UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

LAS ECUACIONES CUADRÁTICAS Y EL TEOREMA FUNDAMENTAL DEL ÁLGEBRA

T E S I S

QUE PARA OBTENER EL TÍTULO DE

MATEMÁTICA

P R E S E N T A

TETIS GISELA CAMACHO ESPINOZA

DIRECTOR DE TESIS:

DR. RODRIGO CAMBRAY NÚÑEZ

2009

1. Datos del alumno.
   Camacho
   Espinoza
   Tetis Gisela
   01 779 7 96 16 12
   Universidad: Universidad Nacional Autónoma de México
   Facultad de Ciencias
   Matemáticas
   099569714

2. Datos del tutor.
   Dr.
   Rodrigo
   Cambray
   Núñez

3. Datos del sinodal 1.
   Dr.
   Alejandro Ricardo
   Garciadiego
   Dantan

4. Datos del sinodal 2.
   Mat.
   Julio César
   Guevara
   Bravo

5. Datos del sinodal 3.
   Mat.
   Enrique
   Vega
   Ramírez

6. Datos del sinodal 4.
   M. en C.
   William José
   Gallardo

7. Datos del trabajo escrito.
   Las ecuaciones cuadráticas y el teorema fundamental del álgebra
   179 p.
   2009

AGRADECIMIENTOS

A mis padres y a mi hermana por el apoyo moral y económico brindado durante mis estudios y generosamente extendido hasta la conclusión de este proyecto.

A mi director de tesis, Dr. Rodrigo Cambray Núñez, por el valioso tiempo dedicado a la revisión de este trabajo.

A los sinodales Dr. Alejandro Garciadiego Dantan, Mat. Julio César Guevara Bravo, Mat. Enrique Vega Ramírez y M.en C. William José Gallardo por su aval para la aprobación de este escrito.

TABLA DE CONTENIDOS

INTRODUCCIÓN

Cuando se estudia la resolución de ecuaciones de segundo grado, debiera ser natural preguntarse por qué estas ecuaciones tienen dos soluciones. Las respuestas que recibimos se relacionan con el llamado *discriminante* de la ecuación cuadrática, reduciéndose la comprensión de lo que sucede al manejo de un algoritmo del que comúnmente se desconoce su procedencia. En la educación escolar, desde la escuela secundaria ese algoritmo se reduce a la aplicación de una fórmula general de resolución: Dada la ecuación $ax^2 + bx + c = 0$, con $a \neq 0$, sus soluciones son

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}\ .$$

En el bachillerato y posteriormente en los cursos superiores, las respuestas a por qué una ecuación cuadrática tiene dos soluciones se complican aún más: las soluciones tienen razón de ser gracias al teorema fundamental del álgebra.

En la carrera de matemáticas de distintas universidades los temas relacionados con el teorema fundamental del álgebra se incluyen en los primeros cursos de álgebra superior (álgebra superior II en la Facultad de Ciencias de la UNAM) y posteriormente en un primer curso de análisis de variable compleja. Particularmente, de acuerdo con el "temario" de álgebra superior II, se pretende que el alumno conozca y maneje el teorema fundamental del álgebra así como sus consecuencias, pero no se proporciona demostración alguna del mismo; se aborda nuevamente el teorema en el apartado "Teoremas de Liouville y fundamental del álgebra" en la asignatura de variable compleja I, reduciéndose el teorema fundamental a la calidad de colorario de otro teorema que no es fundamental.

Bajo esta perspectiva, el propósito central de este trabajo es mostrar la primera demostración válida del teorema cuyo carácter geométrico proporciona las ideas subyacentes de uno de los resultados más importantes de las matemáticas, que los autores modernos han dejado de lado en beneficio de demostraciones formales y artificiosas. Como punto de partida para el estudio del teorema se analizará por qué las ecuaciones de segundo grado tienen dos raíces, lo cual se aborda en el capítulo I. A su vez, esto motivará preguntas como las siguientes:

–¿Cómo se entendían y resolvían en el pasado los problemas que actualmente se representan con ecuaciones?

–¿Qué relación guardaban la práctica de resolución de esos problemas y el entorno social en el que surgieron?

La primera interrogante es muy ambiciosa: una respuesta detallada y profunda requeriría una investigación exhaustiva, lo cual no es la finalidad de este trabajo. Sin embargo, se considera de interés mostrar un panorama general sobre cómo los árabes y algunos otros estudiosos del medievo enfrentaron estas cuestiones, ya que la difusión de sus ideas cimentó los estudios de las matemáticas en el continente europeo hasta mediados del siglo XVI aproximadamente.

La segunda pregunta obedece a que las matemáticas son una creación del hombre influida por lo que le rodea. Estas dos interrogantes se tratan en el capítulo II, abarcando desde el surgimiento de una de las principales obras de las matemáticas árabes, el *Al-jabr wa'l-muqabala*, hasta las aportaciones de los estudiosos de finales de la época medieval.

Durante mucho tiempo las técnicas de resolución de problemas considerados actualmente algebraicos estuvieron ligadas a la percepción. Sin embargo, el interés creciente en "ecuaciones" de grado mayor que 3 suscitó una separación del dominio de lo visible, tendiendo al empleo de algoritmos aritméticos. Estas cuestiones se presentan en el capítulo III, donde también se habla del surgimiento de un "simbolismo" más apropiado que, sin duda, benefició el estudio de la teoría de ecuaciones.

Finalmente, con el surgimiento de este nuevo simbolismo, problemas relacionados con el análisis geométrico y aritmético de la antigua Grecia se vieron desde una nueva perspectiva surgiendo así la geometría analítica, lo cual se muestra en el capítulo IV. El lenguaje de la nueva geometría fue esencial para la demostración del teorema fundamental del álgebra, que se expone en el capítulo V, recurriendo a la acción generadora de las magnitudes complejas de Carl Friedrich Gauss.

Los temas tratados en cada uno de los capítulos muestran el camino de la resolución de problemas relacionados con ecuaciones cuadráticas por medio de un lenguaje retórico, avanzando hacia problemas cada vez más complicados y llegando a ecuaciones polinómicas de mayor grado que en algún tiempo no tendrían referente físico. El desarrollo del simbolismo y la conceptuación de los objetos matemáticos tratados en cada época son de importancia relevante, pues fue finalmente la interpretación que Gauss hizo de aquellos conceptos lo que originó los cimientos de su demostración: la clave fue su vuelta hacia el pasado y su reinterpretación.

CAPÍTULO I

¿POR QUÉ LAS ECUACIONES DE SEGUNDO GRADO TIENEN DOS RAÍCES?

Para responder a la pregunta de por qué las ecuaciones de segundo grado tienen dos soluciones, tenemos que comenzar con otra interrogante: ¿Qué significado geométrico tiene elevar al cuadrado una magnitud? En las matemáticas griegas del siglo V a.C., en la escuela de Platón (h. 427 - h. 348 o 347 a.C.), encontramos en uno de los *Diálogos* un problema en el que está implicado el cuadrado de una magnitud: la duplicación del área de un cuadrado.

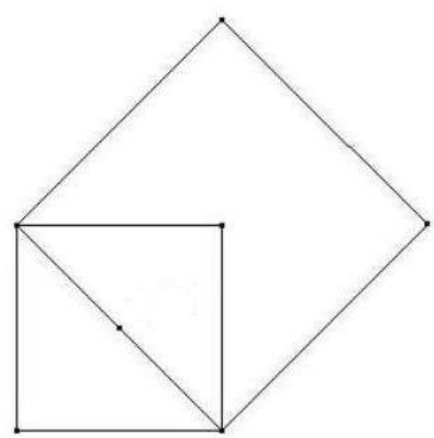

Figura 1.1. Duplicación del área de un cuadrado

En la conversación entre Sócrates, Menón y un esclavo de éste, Sócrates mostró al esclavo un cuadrado cuyo lado medía 2 pies, es decir, un cuadrado con área de 4 pies, y después le preguntó: "¿No podría formarse un espacio doble que éste y del todo semejante, teniendo como él todas sus líneas iguales?" (Platón 2007, p. 301)

Tras conducir al esclavo por el camino de las reminiscencias, Sócrates logró que *recordara* que para duplicar el área se tiene que construir el nuevo cuadrado sobre la diagonal del original, como se muestra en la figura 1.1. La principal dificultad que enfrentó el esclavo para resolver el problema planteado por Sócrates fue su estrecho panorama amarrado a la aritmética de los números enteros. De este modo, en el espacio se mostraron longitudes que no podrían determinarse mediante cuantificación. Sócrates le dijo al esclavo, refiriéndose a la línea que proporciona una superficie de 8 pies: "¿Con qué línea se forma? Procura decírnoslo exactamente, y si no quieres calcularla, muéstranosla" (Platón 2007, p. 303). Así, Sócrates pasó de los números al espacio: "la prueba de Sócrates demuestra que el espacio hace posible lo que los números hacen imposible" (Serrés 1996, p. 234).

Con la idea anterior, consideremos un cuadrado de área 1 y construyamos otro cuya área sea el doble. Si continuamos el proceso de duplicar las áreas de los cuadrados que sucesivamente se van obteniendo, se tendrá un esquema como el de la figura 1.2. Observemos cómo crecen las longitudes de los lados de los cuadrados dibujados. Notemos que en la ilustración aparecen las longitudes 2, 4 y 16, es decir, 2, 2^2 y $(2^2)^2$. El segmento OB de longitud 2 forma un ángulo recto con el segmento OA girando éste en sentido dextrógiro. Para pasar del segmento de longitud 2 al segmento OC, de longitud $2^2 = 4$, ocurre una rotación de un ángulo recto y una extensión de 2 a $2^2 = 4$. Así, el segmento OC de longitud 4 forma con el segmento horizontal OA un ángulo equivalente a dos ángulos rectos.

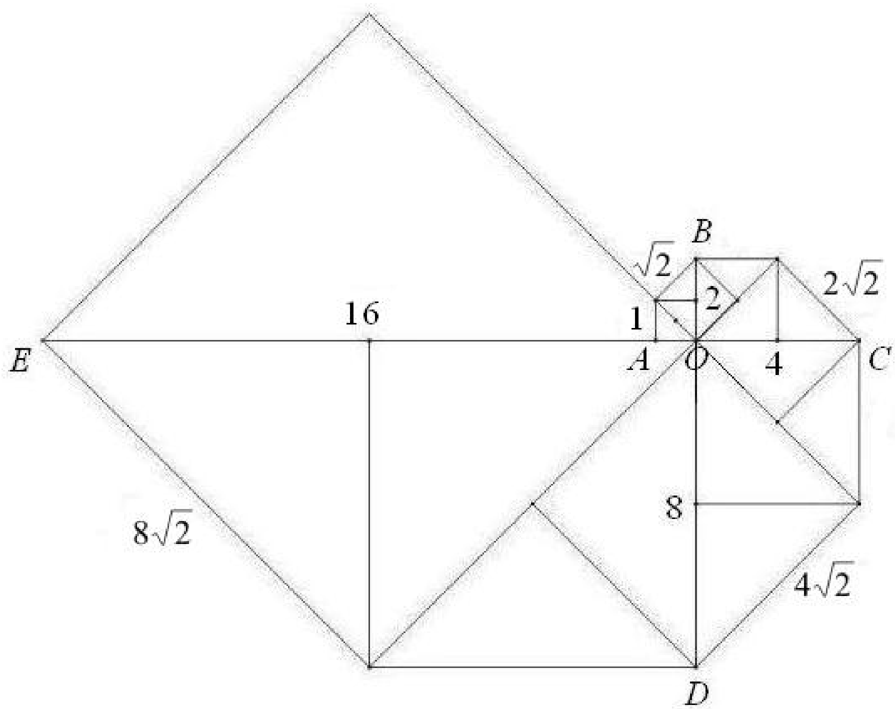

Figura 1.2. Duplicación sucesiva de áreas de cuadrados

Para obtener el segmento de longitud $(2^2)^2 = 16$ se requiere de una rotación de 180° del segmento de longitud $2^2 = 4$ y de la extensión de $4^2 = 16$. De esta manera observamos que la acción de elevar al cuadrado una magnitud se genera de las acciones geométricas de rotar y extender. En este proceso aparecen magnitudes inconmensurables; éstas sólo tenían sentido a través de la geometría, como lo muestra el diálogo entre Sócrates y el esclavo de Menón.

A lo largo de la historia surgió la necesidad de resolver problemas relacionados con lo que actualmente llamamos *ecuaciones de segundo grado*. Los árabes, por ejemplo solucionaron esta clase de problemas mediante la representación geométrica de los 6 tipos de "cuadráticas" que trabajaron. Actualmente englobamos todas esas "ecuaciones" bajo un solo modelo: $ax^2 + bx + c = 0$. Los árabes trabajaron diferentes tipos porque aún no conceptuaban los números negativos.

Los procedimientos de los árabes para resolver problemas relacionados con ecuaciones cuadráticas, interpretados con el simbolismo de nuestro tiempo conducen a las soluciones $x = \frac{-b+\sqrt{b^2-4ac}}{2a}$ y $x = \frac{-b-\sqrt{b^2-4ac}}{2a}$, considerando que $b^2 - 4ac$ no podía ser negativo; de resultar así, tal "solución" se descartaba por carecer de sentido. Una observación digna de mencionar sobre los antecedentes de las matemáticas árabes la hace Katz, al afirmar que éstas provienen de las matemáticas babilonias (Katz 1993, p. 229) y no de Euclides, como se reporta en muchos libros de historia de las matemáticas.

A pesar de que durante mucho tiempo se omitieron las cantidades negativas en los resultados obtenidos al resolver ecuaciones cuadráticas, estas entidades así como los inconmesurables despertaron un interés aritmético en los estudiosos. Es por esto que desde la época árabe algunos de sus sabios surgidos en los siglos XI y XII manipularon cantidades negativas por medio de algoritmos (Katz 1993, p. 236); de igual forma ocurrió con los inconmensurables en una tradición que avanzó hasta bien entrado el siglo XVI, cuando Rafael Bombelli (1526 - 1572) operó con unas extrañas cantidades incomprensibles, aunque muy útiles para resolver ecuaciones: las raíces cuadradas de números negativos.

Un inconveniente que se presentó para el cabal entendimiento de las soluciones de las ecuaciones cuadráticas, así como para la resolución de muchos problemas de la antigüedad, fue el simbolismo. François Viète (1540 - 1603) retomó los conocimientos de los antiguos griegos y los adaptó a las necesidades de su tiempo, creando un nuevo simbolismo matemático. Después Pierre de Fermat (1601 - 1665) y René Descartes (1596 - 1650) crearon la geometría analítica y mejoraron el simbolismo de Viète. Surgió así el lenguaje apropiado que nos permitirá responder las preguntas planteadas al inicio de este capítulo.

Mediante el lenguaje de la geometría analítica y con el significado explicado previamente de elevar al cuadrado una magnitud, partiendo del razonamiento de duplicar un cuadrado, consideremos un segmento de recta ubicado en el plano cartesiano, con uno de sus extremos en el origen y con determinada inclinación con respecto al eje horizontal. Este segmento representará lo que llamaremos una magnitud compleja; para elevarla al cuadrado se necesitará duplicar su ángulo de inclinación y extender su longitud al cuadrado. De este modo, para obtener la raíz cuadrada de una magnitud se sigue un proceso inverso: el ángulo que forma el segmento de recta con el eje de referencia se reduce a la mitad y su longitud se disminuye numéricamente a su raíz cuadrada.

