \sqrt[3]{5})$, i K_1 es pot pensar com l'extensió simple $K_2(\sqrt[3]{5})$ del cos $K_2 = \mathbb{Q}(\sqrt{2})$, el qual és per la seva banda una extensió simple de $\mathbb{Q}$. Per tant, $\mathbb{Q}(\sqrt{2}, \sqrt[3]{5}, \pi)$ es pot obtenir fent la successió d'extensions simples

$$\mathbb{Q} \subset \mathbb{Q}(\sqrt{2}) \subset \mathbb{Q}(\sqrt{2}, \sqrt[3]{5}) \subset \mathbb{Q}(\sqrt{2}, \sqrt[3]{5}, \pi).$$

Més endavant farem servir aquesta idea.

3.6 Grau d'una extensió

Un altre concepte que ens fa falta és el de *grau d'una extensió* qualsevol, no necessàriament simple.

Per entendre quina és la idea, considerem el cas de l'extensió $\mathbb{Q}(\sqrt{2})$ de $\mathbb{Q}$. Tots els elements de $\mathbb{Q}(\sqrt{2})$ són de la forma $a + b\sqrt{2}$ amb $a, b \in \mathbb{Q}$. És a dir, tots són combinació lineal amb coeficients a $\mathbb{Q}$ de només l'1 i $\sqrt{2}$. Es diu que $\{1, \sqrt{2}\}$ són *generadors* de $\mathbb{Q}(\sqrt{2})$ sobre $\mathbb{Q}$. Més encara, 1 i $\sqrt{2}$ són *linealment independents* sobre $\mathbb{Q}$, és a dir, l'única combinació lineal de 1 i $\sqrt{2}$ amb coeficients a $\mathbb{Q}$ que dóna zero és quan s'agafen els dos coeficients iguals a zero. Això és així perquè si fos $a \cdot 1 + b \cdot \sqrt{2} = 0$ amb $a, b \in \mathbb{Q}$ i $b \neq 0$ aleshores tindríem que

$$\sqrt{2} = -\frac{a}{b}.$$

cosa que no pot ser perquè $\sqrt{2}$ és irracional. I en el cas que sigui $b = 0$ necessàriament també ha de ser $a = 0$. Es diu aleshores que el conjunt $\{1, \sqrt{2}\}$ és una *base* de $\mathbb{Q}(\sqrt{2})$ com a *espai vectorial* sobre $\mathbb{Q}$. [4]

Doncs bé, resulta que qualsevol que sigui l'extensió de cossos $K \subset L$, sempre passarà el mateix: existirà una família d'elements de L (en general, pot ser infinita) que són generadors linealment independents sobre K. En particular, tot element de L serà una combinació lineal d'ells amb coeficients a K. Una família així l'anomenarem **base de L sobre K**.

Exemple 27 El cos $\mathbb{C}$ dels complexos és una extensió del cos $\mathbb{R}$ dels reals i una base de $\mathbb{C}$ sobre $\mathbb{R}$ és el conjunt $\{1, i\}$, ja que tot nombre complex és de la forma $a + bi = a \cdot 1 + b \cdot i$ amb $a, b \in \mathbb{R}$ i si $a + bi = 0$ és que $a = b = 0$.

El problema és que per una mateixa extensió $K \subset L$ existiran vàries bases del cos gran sobre el petit. Per exemple, en el cas de l'extensió $\mathbb{Q} \subset \mathbb{Q}(\sqrt{2})$ una altra base de $\mathbb{Q}(\sqrt{2})$ sobre $\mathbb{Q}$ és $\{17, -\sqrt{2}\}$. Un element qualsevol $a + b\sqrt{2}$ de $\mathbb{Q}(\sqrt{2})$ és en aquest cas la combinació lineal

$$a + b\sqrt{2} = \frac{a}{17} \cdot 17 + (-b) \cdot (-\sqrt{2})$$

i fent un raonament similar al d'abans es veu que 17 i $-\sqrt{2}$ són linealment independents sobre $\mathbb{Q}$. De fet, tot conjunt de la forma $\{\alpha, \beta\sqrt{2}\}$ amb α, β racionals qualssevol diferents de zero és una base de $\mathbb{Q}(\sqrt{2})$ sobre $\mathbb{Q}$.

[4] No entrarem en la definició general d'espai vectorial i de base d'un espai vectorial. Es pot trobar, per exemple, al Capítol IV del llibre de M. Castellet i I. Llerena que apareix citat a la Bibliografia.

La bona notícia és que, sigui quina sigui l'extensió $K \subset L$, totes les bases del cos gran sobre el petit tenen el mateix nombre d'elements. [5] Això és el que permet definir el grau de l'extensió.

Definició 28 *S'anomena* **grau d'una extensió** *de cossos $K \subset L$, i es representa per $[L : K]$, al cardinal de qualsevol base de L sobre K. Quan aquest cardinal és finit es parla d'*__extensió finita__*. En cas contrari, es parla d'*__extensió infinita__*.*

En els exemples anteriors tenim que tant $\mathbb{Q} \subset \mathbb{Q}(\sqrt{2})$ com $\mathbb{R} \subset \mathbb{C}$ són totes dues de grau 2, és a dir

$$[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = [\mathbb{C} : \mathbb{R}] = 2.$$

Un exemple d'extensió que no és de grau 2 és la de l'Exemple 21 de més amunt. Més endavant veurem que és de grau 3 (veure Exemple 33 i Teorema 36).

Val la pena remarcar que no totes les extensions de cossos són finites. Per exemple, es pot demostrar que mai es podran obtenir tots els nombres reals només a partir d'un conjunt finit d'ells fent-ne combinacions lineals amb coeficients a $\mathbb{Q}$. Per tant, $[\mathbb{R} : \mathbb{Q}] = \infty$. Ara bé, totes les extensions que considerarem en aquest treball, excepte una (associada al problema de la quadratura del cercle), seran finites.

Una propietat important del grau d'una extensió que necessitarem més endavant és que és "multiplicatiu". És a dir, es té el següent:

Proposició 29 *Si $K \subset L$ i $L \subset M$ són dues extensions de cossos de graus respectius $[L : K]$ i $[M : L]$, l'extensió $K \subset M$ és de grau*

$$[M : K] = [M : L] \cdot [L : K].$$

Prova. Ho raonem només en el cas que totes dues extensions són finites, encara que el resultat és sempre cert. Suposem que $[M : L] = n$ i que $[L : K] = m$. Es tracta de veure que si $\{\alpha_1, \dots \alpha_m\}$ és una base de L sobre K i $\{\beta_1, \dots, \beta_n\}$ una base de M sobre L, una base de M sobre K ve donada pels productes dos a dos dels elements de totes dues bases. El raonament és el següent.

Com que $\{\alpha_1, \dots \alpha_m\}$ és una base de L sobre K i $\{\beta_1, \dots, \beta_n\}$ ho és de M sobre L, qualsevol element de L és de la forma

$$\alpha = a_1\alpha_1 + \cdots + a_m\alpha_m, \tag{3.6.1}$$

amb $a_1, \dots, a_m \in K$, i qualsevol element de M és de la forma

$$\beta = b_1\beta_1 + \cdots + b_n\beta_n, \tag{3.6.2}$$

amb $b_1, \dots, b_n \in L$. En particular, cada b_j serà de la forma (3.6.1). Per tant, existiran $a_{1j}, \dots, a_{mj} \in K$ tals que

$$b_j = a_{1j}\alpha_1 + \cdots + a_{mj}\alpha_m.$$

Substituint aquestes expressions a (3.6.2) i calculant, obtenim que tot element de M és de la forma

$$\begin{aligned}\beta &= (a_{11}\alpha_1 + \cdots + a_{m1}\alpha_m)\beta_1 + \cdots + (a_{1n}\alpha_1 + \cdots + a_{mn}\alpha_m)\beta_n \\ &= a_{11}\alpha_1\beta_1 + \cdots + a_{m1}\alpha_m\beta_1 + \cdots + a_{1n}\alpha_1\beta_n + \cdots + a_{mn}\alpha_m\beta_n,\end{aligned}$$

[5] Això és un cas particular d'un fet més general conegut com a *teorema de la dimensió dels espais vectorials* o també *teorema de Steinitz*. Diu que, qualsevol que sigui l'espai vectorial, totes les seves bases tenen el mateix nombre d'elements. El cas que ens interessa és quan l'espai vectorial és el cos gran L com a espai vectorial sobre el cos petit K. Una demostració d'aquest resultat es pot trobar al llibre de M. Castellet i I. Llerena ja citat.

és a dir, és una combinació lineal dels productes $\{\alpha_1\beta_1, \ldots, \alpha_m\beta_1, \ldots, \alpha_1\beta_n, \ldots, \alpha_m\beta_n\}$ amb coeficients a K. Això prova que aquestes productes són generadors de M sobre K. D'altra banda, suposem que

$$a_{11}\alpha_1\beta_1 + \cdots + a_{m1}\alpha_m\beta_1 + \cdots + a_{1n}\alpha_1\beta_n + \cdots + a_{mn}\alpha_m\beta_n = 0$$

per a certs $a_{ij} \in K$. Això significa que

$$(a_{11}\alpha_1 + \cdots + a_{m1}\alpha_m)\beta_1 + \cdots + (a_{1n}\alpha_1 + \cdots + a_{mn}\alpha_m)\beta_n = 0.$$

Ara, cada suma $a_{1j}\alpha_1 + \cdots a_{mj}\alpha_m$, per a tot $j = 1, \ldots, n$, és un elements de L, i $\beta_1, \ldots, \beta_n$ són linealment independents sobre L. Per tant, totes les sumes anteriors han de ser nul·les. Però si

$$a_{1j}\alpha_1 + \cdots a_{mj}\alpha_m = 0,$$

com que $\alpha_1, \ldots, \alpha_m$ són linealment independents sobre K, resulta que els coeficients $\alpha_1, \ldots, \alpha_m$ han de ser tots zero. Deduïm, doncs, que els productes $\alpha_1\beta_1, \ldots, \alpha_m\beta_1, \ldots, \alpha_1\beta_n, \ldots, \alpha_m\beta_n$ també són linealment independents sobre K. Per tant, són una base de M sobre K. Com que n'hi ha $n \cdot m$ d'aquests productes, el grau de $K \subset M$ és realment el producte dels graus de totes dues extensions. □

3.7 Nombres algebraics i nombres transcendents

Tal i com veurem més endavant (veure Teorema 41), si un nombre real α és constructible a partir d'un cos base K donat, necessàriament compleix una condició: ha de ser el que s'anomena "algebraic" respecte K. En aquesta secció, explicarem que vol dir això i introduirem el concepte important de polinomi mínim d'un nombre algebraic respecte d'un cos K.