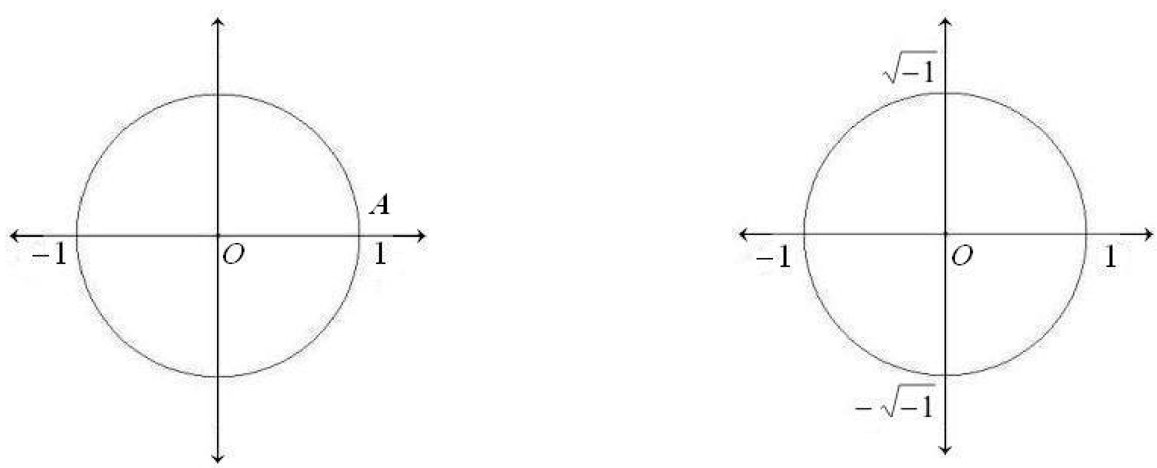

Figura 1.3. Significado geométrico de $\sqrt{-1}$

Ubiquemos entonces las magnitudes 1 y -1 en el plano, caracterizadas por los números reales 1 y -1 respectivamente. Ahora estamos en condiciones de comprender el significado geométrico de $\sqrt{-1}$. El punto señalado como -1 se localiza a 180° en sentido levógiro respecto al eje horizontal $\overrightarrow{OA}$ y a una distancia 1 del origen O. Así, su raíz cuadrada $\sqrt{-1}$, se ubicará a 90° respecto al eje $\overrightarrow{OA}$ y a una distancia de $\sqrt{1} = 1$ del origen (véase la figura 1.3). También podemos ver que el ángulo de 180° es el mismo ángulo que $180° + 360° = 540°$. De este modo, otra raíz cuadrada de -1 se ubicará a $540° \div 2 = 270°$ y a una distancia 1 del origen. Si agregamos 360° más a 540° y extraemos la raíz cuadrada, obtendremos nuevamente la primera raíz. De acuerdo con el esquema cartesiano adoptado para describir este proceso, hemos encontrado $\sqrt{-1}$ y $-\sqrt{-1}$.

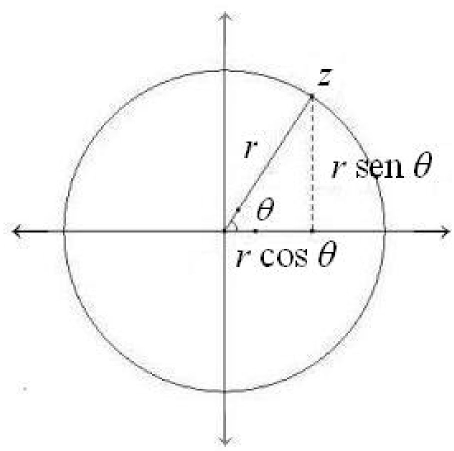

Figura 1.4. Una magnitud compleja z

Luego, hemos comprendido el significado de $\sqrt{-1}$ a través de la acción que genera tal magnitud. Hay que aclarar que aunque se toma prestado el lenguaje de la geometría euclidiana, el principio para elevar al cuadrado no proviene de ella. Al igual que $\sqrt{-1}$ y

$-\sqrt{-1}$, podemos determinar cualquier magnitud en el plano cartesiano por medio de su longitud r y de su ángulo de inclinación θ. Más específicamente, cada magnitud z del plano se distingue de otra por sus coordenadas $(r\cos\theta, r\operatorname{sen}\theta)$, con $0 \leq \theta < 2\pi$ (véase la figura 1.4.)

Consideremos entonces la magnitud compleja z de la figura 1.5, cuyo ángulo de inclinación respecto al eje horizontal es φ y cuya longitud es r. Si elevamos al cuadrado dicha magnitud, se tendrá que z^2 tiene como ángulo de inclinación 2φ y como longitud r^2; similarmente, $(z^2)^2$ tendrá como ángulo de inclinación 4φ y como longitud $(r^2)^2 = r^4$.

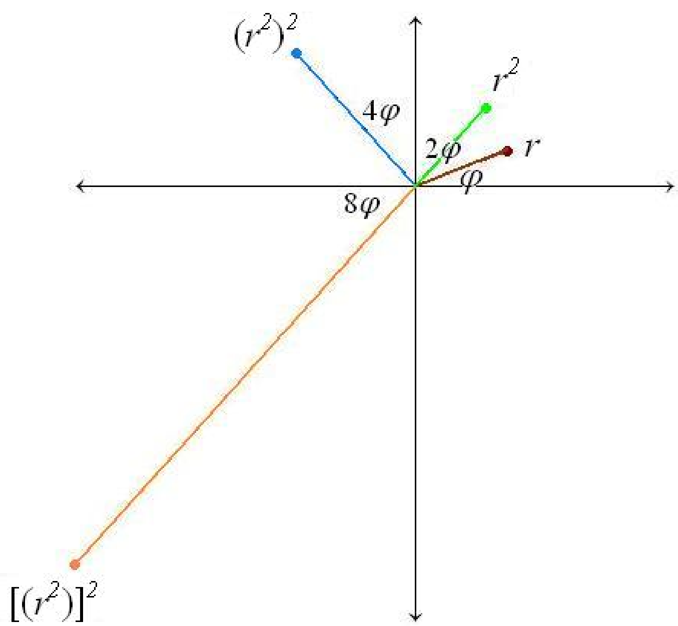

Figura 1.5. Elevación al cuadrado una magnitud

Trabajemos ahora con la ecuación cuadrática $X = x^2 + Ax + M = 0$. Hemos observado que los números complejos pueden descomponerse en términos de su longitud, el coseno de su ángulo de inclinación y el seno de su ángulo de inclinación. Es decir, el número complejo z está representado en el plano por el punto $(r\cos\varphi, r\operatorname{sen}\varphi)$. De este modo las coordenadas de z^2 son $(r^2\cos 2\varphi, r^2\operatorname{sen} 2\varphi)$, ya que sabemos que elevar al cuadrado una magnitud de este tipo implica duplicar su ángulo de inclinación y extender su longitud al cuadrado.

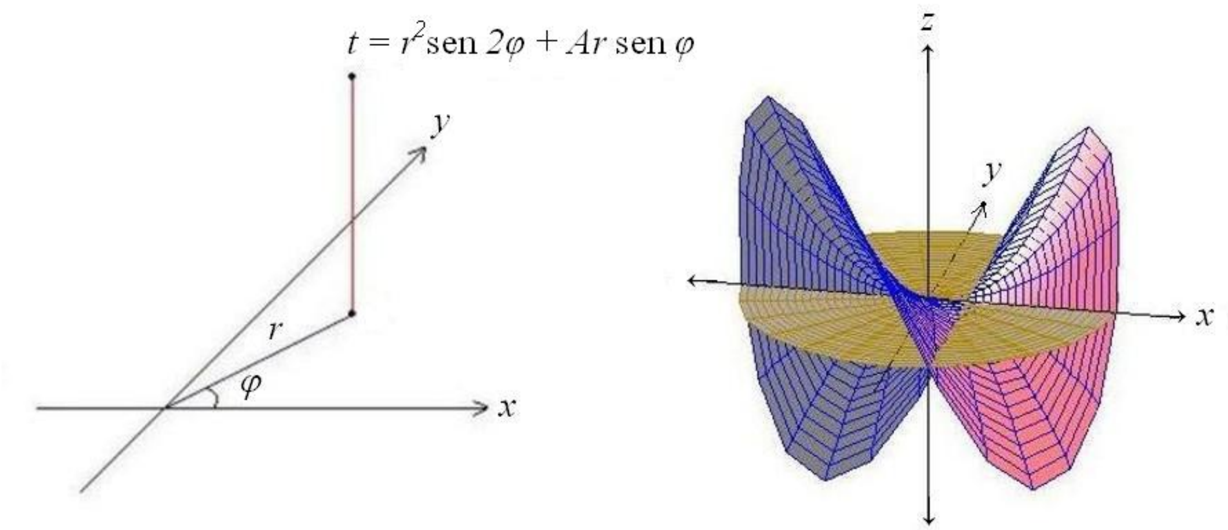

Figura 1.6. Construcción de la superficie $T = r^2\operatorname{sen} 2\varphi + Ar\operatorname{sen}\varphi$

Apoyándonos en esta descomposición de los números complejos, asociemos a la expresión $X = x^2 + Ax + M$ dos superficies. La primera superficie será

$$T = r^2\operatorname{sen} 2\varphi + Ar\operatorname{sen}\varphi + M(r^0\operatorname{sen} 0\varphi) \quad \text{o} \quad T = r^2\operatorname{sen} 2\varphi + Ar\operatorname{sen}\varphi,$$

es decir, sobre cada punto del plano levantaremos una perpendicular cuya longitud será $r^2\operatorname{sen} 2\varphi + Ar\operatorname{sen}\varphi$, como se muestra en la figura 1.6, donde precisamente r es la distancia del origen al punto en el plano y φ es el ángulo de inclinación de la recta que une al punto con el origen respecto al eje de referencia. La segunda superficie se muestra en la figura 1.7 y está representada por la expresión

$$U = r^2 \cos 2\varphi + Ar\cos\varphi + M(r^0 \cos 0\varphi) \quad o \quad U = r^2 \cos 2\varphi + Ar\cos\varphi + M.$$

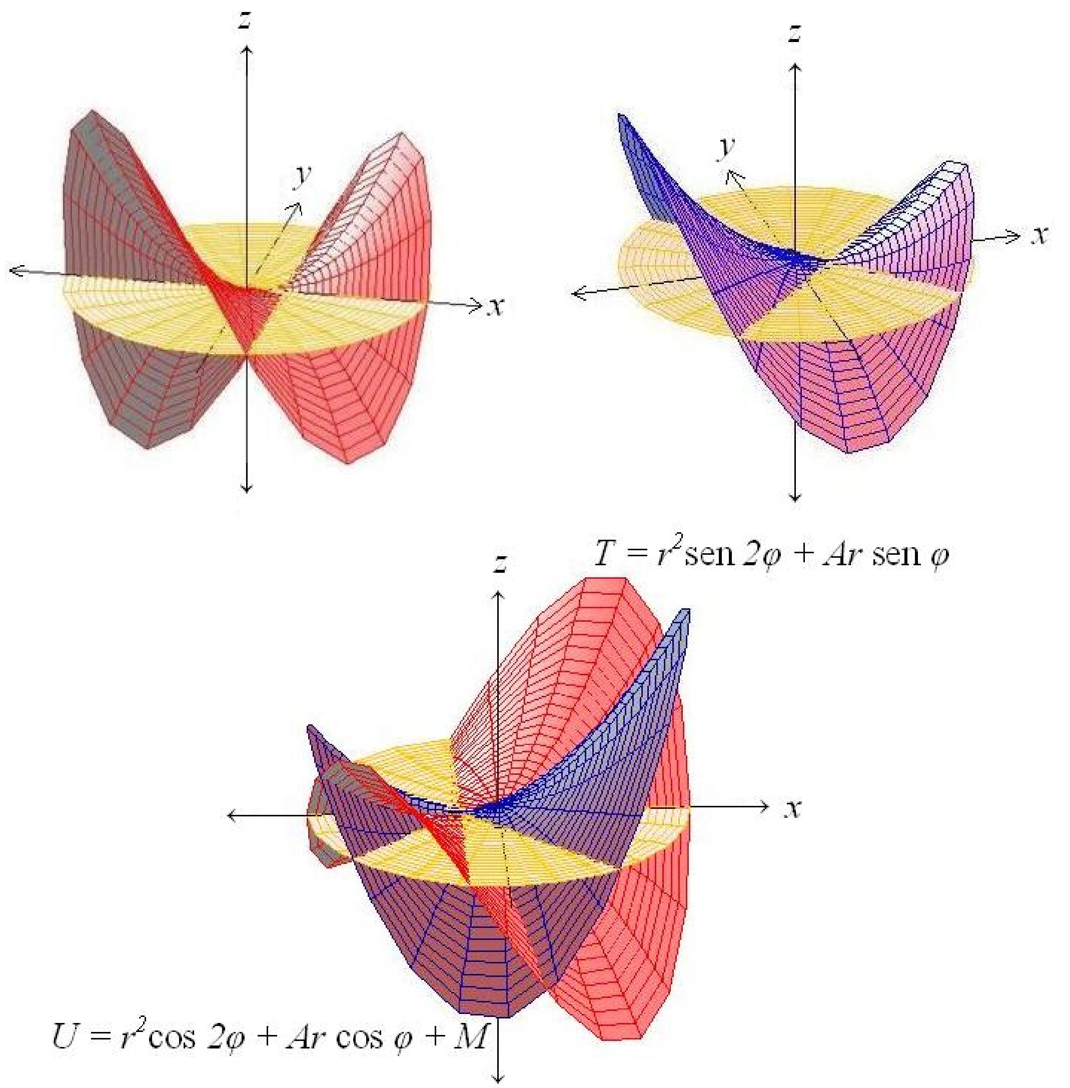

Figura 1.7. Superficies $T = r^2 \operatorname{sen} 2\varphi + Ar \operatorname{sen} \varphi$, $\;U = r^2 \cos 2\varphi + Ar\cos\varphi + M$

y su intersección

Para encontrar las soluciones de la ecuación $X = x^2 + Ax + M = 0$, bastará con mostrar que existe un ángulo φ y una longitud r que satisfacen

$$T = r^2 \operatorname{sen} 2\varphi + Ar \operatorname{sen}\varphi = 0 \quad \text{y} \quad U = r^2 \cos 2\varphi + Ar\cos\varphi + M = 0$$

simultáneamente (véase la figura 1.7). En otras palabras, ubicaremos las intersecciones de cada superficie con el plano y las intersecciones de esas curvas entre sí. De este modo esas intersecciones representarán las soluciones buscadas.

Analicemos primero la superficie $T = r^2 \operatorname{sen} 2\varphi + Ar \operatorname{sen}\varphi$. Si r es infinito, el término $Ar \operatorname{sen}\varphi$ no aporta nada a la expresión comparado con $r^2 \operatorname{sen} 2\varphi$; así, la superficie $T = r^2 \operatorname{sen} 2\varphi + Ar \operatorname{sen}\varphi$ coincidirá con la superficie $S_1(r, \varphi) = r^2 \operatorname{sen} 2\varphi$. Puesto que $F_1(\varphi) = \operatorname{sen} 2\varphi$ es una función continua que varía de 1 a -1, la expresión $T = r^2 \operatorname{sen} 2\varphi$ será algunas veces positiva y otras negativa. En el intervalo $0 \leq \varphi < 2\pi$ la función $F_2(\varphi) = \operatorname{sen}\varphi$ cambia de positiva a negativa una vez, es decir, en el intervalo $0 < \varphi < \pi$ se tiene que $\operatorname{sen}\varphi$ es positiva, y en el intervalo $\pi < \varphi < 2\pi$ se tiene que $\operatorname{sen}\varphi$ es negativa. Observemos entonces que la función $F_1(\varphi) = \operatorname{sen} 2\varphi$ cambiará de signo tres veces: cuando $0 < \varphi < \pi$ se tendrá que $0 < 2\varphi < 2\pi$, y cuando $\pi < \varphi < 2\pi$ tendremos que $2\pi < 2\varphi < 4\pi$; de este modo, si evaluamos la función $F_1(\varphi) = \operatorname{sen} 2\varphi$ en $0 \leq \varphi < 2\pi$, el ángulo φ recorrerá la circunferencia dos veces y obtendremos que la función se anulará 4 veces. Observemos que $T = r^2 \operatorname{sen} 2\varphi + Ar \operatorname{sen}\varphi$ se encontrará sobre el plano cuando $0 < \varphi < \frac{\pi}{2}$ o $\pi < \varphi < \frac{3\pi}{2}$, y se encontrará bajo el plano en los intervalos $\frac{\pi}{2} < \varphi < \pi$ y $\frac{3\pi}{2} < \varphi < 2\pi$.

Por otra parte, podemos escribir

$$T = r^2 \operatorname{sen} 2\varphi + Ar \operatorname{sen}\varphi = 0 \quad \text{como} \quad T = r^2(\operatorname{sen} 2\varphi + \frac{A}{r}\operatorname{sen}\varphi) = 0,$$

y a su vez como $T = \operatorname{sen} 2\varphi + \frac{A}{r}\operatorname{sen}\varphi = 0$. Si r es infinito, la curva $T = \operatorname{sen} 2\varphi + \frac{A}{r}\operatorname{sen}\varphi = 0$ coincidirá con la curva $T = \operatorname{sen} 2\varphi = 0$. Esta ecuación se cumple siempre que $2\varphi = k\pi$, con k en los enteros; es decir, si $\varphi = k\frac{\pi}{2}$. Esto es, $\operatorname{sen} 2\varphi = 0$ cuando $\varphi \in \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$ (el resto de los ángulos posibles coincide con alguno de los de este conjunto). De este modo, la curva $\operatorname{sen} 2\varphi = 0$ está formada por 4 semirrectas cuyos ángulos de inclinación son $\varphi_1 = 0,\ \varphi_2 = \frac{\pi}{2},\ \varphi_3 = \pi,\ \varphi_4 = \frac{3\pi}{2}$ (véase la figura 1.8). Las semirrectas dividen al plano en 4 regiones iguales. Tenemos de este modo que la curva tiene 4 ramas.

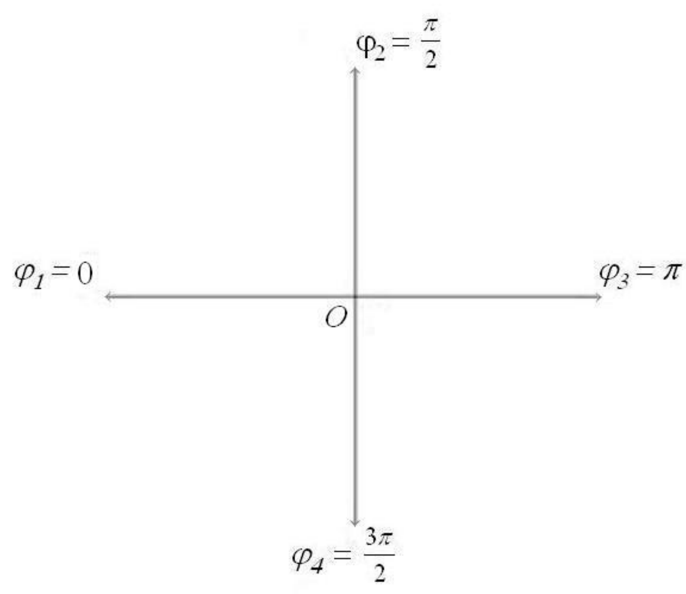

Figura 1.8. Curva $r^2 \operatorname{sen} 2\varphi + Ar \operatorname{sen}\varphi = 0$ cuando r es infinito

Retomemos la expresión $T = r^2 \operatorname{sen} 2\varphi + Ar \operatorname{sen}\varphi$, que puede escribirse como $T = r^2\left(\operatorname{sen} 2\varphi + \frac{A}{r}\operatorname{sen}\varphi\right)$. Procedamos a examinar qué ocurre cuando r es finito. Consideremos una circunferencia de radio r sobre el plano; las condiciones que debe cumplir r se obtendrán en el proceso del análisis. Veamos para qué puntos de la circunferencia la superficie T tiene valores positivos y para cuáles la superficie tiene valores negativos. Para ello dividiremos a la circunferencia: asignemos (1) al punto que se ubica sobre ésta a $\frac{1}{2}(\frac{\pi}{4})$ respecto a $\overrightarrow{OB}$ en sentido dextrógiro (véase la figura 1.9). Etiquetemos con (3) al punto que se encuentra a $\frac{3}{2}(\frac{\pi}{4})$, con (5) al punto que está a $\frac{5}{2}(\frac{\pi}{4})$ y así sucesivamente hasta (15), que se encuentra a $\frac{15}{2}(\frac{\pi}{4})$.

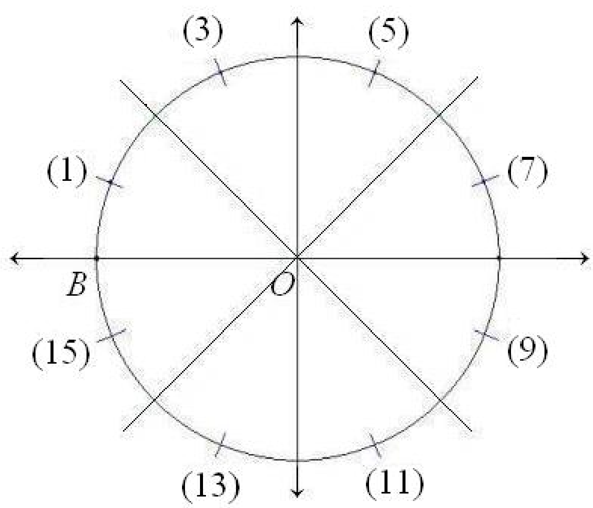

Figura 1.9. Divisiones de una circunferencia de radio r

En el punto (1), en sentido dextrógiro a partir de $\overrightarrow{OB}$, $\varphi = \frac{1}{2}(\frac{\pi}{4}) = \frac{\pi}{8}$; así que $2\varphi = \frac{\pi}{4}$. Sustituyendo el valor de 2φ en la expresión $T = r^2\left(\operatorname{sen} 2\varphi + \frac{A}{r}\operatorname{sen}\varphi\right)$ se tiene

$$T = r^2\left(\operatorname{sen}\frac{\pi}{4} + \frac{A}{r}\operatorname{sen}\varphi\right) = r^2\left(\sqrt{\frac{1}{2}} + \frac{A}{r}\operatorname{sen}\varphi\right).$$

Para que el valor de T sea positivo se necesita que $\left|\frac{A}{r}\operatorname{sen}\varphi\right| < \sqrt{\frac{1}{2}}$. Notemos que

$$\left|\frac{A}{r}\operatorname{sen}\varphi\right| = \left|\frac{A}{r}\right| |\operatorname{sen}\varphi| \leq \left|\frac{A}{r}\right| = |A|\left(\frac{1}{r}\right) < |A|\left(\frac{1}{r-1}\right).$$

La última desigualdad se cumple si $r > 1$, pues de este modo $r > r-1$ y por tanto $\frac{1}{r} < \frac{1}{r-1}$. Luego, si $|A|\left(\frac{1}{r-1}\right) < \sqrt{\frac{1}{2}}$, se tendrá que $\left|\frac{A}{r}\operatorname{sen}\varphi\right| < \sqrt{\frac{1}{2}}$. Así, si $|A|\left(\frac{1}{r-1}\right) < \sqrt{\frac{1}{2}}$ entonces $r > |A|\sqrt{2}+1$. De todo esto se concluye que $\left|\frac{A}{r}\operatorname{sen}\varphi\right| < \sqrt{\frac{1}{2}}$ siempre que el círculo tenga un radio $r > |A|\sqrt{2}+1$. Es decir, la función $T = r^2\left(\operatorname{sen}\frac{\pi}{4} + \frac{A}{r}\operatorname{sen}\varphi\right)$ conservará el mismo signo que $\operatorname{sen}\frac{\pi}{4}$. Consecuentemente, para todos los ángulos φ para los cuales $|\operatorname{sen} 2\varphi| > \sqrt{\frac{1}{2}}$ se tendrá que $T = r^2\left(\operatorname{sen} 2\varphi + \frac{A}{r}\operatorname{sen}\varphi\right)$ conservará el mismo signo que $\operatorname{sen} 2\varphi$.

Regresemos a la circunferencia de radio r. Se ha determinado que $r > |A|\sqrt{2}+1$. Analicemos el valor de T de acuerdo con los valores de 2φ en los puntos que marcamos:

en (1), $\varphi = \frac{1}{2}\left(\frac{\pi}{4}\right) \Rightarrow 2\varphi = \frac{\pi}{4}$ y $\operatorname{sen}\frac{\pi}{4} = \sqrt{\frac{1}{2}}$;

en (3), $\varphi = \frac{3}{2}\left(\frac{\pi}{4}\right) \Rightarrow 2\varphi = \frac{3\pi}{4}$ y $\operatorname{sen}\frac{3\pi}{4} = \sqrt{\frac{1}{2}}$;

en (5), $\varphi = \frac{5}{2}\left(\frac{\pi}{4}\right) \Rightarrow 2\varphi = \frac{5\pi}{4}$ y $\operatorname{sen}\frac{5\pi}{4} = -\sqrt{\frac{1}{2}}$;

en (7), $\varphi = \frac{7}{2}\left(\frac{\pi}{4}\right) \Rightarrow 2\varphi = \frac{7\pi}{4}$ y $\operatorname{sen}\frac{7\pi}{4} = -\sqrt{\frac{1}{2}}$;

en (9), $\varphi = \frac{9}{2}\left(\frac{\pi}{4}\right) \Rightarrow 2\varphi = \frac{9\pi}{4}$ y $\operatorname{sen}\frac{9\pi}{4} = \sqrt{\frac{1}{2}}$;

en (11), $\varphi = \frac{11}{2}\left(\frac{\pi}{4}\right) \Rightarrow 2\varphi = \frac{11\pi}{4}$ y $\operatorname{sen}\frac{11\pi}{4} = \sqrt{\frac{1}{2}}$;

en (13), $\varphi = \frac{13}{2}\left(\frac{\pi}{4}\right) \Rightarrow 2\varphi = \frac{13\pi}{4}$ y $\operatorname{sen}\frac{13\pi}{4} = -\sqrt{\frac{1}{2}}$; y

en (15), $\varphi = \frac{15}{2}\left(\frac{\pi}{4}\right) \Rightarrow 2\varphi = \frac{15\pi}{4}$ y $\operatorname{sen}\frac{15\pi}{4} = -\sqrt{\frac{1}{2}}$.

Podemos deducir entonces que T tendrá un valor positivo cuando 2φ caiga entre (1) y (3) y entre (9) y (11); T tendrá un valor negativo cuando 2φ caiga entre (5) y (7) y entre (13) y (15), como se muestra en la figura 1.10; la superficie T nunca tendrá como valor 0 en cualquiera de estos intervalos, pues $\operatorname{sen} 2\varphi$ nunca es nulo en ellos.

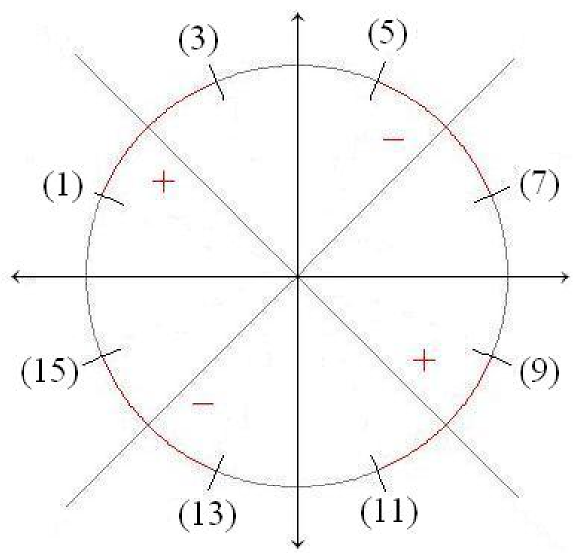

Figura 1.10. Puntos de la circunferencia donde T es positiva o negativa

Como el valor de T es positivo en (3) y negativo en (5), y puesto que la superficie es continua, se sigue que en algún punto entre (3) y (5) T será igual a 0 (véase la figura 1.11). Lo mismo ocurrirá para algún punto entre (7) y (9), para otro entre (11) y (13), y uno más entre (15) y (1). Veamos que $T = 0$ en únicamente un punto de cada intervalo señalado. Para ello, notemos que si existieran dos puntos con tal característica forzosamente existiría un máximo o un mínimo de la función en dichos intervalos.

Obtengamos la derivada de $T = r^2 \operatorname{sen} 2\varphi + Ar \operatorname{sen} \varphi$. Se tiene que

$$\frac{dT}{d\varphi} = 2r^2 \cos 2\varphi + Ar \cos \varphi.$$

Podemos escribir esta expresión como $\frac{dT}{d\varphi} = 2r^2\left(\cos 2\varphi + \frac{A}{2r}\cos\varphi\right)$. Observemos que en el punto etiquetado con (3), $\varphi = \frac{3}{2}\left(\frac{\pi}{4}\right)$; es decir, $2\varphi = \frac{3\pi}{4}$ y $\cos\frac{3\pi}{4} = -\sqrt{\frac{1}{2}}$. Para que $\frac{dT}{d\varphi} = 2r^2\left(-\sqrt{\frac{1}{2}} + \frac{A}{2r}\cos\varphi\right)$ conserve el mismo signo que $\cos 2\varphi = -\sqrt{\frac{1}{2}}$ se necesita que $\left|\frac{A}{2r}\cos\varphi\right| < \sqrt{\frac{1}{2}}$. Pero

$$\left|\frac{A}{2r}\cos\varphi\right| = \left|\frac{A}{2r}\right| |\cos\varphi| \leq \left|\frac{A}{2r}\right| = \frac{1}{2}\left|\frac{A}{r}\right| < \left|\frac{A}{r}\right| = |A|\left(\frac{1}{r}\right) < |A|\left(\frac{1}{r-1}\right).$$

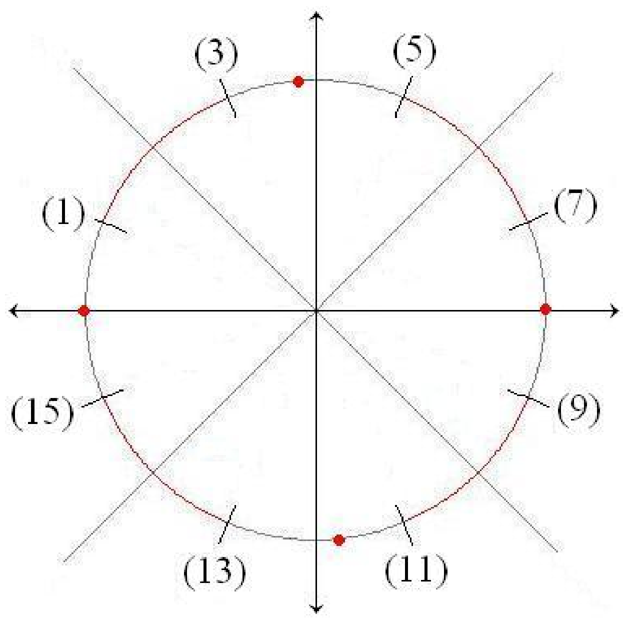

Figura 1.11. Puntos de la circunferencia donde $T = 0$

La última desigualdad es cierta siempre que $r > 1$. De este modo, si $|A|\left(\frac{1}{r-1}\right) < \sqrt{\frac{1}{2}}$ se tendrá que $\left|\frac{A}{2r}\cos\varphi\right| < \sqrt{\frac{1}{2}}$; luego, si esto ocurre, $r > |A|\sqrt{2} + 1$, que es la misma condición que se obtuvo para que la superficie $T = \left(\operatorname{sen} 2\varphi + \frac{A}{r}\operatorname{sen}\varphi\right)$ conservara el mismo signo que $\operatorname{sen} 2\varphi$. De este modo, sobre la misma circunferencia de radio

$r > |A|\sqrt{2}+1$ la derivada $\frac{dT}{d\varphi} = 2r^2\left(\cos 2\varphi + \frac{A}{2r}\cos\varphi\right)$ tendrá el mismo signo que $\cos 2\varphi$ para todos los ángulos φ tales que $|\cos 2\varphi| > \sqrt{\frac{1}{2}}$.

Así, $\frac{dT}{d\varphi}$ es negativa cuando 2φ cae entre (3) y (5) y entre (11) y (13). La derivada es positiva cuando 2φ cae entre (7) y (9) y entre (15) y (1). Es decir, $\frac{dT}{d\varphi}$ no es nula en ningún punto de esos intervalos; consecuentemente, en ellos no hay máximos ni mínimos. Por lo tanto, $T = 0$ en únicamente un punto de cada intervalo.

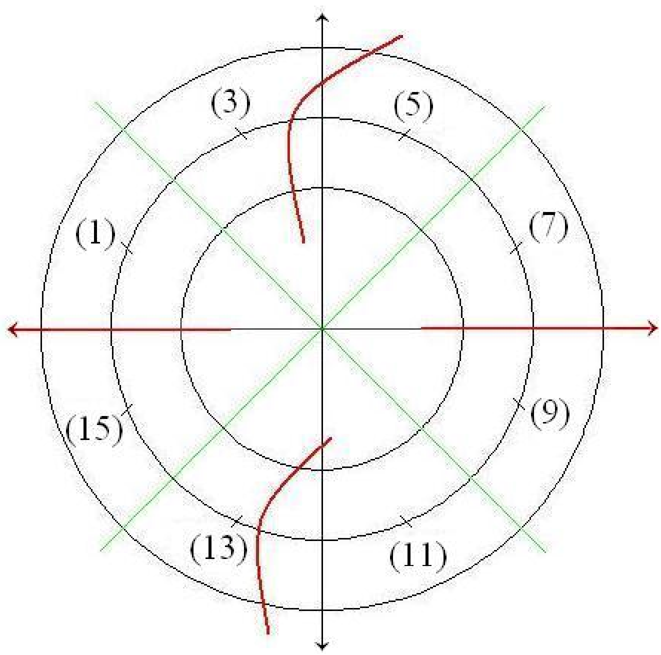

Figura 1.12. Primera curva

Si trazamos otra circunferencia de radio menor que r pero mayor que $|A|\sqrt{2}+1$ y la dividimos de la misma manera, encontraremos también un punto entre (3) y (5) donde T es igual a 0, así como uno entre (7) y (9), (11) y (13) y entre (15) y (1). Lo mismo ocurrirá en una circunferencia de radio mayor que r. Mientras más próximas sean las circunferencias, más próximos estarán entre sí los puntos entre las marcas (3) y (5) donde $T = 0$; de igual modo ocurrirá con los demás puntos.