Els nombres reals s'acostumen a dividir en racionals i irracionals, segons que siguin decimals periòdics o no. Però també es poden considerar altres criteris de classificació. Un d'ells és el que distingeix entre *nombres algebraics* i *nombres transcendents*. Un exemple del primer grup és $\sqrt{2}$ i un del segon és π. La diferència entre uns i altres està en el fet de ser o no arrel d'algun polinomi amb coeficients a $\mathbb{Q}$. En el cas de $\sqrt{2}$ és algebraic perquè és arrel del polinomi $x^2 - 2 \in \mathbb{Q}[x]$. En canvi, tot i que no és gens evident, es pot demostrar que π no és arrel de cap polinomi amb coeficients a $\mathbb{Q}$ (això serà un dels punts clau per provar que no es pot quadrar un cercle amb regle i compàs).

De fet, la idea de nombre algebraic o transcendent es pot referir a qualsevol subcòs K de $\mathbb{R}$, no necessàriament $\mathbb{Q}$. La definició general és la següent.

Definició 30 *Sigui $K \subset \mathbb{R}$ un subcòs qualsevol. Un nombre real α es diu que és* **algebraic respecte** *K si existeix un polinomi no nul amb coeficients a K tal que α n'és arrel. En cas contrari, es diu que és* **transcendent respecte** *K. Quan $K = \mathbb{Q}$, parlarem simplement de* **nombre algebraic** *i de* **nombre transcendent**.

Observar que tots els nombres racionals són algebraics, ja que donats $a, b \in \mathbb{Z}$ qualssevol, amb $b \neq 0$, el nombre racional a/b és arrel del polinomi $bx - a \in \mathbb{Q}[x]$. Així doncs, la distinció entre algebraics i transcendents només és rellevant en el cas dels nombres irracionals.

Una altra observació important és que si un nombre α és algebraic respecte K, no serà arrel d'un únic polinomi amb coeficients a K. Per ex., si α és arrel d'un polinomi $p(x) \in K[x]$, també ho és del polinomi $p(x)q(x)$ per qualsevol altre $q(x) \in K[x]$. És per aquest motiu que s'introdueix la definició següent.

Definició 31 *Si $\alpha \in \mathbb{R}$ és algebraic respecte K, s'anomena* **polinomi mínim** *de α respecte K al polinomi mònic* [6] *de $K[x]$ de grau més petit que té α per arrel. El denotarem per $m_{\alpha,K}(x)$. Al seu grau se l'anomena* **grau de α respecte** *K i el representarem $deg_K(\alpha)$.*

La condició de mònic només s'imposa per tal que cada nombre algebraic tingui un únic polinomi mínim. Si no la imposèssim, qualsevol polinomi mínim multiplicat per una constant diferent de zero seria també un polinomi mínim.

Una propietat important del polinomi $m_{\alpha,K}(x)$ és que és irreductible a $K[x]$, és a dir, no es pot descomposar com a producte de dos polinomis no constants de $K[x]$ de graus menors. Això és fàcil de veure. Si descomposés com a producte de dos polinomis $p(x), q(x) \in K[x]$ de graus menors, almenys un dels dos factors s'hauria d'anul·lar quan $x = \alpha$ i, per tant, seria un polinomi de grau menor que el de $m_{\alpha,K}(x)$ que també tindria α per arrel. Però això es contradiu amb el fet que $m_{\alpha,K}(x)$ és el polinomi mínim de α. En realitat, aquesta propietat caracteritza el polinomi mínim en el sentit següent.

Proposició 32 *Sigui α un nombre algebraic respecte K i sigui $p(x) \in K[x]$ un polinomi mònic que té α per arrel. Aleshores, $p(x)$ és el polinomi mínim de α respecte K si i només si és irreductible a $K[x]$.*

Prova. Ja hem vist que si és el polinomi mínim respecte K és irreductible a $K[x]$. Ens falta veure que si és irreductible a $K[x]$, és el polinomi mínim respecte K. En efecte, com que té α per arrel, el seu grau és més gran o igual que el de $m_{\alpha,K}(x)$. Per tant, podem dividir $p(x)$ per $m_{\alpha,K}(x)$ (dividim a $K[x]$). Representem per $q(x), r(x) \in K[x]$ els polinomis quocient i resta, respectivament, de manera que

$$p(x) = m_{\alpha,K}(x)q(x) + r(x).$$

En particular, $r(x)$ és de grau menor que $m_{\alpha,K}(x)$. Fent $x = \alpha$ deduïm que $r(\alpha) = 0$, ja que α és arrel tant de $p(x)$ com de $m_{\alpha,K}(x)$. Però això implica que $r(x) = 0$, ja que sinó tindríem un polinomi de $K[x]$ i de grau menor que $m_{\alpha,K}(x)$ que tindria α per arrel. D'altra banda, com que $p(x)$ és irreductible, $q(x)$ ha de ser una constant. Però tant $p(x)$ com $m_{\alpha,K}(x)$ són mònics, de manera que $q(x)$ ha de ser necessàriament igual a 1 i $p(x) = m_{\alpha,K}(x)$. □

L'interès d'aquest resultat és que porporciona una manera en alguns casos més simple de veure si un polinomi és o no el polinomi mínim d'un nombre algebraic α. Enlloc d'haver de mirar si existeixen o no polinomis de $K[x]$ de grau menor que tenen α per arrel, només cal comprovar si és o no irreductible a $K[x]$.

Exemple 33 $\sqrt[3]{2}$ és algebraic (respecte $\mathbb{Q}$) de grau 3 perquè és arrel del polinomi $x^3 - 2$ i aquest polinomi és irreductible a $\mathbb{Q}[x]$. Per a provar que és irreductible, en aquest cas només cal observar que si no fos irreductible a $\mathbb{Q}[x]$, hauria de tenir almenys un factor de grau 1 a la seva descomposició i, per tant, almenys una arrel a $\mathbb{Q}$. Però com que el polinomi és mònic, sabem que els únics racionals que poden ser arrel d'aquest polinomi són els divisors de 2 i cap d'ells ho és.

[6] Un polinomi es diu que és *mònic* quan el coeficient del monomi de grau més grau és 1.

Exemple 34 Si $K = \mathbb{Q}(\sqrt{2})$, aleshores $\sqrt[3]{3}$ és algebraic respecte K de grau 3 perquè és arrel del polinomi $x^3 - 3$ i aquest polinomi també és irreductible a $K[x]$. Com en l'Exemple 33, per a provar que és irreductible es tracta d'observar que si no fos irreductible a $K[x]$, hauria de tenir almenys un factor de grau 1 a la seva descomposició i, per tant, almenys una arrel a K. Però les úniques arrels d'aquest polinomi són, a més de $\sqrt[3]{3}$, els nombres complexos (en forma polar) $\sqrt[3]{3}_{2\pi/3}$ i $\sqrt[3]{3}_{4\pi/3}$, i cap d'ells és de K. De fet, $x^3 - 3$ també és irreductible a $\mathbb{Q}[x]$ i, per tant, és el polinomi mínim de $\sqrt[3]{3}$ tant respecte $\mathbb{Q}$ com respecte K, així que passar de $\mathbb{Q}$ a l'extensió $\mathbb{Q}(\sqrt{2})$ no afecta gens al grau de $\sqrt[3]{3}$.

En els dos exemples anteriors, el candidat a polinomi mínim era molt simple de veure, però en general no és tan fàcil. Una manera de trobar un candidat consisteix en calcular les successives potències de α fins que se'n troba una que s'aconsegueix expressar com a combinació lineal de les anteriors amb coeficients a K. La potència més petita per la qual això passa donarà un polinomi de $K[x]$ que tindrà α per arrel i, per tant, un candidat a polinomi mínim. D'acord amb el resultat anterior, només caldrà comprovar que és irreductible a $K[x]$ per confirmar que ho és, i la comprovació d'això pot ser delicada. De fet, en general, provar que un polinomi és irreductible respecte un cos K (és a dir, a $K[x]$) no és una cosa fàcil i no entrarem en aquest tema.

Exemple 35 $\sqrt{2}+\sqrt{3}$ és algebraic de grau 4. Per a veure-ho, calculem les successives potències de $\sqrt{2} + \sqrt{3}$. Tenim que

$$\begin{aligned}(\sqrt{2} + \sqrt{3})^2 &= 5 + 2\sqrt{6},\\ (\sqrt{2} + \sqrt{3})^3 &= 11\sqrt{2} + 9\sqrt{3},\\ (\sqrt{2} + \sqrt{3})^4 &= 49 + 20\sqrt{6} = (-1)\cdot 1 + 10(\sqrt{2} + \sqrt{3})^2.\end{aligned}$$

Per tant, $\sqrt{2} + \sqrt{3}$ és solució de l'equació de grau 4

$$x^4 - 10x^2 + 1 = 0.$$

Per provar que aquest és el polinomi mínim hem de veure que és irreductible a $\mathbb{Q}$. En aquest cas, una manera de raonar és la següent. Si fos reductible, necessàriament descomposaria en alguna de les 4 maneres següents: (1) com a producte de quatre polinomis de grau 1, (2) com a producte de dos de grau 1 i un de grau 2, (3) com a producte d'un de grau 1 i un de grau 3 i (4) com a prodcute de dos de grau 2. Els tres primers casos implicarien que ha de tenir arrels a $\mathbb{Q}$, i no és el cas, ja que resolent l'equació es comprova que les arrels són

$$\pm\sqrt{5 + 2\sqrt{6}}, \quad \pm\sqrt{5 - 2\sqrt{6}}$$

i cap d'elles és racional. En l'últim cas, hauríem de poder descomposar el polinomi de la forma

$$(x^2 + ax + b)(x^2 + a'x + b')$$

amb $a, b, a', b' \in \mathbb{Q}$. Per tal que això sigui possible, hauria de passar que quan es fes el producte $(x-\alpha)(x-\beta)$, amb α, β dues de les quatre arrels anteriors, el resultat hauria de ser un polinomi amb coeficients a $\mathbb{Q}$. Però això no és mai cert perquè

$$(x-\alpha)(x-\beta) = x^2 - (\alpha+\beta)x + \alpha\beta,$$

i, triem com triem α i β d'entre les quatre arrels anteriors, és fàcil comprovar que o bé $\alpha+\beta$ o bé $\alpha\beta$ (o tots dos) són irracionals. Per tant, $x^4 - 10x^2 + 1$ és irreductible a $\mathbb{Q}$ i és el polinomi mínim buscat.