El conjunto de todos esos puntos en las diferentes circunferencias forma la curva $T = r^2 \operatorname{sen} 2\varphi + r \operatorname{sen} \varphi = 0$, que llamaremos *primera curva*, como se ilustra en la figura 1.12. En otras palabras, estas circunferencias se intersecan con las 4 ramas de la curva $T = 0$. Es importante dejar claro que el eje horizontal siempre formará parte de la *primera curva*, pues evaluando $\varphi = 0$ y $\varphi = \pi$ en $T = r^2 \operatorname{sen} 2\varphi + Ar \operatorname{sen} \varphi$ obtendremos $r^2 \operatorname{sen} 0 + Ar \operatorname{sen} 0 = 0$ y $r^2 \operatorname{sen} \pi + Ar \operatorname{sen} \pi = 0$, respectivamente.

Analicemos ahora la superficie $U = r^2 \cos 2\varphi + Ar \cos \varphi + M$. Esta expresión puede escribirse como $U = r^2 \left(\cos 2\varphi + \frac{A}{r} \cos \varphi + \frac{M}{r^2}\right)$. Para el caso en que r es infinito se obtendrá que U coincide con la superficie $S_2(r, \varphi) = r^2 \cos 2\varphi$; y puesto que $\cos 2\varphi$ varía de 1 a -1, en algún valor de φ se tendrá que U es igual a 0. En el intervalo $0 \leq \varphi < 2\pi$ la función $f(\varphi) = \cos \varphi$ cambia de signo 2 veces; cuando $0 \leq \varphi < \frac{\pi}{2}$ se tiene que $\cos \varphi > 0$; si $\frac{\pi}{2} < \varphi < \frac{3\pi}{2}$ se tiene que $\cos \varphi < 0$, y si $\frac{3\pi}{2} < \varphi < 2\pi$ entonces $\cos \varphi > 0$. Luego, $\cos 2\varphi$ cambia de signo 4 veces:

$$0 \leq \varphi < \tfrac{\pi}{4} \Rightarrow 0 \leq 2\varphi < \tfrac{\pi}{2} \;\text{ y }\; \cos 2\varphi > 0;$$

$$\tfrac{\pi}{4} < \varphi < \tfrac{3\pi}{4} \Rightarrow \tfrac{\pi}{2} < 2\varphi < \tfrac{3\pi}{2} \;\text{ y }\; \cos 2\varphi < 0;$$

$$\tfrac{3\pi}{4} < \varphi < \tfrac{5\pi}{4} \Rightarrow \tfrac{3\pi}{2} < 2\varphi < \tfrac{5\pi}{2} \;\text{ y }\; \cos 2\varphi > 0;$$

$$\tfrac{5\pi}{4} < \varphi < \tfrac{7\pi}{4} \Rightarrow \tfrac{5\pi}{2} < 2\varphi < \tfrac{7\pi}{2} \;\text{ y }\; \cos 2\varphi < 0, \;\text{ y}$$

$$\tfrac{7\pi}{4} < \varphi \leq 2\pi \Rightarrow \tfrac{7\pi}{2} < 2\varphi \leq 4\pi \;\text{ y }\; \cos 2\varphi > 0.$$

Así, $U = r^2 \cos 2\varphi$ se encontrará arriba del plano cuando $0 \leq \varphi < \frac{\pi}{4}$, $\frac{3\pi}{4} < \varphi < \frac{5\pi}{4}$ o $\frac{7\pi}{4} < \varphi < 2\pi$. La superficie U se encontrará bajo el plano cuando $\frac{\pi}{4} < \varphi < \frac{3\pi}{4}$ o

$\frac{5\pi}{4} < \varphi < \frac{7\pi}{4}$. $U = r^2 \cos 2\varphi + Ar\cos\varphi + M = 0$ se puede rescribir como $U = \cos 2\varphi + \frac{A}{r}\cos\varphi + \frac{M}{r^2} = 0$ cuando r es infinito; la curva que obtenemos en este caso coincide con $\cos 2\varphi = 0$. Esta ecuación se cumple siempre que $2\varphi = k\frac{\pi}{2}$, con k en los enteros impares; es decir, si $\varphi = k\frac{\pi}{4}$; esto es, $\cos 2\varphi = 0$ cuando $\varphi \in \{\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}\}$. Así, la curva $\cos 2\varphi = 0$ está formada por las semirrectas que parten del origen y cuyos ángulos de inclinación son los anteriormente dichos. Esto nos lleva a concluir que la curva $U = r^2\cos 2\varphi + Ar\cos\varphi + M = 0$ tiene 4 ramas.

Analizando U de manera similar a la superficie T cuando r es finito, consideremos una circunferencia de radio r dividida igual que antes. En el punto etiquetado como (3), $\varphi = \frac{3}{2}\left(\frac{\pi}{4}\right)$; así, $2\varphi = \frac{3\pi}{4}$. Sustituyendo el valor de 2φ en $U = r^2\left(\cos 2\varphi + \frac{A}{r}\cos\varphi + \frac{M}{r^2}\right)$ se tiene que

$$U = r^2\left(\cos\frac{3\pi}{4} + \frac{A}{r}\cos\varphi + \frac{M}{r^2}\right) = r^2\left(-\sqrt{\frac{1}{2}} + \frac{A}{r}\cos\varphi + \frac{M}{r^2}\right).$$

Veamos en qué intervalos de la circunferencia conserva U el mismo signo que $\cos\frac{3\pi}{4}$; para que esto ocurra se necesita que $\left|\frac{A}{r}\cos\varphi + \frac{M}{r^2}\right| < \sqrt{\frac{1}{2}}$; se tiene que

$$\begin{aligned}
\left|\tfrac{A}{r}\cos\varphi + \tfrac{M}{r^2}\right| &\leq \left|\tfrac{A}{r}\cos\varphi\right| + \left|\tfrac{M}{r^2}\right| \\
&= \left|\tfrac{A}{r}\right|\left|\cos\varphi\right| + \left|\tfrac{M}{r^2}\right| \\
&\leq \left|\tfrac{A}{r}\right| + \left|\tfrac{M}{r^2}\right| \\
&= |A|\left(\tfrac{1}{r}\right) + |M|\left(\tfrac{1}{r^2}\right) \\
&< g\left(\tfrac{1}{r} + \tfrac{1}{r^2}\right)
\end{aligned}$$

$$< g\left(\tfrac{1}{r} + \tfrac{1}{r^2} + \tfrac{1}{r^3} + \cdots + \tfrac{1}{r^{n-1}} + \dots\right)$$
$$= g\left(\frac{\frac{1}{r}}{1-\frac{1}{r}}\right)$$
$$= g\left(\tfrac{1}{r-1}\right),\text{ siendo } g = \text{máx}\,\{|A|, |M|\} \text{ y siempre que } \tfrac{1}{r} < 1.$$

Así, $\left|\frac{A}{r}\cos\varphi + \frac{M}{r^2}\right| < g\left(\frac{1}{r-1}\right)$; de este modo, si $g\left(\frac{1}{r-1}\right) < \sqrt{\frac{1}{2}}$ entonces $\left|\frac{A}{r}\cos\varphi + \frac{M}{r^2}\right| < \sqrt{\frac{1}{2}}$. Pero $g\left(\frac{1}{r-1}\right) < \sqrt{\frac{1}{2}}$ implica que $r > g\sqrt{2} + 1$. Luego, $U = r^2\left(\cos\frac{3\pi}{4} + \frac{A}{r}\cos\varphi + \frac{M}{r^2}\right)$ conservará el mismo signo que $\cos\frac{3\pi}{4}$ siempre que la circunferencia sobre la que trabajamos tenga radio $r > g\sqrt{2} + 1$, con $g = \text{máx}\,\{|A|, |M|\}$. De igual modo, para todos los ángulos φ para los cuales $|\cos 2\varphi| > \sqrt{\frac{1}{2}}$ se tendrá que $U = r^2\left(\cos 2\varphi + \frac{A}{r}\cos\varphi + \frac{M}{r^2}\right)$ conservará el mismo signo que $\cos 2\varphi$.

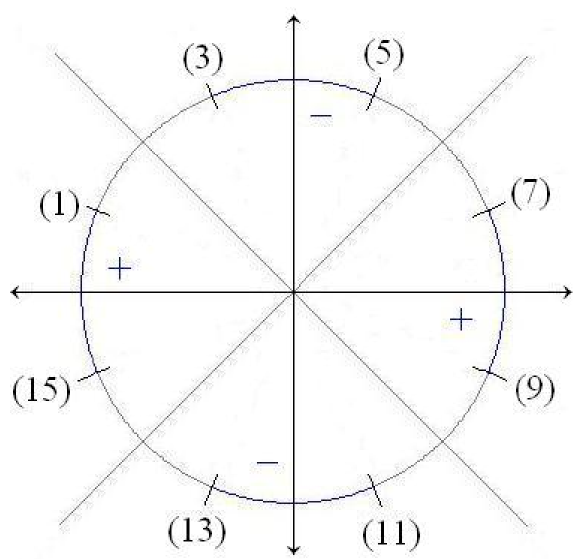

Figura 1.13. Puntos de la circunferencia donde U es positivo o negativo

Por un razonamiento análogo al empleado con la superficie T, se concluye que la superficie U tiene valores negativos del punto (3) al (5) y del (11) al (13); y U tiene valores positivos del (7) al (9) y del (15) al (1), como se muestra en la figura 1.13. Además, existe un único punto entre (1) y (3) donde $U = 0$, del mismo modo que entre (5) y (7), (9) y (11) y entre (13) y (15). De aquí que si tomamos como radio de la circunferencia $r = g\sqrt{2} + 1$, encontraremos puntos donde $T = 0$ y puntos donde $U = 0$ ubicados de forma alternada alrededor de la circunferencia.

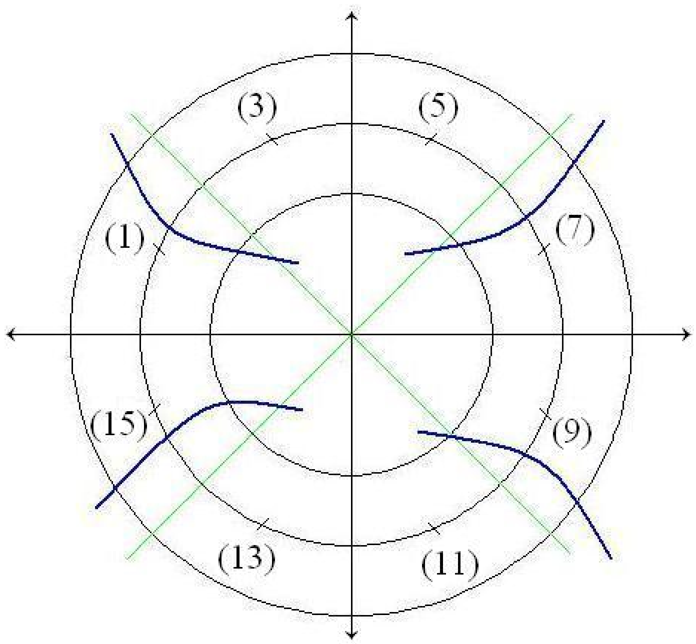

Figura 1.14. Segunda curva

Como se mencionó anteriormente, si tomamos circunferencias con radio mayor que $g\sqrt{2} + 1$, y las dividimos de la misma manera, los puntos donde $U = 0$ entre las marcas (1) y (3), por ejemplo, de cada circunferencia estarán tan próximos como lo estén las propias circunferencias. Así, estos puntos formarán la curva $U = r^2 \cos 2\varphi + |A| r \cos \varphi + M = 0$, que llamaremos *segunda curva*, como se muestra en la figura 1.14.

En un círculo de radio $r > g\sqrt{2}+1$ designemos con el número 0 la parte izquierda del eje horizontal, que es la primera rama de la *primera curva* (véase la figura 1.15). Asignemos el 1 a la rama inmediata que entra al círculo y que pertenece a la *segunda curva*; con el número 2 a la segunda rama de la *primera curva*; con el número 3 a la segunda rama de la *segunda curva*, y así sucesivamente hasta el número 7, que identifica a la cuarta rama de la *segunda curva*.

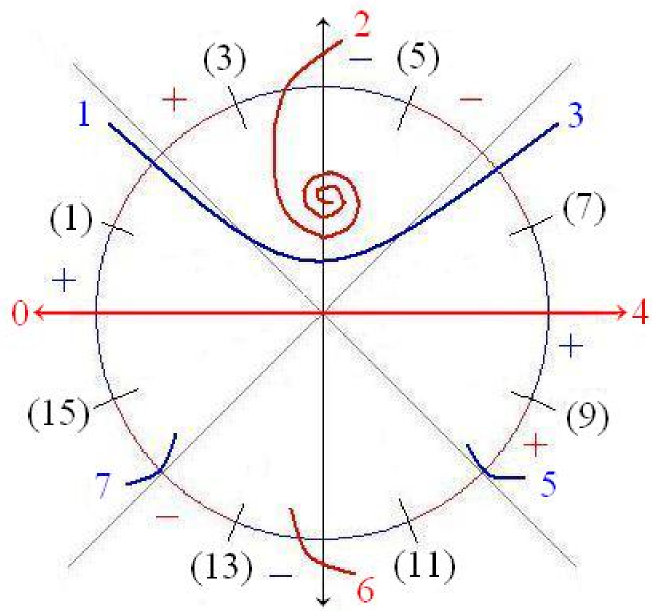

Figura 1.15. Rama de la curva $T=0$ que no sale del círculo de radio $r > g\sqrt{2}+1$

Veamos que cada rama de la primera curva que entra al círculo saldrá de él por medio de otra rama de la misma curva y en su camino se intersecará con una rama de la segunda curva. Supongamos, sin pérdida de generalidad, que la rama 2 que corresponde a la primera curva no sale del círculo perdiéndose dentro de él, como se muestra en la figura 1.15. La rama 1 correspondiente a la segunda curva entra al círculo entre (1) y (3); hagamos que dé un rodeo a la rama 2 para no tocarla, y de ese modo salga por la rama 3, pues si saliera por la rama 5 o la 7 cortaría al eje horizontal, que siempre forma parte de la primera curva.

Notemos ahora que la rama 1 tiene a la superficie $T = r^2 \operatorname{sen} 2\varphi + Ar \operatorname{sen} \varphi$ sobre ella, pues de (1) a (3) la superficie siempre tiene valores positivos. Observemos además que la rama 3 tiene a la superficie T debajo de ella, pues de (5) a (7) la superficie tiene valores negativos. Luego, puesto que la rama 1 se encuentra en el plano y como T es una superficie continua, en algún momento la rama 1 está sobre la superficie T, y esto ocurre justo cuando T es igual a 0, es decir, la rama 1 tocó a la rama 2. Hemos llegado a una contradicción. De lo anterior se concluye que toda rama que entra al círculo necesariamente sale de él.

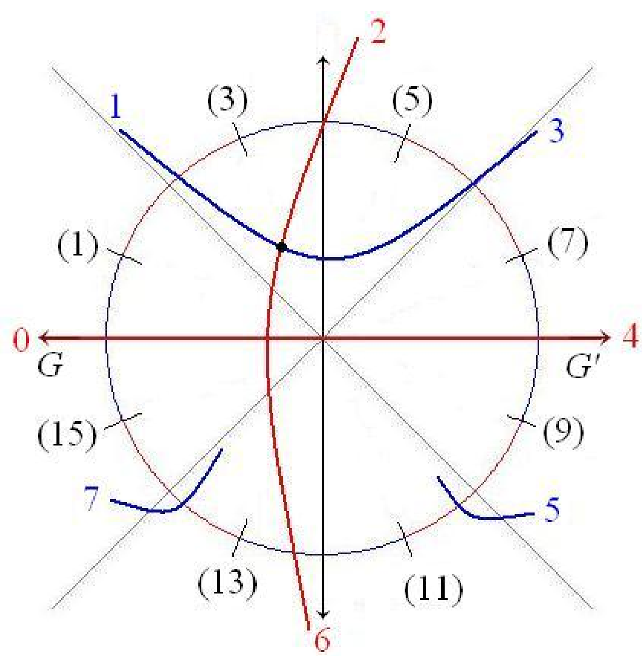

Figura 1.16. Intersección de las curvas $T = 0$ y $U = 0$

Otra consecuencia de la observación anterior es que toda rama etiquetada con un número par perteneciente a la *primera curva* está conectada con otra rama par; del mismo modo, toda rama etiquetada con un número impar que forma parte de la *segunda curva* está conectada con otra de su misma naturaleza. Veamos que independientemente de la conexión entre ramas pares y ramas impares siempre ocurrirá una intersección de la primera curva con la segunda.

Supongamos que se pueden unir dos ramas impares de tal manera que no se intersequen con la *primera curva*. Puesto que el eje GG' forma parte de la primera curva necesariamente tendremos que unir la rama 1 con la rama 3, como se muestra en la figura 1.16. Por otra parte, la rama 2 tendría que unirse con la rama 6 puesto que no puede perderse dentro del círculo; de aquí que la rama de la *primera curva*, resultado de unir 2 con 6, se interseca con la rama de la segunda curva que se generó al unir 1 con 3. Luego, siempre ocurrirá una intersección de la *primera curva* con la *segunda curva*. Así, las intersecciones de la curva $T = 0$ con la curva $U = 0$ serán *las* soluciones de la ecuación $X = x^2 + Ax + M = 0$.

Veamos las posibles formas en que se pueden intersecar las ramas de $T = 0$ y $U = 0$ dentro del círculo. Puede ocurrir que ambas componentes de la segunda curva se intersequen con la misma componente de la primera curva. En este caso la segunda curva se puede intersecar con el eje de referencia dando lugar a dos raíces reales, o con la segunda componente de la primera curva, generandose dos raíces complejas (véase la figura 1.17 *a* y *b*). La tercera posibilidad es que las todas la componentes se toquen en un solo punto. Es decir, las dos componentes de $T = 0$ y las dos de $U = 0$ coinciden en un punto que representa una única raíz de doble multiplicidad (véase la figura 1.17 *c*).

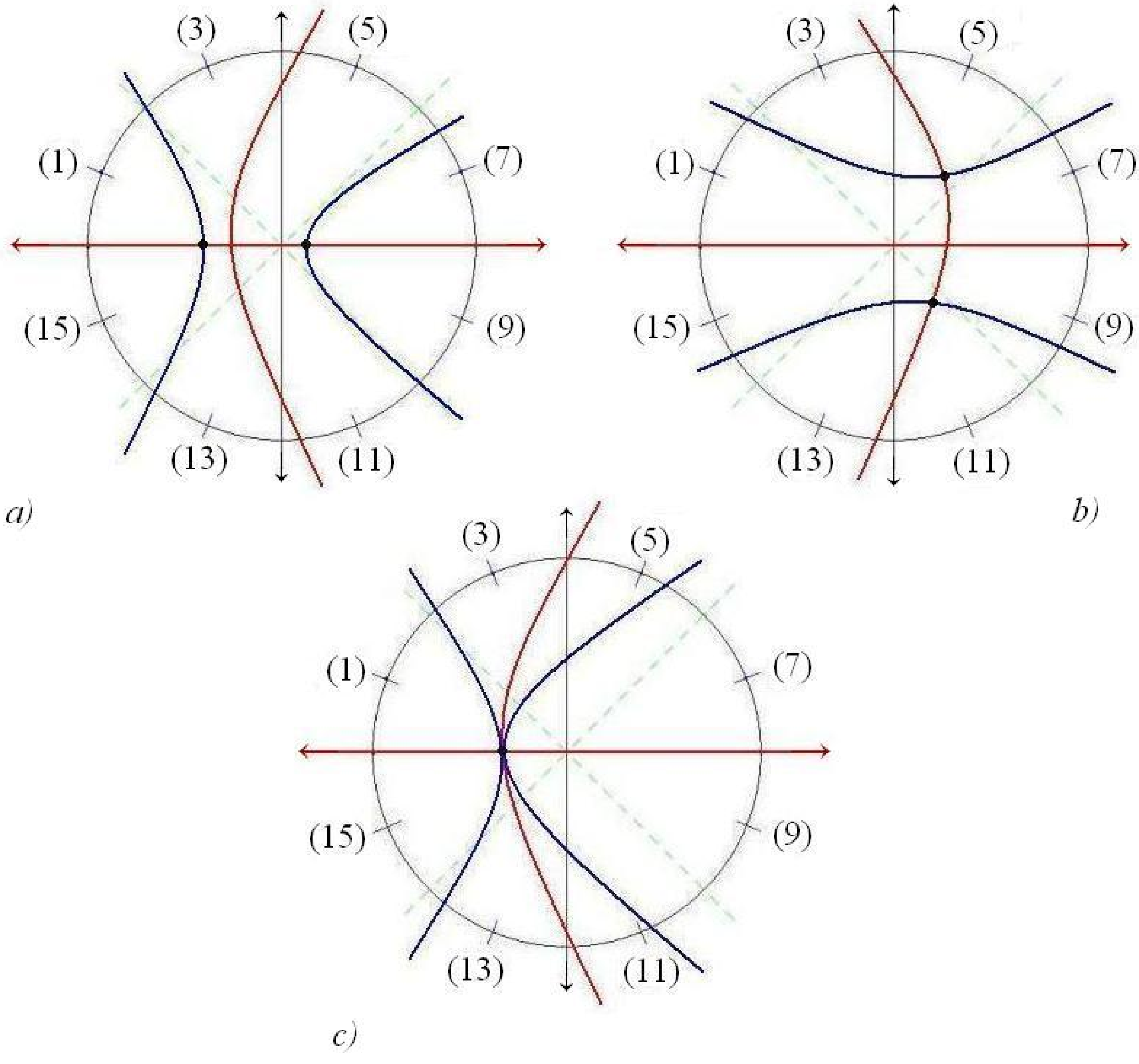

Figura 1.17. Intersecciones de $T = 0$ y $U = 0$ que generan 2 raíces

Mediante este análisis podemos comprender por qué las ecuaciones de segundo grado tienen 2 raíces, haciendo explícita la manera en que estas soluciones se generan. Los siguientes ejemplos muestran los tres diferentes tipos de soluciones. El primero es de la ecuación cuadrática $x^2 - 2x - 35 = 0$, que tiene dos soluciones reales (véase la figura 1.18).

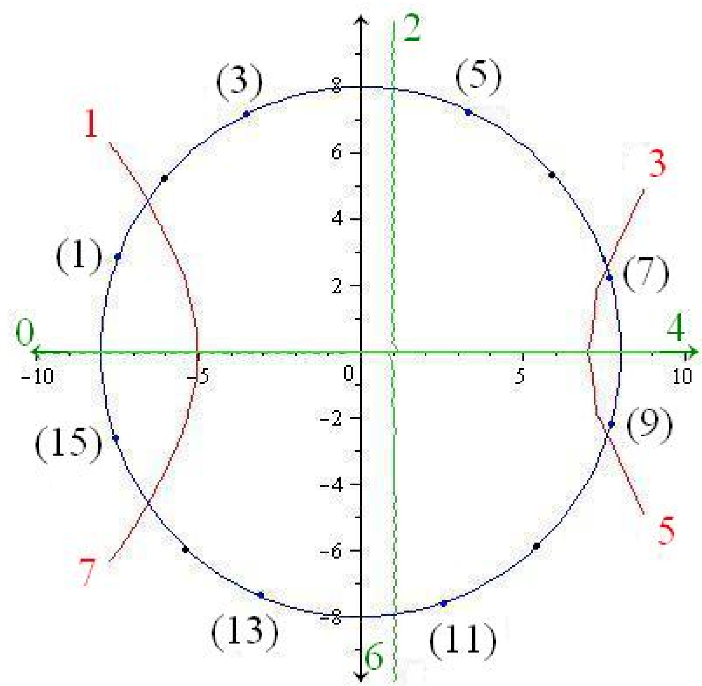

Figura 1.18. Geometría de las soluciones de $x^2 - 2x - 35 = 0$

El segundo ejemplo es el de la ecuación $x^2 - 4x + 13 = 0$, que tiene dos soluciones complejas (véase la figura 1.19).

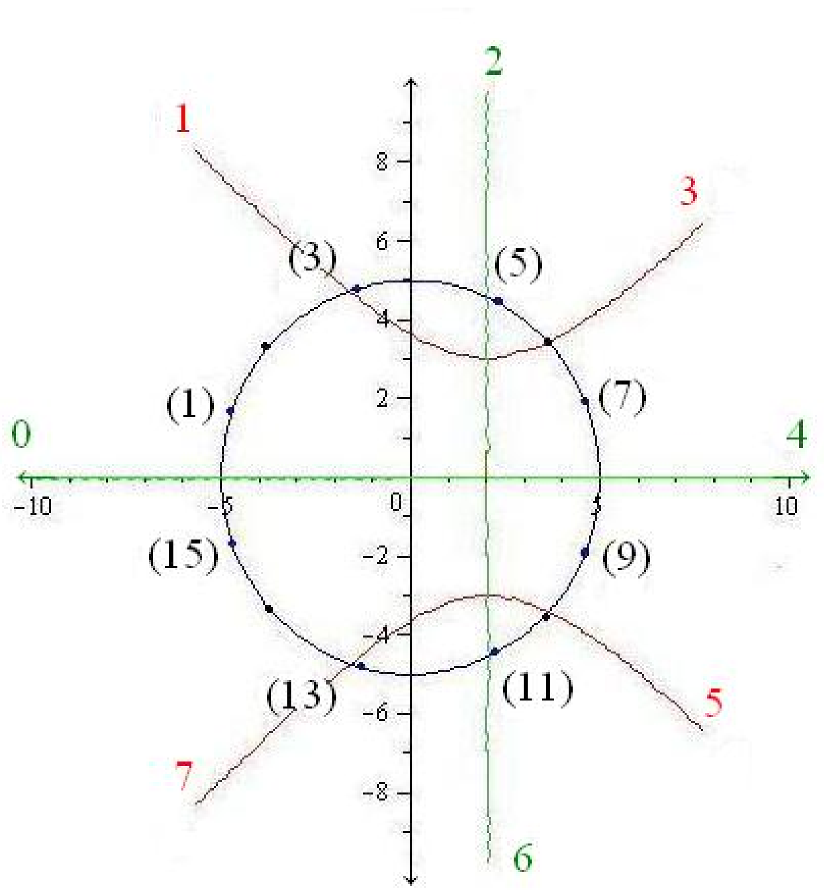

Figura 1.19. Geometría de las soluciones de $x^2 - 4x + 13 = 0$

Finalmente, el tercer ejemplo es el de la ecuación $x^2 - 4x + 4 = 0$, la cual tiene una solución real de multiplicidad 2 (véase la figura 1.20).

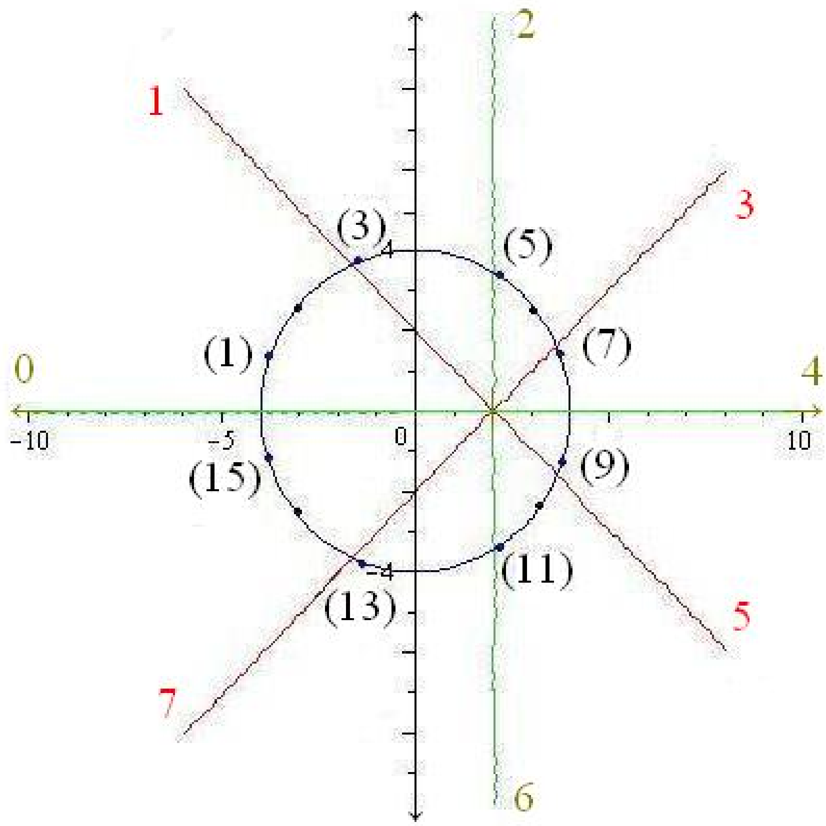

Figura 1.20. Geometría de las soluciones de $x^2 - 4x + 4 = 0$

CAPÍTULO II

LOS ESTUDIOSOS DEL MEDIEVO Y LAS EXPRESIONES CUADRÁTICAS

El arte del al-jabr y el al-muqabala

El Al-jabr wa'l-muqabala de al-Khwarizmi y la obra de ibn Turk

Uno de los libros más antiguos del mundo árabe relacionado con lo que actualmente conocemos como álgebra es el *al-Kitab al-muhtasar fi hisab al-jabr wa'l-muqabala* [El libro condensado sobre cálculo mediante *al-Jabr* y *al-Muqabala*]. Este tratado fue desarrollado por Muhammad ibn Musa al-Khwarizmi, alrededor del año 830 d.C., quien fue un estudiante de la Casa de la Sabiduría, establecida en Bagdad en el periodo 786 - 809.

El título *al-Kitab al-muhtasar fi hisab al-jabr wa'l-muqabala* contiene los términos que representan los principales rasgos de las matemáticas desarrolladas en esta obra: *al-jabr wa'l-muqabala*. "Al-jabr" es eliminar las cantidades que aparecen restando en una expresión, y "al-muqabala" es reducir los términos semejantes de una expresión. Por ejemplo, empleando el simbolismo actual, la expresión $x = 10 - 3x$ se transforma mediante "al-jabr" en $4x = 10$ y $10x + 64 = 5x + 36 + x^2$ se transforma mediante "al-muqabala" en $5x + 28 = x^2$.

Al-Khwarizmi elaboró este tratado como un libro de texto elemental de matemáticas prácticas y fue escrito de manera totalmente retórica, es decir, no se manejaban símbolos, y las manipulaciones se hacían verbalmente, como era usual en los trabajos árabes sobre estas cuestiones matemáticas. Comienza con una discusión sobre las expresiones que actualmente conocemos como ecuaciones de primer y segundo grados y al final de la obra se incluyen dos partes sobre aplicaciones prácticas a los negocios y asuntos de medidas y herencias (Parshall 1988, p. 132).

De acuerdo con lo comúnmente escrito en los libros de historia de las matemáticas, cuando al-Khwarizmi elaboró su obra, tenía conocimientos de matemáticas griegas gracias a la introducción de éstas en la Casa de la Sabiduría; conoció los *Elementos* de Euclides (300 a.C. aproximadamente), cuya traducción al árabe existía en dos distintas versiones y esto lo motivó a dar una demostración geométrica de todos los procedimientos algebraicos que abordó. Sin embargo, Katz afirmó que estas demostraciones no se tomaron de los textos griegos, sino que muestran gran similitud con los argumentos de los babilonios, tomando como evidencia más clara el desarrollo de las expresiones actualmente asociadas con ecuaciones de segundo grado (Katz 1993, p. 229).

Al-Khwarizmi distinguió seis tipos diferentes de expresiones sobre cuadrados y trabajó con tres clases de cantidades: una cantidad desconocida o raíz, el cuadrado de la misma y los números absolutos o las constantes. Propuso entonces la siguiente clasificación, en donde los términos de las expresiones y las raíces de las mismas siempre se tomaron positivos:

1.- *Cuadrados iguales a raíces* (en lenguaje moderno: $ax^2 = bx$).

2.- *Cuadrados iguales a números* (es decir, $ax^2 = c$).

3.- *Raíces iguales a números* (es decir, $bx = c$).