3.8 Càlcul del grau d'una extensió algebraica finita

L'última cosa que ens fa falta, en aquest cas no per entendre què diu el criteri de constructibilitat sinó per tal de provar-lo, és un mètode de calcular el grau d'una extensió simple $K \subset K(\alpha)$ quan α és algebraic respecte K. Això també ens permetrà saber si el nombre real α té possibilitats o no de ser constructible a partir de K (veure Teorema 41). Com que costa pràcticament el mateix, considerarem el cas general d'una extensió finita qualsevol de les anomenades *algebraiques*.

Si $K \subset \mathbb{R}$ és un subcòs qualsevol de $\mathbb{R}$, s'anomena **extensió algebraica finita** de K tota extensió de K de la forma $K(\alpha_1, \ldots, \alpha_n)$, amb tots els nombres $\alpha_1, \ldots, \alpha_n$ algebraics respecte K. [7] Quan $n = 1$, es parla d'**extensió algebraica simple** (el caràcter finit de l'extensió en aquest cas és automàtic per la condició de ser algebraica).

La primera cosa que hem d'observar és que el càlcul del grau d'una extensió algebraica finita qualsevol $K(\alpha_1, \ldots, \alpha_n)$ sempre es pot reduir a calcular graus d'extensions algebraiques simples. Només cal recordar que $K(\alpha_1, \ldots, \alpha_n)$ es pot obtenir fent un nombre finit d'extensions simples, totes algebraiques, i que el grau d'una extensió és multiplicatiu (veure Proposició 29). Per tant, tindrem que

$$\begin{aligned}[K(\alpha_1, \ldots, \alpha_n) : K] = [K(\alpha_1, \ldots, \alpha_n) : K(\alpha_1, \ldots, \alpha_{n-1})] \cdot [K(\alpha_1, \ldots, \alpha_{n-1}) : K(\alpha_1, \ldots, \alpha_{n-2})] \\ \cdots [K(\alpha_1, \alpha_2) : K(\alpha_1)] \cdot [K(\alpha_1) : K].\end{aligned}$$

Tenint en compte això, n'hi ha prou amb saber calcular els graus de les extensions algebraiques simples.

La resposta en aquest cas és molt senzilla, i ve donada pel resultat següent. La prova no és fàcil i fa servir l'anomenada *identitat de Bezout*, segons la qual el màxim comú divisor de dos polinomis és una certa combinació d'ells dos (l'enunciat exacte i la prova de la identitat de Bezout es poden trobar al § 5.3 de l'Apèndix). Però degut a la importància que el resultat té per aquest treball hem decidit incloure-la.

Teorema 36 *Si α és algebraic respecte K, el grau de l'extensió algebraica simple $K \subset K(\alpha)$ és*

$$[K(\alpha) : K] = deg_K(\alpha).$$

Més concretament, si $deg_K(\alpha) = n$, una base de $K(\alpha)$ sobre K és $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$.

[7] No considerem el cas que algun α_i és transcendent respecte K perquè en aquest cas es pot veure que el grau sempre és infinit.

Prova. Hem de comprovar les dues condicions de base: (1) que tot element de $K(\alpha)$ és combinació lineal de $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ amb coeficients a K, i (2) que $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ són linealment independents sobre K.

- Prova de (1): Equival a provar que tot element de $K(\alpha)$ és de la forma $p(\alpha)$ per algun polinomi $p(x) \in K[x]$ de grau menor que n. Per tal de provar-ho, recordem que $K(\alpha)$ és el subcòs més petit de $\mathbb{R}$ que conté K i α. Com que és estable per sumes, productes i inversos, també contindrà tots els reals de la forma $s(\alpha)/t(\alpha)$ per a qualsevol parell de polinomis $s(x), t(x) \in K[x]$ amb $t(\alpha) \neq 0$. Diguem-li S al conjunt de tots els reals d'aquesta forma, és a dir
$$S = \left\{ \frac{s(\alpha)}{t(\alpha)}, \text{ per a tot } s(x), t(x) \in K[x] \text{ amb } t(\alpha) \neq 0 \right\}.$$
Tenim, doncs, que $S \subseteq K(\alpha)$. Ara bé, es pot comprovar que S és un subcòs de $\mathbb{R}$, i clarament conté K i α. Com que $K(\alpha)$ és el subcòs més petit que conté K i α, deduïm que també $K(\alpha) \subseteq S$ i, per tant, que $K(\alpha) = S$. Això prova que tot element de $K(\alpha)$ és el valor en α del quocient de dos polinomis $s(x), t(x) \in K[x]$ amb $t(x)$ tal que $t(\alpha) \neq 0$. Provem ara que si $t(x) \in K[x]$ és tal que $t(\alpha) \neq 0$, aleshores l'invers del nombre real $t(\alpha)$ és el valor en α d'algun altre polinomi $a(x) \in K[x]$. En efecte, si $t(\alpha) \neq 0$, el polinomi $t(x)$ no és divisible pel polinomi mínim $m_{\alpha,K}(x)$ de α. A més, $m_{\alpha,K}(x)$ és irreductible a K, de manera que $t(x)$ i $m_{\alpha,K}(x)$ són polinomis relativament primers (no tenen factors comuns a la seva descomposició com a producte de polinomis irreductibles de $K[x]$). I és aquí on utilitzem la *identitat de Bezout*, segons la qual existiran aleshores polinomis $a(x), b(x) \in K[x]$ tals que
$$a(x)t(x) + b(x)m_{\alpha,K}(x) = 1$$
(veure § 5.3 de l'Apèndix). Si avaluem el membre de l'esquerra a $x = \alpha$, deduïm que $a(\alpha)t(\alpha) = 1$ i, per tant, que l'invers de $t(\alpha)$ és efectivament el valor en α d'un cert polinomi $a(x) \in K[x]$. Per tant, obtenim que tot element de $K(\alpha)$ és de la forma
$$\frac{s(\alpha)}{t(\alpha)} = s(\alpha)a(\alpha) = p(\alpha),$$
és a dir, és el valor en α d'algun polinomi $p(x) \in K[x]$. Queda veure que podem triar $p(x)$ que sigui de grau menor que n. Però això ja és fàcil, ja que si $p(x)$ no és de grau menor que n, només cal que fer la divisió de $p(x)$ per $m_{\alpha,K}(x)$ i quedar-nos amb el residu $r(x)$ de la divisió. Si $q(x)$ és el polinomi quocient, tindrem que
$$p(x) = m_{\alpha,K}(x)q(x) + r(x).$$
Com que $m_{\alpha,K}(\alpha) = 0$, $r(x)$ val en α el mateix que $p(x)$. Queda, per tant, provat que tot element de $K(\alpha)$ és efectivament una combinació lineal de $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$.

- Prova de (2): Per tal de provar la independència lineal sobre K de $1, \alpha, \ldots, \alpha^{n-1}$, suposem que fossin linealment dependents, és a dir, que existissin $a_0, a_1, a_2, \ldots, a_{n-1} \in K$ no tots iguals a zero tals que
$$a_0 + a_1\alpha + a_2\alpha^2 + \ldots + a_{n-1}\alpha^{n-1} = 0.$$

Si $i \leq n-1$ és el subíndex més gran tal que $a_i \neq 0$, vol dir que en realitat tenim que

$$a_0 + a_1\alpha + a_2\alpha^2 + \ldots + a_i\alpha^i = 0.$$

Com que $a_i \neq 0$, podem dividir-ho tot per a_i i obtenim que α és arrel del polinomi de $K[x]$

$$\frac{a_0}{a_i} + \frac{a_1}{a_i}x + \ldots + \frac{a_{i-1}}{a_i}x^{i-1} + x^i = 0,$$

de grau $i < n$, la qual cosa es contradiu amb el fet que n és el grau del polinomi mínim.

Queda, doncs, provat que $\{1, \alpha, \ldots, \alpha^{n-1}\}$ és una base de $K(\alpha)$ sobre K. □

Aquest resultat permet determinar ràpidament de quina forma són els elements de qualsevol extensió algebraica simple $K(\alpha)$ un cop se sap el grau de α respecte K: serà una certa combinació lineal de totes les potències de α fins al grau de α respecte K menys 1.

Exemple 37 Més enrere, a l'Exemple 21, raonàvem d'una manera més o menys complicada que $\mathbb{Q}(\sqrt[3]{2})$ no podia constar només dels reals de la forma $a + b\sqrt[3]{2}$, amb $a, b \in \mathbb{Q}$. Ara, en canvi, sabem per l'Exemple 33 que $\sqrt[3]{2}$ és de grau 3 respecte $\mathbb{Q}$. Per tant, pel Teorema 36, $\mathbb{Q}(\sqrt[3]{2})$ és una extensió de $\mathbb{Q}$ de grau 3, i una base de $\mathbb{Q}(\sqrt[3]{2})$ sobre $\mathbb{Q}$ és $\{1, \sqrt[3]{2}, \sqrt[3]{4}\}$. Els elements de $\mathbb{Q}(\sqrt[3]{2})$ són, doncs, tots els de la forma $a + b\sqrt[3]{2} + c\sqrt[3]{4}$ amb $a, b, c \in \mathbb{Q}$ i només aquests.

Exemple 38 Per tal de calcular el grau de l'extensió $\mathbb{Q}(\sqrt{2}, \sqrt[3]{3})$ de $\mathbb{Q}$ pensem en les dues extensions simples

$$\mathbb{Q} \subset \mathbb{Q}(\sqrt{2}) \subset \mathbb{Q}(\sqrt{2}, \sqrt[3]{3}).$$

La primera ja sabem que és de grau 2, i que una base de $\mathbb{Q}(\sqrt{2})$ sobre $\mathbb{Q}$ és $\{1, \sqrt{2}\}$. Pel que fa a la segona, sabem per l'Exemple 34 i pel resultat anterior que és de grau 3, i que una base de $\mathbb{Q}(\sqrt{2}, \sqrt[3]{3})$ sobre $\mathbb{Q}(\sqrt{2})$ és $\{1, \sqrt[3]{3}, \sqrt[3]{9}\}$. Per tant, de la Proposició 29 deduïm que $\mathbb{Q}(\sqrt{2}, \sqrt[3]{3})$ és una extensió de $\mathbb{Q}$ de grau

$$[\mathbb{Q}(\sqrt{2}, \sqrt[3]{3}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{2}, \sqrt[3]{3}) : \mathbb{Q}(\sqrt{2})] \cdot [\mathbb{Q}\sqrt{2}) : \mathbb{Q}] = 6,$$

i que una base de $\mathbb{Q}(\sqrt{2}, \sqrt[3]{3})$ sobre $\mathbb{Q}$ és $\{1, \sqrt{2}, \sqrt[3]{3}, \sqrt[3]{9}, \sqrt{2}\sqrt[3]{3}, \sqrt{2}\sqrt[3]{9}\}$. Així, un element genèric de $\mathbb{Q}(\sqrt{2}, \sqrt[3]{3})$ és de la forma

$$a + b\sqrt{2} + c\sqrt[3]{3} + d\sqrt[3]{9} + e\sqrt{2}\sqrt[3]{3} + f\sqrt{2}\sqrt[3]{9}$$

amb $a, b, c, d, e, f \in \mathbb{Q}$.