4.- *Cuadrados y raíces iguales a números* (es decir, $ax^2 + bx = c$).

5.- *Cuadrados y números iguales a raíces*: $ax^2 + c = bx$.

6.- *Raíces y números iguales a cuadrados*: $bx + c = ax^2$.

(En lo subsecuente emplearemos el simbolismo moderno para referirnos a las expresiones anteriores, dejando claro que las ecuaciones como tales no existían en aquella época, pues no se utilizaba el lenguaje simbólico empleado en nuestros días.)

A continuación se describirá cómo al-Khwarizmi resolvió problemas del tipo cuadrados y números iguales a raíces (esto es, del tipo 5: $ax^2 + c = bx$). Para comenzar, la expresión puede ser llevada a la forma $x^2 + c = bx$. En la figura 2.1 se representa a x^2 con el cuadrado $ABCD$; el rectángulo $AHNB$ representa a c.

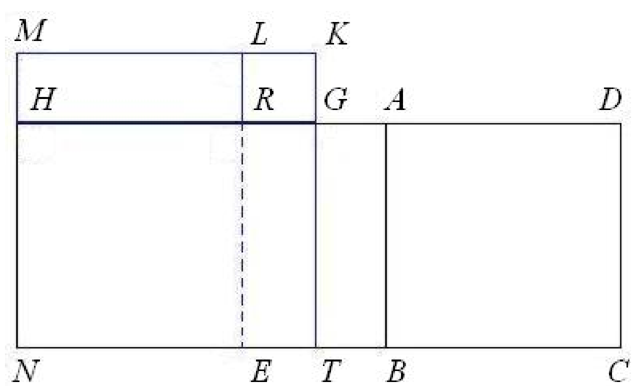

Figura 2.1. Esquema geométrico para la resolución de $x^2 + c = bx$

Como $x^2 + c$ es lo mismo que bx, la longitud de $\overline{HD}$ es b. Se localiza G, el punto medio de $\overline{HD}$, y se traza $\overline{GT}$ paralelo a $\overline{AB}$ prolongándolo hasta K, de modo que $GK = GA$; luego se traza el rectángulo $GKMH$. Se determina L de modo que $GK = KL$, y se completa así el cuadrado $KLRG$. Como G es punto medio de $\overline{HD}$ y $GA = LR = RG$, se tiene que $AD = HR$, por lo que el rectángulo $LMHR$ es igual al rectángulo $GTBA$. Luego, el área del cuadrado $MNTK$ es $\left(\frac{b}{2}\right)^2$, siendo el área de éste menos el área del cuadrado $KLRG$ igual al área del rectángulo $AHNB$, que es c. Así, el cuadrado $KLRG$ tiene como área $\left(\frac{b}{2}\right)^2 - c$. Como el lado de $KLRG$ es igual a AG, se sigue que $x = AD = GD - AG$, es decir, el procedimiento nos lleva al resultado que en notación moderna escribimos como $x = \frac{b}{2} - \sqrt{\left(\frac{b}{2}\right)^2 - c}$. Al-Khwarizmi aclaró, respecto a la cantidad dentro del radical, que

> . . . si el producto [de la mitad del número de “raíces” $\left(\frac{b}{2}\right)$ con ella misma] es menor que el número relacionado con el cuadrado [es decir c] entonces el caso es imposible; y si el producto $\left[\left(\frac{b}{2}\right)^2\right]$ es igual al número mismo $[c]$, la raíz $[x]$ es igual a la mitad del número de raíces $\left[\frac{b}{2}\right]$ únicamente, sin la adición o sustracción. (Citado en Katz 1993, p. 231)

Para al-Khwarizmi no tenía sentido una cantidad negativa, ya que sus procedimientos se basaban en la representación de áreas y segmentos, y no se tenía la conceptuación de áreas negativas ni de segmentos negativos.

Otro libro sobre *al-jabr* y *al-muquabala* fue escrito por Abd al-Hamid ibn Wasi ibn Turk al-Jili (fl. 830), contemporáneo de Al-Khwarizmi: *Kitab al-jabr wa’l muqabala*. En su libro, ibn Turk trata con expresiones de los tipos 1, 4, 5 y 6 de la clasificación de al-Khwarizmi e incluye más detalles sobre las demostraciones. Por ejemplo, en la resolución de expresiones de la forma $x^2 + c = bx$, al-Khwarizmi colocó G, el punto medio de $\overline{HD}$, en $\overline{HA}$ (véase la figura 2.1); ibn Turk puso énfasis en que G puede estar en el segmento $\overline{AD}$, lo que lleva a que $x = \frac{b}{2} + \sqrt{\left(\frac{b}{2}\right)^2 - c}$. En este caso, el procedimiento con base en la figura 2.2 se muestra a continuación.

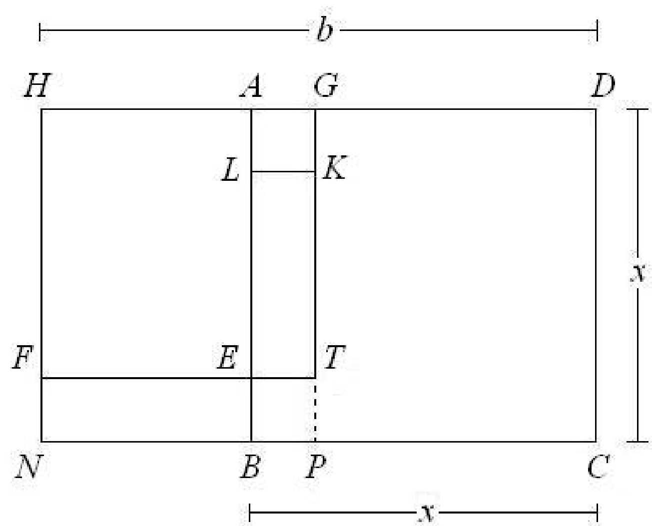

Figura 2.2 Esquema utilizado por ibn Turk

El cuadrado $ABCD$ representa a x^2 y el rectángulo $AHNB$ representa a c; así que la longitud DH es igual a b. Se encuentra el punto medio de $\overline{DH}$ y se denota con G (nótese que G está en $\overline{AD}$). Se construye el cuadrado $GALK$ y se prolonga $\overline{GK}$ hasta T, de modo que KT tenga la misma longitud que AH. Así, se tiene el cuadrado $GHFT$. Ahora, como $DG + GA = x$ y $AE + EB = x$, siendo $DG = AE$ $(DG = GH = GT = AE)$, al sustituir en $AE + EB = x$ se obtiene que $DG + EB = x$. Luego, $GA = EB$ (conviene recordar que $KL = GA$). Por otra parte, por construción se tiene que $KT = AH = EF$.

Así, el rectángulo $KLET$ es igual al rectángulo $EFNB$. Luego, $GHFT - GALK = AHNB = c$. Siendo $GHFT = \left(\frac{b}{2}\right)^2$, el cuadrado $GALK = \left(\frac{b}{2}\right)^2 - c$, por lo que $GA = AL = \sqrt{\left(\frac{b}{2}\right)^2 - c}$. Como $x = DG + GA$, se tiene que $x = \frac{b}{2} + \sqrt{\left(\frac{b}{2}\right)^2 - c}$ (Katz 1993, p. 231).

En sus procedimientos al-Khwarizmi e ibn Turk emplearon la geometría para interpretar las incógnitas y las constantes de las expresiones con que trabajaron. Lo que hoy escribimos como x^2 fue representado con el área de un cuadrado de lado x.

Abu-Kamil y Thabit ibn Qurra

A fines del siglo IX y principios del X aparecieron nuevas traducciones de los *Elementos* realizadas por contemporáneos cercanos a Abu-Kamil (850 - 930), un egipcio que escribió su propio *Kitab fi al-jabr wa'l-muqabala* basado en el trabajo de al-Khwarizmi. Abu-Kamil empleó las justificaciones formales y deductivas de Euclides en el trabajo que desarrolló al-Khwarizmi, introduciendo así el rigor euclidiano en las bases de la obra de al-Khwarizmi (Parshall 1988, pp. 134 - 135).

Abu-Kamil encontró el valor de x en la expresión $x^2 + 10x = 39$ empleando la proposición 6 del libro II de los *Elementos* de Euclides:

> Si se divide en dos partes iguales una línea recta y se le añade, en línea recta, otra recta, el rectángulo comprendido por la (recta) entera con la (recta) añadida y la (recta) añadida junto con el cuadrado de la mitad es igual al cuadrado de la (recta) compuesta por la mitad y la (recta) añadida. (Euclides 2000, p. 94) (Obsérvese la figura 2.3.)

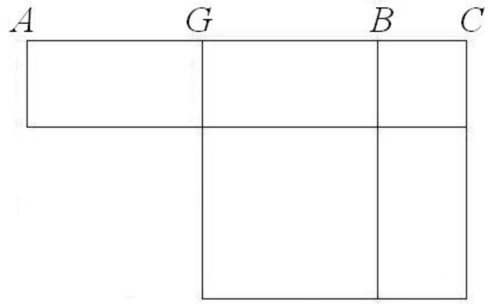

Figura 2.3 Proposición 6 del libro II de los *Elementos* de Euclides

Representó $x^2 + 10x = 39$ partiendo del cuadrado $ABGD$ de la figura 2.4, donde la longitud de $\overline{AB}$ es x. El procedimiento es el siguiente: se construye el rectángulo $WHBA$ en el que la longitud de $\overline{BH}$ es 10. Así, la superficie $WHGD$ representa $x^2 + 10x,$ por lo que $WHGD$ es igual a 39. Se divide $\overline{BH}$ por la mitad, siendo F su punto medio. Se traza $\overline{FM}$ paralelo a $\overline{BA}$ y se prolonga hasta K, de modo que $MK = BF$, y se completa el cuadrado $KMAE$ así como el rectángulo $ADNE$.

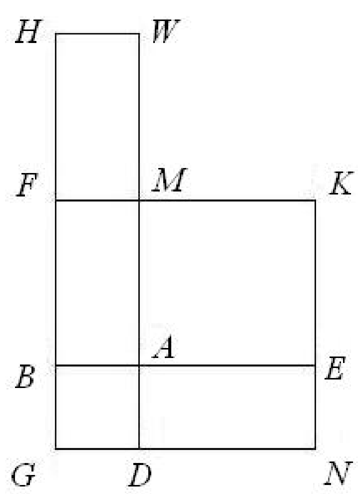

Figura 2.4. Procedimiento geométrico de Abu-Kamil

Por la proposición 6 del libro II de los *Elementos* de Euclides, se tiene que el cuadrado $FGNK$ es igual al rectángulo $WHGD$ más el cuadrado $MAEK$. Siendo el producto de HG por GD igual a 39 y el producto de FB por sí mismo igual a 25, es decir, $(FB)^2 = 25$, la suma de ambos es 64. Así, FG tiene como longitud 8 puesto que $\sqrt{64} = 8$; y sabiendo que FB tiene como longitud 5, se tiene que el segmento restante BG mide 3, que es la raíz buscada (Parshall 1988, p. 135).

Otro estudioso de este tipo de problemas fue Thabit ibn Qurra. Sus escritos y los de Abu-Kamil aparecieron aproximadamente 50 años después de los trabajos de al-Khwarizmi. Thabit ibn Qurra fue un estudiante de la Casa de la Sabiduría nacido en Harran, región que actualmente se sitúa en Turquía. Ibn Qurra y Abu-Kamil, según las versiones estándares de los historiadores de las matemáticas, pensaron que las justificaciones para resolver los problemas sobre *al-jabr* y *al-muqabala* debían emanar de los *Elementos* de Euclides (Katz 1993, p. 233) por su carácter sintético.

Uno de los trabajos de Thabit ibn Qurra fue *Qwal fi tashih masa-il al-jabr bi l-barahin al-handasiya*. Como un ejemplo de su obra, se resolverá la ecuación $x^2 + bx = c$. En la figura 2.5, $\overline{AB}$ representa a x. Así, el cuadrado $ABCD$ es x^2; $\overline{BE}$ representa a b, por lo que el rectángulo $BEFC$ es bx. Luego, el rectángulo $AEFD = AB \times AE$ es c.

Sea W el punto medio de $\overline{BE}$; de acuerdo con la proposición 6 del libro II de los *Elementos* de Euclides, se tiene que $AB \times AE + (BW)^2 = (AW)^2$. Pero $AB \times AE = c$ y $(BW)^2 = \left(\frac{b}{2}\right)^2$. Así, $(AW)^2 = c + \left(\frac{b}{2}\right)^2$. Como $x = AB = AW - BW$, $x = \sqrt{c + \left(\frac{b}{2}\right)^2} - \frac{b}{2}$. Según el relato de Katz (1993, pp. 233 - 234), Thabit notó que la geometría de Euclides era "análoga" a la geometría utilizada por al-Khwarizmi, la cual a su vez fue tomada de los babilonios, y la empleó como la justificación que consideraba necesaria; de este modo Euclides aparece como sintetizador de los babilonios y su trabajo sirvió a Thabit para dotar de rigor a sus aportaciones. Este planteamiento lleva a preguntar si existe alguna relación entre el trabajo desarrollado por los babilonios y los *Elementos* de Euclides que conduzca a tal similitud.

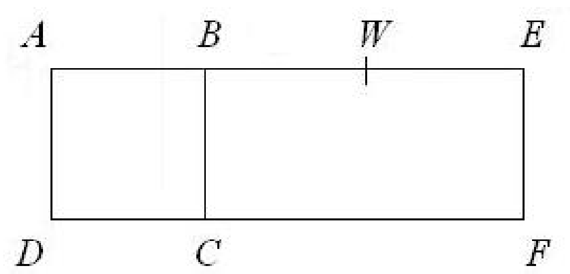

Figura 2.5. Representación geométrica de $x^2 + bx = c$

Los trabajos sobre exponentes de al-Karaji y al-Samaw'al

Abu Bakr al-Karaji (aprox. 953 - 1029) y al-Samaw'al (aprox. 1130 - 1180) continuaron la línea matemática árabe utilizando técnicas de aritmética en los temas relacionados con el al-jabr y al-muqabala. En la primera década del siglo XI al-Karaji escribió su obra *al-Fakhri* (El Maravilloso), cuyo propósito fue "la determinación de desconocidas partiendo de conocidas", del mismo modo que el arte del al-jabr y al-muqabala en general según el autor. Al-Karaji trabajó en Bagdad alrededor del año 1000 y escribió diversas obras matemáticas y tratados relacionados con ingeniería. Comenzó su obra *al-Fakhri* estudiando lo que actualmente conocemos como el álgebra de exponentes, aunque este tema ya había sido abordado por escritores más antiguos, entre ellos Diofanto (fl. 250 d.C.).

En su obra, al-Karaji planteó que las potencias pueden extenderse infinitamente, definió cada potencia de forma recursiva, es decir multiplicando por x la potencia anterior, desarrollando una serie infinita de proporciones (Katz 1993, p. 235); así construye las potencias x^n y sus recíprocas, $\frac{1}{x^n}$:

$$1 : x = x : x^2 = x^2 : x^3 = \cdots$$

$$\frac{1}{x} : \frac{1}{x^2} = \frac{1}{x^2} : \frac{1}{x^3} = \frac{1}{x^3} : \frac{1}{x^4} = \cdots$$

Al-Karaji estableció procedimientos generales para sumar, restar y multiplicar monomios y polinomios; también dividió monomios y polinomios entre monomios.

Otro trabajo relacionado con estos temas fue el de al-Samaw'al, quien consideró coeficientes negativos en las expresiones que utilizó. Las reglas para trabajar con estos coeficientes aparecen en el libro de texto *Al-Bahir fi'l-hisab* (El libro brillante de cálculo), mediante las cuales al-Samaw'al pudo sumar y restar polinomios fácilmente combinando los términos de los mismos; formuló además la ley de los exponentes que en notación moderna establece que $x^m x^n = x^{m+n}$, aunque esta ley ya había sido utilizada por otros como al-Karaji y Abu-Kamil. En aquellos tiempos, a causa del estilo retórico empleado en la elaboración de escritos algebraicos, la propiedad de sumar los exponentes cuando se multiplican potencias de una misma base no fue observada; como ejemplo del estilo retórico, se tiene que para expresar el producto de una potencia cúbica por una cuadrada se utilizaban las palabras *cuadrado-cubo* (Katz 1993, pp. 236 - 237).

Al-Samaw'al explicó la ley de los exponentes utilizando una tabla, en la que cada columna, encabezada por un número arábigo, representaba una potencia diferente de la "desconocida". Considerando que la columna central de la tabla está encabezada por el 0, hacia la izquierda de ésta se encuentran las potencias de x y hacia la derecha se localizan las potencias de $\frac{1}{x}$. La tabla, con simbolismo actual, es la siguiente.

…7	6	5	4	3	2	1	0	1	2	3	4	5	6	7…
$\ldots x^7$	x^6	x^5	x^4	x^3	x^2	x	1	x^{-1}	x^{-2}	x^{-3}	x^{-4}	x^{-5}	x^{-6}	$x^{-7}\ldots$
…128	64	32	16	8	4	2	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}\ldots$

La columna encabezada por el 2 a la izquierda del 0 es llamada "cuadrado"; la encabezada por el 5 a la izquierda es llamada "cuadrado-cubo", y la encabezada por el 3 a la derecha es llamada "parte de cubo". En el tercero y último renglón de la tabla colocó las potencias de 2. Al-Samawa'l utilizó la tabla para explicar que $x^m x^n = x^{m+n}$ de la siguiente manera:

> La distancia del orden del producto de los factores a uno de los factores es igual a la distancia del orden del otro factor a la unidad. Si los factores están en diferentes direcciones entonces contamos (la distancia) del orden del primer factor a la unidad; pero si están en la misma dirección, contamos del más lejano a la unidad. (Katz 1993, p. 237)

Como ejemplo, para multiplicar x^2 por x^4 se cuentan cuatro ordenes a la izquierda de la columna 2, obteniendo como resultado x^6; para multiplicar x^{-5} por x^3 se cuentan 5 ordenes a la derecha de la columna 3, obteniendo x^{-2}.

Los trabajos de al-Karaji y al-Samawa'l mostraron un interés en la manipulación de potencias mayores mayores que 2 aunque no se les asoció ningún referente físico.

Las expresiones cuadráticas después de los árabes

Durante la Edad Media (del siglo V hasta mediados del siglo XV), drásticos cambios sociales se relacionaron íntimamente con el empleo y el desarrollo de la aritmética. La producción e intercambio de bienes llevaron a la necesidad de asociarles un valor; el pensamiento filosófico, religioso y político de los productores forjó el concepto de valor.

Un hecho relevante en las actividades de producción, intercambio y valuación de los bienes fue la invasión de España por los árabes en el año 711 y su posterior dominio durante tres siglos. Como consecuencia surgió un nuevo imperio dentro de un continente con gran intercambio comercial y se estableció la moneda en circulación para efectos de intercambio. Cuando el régimen árabe cayó, las bibliotecas fueron destruidas y los libros que sobrevivieron se colocaron en las bibliotecas de las nuevas cortes. Fue así como España se convirtió en uno de los principales centros de traducción del árabe al latín, pues resguardó textos como el *Al-jabr wa'l-muqabala* que fue traducido en el siglo XII por Robert de Chester (fl. 1141 - 1150) y también por Gerard de Cremona (1114 - 1187) (Parshall 1988, p. 137).

Como resultado de las nuevas condiciones sociales que dieron al comercio un lugar clave, aparecieron nuevas formas de cálculo. Pronto el manejo de la aritmética se convirtió en requisito indispensable para una carrera exitosa en la burocracia. Las ganancias y pérdidas que se midieron en dinero, se pudieron predecir aritméticamente. De este modo, la relación producto-valor-intercambio inició una gran tradición algebraica con bases aritméticas, pero sobre todo con raíces sociales (Hadden 1994, p. 88).

Leonardo de Pisa

Uno de los principales exponentes de las matemáticas medievales fue Leonardo de Pisa (1170 - 1240), conocido como Fibonacci. Gracias a las prácticas comerciales de su padre recorrió Egipto, Siria, Grecia, Sicilia y Provenza. Durante sus viajes aprendió reglas para realizar cálculos con los números hindo-arábigos, cómo resolver problemas prácticos relacionados con ganacias, conversiones de moneda y mediciones, además de problemas que actualmente consideramos de carácter algebraico relacionados con movimiento, el residuo Chino, problemas de contenedores y con lo que ahora llamamos ecuaciones cuadráticas.

En 1202 escribió el *Liber abaci* (*Libro de ábaco*), revisado y aumentado en 1228. Las fuentes del *Liber abaci* son árabes en su mayoría, obtenidas en los viajes de Fibonacci, pero adaptadas y complementadas por él mismo (Katz 1993, p. 282). La obra contiene soluciones a problemas prácticos, así como aritmética teórica y geometría. Algo importante es que trabajó con las demostraciones de los libros II y IX de los *Elementos* de Euclides dando demostraciones numéricas (Hadden 1994, p. 88).

Una de los problemas más famoso del *Liber abaci* es el problema de los conejos:

> Un hombre tiene un par de conejos juntos en un lugar cerrado, y desea saber cuántos serán procreados por este par en un año, cuando es de su naturaleza que en un mes nace otro par, y en el segundo mes, esos nacidos procrearán también. Por lo escrito anteriormente, en el primer mes nacerá un par, [así] habrá dos pares en un mes. Uno de éstos, a saber, el primero procreará en el segundo mes, y así en el segundo mes habrá 3 pares; de éstos en un mes dos estarán preñados, y en el tercer mes nacerán 2 pares de conejos, y así habrá 5 pares en el mes; en este mes 3 pares estarán preñados, y en el cuarto mes habrá 8 pares, de los cuales 5 pares procrearán otros 5 pares, éstos son añadidos a los 8 pares haciendo 13 pares en el quinto mes; estos 5 pares que nacerán en este mes no se aparearán, pero los otros 8 pares estarán preñados, y así habrá en el sexto mes 21 pares; a éstos son añadidos los 13 pares que nacerán en el séptimo mes, habrá 34 pares en este mes; a éstos se añaden los 21 pares que nacerán en el octavo mes, habrá 55 pares en este mes; a éstos son añadidos los 34 pares que nacerán en el noveno mes, habrá 89 pares en este mes; a éstos son añadidos los 55 pares que nacerán en el décimo mes, habrá 144 pares en este mes; a éstos son añadidos otra vez los 89 pares que nacerán en el onceavo mes, habrá 233 pares en este mes. A éstos son añadidos aún los 144 pares que nacerán en el último mes, habrá 377 pares; y estos muchos pares serán producidos a partir del par anteriormente escrito en el mencionado lugar al fin de un año. (Fibonacci 2002, pp. 404 - 405)

En el margen de sus notas listó la sucesión

$$1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, \ldots$$

en la que se observa que cada número se obtiene de sumar los dos anteriores. Esta sucesión recursiva es conocida en nuestros días como sucesión de Fibonacci.

En el capítulo XV del *Liber abaci* Leonardo de Pisa introdujo los conocimientos de Al-Khwarizmi sobre la resolución de lo que ahora llamamos ecuaciones de segundo grado. La sección lleva por título “Aquí comienza la parte tres sobre la solución de determinados problemas de acuerdo con el método de álgebra y almucabala, a saber proporción y reintegración”. Aquí Fibonacci expuso tres propiedades inherentes a los números, escribió: “En la composición del álgebra y almucabala tres propiedades que están en cualquier número son consideradas; son raíces, cuadrados y números simples ” (Fibonacci 2002, p. 554), llamando *raíz* y *census* a la desconocida del problema y a su cuadrado, respectivamente. En este apartado trabajó con los seis tipos de expresiones propuestas por Al-Khwarizmi en su *Al-jabr wa'l-muqabala*, dio la regla general para resolver cada una y las justificó por medio de argumentos geométricos. Para el caso *census* más número igual a un número de raíces (en notación moderna, $x^2 + c = bx$) escribió:

> Y cuando ocurra que el census más un número sea igual a un número de raíces, entonces se puede operar con cualquier número igual o menor que el cuadrado de la mitad del número de raíces, si es igual entonces la mitad del número de raíces es la raíz del census, y si el número con el cual el census es

> igual al número de raíces es menor que el cuadrado de la mitad del número de raíces entonces sustraiga el número del cuadrado [de la mitad del número de raíces], y [la raíz cuadrada de] eso que queda se sustrae de la mitad del número de raíces; y si eso que queda no es la raíz del census buscado entonces se suma eso que ha sido sustraido al número desde el que se sustrajo, y se tendrá la raíz del census buscado. (Fibonacci 2002, p. 557)

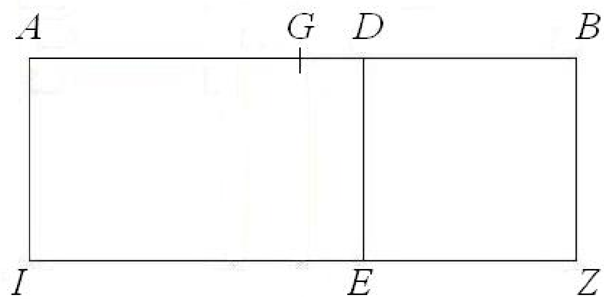

Figura 2.6. Representación geométrica de $x^2 + 40 = 14x$

Para argumentar la regla resolvió la ecuación census más 40 igual a 14 raíces ($x^2 + 40 = 14x$) con el siguiente procedimiento. Emplearemos la figura 2.6. Sea el segmento de línea AB igual a 14, se divide tal segmento en dos partes iguales por el punto G y en dos partes desiguales por el punto D, de tal forma que AD sea mayor que DB. Sobre la porción menor $\overline{DB}$, se construye el cuadrado DZ, se extiende $\overline{ZE}$ hasta el punto I, de modo que el segmento ZI sea igual al segmento AB, y se traza $\overline{IA}$. Como ZB es la raíz del census DZ, y AB es igual a 14 se tiene que AZ son 14 raíces del census. Como $x^2 + 40 = 14x$, se tiene que el área AE es igual a 40, pero $AE = DA \times DE$, esto es, $AE = DA \times BD$ (ya que $DE = DB$); si el cuadrado del segmento DG es agregado a 40, se tendrá como total 49 (esto se argumenta perfectamente cuando se completa el cuadrado de AG según el procedimiento de Al-Khwarizmi, como en la figura 2.7), es decir, el cuadrado del segmento GB, de aquí que $DG^2 = 9$ y $GD = 3$. Luego,

$GD + GA = 10 = DA$ y $GB - GD = 4 = DB$, que es la raíz del census buscado.

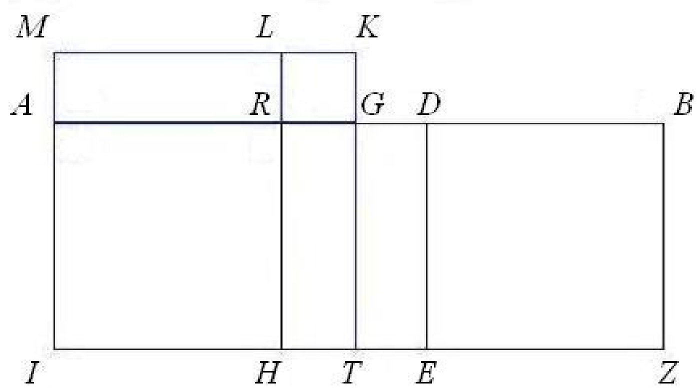

Figura 2.7. Justificación geométrica para resolver $x^2 + 40 = 14x$

Y si el census se construye sobre el segmento AD (véase la figura 2.8), es decir, $x^2 = AE$, entonces el área DZ es igual a 40 y $DZ = DE \times DB = AD \times DB$; y si se resta 40 del cuadrado del segmento AG, que es lo mismo que restar 40 del cuadrado de $\overline{GB}$, el residuo será 9, cuya raíz, a saber 3, es el segmento de línea GD (la justificación geométrica se da por el procedimiento de ibn-Turk, como en la figura 2.8), así $AD = 10$, que es la raíz del census 100.

Leonardo de Pisa escribió cuatro trabajos más. *Practica geometriae* trata con problemas de geometría de figuras planas y sólidas, del triángulo a los poliedros; el tercero fue *Flos super solutionibus quarumdam questionum ad numerum et ad geometriam pertinentium* [*Ramillete de soluciones de determinadas cuestiones relativas al número y a la geometría*]; su cuarta obra fue una carta a Teodoro, filósofo de la corte de Federico II (1194 - 1250), que contiene dos problemas, uno relacionado con proporciones y otro sobre geometría. El quinto trabajo fue el *Liber quadratorum* o libro de los números cuadrados (De Pisa 1973, pp. 8 - 13).

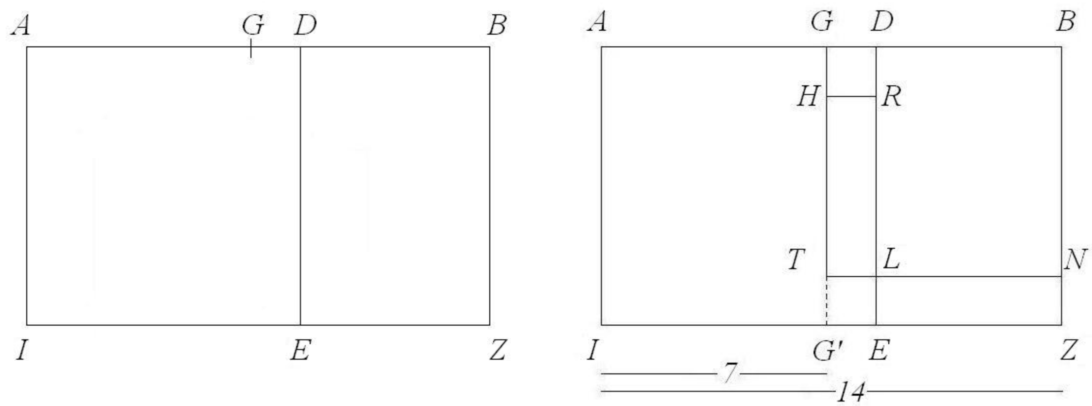

Figura 2.8. Representación geométrica de $x^2 + 40 = 14x$

El *Liber quadratorum* fue escrtito como respuesta a un par de retos formulados por Juan de Palermo, un estudioso de la corte de Federico II, y por Teodoro. Palermo pidió a Fibonacci "encontrar un número cuadrado al que, cuando se le agreguen o sustraigan cinco unidades, quede un número cuadrado". Teodoro le planteó

> [...] encontrar tres número los cuales, sumados al cuadrado del primer número den un número cuadrado. Además, si este cuadrado es sumado al cuadrado del segundo número se produzca también un número cuadrado. Si a este cuadrado se suma el cuadrado del tercer número, resulte similarmente un número cuadrado. (Katz 1993, p. 285)

La obra está formada por 20 proposiciones, todas son propiedades de los números cuadrados que conducen a la solución de los problemas indeterminados propuestos al autor. Algunas proposiciones fueron desarrolladas anteriormente por Nicómaco de Gerasa

(60 - 120 aprox.) en su *Introductio Arithmetica*, pero esto no resta reconocimiento al trabajo de Fibonacci, pues demostró y justificó numéricamente otras proposiciones que Nicómaco nunca demostró.

Los problemas desarrollados en el trabajo de Leonardo de Pisa guardan semejanza con los problemas que solucionó Diofanto en su *Arithmetica* durante el siglo III. La influencia de la obra diofantina no fue directa, pues llegó a Italia a finales del siglo XV y fue publicada por primera vez en 1575 por Xilander (1532 - 1576). Los árabes estuvieron en contacto con la *Arithmetica* desde el siglo X; así, Fibonacci tuvo acceso a los conocimientos de Diofanto (fl. 250 d.C.) por medio de las obras árabes (De Pisa 1973, pp. 17 - 20).

Los maestros de ábaco

En el periodo de 1200 a 1600, en Europa nuevos estudiosos de las matemáticas asimilaron los conocimientos griegos y árabes que penetraron en el continente por medio de las conquistas árabes, adaptándolos a las necesidades de su entorno social. Los maestros de ábaco o abacistas pertenecieron a este grupo. El periodo abacista comprende desde la aparición del *Liber abaci* de Fibonacci en 1202, hasta la publicación del *Summa de arithmetica geometria proportioni et proportionalita* de Luca Pacioli (1445 - 1517) en 1494 (Hadden 1994, p. 89).

Los algoritmos del sistema hindo-arábigo que los abacistas enseñaron a sus estudiantes, así como los diferentes métodos para resolver problemas, fueron tomados de las matemáticas árabes. Los textos que emplearon fueron escritos por ellos mismos en lenguaje vernáculo y constaban de problemas con soluciones, algunos eran problemas auténticos que los estudiantes enfrentaban en los negocios de sus padres, otros eran problemas de recreación, y había algunos relacionados con el calendario y la astrología. De esta manera, los textos de las escuelas de ábaco no sólo servían a los alumnos: eran además una buena referencia para los comerciantes, pues los procedimientos de resolución de los problemas podían seguirse fácilmente sin necesidad de comprender la teoría existente detrás de ellos (Katz 1993, p. 315).