Observem que obtenir les potències α^q per a $q \geq n$ com a combinacions lineals de $1, \alpha, \ldots, \alpha^{n-1}$ és molt fàcil. Només cal utilitzar el polinomi mínim $m_{\alpha,K}(x)$, que sabem que és de la forma

$$m_{\alpha,K}(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_2x^2 + a_1x + a_0,$$

amb $a_0, \ldots, a_{n-1} \in K$. Com que $m_{\alpha,K}(\alpha) = 0$, tindrem que

$$\alpha^n = -a_0 - a_1\alpha - a_2\alpha^2 - \ldots - a_{n-1}\alpha^{n-1}, \tag{3.8.1}$$

que expressa α^n com a combinació lineal de $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$ amb coeficients a K. Per tal d'obtenir les potències més grans només cal multiplicar a les dues bandes de (3.8.1) per α tantes vegades com faci falta i anar substituint cada potència α^q per $q \geq n$ pel que val. Per exemple

$$\begin{aligned}
\alpha^{n+1} &= -a_0\alpha - a_1\alpha^2 - \ldots - a_{n-2}\alpha^{n-1} - a_{n-1}\alpha^n \\
&= -a_0\alpha - a_1\alpha^2 - \ldots - a_{n-2}\alpha^{n-1} - a_{n-1}(-a_0 - a_1\alpha - a_2\alpha^2 - \ldots - a_{n-1}\alpha^{n-1}) \\
&= a_{n-1}a_0 + (a_{n-1}a_1 - a_0)\alpha + (a_{n-1}a_2 - a_1)\alpha^2 + \cdots + (a_{n-1}a_{n-1} - a_{n-2})\alpha^{n-1}.
\end{aligned}$$

Anàlogament, podem obtenir α^{n+2} multiplicant (3.8.1) per α^2 i substituint les expressions ja obtingudes per α^n i α^{n+1}, i així successivament.

3.9 Criteri de constructibilitat d'un nombre real amb regle i compàs

Ara ja estem en condicions d'enunciar i entendre què diu el criteri de constructibilitat d'un nombre real a partir d'un $\mathcal{B}$ donat. No només és una condició que ha de complir qualsevol nombre per poder ser constructible amb regle i compàs a partir de $\mathcal{B}$, sinó que a més, si un nombre la compleix, és segur que és constructible amb regle i compàs a partir de $\mathcal{B}$. La condició és el següent:

Teorema 39 *Un nombre real t és constructible a partir $\mathcal{B}$ (és a dir, $t \in \mathfrak{C}_\mathcal{B}$) si i només si existeix un enter $p \geq 1$ i una successió de subcossos $L_0, L_1, \ldots, L_p$ de $\mathbb{R}$ tals que:*

(1) $L_0 = \mathbb{Q}(X_\mathcal{B})$,

(2) per tot $i \in \{1, 2, \ldots, p\}$, L_i *és una extensió de* L_{i-1} *de grau 2, i*

(3) $t \in L_p$.

Prova. Suposem primer que t és constructible a partir de $\mathcal{B}$ i provem que aleshores existeix la successió de subcossos de la que parla l'enunciat. D'acord amb el Lema 14, si t és un nombre constructible a partir de $\mathcal{B}$, el punt M de coordenades $(t, 0)$ també ho és. Representem per $M_1, M_2, ..., M_n$, amb $M_n = M$, el seguit de punts que s'han de construir prèviament fins arribar a construir M, i siguin $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ les coordenades respectives. Considerem aleshores la successió d'extensions de $\mathbb{Q}$ següent, obtinguda afegint successivament les

coordenades dels punts M_i que es van construint:

$$\begin{aligned}
K_0 &= \mathbb{Q}(X_{\mathcal{B}}) \\
K_1 &= K_0(x_1, y_1) = \mathbb{Q}(X_{\mathcal{B}}, x_1, y_1) \\
K_2 &= K_1(x_2, y_2) = \mathbb{Q}(X_{\mathcal{B}}, x_1, y_1, x_2, y_2) \\
&\cdots \\
K_i &= K_{i-1}(x_i, y_i) = \mathbb{Q}(X_{\mathcal{B}}, x_1, y_1, \ldots, x_i, y_i) \\
&\cdots \\
K_n &= K_{n-1}(x_n, y_n) = \mathbb{Q}(X_{\mathcal{B}}, x_1, y_1, \ldots, x_n, y_n)
\end{aligned}$$

Clarament, $K_0 \subseteq K_1 \subseteq \cdots \subseteq K_{i-1} \subseteq K_i \subseteq \cdots \subseteq K_n$. Volem veure que per qualsevol $i \in \{1, \ldots, n\}$ o bé $K_i = K_{i-1}$ o bé K_i és una extensió de grau 2 de K_{i-1}. En efecte, poden passar tres coses, segons que el nou punt M_i que afegim sigui la intersecció de dues rectes, de dues circumferències o d'una recta i una circumferència. Analitzem els tres casos per separat. Observar que en tots ells les equacions d'aquestes rectes i circumferències tindran totes elles els coeficients a K_{i-1} perquè estan definides a partir de punts base i/o dels punts $M_1, \ldots, M_{i-1}$ construïts prèviament.

(1) Si M_i és la intersecció de dues rectes, el parell (x_i, y_i) serà la solució d'un sistema d'equacions lineals compatible determinat de la forma

$$\begin{cases} \alpha x_i + \beta y_i + \gamma = 0 \\ \alpha' x_i + \beta' y_i + \gamma' = 0 \end{cases}$$

amb $\alpha, \beta, \gamma, \alpha', \beta', \gamma' \in K_{i-1}$. Tant x_i com y_i continuaran estant dins de K_{i-1} perquè K_{i-1} és estable per sumes, productes, oposats i inversos, que són les úniques operacions que cal fer per trobar-les. Per tant, en passar de K_{i-1} a K_i no afegim res que ja no hi pertanyi i tenim que $K_i = K_{i-1}$.

(2) Si M_i és la intersecció d'una recta i d'una circumferència, el parell (x_i, y_i) és solució d'un sistema d'equacions de la forma

$$\begin{cases} \alpha x_i + \beta y_i + \gamma = 0 \\ x_i^2 + y_i^2 - 2\alpha' x_i - 2\beta' y_i + \gamma' = 0 \end{cases}$$

amb $\alpha, \beta, \gamma, \alpha', \beta', \gamma' \in K_{i-1}$. Aïllem per exemple y_i (si $\beta \neq 0$) de la primera equació:

$$y_i = \frac{-\alpha x_i - \gamma}{\beta}. \tag{3.9.1}$$

Substituint a la segona i agrupant termes, s'obté que x_i és arrel del polinomi mònic

$$P(x) = x^2 + \frac{\beta^2}{\alpha^2 + \beta^2}\left(\frac{2\alpha\gamma}{\beta^2} - 2\alpha' + \frac{2\beta'}{\beta}\right)x + \frac{\beta^2}{\alpha^2 + \beta^2}\left(\frac{\gamma^2}{\beta^2} + \frac{2\beta'\gamma}{\beta} + \gamma'\right) \in K_{i-1}[x].$$

Per tant, o bé x_i és de K_{i-1} o si no, és de grau 2 respecte de K_{i-1}, ja que $P(x)$ en serà el polinomi mínim respecte de K_{i-1}. Si és $x_i \in K_{i-1}$, també y_i serà de K_{i-1} (veure

Equació (3.9.1)) i, per tant, tindrem que $K_i = K_{i-1}$. En canvi, si $x_i \notin K_{i-1}$, com que $y_i \in K_{i-1}(x_i)$ (una altra vegada per l'Equació (3.9.1)), tindrem que

$$K_i = K_{i-1}(x_i, y_i) = K_{i-1}(x_i)$$

i, per tant, K_i és una extensió algebraica simple de grau 2 de K_{i-1} d'acord amb el Teorema 36.

(3) Si M_i és la intersecció de dues circumferències, el parell (x_i, y_i) és solució d'un sistema d'equacions de la forma

$$\begin{cases} x^2 + y^2 - 2\alpha x - 2\beta y + \gamma = 0 \\ x^2 + y^2 - 2\alpha' x - 2\beta' y + \gamma' = 0 \end{cases}$$

amb $\alpha, \beta, \gamma, \alpha', \beta', \gamma' \in K_{i-1}$. Restant una equació de l'altra, aquest sistema és equivalent al sistema

$$\begin{cases} x^2 + y^2 - 2\alpha x - 2\beta y + \gamma = 0 \\ 2(\alpha - \alpha')x + 2(\beta - \beta')y - (\gamma - \gamma') = 0 \end{cases}$$

i a partir d'aquí el raonament és el mateix del cas anterior.

La successió de subcossos $L_0, L_1, \ldots, L_p$ de l'enunciat s'obté aleshores suprimint tots els subcossos que es repeteixin a la seqüència $K_0, K_1, \ldots, K_n$. Que $t \in L_p$ és perquè $M_n = M$.

Per demostrar la implicació contrària el que es fa es veure que per qualsevol successió de subcossos $L_0, L_1, \ldots, L_p$ com la de l'enunciat en realitat es té que tots ells estan dintre de $\mathfrak{C}_\mathcal{B}$ (en particular, el nombre $t \in L_p$ serà constructible). I això es prova fent el que s'anomena *inducció*. La idea és provar primer que L_0 està dintre de $\mathfrak{C}_\mathcal{B}$ i després veure que si un qualsevol dels subcossos L_i està dintre de $\mathfrak{C}_\mathcal{B}$ també ho està el següent L_{i+1}. Com que aquesta implicació en realitat no la necessitarem, no en fem els detalls. La demostració completa es pot trobar al llibre de J.C. Carrega que apareix citat a la Bibliografia. □

Exemple 40 Ja sabem per l'Exemple 24 que el nombre real $\sqrt[4]{2}$ és constructible a partir de $\mathbb{Q}$ i volem trobar la cadena d'extensions de cossos de la qual parla el criteri anterior. D'acord amb la demostració, es tracta de trobar la seqüència de punts que s'han d'anar construint fins arribar al punt que volem. Aplicant els passos descrits a l'Exemple 24, es pot comprovar que s'arriba al punt $(1, \sqrt[4]{2})$ després de construir els punts següents amb regle i compàs (detallem els punts involucrats en cada pas tal i com estan descrits a l'Exemple):

Pas 1. $(2, 0)$.
Pas 2. $(3, 0)$, $(3/2, 3\sqrt{3}/2)$, $(3/2, -3\sqrt{3}/2)$, $(3/2, 0)$, $(1, \sqrt{3})$, $(1, -\sqrt{3})$, $(1, \sqrt{2})$.
Pas 3. $(0, 2\sqrt{2})$, $(\sqrt{6}, \sqrt{2})$, $(\sqrt{6}, -\sqrt{2})$, $(0, \sqrt{2})$.
Pas 4. $(\sqrt{2}, 0)$.
Pas 5. $(1 + \sqrt{2}, 0)$, $((1 + \sqrt{2})/2, (\sqrt{3} + \sqrt{6})/2)$, $((1 + \sqrt{2})/2, -(\sqrt{3} + \sqrt{6})/2)$, $((1 + \sqrt{2})/2, 0)$, $(1, \sqrt[4]{2})$.

Per a obtenir tots aquests punts només cal anar trobant les equacions de les rectes i circumferències involucrades i calcular-ne les interseccions que interessen. A partir

de la llista de punts anterior, deduïm que la successió d'extensions per les que passem fins arribar a construir $\sqrt[4]{2}$ és la següent:

$$\mathbb{Q} \subset \mathbb{Q}(\sqrt{3}) \subset \mathbb{Q}(\sqrt{3}, \sqrt{2}) \subset \mathbb{Q}(\sqrt{3}, \sqrt{2}, \sqrt[4]{2}) = \mathbb{Q}(\sqrt{3}, \sqrt[4]{2})$$

(només indiquem les extensions que realment ho són, ja que moltes de les extensions en realitat no ho són perquè les coordenades del nou punt construït ja pertanyen al cos de partida).

Observar que l'extensió final de l'exemple anterior és més gran que el subcòs més petit que conté $\sqrt[4]{2}$, és a dir, que $\mathbb{Q}(\sqrt[4]{2})$. En general, passarà que per tal de construir un real t amb regle i compàs serà necessari anar a una extensió més gran del que en principi caldria. Això és perquè el punt que es voldrà construir en general no es podrà construir en un sol pas a partir dels punts base, sinó que caldrà construir-ne uns quants de previs, i alguns d'aquests normalment ja suposaran fer extensions del cos base.

Capítol 4

Impossibilitat de les tres construccions clàssiques

Amb tots els conceptes i resultats de què disposem, en aquest capítol ja podem per fi explicar per què les tres construccions amb regle i compàs que es plantejaven els grecs fa més de 2000 anys eren impossibles. Ells no van poder-ho demostrar perquè feia falta una manera radicalment diferent d'abordar el problema (la manera algebraica). Però també s'ha de dir que, no sent capaços de fer les construccions que es plantejaven amb regle i compàs (eren, de fet, impossibles), sí que van ser capaços de fer-les per altres mètodes. Acabarem el capítol comentant molt breument alguns d'aquests mètodes.

4.1 El resultat de Wantzel

Per tal de provar que els tres problemes clàssics no tenen solució, només necessitem una conseqüència del criteri de constructibilitat que hem vist al capítol anterior. L'anomenarem *resultat de Wantzel* perquè va ser el matemàtic francés Pierre Wantzel qui primer el va demostrar l'any 1837. El resultat diu el següent:

Teorema 41 (Resultat de Wantzel) *Sigui $\mathcal{B}$ el conjunt de punts base i $K = \mathbb{Q}(X_{\mathcal{B}})$ el cos base corresponent. Si un nombre real α és constructible a partir de K, aleshores és algebraic respecte K i el grau de l'extensió $K \subset K(\alpha)$ és una potència de 2.*

Prova. D'acord amb el criteri de constructibilitat (Teorema 39), sabem que si α és constructible a partir de K existeix una successió de subcossos $L_0 \subset L_1 \subset \cdots \subset L_p$ de $\mathbb{R}$ tal que $L_0 = K$, $\alpha \in L_p$ i $[L_{j+1} : L_j] = 2$ per a tot $0 \leq j \leq p-1$. Per tant, per la propietat multiplicativa dels graus (Proposició 29) tenim que

$$[L_p : K] = [L_p : L_{p-1}] \cdot [L_{p-1} : L_{p-2}] \cdots [L_1 : K] = 2^p$$

Ara bé, tal i com hem vist a l'Exemple 40, L_p pot no coincidir amb $K(\alpha)$. Però el que sí que és segur és que $K(\alpha) \subseteq L_p$, és a dir, L_p és una extensió de $K(\alpha)$, perquè $K(\alpha)$ és el subcòs més petit de $\mathbb{R}$ que conté K i α i L_p és un subcòs que conté K i α. Aplicant aleshores un altre cop la propietat multiplicativa dels grau deduïm que

$$[L_p : K(\alpha)] \cdot [K(\alpha) : K] = [L_p : K] = 2^p$$

i, per tant, $[K(\alpha) : K]$ és necessàriament una potència de 2 (no poden aparèixer altres factors primers a la seva descomposició perquè sinó també hi serien a descomposició de $[L_p : K]$). Suposem que és $[K(\alpha) : K] = 2^q$ $(g \geq 1)$. Falta provar que α és algebraic respecte K. Per tal de veure-ho, considerem la família de reals $1, \alpha, \alpha^2, \ldots, \alpha^{2^q}$. Tots ells pertanyen a $K(\alpha)$ per estabilitat. En total n'hi ha $2^q + 1 > 2^q$ i, per tant, no poden ser linealment independents sobre K. Per tant, existeixen $a_0, a_1, \ldots, a_{2^q} \in K$ tals que

$$a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_{2^q}\alpha^{2^q} = 0$$

i α és efectivament algebraic respecte K. □

Amb aquest resultat ja podem entendre per què són impossibles les tres construccions. Començarem amb el cas més simple, el de la duplicació del cub. El cas de la trisecció d'un angle és una mica més complicat, perquè hi ha angles per als quals és possible i angles per als quals no ho és. Finalment, el cas de la quadratura del cercle és de llarg el més difícil dels tres. De fet, no en donarem una demostració completa de la impossibilitat, perquè és necessari provar que el nombre π és transcendent i la prova d'això és difícil.

4.2 Impossibilitat de la duplicació del cub

Recordem que aquest problema consisteix en construir amb regle i compàs el costat d'un cub de volum el doble del d'un cub donat. Per tal de simplificar el problema suposarem que el cub original és de volum 1 i, per tant, el que hem de construir és de volum 2. Tal i com ja hem explicat a la Introducció, això significa que partim d'un segment de longitud 1 i que n'hem de construir un de longitud $\sqrt[3]{2}$.

Triem el sistema de coordenades de manera que els extrems del segment de partida són els punts $(0, 0)$ i $(1, 0)$. El conjunt de punts base és, doncs, $\mathcal{B} = \{(0, 0), (1, 0)\}$ i el cos base és $\mathbb{Q}$. Amb regle i compàs es fàcil construir una còpia d'un segment qualsevol del pla en qualsevol altra posició. Per tant, suposant que el nou segment el dibuixem amb origen al $(0, 0)$, el problema es redueix a saber construir amb regle i compàs el punt $(0, \sqrt[3]{2})$. Això només és possible si el nombre $\sqrt[3]{2}$ és constructible amb regle i compàs a partir de $\mathbb{Q}$.

Ara bé, ja hem vist a l'Exemple 37 que $\sqrt[3]{2}$ és de grau 3 respecte $\mathbb{Q}$ (el polinomi mínim és $x^3 - 2 \in \mathbb{Q}[x]$) i, per tant, que

$$[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3.$$

Pel criteri de Wantzel, doncs, $\sqrt[3]{2}$ no és constructible a partir de $\mathbb{Q}$, ja que el grau d'aquesta extensió no és una potència de 2.

Encara que no es pugui duplicar un cub de l'espai ordinari de tres dimensions, observem que sí que es pot duplicar un *hipercub* d'un espai de quatre dimensions. Un hipercub no és més que la versió en dimensió 4 del que és el quadrat en dimensió 2 i el cub ordinari en dimensió 3. El seu "hipervolum", si és de costat a, serà a^4. Per tant, l'hipercub de volum doble tindrà un costat de longitud $a\sqrt[4]{2}$ i la seva construcció correspon en el fons a construir amb regle i compàs el nombre real $\sqrt[4]{2}$, que sabem que sí que és constructible (veure Exemple 24). En general, es podran duplicar els hipercubs de tots els espais que siguin de dimensió una potència de dos qualsevol.

4.3 Impossibilitat de la trisecció d'un angle genèric

Recordem que aquest problema consisteix en construir amb regle i compàs les dues semirectes que divideixen en tres parts iguals un angle α donat. De fet, n'hi ha prou amb saber-ne construir una, ja que l'altra s'obtindrà construint la bisectriu de l'angle més gran que quedi.

Tal i com hem explicat a la Introducció, donar un angle equival a donar-ne el vèrtex O i dos punts A i B, un de cada costat, que suposarem a la mateixa distància de O. Agafem el sistema de coordenades de manera que $O = (0,0)$ i $A = (1,0)$. El punt B és aleshores el de coordenades $B = (\cos\alpha, \sin\alpha)$. Representant per $\mathcal{B}_\alpha$ el conjunt de punts base i per K_α el cos base corresponent, tenim doncs que $\mathcal{B}_\alpha = \{(0,0),(1,0),(\cos\alpha,\sin\alpha)\}$ i $K_\alpha = \mathbb{Q}(\cos\alpha,\sin\alpha)$. En general, tindrem que

$$\mathbb{Q} \subsetneq \mathbb{Q}(\cos\alpha) \subsetneq K_\alpha,$$

encara que per alguns angles α alguna de les inclusions pot ser una igualtat. Per exemple, tots tres cossos són $\mathbb{Q}$ si $\alpha = \pi/2$. En canvi, si $\alpha = \pi/3$ els dos primers són $\mathbb{Q}$ però el tercer és $\mathbb{Q}(\sqrt{3})$, ja que $\cos(\pi/3) = 1/2$ i $\sin(\pi/3) = \sqrt{3}/2$.

El primer que hem d'observar és que la condició que l'angle $\alpha = \angle AOB$ sigui trisecable equival a què el punt $M = (\cos(\alpha/3), \sin(\alpha/3))$ és constructible a partir de $\mathcal{B}_\alpha$. En efecte, si és trisecable, M és constructible perquè és la intersecció de la "trisectriu" amb la circumferència centrada a O i de radi OA, i a l'inrevés, si M és constructible, la "trisectriu" també ho és perquè no és més que la recta OM.

Per altra banda, M és constructible a partir de $\mathcal{B}_\alpha$ si i només si el nombre real $t = \cos(\alpha/3)$ és constructible a partir de K_α, ja que si M és constructible a partir de $\mathcal{B}$, les seves dues coordenades ho són a partir de K_α, i a l'inrevés, si t és constructible a partir de K_α, també ho és $\sin(\alpha/3)$ perquè és l'arrel quadrada de $1 - t^2$.

Resumint, α serà trisecable si i només si t és constructible a partir de K_α. Tenim aleshores el criteri següent, que ens diu quins angle seran trisecables i quins no.

Teorema 42 *L'angle α és trisecable a partir de $\mathcal{B}_\alpha = \{(0,0),(1,0),(\cos\alpha,\sin\alpha)\}$ si i només si el polinomi $p_\alpha(x) = 4x^4 - 3x - \cos\alpha$ és reductible a $\mathbb{Q}(\cos\alpha)[x]$.*

Observem que és la reductibilitat a $\mathbb{Q}(\cos\alpha)[x]$ la que cal estudiar, i no la reductibilitat a $K_\alpha[x]$. En general, com que són cossos diferents, són condicions diferents.

Prova. La part més difícil és veure que si α és trisecable a partir de $\mathcal{B}_\alpha$ es compleix la condició de l'enunciat. Suposem, doncs, que α és trisecable a partir de $\mathcal{B}_\alpha$, és a dir, que t és constructible a partir de K_α. Pel resultat de Wantzel, això implica que t és algebraic respecte K_α i que $[K_\alpha(t) : K_\alpha] = 2^q$ per algun $q \geq 0$. Ara bé, com que $\sin\alpha = 1 - t^2$, això implica que si t és algebraic respecte K_α també ho serà respecte $\mathbb{Q}(\cos\alpha)$. Un polinomi de $\mathbb{Q}(\cos\alpha)[x]$ que tingui t per arrel es pot obtenir de la manera següent. Partim del polinomi mínim $m_{t,K_\alpha}(x)$ i substituim totes les potències parelles de $\sin\alpha$ que hi apareixen per la potència que correspongui de $1 - t^2$, i totes les potències senars pel producte de $\sin\alpha$ i la potència que toqui de $1 - t^2$. A continuació, aïllem una aparició qualsevol de $\sin\alpha$ de la igualtat $m_{t,K_\alpha}(t) = 0$, elevem al quadrat i substituim $\sin^2\alpha$ per $1 - t^2$. Això ens donarà un nou polinomi que tindrà t per arrel però una aparició menys de $\sin\alpha$. Repetint això tantes vegades com calgui, farem desaparèixer tots els $\sin\alpha$ i ens quedarà un polinomi de $\mathbb{Q}(\cos\alpha)[x]$ que tindrà t per arrel. Per tant, ja sabem que si α és trisecable, t és algebraic respecte $\mathbb{Q}(\cos\alpha)$.

Per altra banda, $\sin\alpha$ és arrel de $x^2+\cos^2\alpha-1\in\mathbb{Q}(\cos\alpha)[x]$ i, per tant, també és algebraic respecte $\mathbb{Q}(\cos\alpha)$ i $[K_\alpha:\mathbb{Q}(\cos\alpha)]$ val 1 o 2. Per la multiplicativitat dels graus (Proposició 29) tenim doncs que

$$[K_\alpha(t):\mathbb{Q}(\cos\alpha)]=[K_\alpha(t):K_\alpha]\cdot[K_\alpha:\mathbb{Q}(\cos\alpha)]=2^{q'},$$

amb $q'=q$ o bé $q'=q+1$. Però

$$[K_\alpha(t):\mathbb{Q}(\cos\alpha)]=[K_\alpha(t):\mathbb{Q}(\cos\alpha)(t)]\cdot[\mathbb{Q}(\cos\alpha)(t):\mathbb{Q}(\cos\alpha)].$$

Per tant, si aquest producte és una potència de 2, també ho ha de ser $[\mathbb{Q}(\cos\alpha)(t):\mathbb{Q}(\cos\alpha)]$, però aquest grau és el grau de $m_{t,\mathbb{Q}(\cos\alpha)}(x)$, el polinomi mínim de t respecte $\mathbb{Q}(\cos\alpha)$. Finalment, per trigonometria sabem que

$$4\cos^3(\alpha/3)=3\cos(\alpha)+\cos\alpha.$$

Això ens diu que t és arrel de $p_\alpha(x)=4x^3-3x-\cos\alpha\in\mathbb{Q}(\cos\alpha)[x]$. Però això implica que aquest polinomi $p_\alpha(x)$ no pot ser irreductible a $\mathbb{Q}(\cos\alpha)[x]$ perquè si ho fos, seria el polinomi mínim $m_{t,\mathbb{Q}(\cos\alpha)}(x)$ (veure Proposició 32), i acabem de veure que $m_{t,\mathbb{Q}(\cos\alpha)}(x)$ és de grau una potència de 2.

Veure que α és trisecable si $p_\alpha(x)$ és reductible a $\mathbb{Q}(\cos\alpha)[x]$ és molt més fàcil. Sabem ja que t és arrel de $p_\alpha(x)$. Com que estem suposant que $p_\alpha(x)$ és reductible a $\mathbb{Q}(\cos\alpha)[x]$, això vol dir que t és arrel d'un dels factors de la descomposició de $p_\alpha(x)$ i, per tant, arrel d'un polinomi de $\mathbb{Q}(\cos\alpha)[x]$ de grau 1 o 2. Això significa que t es pot obtenir a partir d'elements de $\mathbb{Q}(\cos\alpha)$ fent operacions aritmètiques elementals i, com a molt, arrels quadrades. Però $\mathbb{Q}(\cos\alpha)\subset K_\alpha$ i, d'acord amb el Teorema 23, el cos dels nombres constructibles a partir de K_α és estable per arrels quadrades, així que t és constructible a partir de K_α i, per tant, α és trisecable. □

Exemple 43 L'angle $\alpha=\pi/3$ no és trisecable amb regle i compàs perquè el seu cosinus és $1/2$ i el polinomi $p(x)=4x^3-3x-1/2$ o, el que és el mateix, el polinomi $q(x)=8x^3-6x-1$ és irreductible a $\mathbb{Q}[x]$ (observar que $\mathbb{Q}(1/2)=\mathbb{Q}$). En efecte, si fos reductible, hauria de tenir almenys una arrel racional, i per Ruffini es comprova fàcilment que cap de les possibles arrels racionals $\pm1,\pm1/2,\pm1/4,\pm1/8$ és realment arrel.

Exemple 44 L'angle $\alpha=\pi/4$ és trisecable amb regle i compàs perquè el seu cosinus és $\sqrt{2}/2$ i el polinomi $p(x)=4x^3-3x-\sqrt{2}/2$ és reductible a $\mathbb{Q}(\sqrt{2}/2)[x]$. En efecte, si provem una arrel del tipus $a\sqrt{2}$, amb $a\in\mathbb{Q}$, resulta que a ha de complir l'equació $16a^3-6a-1=0$. Les possibles arrels racionals d'aquesta equació sabem que són $\pm1,\pm1/2,\pm1/4,\pm1/8,\pm1/16$ i és fàcil comprovar que $-1/2$ n'és una. Per tant, $\sqrt{2}/2$ és arrel del polinomi $p(x)$. De fet, es té que $p(x)=(x+\sqrt{2}/2)(4x^2-2\sqrt{2}x-1)$.

4.4 Impossibilitat de la quadratura del cercle

Recordem que aquest problema consisteix en constuir un quadrat que tingui la mateixa àrea que un cercle donat. Per fer-ho més fàcil, suposarem que el cercle és de radi 1, així que la seva àrea és π i el costat del quadrat que s'ha de construir és de longitud $\sqrt{\pi}$.

Triem el sistema de coordenades de manera que el centre del cercle de partida és el punt $(0,0)$ i com a punt del cercle agafem el $(1,0)$. El conjunt de punts base és doncs $\mathcal{B} = \{(0,0),(1,0)\}$ i el cos base és $\mathbb{Q}$. Com que un cop es té un costat del quadrat, la resta es pot construir fàcilment amb regle i compàs, el problema és construir amb regle i compàs un segment de longitud $\sqrt{\pi}$. Pel mateix argument que en el cas de la duplicació del cub, això equival a construir amb regle i compàs el punt $(0,\sqrt{\pi})$ a partir del $\mathcal{B}$ anterior, és a dir, a construir el nombre real $\sqrt{\pi}$ a partir de $\mathbb{Q}$. Ara bé, pel criteri de Wantzel, si $\sqrt{\pi}$ és constructible a partir de $\mathbb{Q}$ és algebraic respecte $\mathbb{Q}$. I tot i que no és gens fàcil de provar, es pot demostrar que π és transcendent. Per tant, el cercle no es pot quadrar.

Com ja hem dit a la introducció, la primera prova de la transcendència de π data del 1882 i és deguda al matemàtic alemany Carl Louis Ferdinand von Lindemann. En aquest sentit, la prova definitiva de la impossibilitat de quadrar el cercle no es va tenir fins a l'any 1882. Una demostració del fet que π és transcendent es pot trobar al Capítol VI del llibre de I. Stewart que apareix citat a la Bibliografia, però és realment molt complicada.

4.5 Solució dels tres problemes per altres mètodes

No voldria acabar aquest treball sense explicar que els matemàtics grecs, no trobant la manera de resoldre els tres poblemes anteriors només amb regle i compàs, van buscar mètodes alternatius de fer-ho. Quins altres mètodes podien utilitzar?

Tal i com hem explicat al Capítol 2, utilitzar només un regle i un compàs vol dir fer anar només rectes i circumferències per construir nous punts a partir dels que ja es tenen. Ara bé, en principi es podria plantejar exactament el mateix problema de construcció de punts a partir d'uns de donats però utilitzant altres famílies de corbes.

Per exemple, sabem que una paràbola qualsevol queda determinada pel seu vèrtex i el seu focus. De fet, agafant coordenades de manera que l'origen és el vèrtex i el focus és el punt de coordenades $(0,b/2)$, la paràbola corresponent està determinada i és la d'equació $2by = x^2$. Podem aleshores imaginar un aparell, que es podria anomenar *compàs parabòlic*, que permetés dibuixar qualsevol paràbola coneguts el seu vèrtex i el seu focus. Podríem doncs fer construccions amb regle, compàs i compàs parabòlic, que voldrà dir construir nous punts traçant rectes, circumferències i paràboles (aquestes últimes amb el vèrtex i el focus en punts que es tenen prèviament), i es plantejaria aleshores el problema de quines construccions es poden fer amb regle, compàs i compàs parabòlic i quines no.

Bàsicament, això és el que van fer els matemàtics grecs per resoldre els tres problemes. Com que només amb rectes i circumferències no se'n sortien, van intentar-ho fer utilitzant altres tipus de corbes addicionals. En tots tres casos se'n van sortir. En realitat, van ser capaços de trobar diverses maneres de resoldre cadascun dels tres problemes utilitzant diferents tipus de corbes. En aquesta secció descrivim breument només un d'aquests mètodes en cada cas. Però se'n poden trobar d'altres en el llibre de J.C. Carrega que apareix a la Biliografia (Capítol V).

4.5.1 Duplicació del cub

El matemàtic grec Menechme (s.IV a.C.) se'n va adonar que podia utilitzar les dues paràboles $y = x^2$ i $y^2 = 2x$ per tal de resoldre el problema de la duplicació del cub. Si es busquen els punts de tall d'aquestes dues paràboles es troben dos punts: un és l'origen i l'altre és, efectivament, un punt que té per abscissa $\sqrt[3]{2}$. Per descomptat, Menechme no raonava d'aquesta manera perquè encara no s'havia introduït el mètode de les coordenades i, per tant, no sabia descriure les paràboles amb paràboles. En qualsevol cas, amb regle, compàs i això que anomenem compàs parabòlic es pot duplicar el cub seguint els passos següents.

El conjunt de punts base és $\mathcal{B} = \{(0,0),(1,0)\}$ (són els dos extrems d'un costat del cub original) i hem de veure que a partir d'aquest conjunt es poden construir les dues paràboles anteriors. Com que disposem d'un compàs parabòlic, només cal assegurar-se que es poden construir els vèrtexs i focus corresponents. Els vèrtexs ja els tenim, perquè tots dos són l'origen. Pel que fa als focus, el de la primera paràbola és el punt $F_1 = (0, 1/4)$, com es dedueix comparant la seva equació amb l'equació general $2by = x^2$ d'una paràbola qualsevol amb vèrtex a l'origen i focus al punt $(0, b/2)$. Per tant, F_1 és constructible amb regle i compàs: només cal traçar la circumferència de radi 1 centrada a l'origen i dividir en quatre parts iguals el segment de longitud 1 sobre l'eix OY que s'obté. El focus de la segona és el punt també constructible amb regla i compàs $F_2 = (1/2, 0)$. Això ho deduïm del fet que l'equació d'una paràbola de vèrtex a l'origen i focus en el punt $(b/2, 0)$ és $2bx = y^2$ (només cal intercanviar els papers de la x i la y en el cas anterior, ja que ara l'eix de simetria és l'eix OX). Per tant, les dues paràboles són constructibles amb regle, compàs i compàs parabòlic a partir de $\mathcal{B}$ i el cub es pot duplicar.

4.5.2 Trisecció de l'angle

Un altre matemàtic grec, Nicomedes (s.II a.C.), va ser capaç de resoldre el problema de la trisecció d'un angle qualsevol utilitant les corbes que avui es coneixen com a *concoides de Nicomedes*. Una concoide de Nicomedes queda determinada per un punt P (el pol de la concoide), una recta r (l'eix de la concoide) i un nombre real a. La corba es construeix traçant des del pol P semirectes que intersequen l'eix r. Si M és el punt d'intersecció d'una d'aquestes semirectes amb r, el punt M' sobre la mateixa semirecta però a distància a de M és un punt de la concoide. [1] La concoide en si és la corba que descriuen tots els punts M' quan M recorre tot l'eix r.

Doncs bé, si ens imaginem que tenim un aparell, que podríem anomenar *compàs de Nicomedes*, que ens permet dibuixar qualsevol concoide de Nicomedes a partir del pol P, l'eix r i la distància a, es pot veure que qualsevol angle es pot trisecar amb regle, compàs i compàs de Nicomedes. Per raons d'espai no descrivim com es fa. Els detalls es poden trobar al llibre de J.C. Carrega que ja hem esmentat abans (§ 2.1 del Capítol V). En aquest mateix llibre s'explica també com es poden utilitzar les concoides de Nicomedes per resoldre el problema de la duplicació un cub.

4.5.3 Quadratura del cercle

El germà de Menechme, Dinostrat, se'n va adonar que podia utilitzar una corba que havia introduït anys abans Hippies d'Elis (s. V a.C) per resoldre el problema de la trisecció de l'angle però per quadrar el cercle. La corba es coneix amb el nom de *trisectriu d'Hippies* o *quadratriu de Dinostrat* i es diu que és una corba "mecànica", perquè és la corba que descriu el punt

[1] De fet, hi haurà dos punts M_1' i M_2' a distància a de M, un a cada banda de r, així que la concoide corresponent constarà de dues parts.

d'intersecció de dues rectes que es mouen d'una determinada manera. Concretament, es construeix de la manera següent. Hem de considerar un quadrat, els vèrtexs consecutius del qual els anomenarem A, B, C, D. Tracem l'arc de circumferència de centre A que passa per B i D i considerem dues rectes mòbils. Una parteix de la posició que ocupa la recta AD i gira entorn del punt A i a velocitat constant fins a la posició que ocupa la recta AB. L'altra parteix de la posició que ocupa la recta CD i es desplaça paral·lelament a ella mateixa i a velocitat constant també fins a la posició que ocupa la recta AB. Si suposem que les dues rectes mòbils comencen a moure's justament en el mateix instant des de les posicions inicials respectives, la quadratriu de Dinostrat és la corba descrita pel punt d'intersecció de totes dues. Observar que la quadratiu queda totalment determinada per només dos punts, els vèrtexs A i B, per exemple.

Doncs bé, si ens imaginem que tenim un aparell, que podríem anomenar *compàs de Dinostrat*, que ens permet dibuixar la quadratriu de Dinostrat corresponent a dos punts donats qualssevol A i B, es pot veure que qualsevol cercle es pot quadrar amb regle, compàs i compàs de Dinostrat. També per raons d'espai no descrivim el mètode de fer-ho. Els detalls un altre cop es poden trobar al llibre de J.C. Carrega ja esmentat (§ 5.2 del Capítol V). A més, una descripció visual del mètode es pot trobar a l'adreça web següent:

http://docentes.educacion.navarra.es/msadaall/geogebra/figuras/trisectriz.htm

Capítol 5

Apèndix

En aquest Apèndix recollim les proves d'uns pocs resultats que hem fet anar al llarg del treball.

5.1 Irracionalitat de $\sqrt[3]{2}$ i de $\sqrt[3]{4}$

El raonament és semblant al que es fa per provar la irracionalitat de $\sqrt{2}$ i utilitza l'anomenat *mètode de reducció a l'absurd*. Consisteix en suposar que és racional i veure que aleshores es pot arribar a una contradicció, és a dir, una situació on resulta que al mateix temps són certes una cosa i la contrària. Com que això no pot ser, el nombre en qüestió no pot ser racional. La manera d'arribar a la contradiccó és el següent.

Suposem que $\sqrt[3]{2}$ és racional. Això vol dir que existeixen $p, q \in \mathbb{Z}$, amb $q \neq 0$, tals que

$$\sqrt[3]{2} = p/q.$$

Sempre podem suposar, si cal simplificant, que p/q és una fracció irreductible, de manera que p i q no tenen divisors comuns. Aquest és un punt clau en el raonament. Elevant al cub l'expressió anterior de $\sqrt[3]{2}$, tindrem que

$$2q^3 = p^3$$

i, per tant, deduïm que p^3 és parell. Però això implica que el propi p és parell, ja que la descomposició de p^3 com a producte de primers no és més que elevar al cub la descomposició de p i, per tant, si a la de p^3 hi apareix el 2, és que ja hi és a la de p. Posem aleshores que $p = 2k$, on $k \in \mathbb{Z}$. Substituint a la igualtat $2q^3 = p^3$ obtenim que $2q^3 = 2^3k^3$ i per tant, que $q^3 = 2^2k^3$, és a dir, q^3 també és parell. Per la mateixa raó d'abans, això implica que el propi q és parell. Per tant, p i q són tots dos divisibles per 2. Però això es contradiu amb la nostra suposició d'entrada que p/q és irreductible. En definitiva, admetre que $\sqrt[3]{2}$ és racional porta a una contradicció i, per tant, no pot ser-ho. Exactament el mateix raonament prova que tampoc és racional $\sqrt[3]{4}$.

5.2 Dos resultats de geometria elemental

5.2.1 Segon teorema de Tales

Diu que qualsevol triangle inscrit en una circumferència que tingui per un dels costats un diàmetre de la circumferència és rectangle en el vèrtex oposat al diàmetre. En efecte, sigui

ABC el triangle inscrit, amb AC un diàmetre de la circumferència. Si O és el centre de la circumferència, els triangles OAB i OBC són isòsceles, ja que tenen dos dels costats iguals al radi. Els angles $\angle OAB$ i $\angle OBA$, d'una banda, i els angles $\angle OBC$ i $\angle OCB$, de l'altra, són doncs iguals (veure Figura 5.1). Per tant, si representem per α el primer angle i per β el segon, tenim que $2\alpha + 2\beta = \pi$ i d'aquí que $\alpha + \beta = \pi/2$.

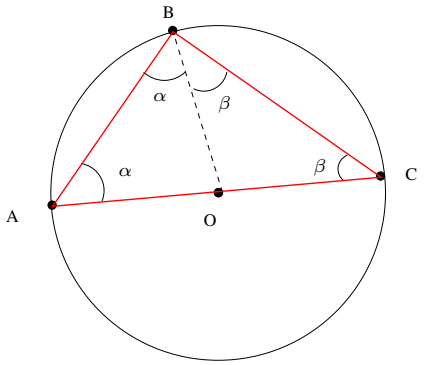


Figura 5.1: Segon teorema de Tales

5.2.2 Teorema de l'altura

Diu que si MNP és un triangle rectangle, amb l'angle recte a P, h és l'altura des de P i X el peu d'aquesta altura, aleshores $h^2 = MX \cdot XN$ (veure Figura 5.2). Es demosta ràpidament fent trigonometria. En efecte, si $\alpha = \angle XMP$ i $\beta = \angle XNP$, sabem que són angles complementaris perquè l'angle a P és recte. Això vol dir que la tangent trigonomètrica d'un és la inversa de la de l'altre. D'altra banda, $\tan\alpha = h/MX$ i $\tan\beta = h/XN$. Per tant

$$h^2 = (MX \cdot \tan\alpha) \cdot (XN \cdot \tan\beta) = MX \cdot XN \cdot \tan\alpha \cdot \frac{1}{\tan\alpha} = MX \cdot XN.$$

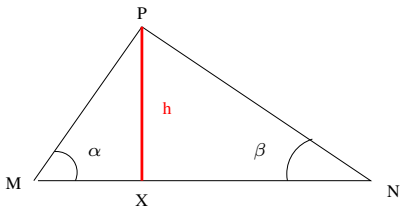


Figura 5.2: Teorema de l'altura

5.3 Identitat de Bezout

La identitat de Bezout és una propietat que ja es compleix en el cas dels enters, no només en el cas de polinomis amb coeficients en un cos. La prova és essencialment la mateixa en tots dos casos, així que pensarem en enters. L'enunciat és el següent.

Teorema 45 (Identitat de Bezout) *Siguin a, b dos enters qualssevol i sigui d el seu màxim comú divisor. Aleshores, existeixen enters x, y tals que $d = ax + by$.*

La demostració només l'expliquem per sobre. Es basa en l'anomenat **algorisme d'Euclides** per al càlcul del màxim comú divisor de dos enters a i b (o de dos polinomis $p(x)$ i $q(x)$ amb coeficients en un cos). L'algorisme simplement consisteix en fer divisions successives. Primer divideixo a entre b, i obtinc un quocient q_1 i un residu r_1, que sabem que compleix que $0 \leq r_1 < b$. A continuació, divideixo b entre el residu r_1 obtingut abans, i obtinc un nou quocient q_2 i un nou residu r_2, que complirà que $0 \leq r_2 < r_1$. Després divideixo el primer residu r_1 entre el nou residu r_2, i obtinc un quocient q_3 i un residu r_3 que complirà que $0 \leq r_3 < r_2$. Repetint el procés tan com calgui (divideixo r_2 entre r_3, r_3 entre r_4, etc), obtindré una seqüència decreixent de residus $r_1, r_2, r_3, \ldots$ tots positius i enters i, per tant, que necessàriament acabarà en el 0. Tindré doncs la successió de divisions següent:

$$\begin{array}{ll}
a = bq_1 + r_1, & 0 \leq r_1 < b \\
b = r_1q_2 + r_2, & 0 \leq r_2 < r_1 \\
r_1 = r_2q_3 + r_3, & 0 \leq r_3 < r_2 \\
r_2 = r_3q_4 + r_4, & 0 \leq r_4 < r_3 \\
\ldots & \\
r_{n-2} = r_{n-1}q_n + r_n, & 0 \leq r_r < r_{n-1} \\
r_{n-1} = r_nq_{n+1} &
\end{array}$$

Es pot aleshores provar que r_n, és a dir, l'últim residu diferent de zero, és el màxim comú divisor dels dos enters de partida a i b. De fet, el que passa és que, a mesura que anem fent divisions, en tot moment el màxim comú divisor del dividend i el divisor és igual al màxim comú divisor del divisor i el residu de la divisió. És a dir, tenim que

$$m.c.d.(a, b) = m.c.d.(b, r_1) = m.c.d.(r_1, r_2) = \cdots = m.c.d.(r_{n-1}, r_n)$$

i és evident que el màxim comú divisor de r_{n-1}i r_n és r_n, ja que r_{n-1} és un múltiple de r_n.

Què té a veure això amb la identitat de Bezout? La relació és que les successives divisions anteriors permeten trobar els enters x, y als que es refereix el teorema. Es tracta simplement de fer marxa enrere en l'algorisme. Primer aïllo de la penúltima divisió r_n, que ja sé que és el màxim comú divisor de a i b, i obtinc

$$d = r_n = r_{n-2} - r_{n-1}q_n$$

és a dir, una expressió de d com a combinació lineal de r_{n-2} i r_{n-1}. A continuació, aïllo de la divisió anterior r_{n-1} en termes de r_{n-2} i r_{n-3} i ho substitueixo a la igualtat anterior. S'obté que

$$d = r_{n-2} - (r_{n-3} - r_{n-2}q_{n-1})q_n = r_{n-2}(1 + q_{n-1}q_n) - r_{n-3}q_n$$

és a dir, una expressió de d com a combinació lineal de r_{n-3}, r_{n-2}. Fent així marxa enrere, al final s'acaba obtenint d com a combinació lineal de a i de b, que és el que volíem.

Capítol 6

Consideracions finals personals

És cert que aquest treball no conté cap resultat original. Tot el que hi explico i en realitat molt més es pot trobar en els llibres. Per exemple, he parlat molt poc dels diferents mètodes com els grecs van acabar resolent els tres problemes. Tampoc he parlat del problema de la construcció amb regle i compàs dels polígons regulars (excepte per un breu comentari a l'Exemple 25) quan, de fet, ja des de Gauss se sap quins es poden construir i quins no. Concretament, des de l'any 1801, més de 30 anys abans que Wantzel provés el seu criteri. És l'any en què Gauss va demostrar que són constructibles amb regle i compàs exactament aquells que tenen un nombre de costats n que és o bé de la forma $n = 2^m$, per algun $m \geq 2$, o bé de la forma $n = 2^m p_1 p_2 \cdots p_r$, amb $m \geq 0$ i $p_1, p_2, \ldots, p_r$ nombres primers diferents d'un tipus especial. [1] Per exemple, es poden construir amb regle i compàs els polígons de 3, 4, 5, 6, 8, 10, 12, 15, 16 i 17 costats i, en canvi, no es poden construir els de 7, 9, 12, 14, 18 i 19 costats. Al Capítol IV del llibre de J.C. Carrega (veure Bibliografia) se'n poden trobar més detalls.

Tot i així, fer aquest treball ha estat per a mi una autèntica recerca perquè m'ha obligat a endinsar-me de ple en el món de les matemàtiques, i m'ha portat a descobrir moltes coses que, encara que ja se saben des de fa molt de temps, jo desconeixia totalment.

He de dir que la meva intenció inicial no era fer un treball de matemàtiques. De fet, quan al principi em vaig plantejar la qüestió de sobre quin tema podia fer-lo, tenia clar que volia fer-lo en alguna cosa de biologia. Concretament, sobre alguna cosa que tingués a veure amb el cos humà. Després d'un temps donant-li voltes, i parlant-ne amb el meu pare, vaig pensar que un bon tema seria fer un treball sobre el funcionament del cervell, en particular, sobre com es comuniquen les neurones. La idea em va venir d'un llibret petit que vaig trobar a la biblioteca (no en recordo l'autor), i que es titula precisament així: *Como se comunican las neuronas* (publicat pel CSIC). Pel que vaig poder veure fullejant aquest llibre i d'altres, en el tema hi apareixia bastanta química i també física i matemàtiques. Per exemple, vaig veure que es parlava del *model de Hodgkin-Huxley*, un model matemàtic que van introduir als anys 50 aquests dos investigadors i que és una descripció de com s'inicien i es propaguen els anomenats "potencials d'acció" de les neurones. El tema no semblava fàcil, però semblava interessant. A més, el treball havia fet que els dos investigadors guanyessin el premi Nobel de Fisiologia i Medicina l'any 1963. Però després

[1] Han de ser el que s'anomenen *nombres primers de Fermat*, és a dir nombres primers de la forma $p = 2^{2^k} + 1$ per algun $k \geq 0$. Només per determinats valors de k aquests nombres són realment primers. Per ex., si $k = 0, 1, 2, 3, 4$ s'obtenen els nombres $3, 5, 17, 257, 65537$ que es pot comprovar que són primers. En canvi, per $k = 5$ s'obté el nombre 4294967297, i l'any 1732 el matemàtic suís Leonard Euler va provar que no és primer. És el producte $641 \cdot 6700417$.

de parlar-ne amb el seminari de Biologia em van convèncer que el tema era massa complicat i massa teòric, i que en un treball de Biologia sempre era interessant si contenia alguna part pràctica. Al final, vaig acabar abandonant la idea i va ser quan vaig començar a pensar en fer un treball d'alguna cosa de matemàtiques. Vaig aleshores parlar amb el meu pare, que em va suggerir uns quants temes, entre ells el que al final ha estat el tema d'aquest treball.

La veritat és que, en aquestes alçades, ja sé que no em dedicaré a les matemàtiques. Continua sent la biologia i, sobretot, la medicina el que més m'agrada. Però he d'admetre que l'experiència de fer aquest treball m'ha fet comprendre una mica més com és realment el món de les matemàtiques i com treballen els matemàtics. Per exemple, m'ha permès descobrir coses com ara la importància que en matemàtiques té definir els conceptes de manera precisa, el rigor amb què els matemàtics es plantegen i afronten els problemes i la idea de construir pas a pas i amb paciència tota una teoria amb la finalitat de resoldre un problema concret (que sovint només està a la ment del matemàtic). I a més, m'ha permès comprendre que realment es pot *demostrar* que hi ha problemes impossibles de resoldre.

Sigui com sigui, fer aquest treball m'ha resultat enriquidor. I tot i que no em dedicaré a les matemàtiques, sempre tindré un respecte especial per elles. No tant per la seva dificultat, sinó més aviat pel geni i la gran imaginació que he pogut veure que amaguen. De fet, el meu pare fins i tot m'ha volgut convèncer que les matemàtiques no són una simple creació de la ment humana, sinó que són l'essència de l'Univers on vivim. Però això ja és una altra història.

Bibliografia

Per fer aquest treball, m'he basat sobretot en els dos llibres següents:

- J.C. CARREGA: *Théorie des corps. La règle et le compas*, Hermann (1989).
- A. JONES, S.A. MORRIS, K.R. PEARSON: *Abstract algebra and famous impossibilities*, Universitext, Springer-Verlag (1991).

Altres llibres que en algun moment també he mirat són:

- M. CASTELLET, I. LLERENA: *Àlgebra lineal i geometria*, Manuals de la UAB (1988).
- I. STEWART: *Galois theory*, 3a. ed., Chapman Hall (2004).