Con el crecimiento del comercio, el uso de las proporciones se convirtió en un recurso indispensable. Los cálculos aritméticos se volvieron más complicados, algunos de los mercaderes intercambiaron sus mercancías, pero otros comenzaron a utilizar dinero (Hadden 1994, p. 89). En estas condiciones, aparecen en el siglo XIV las primeras casas bancarias en Italia. El término *ragione* fue utilizado entonces para referirse a los informes de contabilidad de alguna firma, otros derivados del término se atribuyeron a las prácticas contables o a los encargados de éstas. En 1328 Paolo Gerardi escribió el *Libro di ragioni*, un libro de ábaco con el primer intento de resolución de una expresión que en la actualidad relacionamos con una ecuación cúbica (Hadden 1994, p. 95). La obra contiene 193 problemas escritos de forma retórica, todos de naturaleza comercial, con excepción de los últimos 15 que resuelve sin demostración, a saber, las 6 expresiones cuadráticas que trabajó al-Khwarizmi y 9 expresiones cúbicas de las cuales 5 eran irreducibles. De esta manera, aparecen por primera vez en la literatura de occidente las soluciones, aunque erróneas, de las que hoy conocemos como ecuaciones cúbicas irreducibles; en notación moderna, $ax^3 = bx + N$, $ax^3 = bx^2 + N$ y $ax^3 = bx^2 + cx + N$, lo cual indica que la búsqueda de tales soluciones no comienza en el siglo XVI con la tan mencionada controversia entre Cardano y Tartaglia (Parshall 1988, p. 139).

La propuesta de Gerardi para resolver la cúbica irreducible $ax^3 = bx + N$ fue una aplicación de la ya conocida solución para la ecuación cuadrática; afirmó que la respuesta era

$$x = \sqrt{\frac{N}{a} + \left(\frac{b}{2a}\right)^2} + \frac{b}{2a}\,.$$

A mediados del siglo XIV la ecuación irreducible de cuarto grado apareció en dos nuevos libros de la escuela de ábaco, en el anónimo *Trattato dell'alcibra amuchabile* (1340) y en el *Aliabraa argibra*, del Maesrtro Dardi de Pisa, escrito posiblemente en 1344, constituyendo evidencia de que la investigación de ecuaciones de orden superior tiene una larga tradición (Parshall 1988, p. 139).

Los abacistas crearon términos especiales para referirse a cantidades desconocidas, éstos no se relacionan con los términos que aparecen en los trabajos de Diofanto, pues estuvieron fuera de su alcance. Probablemente los maestros de ábaco tuvieron actividades y tradiciones similares a las de Diofanto que los llevaron a la construcción de un "simbolismo" (Hadden 1994, pp. 92 - 93).

En 1463 aparece el *Trattato di praticha d' arismetrica* del Maestro Benedetto, en el que se introducen abreviaciones de las varias potencias de la incógnita, a saber *cosa* $= \rho$, lo que en nuestro síbolismo es la incógnita x; *c* $=$ *census*, en notación moderna x^2; *b* $=$ *cubo*, es decir, x^3; *cc* $=$ *censo di censo*, en nuestros días x^4; *bc* $=$ *cubo relato cosa*, es decir, x^5 y *bb* $=$ *cubo di cubo cosa* que representaba x^6 (Parshall 1988, p. 139). En 1494 se imprime el primer libro de álgebra y contabilidad, el *Summa de arithmetica geometria* *proportioni et proportionalita*, escrito por Luca Pacioli (Hadden 1994, p. 93). El autor introdujo en la obra abreviaciones de los términos desconocidos así como de sus potencias. La palabra *cosa* (cosa) cambió por *C*, *censo* (cuadrado) cambió por *Ce*, *cubo* (cubo) y *radice* (raíz) fueron sustituidas por *Cu* y *R* respectivamente. Agrega a estas abreviaciones $\bar{p}$ que significa *più* (más) y $\bar{m}$ que quiere decir *meno* (menos) (Parshall 1988, p. 141).

Así, se lograron avances respecto a la forma de nombrar los diferentes elementos empleados en estas cuestiones matemáticas, pero los adelantos respecto a la resolución de los problemas relacionados con las ecuaciones de tercero y cuarto grados tuvieron que esperar. El maestro Benedetto reconoció que la supuesta solución manejada por varias generaciones de abacistas era errónea, y que la solución correcta continuaba siendo eludida.

Scipione del Ferro (1465 - 1526), profesor de la Universidad de Bologna, movido por las ideas de Luca Pacioli resolvió la ecuación cúbica $ax^3 + bx = c$ entre 1500 y 1515. Del Ferro guardó su descubrimiento en secreto según las costumbres de su tiempo, revelándolo sólo a un grupo selecto entre los que se encontraban su yerno Annibale della Nave y su alumno Antonio Maria Fiore. Del Ferro muere en 1526, y en 1530 Fiore lanza un desafío a Zuannnin de Tonini da Coi, un matemático de Brescia, pidiéndole resolver una serie de problemas que implicaban la solución de la ecuación cúbica. Tales desafíos eran muy comunes entre los matemáticos de esa época, pues además de obtener renombre, el ganador de tal prueba pública aseguraba una buena cantidad de alumnos que ofrecían buena paga por clases de alto nivel; de aquí que el poseer un descubrimiento era una ventaja especial para cualquier competidor y por tanto tal hallazgo era guardado celosamente (Parshall 1988, p. 142).

De Tonini, incapaz de resolver los problemas, los proporciona a Niccolò Tartaglia (1499 - 1557), quien concluye que es imposible dicha tarea; en 1535 Tartaglia recibe el reto nuevamente pero de manos de Fiore. La prueba consistía de 30 problemas que requerían la fórmula para resolver la ecuación cúbica. Tartaglia resuelve de manera independiente la ecuación y gana la contienda. Al saber Girolamo Cardano del descubrimiento de Tartaglia, le propone incluir el hallazgo en la obra que preparaba. Tartaglia se niega, pero Cardano publica el resultado en 1545 (Parshall 1988, p. 143).

Asignación del valor de las mercancías como proceso de abstracción

Para comprender mejor por qué las matemáticas actuales presentan la estructura que conocemos es importante mirar hacia atrás, hacia el proceso que las originó. Retomando el estudio de Hadden (1994) sobre cómo se desarrollaron los conceptos necesarios para la creación de la mecánica matemática a partir de las experiencias artesanales, nos guiaremos en obras marxistas como un recurso especial de este historiador para comprender cómo las relaciones sociales producen una apropiaciación particular del conocimiento. En especial dirigiremos nuestra atención al proceso de abstracción que el ser humano realizó en busca de la asignación del valor de los bienes que producía con fines de intercambio, que más tarde daría sentido a la reinterpretación del conocimiento matemático anterior para conformar el álgebra que conocemos actualmente.

Desaparecer los grilletes de la teología y encontrar los fundamentos de las nuevas formas de adueñarse de la naturaleza en beneficio de la sociedad requirió de un nuevo conjunto de relaciones. Las premisas sobre las cuales Karl Marx (1818 - 1883) realizó su investigación dentro de las bases de la sociedad burguesa fueron llamadas por él *premisas reales*. Marx no basó sus estudios en alguna concepción de las cualidades que hacen humanos a los humanos: partió de los humanos mismos afirmando que la presuposición básica de toda la historia es la existencia real y viva (Hadden 1994, p. 35). En este contexto, los seres humanos no conceptualizan las diferencias entre ellos y los animales, ni concluyen tales diferencias bajo observación: los seres humanos hacen la diferencia en virtud de su actividad productiva. La habilidad de producir un excedente conduce a la división del trabajo, de clases y del estado. El acto histórico de producir nuestros medios de subsistencia es lo que nos permite tener una historia (Hadden 1994, p. 36).

Cuando nos preguntamos por qué un principio se manifestó en un siglo o en otro nos estamos forzando a examinar con detalle qué clase de hombres vivieron en aquel siglo, cuáles fueron sus necesidades, sus fuerzas productivas, su modo de producción, sus materias primas; en resumen, necesitamos saber qué relaciones existían entre los hombres y cuáles resultaron de aquellas condiciones en las que vivían. El análisis marxista de la economía moderna parte de la mercancía, el valor de uso, el valor de intercambio, el trabajo y el dinero. Para su estudio, se investigó la forma, la substancia y la magnitud del valor.

El proceso comenzó con el análisis de la mercancía, un producto que satisface una determinada necesidad del hombre y que no se produce para el consumo propio sino con fines de cambio. Así, la cara concreta de la sociedad capitalista, la mercancía, presentó un doble aspecto que debe considerarse: el valor de uso y el valor de cambio. Examinemos un poco más; en *El Capital*, Karl Marx escribió:

> La mercancía es en primer término un objeto externo, una cosa apta para satisfacer necesidades humanas, de cualquier clase que ellas sean. El carácter de estas necesidades, el que broten por ejemplo del estómago o de la fantasía, no interesa en lo más mínimo para estos efectos. Ni interesa tampoco desde este punto de vista, cómo ese objeto satisface las necesidades humanas, si directamente como medio de vida, o indirectamente como medio de producción. (Marx 1982, p. 3)

La propiedad que tiene la mercancía de satisfacer una necesidad humana se denomina *valor de uso*. Por otra parte, en la producción mercantil se realiza un constante cambio de unos valores de uso por otros en función de una determinada relación cuantitativa, dicha relación respecto a la que un valor de uso se cambia por otro representa el *valor de cambio* de la mercancía. En su obra *Miseria de la filosofía* Marx explicó:

> El intercambio tiene su historia. Ha atravesado diferentes fases. Hubo un tiempo, como por ejemplo en la Edad Media, en que no se cambiaba más que lo superfluo, el excedente de la producción sobre el consumo. Hubo luego un tiempo en que no solamente lo superfluo sino todos los productos, toda la vida industrial, pasaron a la esfera del comercio, un tiempo en que la producción entera dependía del cambio [. . .] Por último, llegó un tiempo en que todo lo que los hombres habían venido considerando como inalienable se hizo objeto

> de cambio, de tráfico y podía enajenarse. Es el tiempo en que incluso las cosas que hasta entonces se transmitían, pero nunca se intercambiaban; se donaban, pero nunca se vendían; se adquirían, pero nunca compraban: virtud, amor, opinión, ciencia, conciencia, etc., todo en suma, pasó a la esfera del comercio. Es el tiempo de la corrupción general, de la vanalidad universal, o para expresarnos en términos de economía política, el tiempo en que cada cosa, moral o física, convertida en valor de cambio, es llevada al mercado para ser apreciada en su más justo valor. (Marx 1847, p. 29)

Aunque el valor de intercambio aparezca como algo meramente accidental y cuantitativo, pues por una determinada mercancía se recibe en intercambio otra cantidad de otra mercancía, Marx refiere al respecto:

> [. . .] primero, que los diversos valores de cambio de la misma mercancía expresan todos ellos algo igual; segundo, que el valor de cambio no es ni puede ser más que la *expresión* de un contenido diferenciable de él, su “forma de manifestarse”. (Marx 1982, p. 4)

Pero, ¿qué propiedad de las mercancías determina su valor de intercambio? Sin duda nos referimos a aquello contenido en la mercancía, a la substancia social que encontrará la “forma de manifestarse” en dicho valor (Hadden 1994, p. 41).

> Ahora bien, si prescindimos del valor de uso de las mercancías, éstas sólo conservan una cualidad: la de ser productos del trabajo. Pero no productos de

> un trabajo real y concreto. Al prescindir de su valor de uso, prescindimos también de los elementos materiales y de las formas que los convierten en tal valor de uso. Dejarán de ser una mesa, una casa, una madeja de hilo o un objeto útil cualquiera. Todas sus propiedades se habrán evaporado. Dejarán de ser también productos del trabajo del ebanista, del carpintero, del tejedor o de otro trabajo productivo concreto cualquiera. Con el carácter útil de los productos del trabajo desaparecerá el carácter útil de los trabajos que representan y desaparecerán también, por tanto, las diversas formas concretas de esos trabajos, que dejarán de distinguirse unos de otros para reducirse todos ellos al mismo trabajo humano, al trabajo humano abstracto. (Marx 1982, pp. 5 - 6)

Marx definió el doble carácter del trabajo: trabajo concreto y trabajo abstracto o trabajo humano, lo que justificó el doble carácter de las mercancías y por tanto las bases del valor de las mismas. Sobre el trabajo concreto escribió:

> El trabajo cuya utilidad viene a materializarse así en el valor de uso es lo que llamamos, resumiendo todo eso, *trabajo útil* [. . .] Vemos pues, que el valor de uso de toda mercancía representa una determinada actividad productiva encaminada a un fin o lo que es lo mismo, un determinado trabajo útil. Los valores de uso no pueden enfrentarse los unos con los otros como *mercancías* si no encierran trabajos útiles cualitativamente distintos. (Marx 1982, p. 9)

El trabajo que produce un valor de uso es de tipo diferente al que produce un valor de intercambio. Para producir un valor de uso se aplican movimientos particulares sobre un material particular para transformarlo en satisfactor de una necesidad humana. Para producir un valor de intercambio se requiere únicamente del gasto de trabajo humano abstracto. Para explicarnos qué es el trabajo abstracto, Marx escribió:

> Si prescindimos del carácter concreto de la actividad productiva y, por tanto, de la utilidad del trabajo, ¿qué queda en pie de él? Queda simplemente, el ser un *gasto de fuerza humana de trabajo*. El trabajo del sastre y del tejedor, aun representando actividades productivas cualitativamente distintas, tiene de común el ser un gasto productivo de cerebro humano, de músculo, de nervios, de brazo, etcétera; por tanto, en ese sentido, ambos son trabajo humano [. . .] el valor de la mercancía sólo representa trabajo humano, gasto de *trabajo humano* pura y simplemente. (Marx 1982, p. 11)

Además escribió:

> Recordemos, sin embargo, que las mercancías sólo se materializan como valores en cuanto son expresión de la misma unidad social: trabajo humano, que por tanto, su materialidad como valores es puramente social, y comprenderemos sin ningún esfuerzo que esa materialidad como valores sólo puede revelarse en la relación social de unas mercancías con otras. En efecto, en nuestra investigación comenzamos estudiando el valor de cambio o relación de cambio de las mercancías, para descubrir, encerrado en esta relación, su

> valor. Ahora, no tenemos más remedio que retrotraernos nuevamente a esta forma o manifestación de valor. (Marx 1982, pp. 14 - 15)

La magnitud del valor de una mercancía está determinada por la cantidad de trabajo que ésta encierra. A su vez la inversión de trabajo se mide en tiempo: horas, días, semanas. No todos los productores de determinada mercancía emplean el mismo tiempo en su elaboración, de aquí que no puede medirse la magnitud del valor de un producto de acuerdo con el tiempo que cada productor individual invierta en él. La magnitud del valor de una mercancía se mide por el tiempo de trabajo socialmente necesario para la producción en cuestión. Por tiempo de trabajo socialmente necesario se entiende la cantidad de trabajo que se emplea para producir una mercancía en las condiciones sociales medias de producción (el nivel técnico, la habilidad de los productores, la intensidad de trabajo).

El valor de las mercancías es producto del trabajo invertido en su elaboración; dicho valor sólo puede manifestarse mediante la equiparación de unas mercancías con otras mediante el proceso de intercambio. Pensemos, por ejemplo, en que 20 metros de lino equivalen a un abrigo, es decir, 20 metros de lino = un abrigo. El abrigo sirve de medio de expresión para el valor del lino, la igualdad muestra que en la producción del lino y en la elaboración de un abrigo se ha invertido igual cantidad de trabajo. La mercancía cuyo valor se expresa en otra, en nuestro ejemplo el lino, representa la forma relativa del valor; la mercancía cuyo valor sirve de medio para expresar el valor de otra, en nuestro ejemplo el abrigo, representa la forma equivalente del valor. Marx escribió al respecto:

> Forma relativa del valor y forma equivalente son dos aspectos de la misma relación, aspectos inseparables y que se condicionan mutuamente, pero también y a la par dos *extremos opuestos y antagónicos, los dos polos de la misma expresión del valor*; estos dos términos se *desdoblan* constantemente entre las *diversas* mercancías relacionadas entre sí por la expresión del valor. (Marx 1982, pp. 15 - 16)

El valor del lino en nuestro ejemplo sólo puede expresarse en términos relativos, es decir, recurriendo a otra mercancía. La forma relativa del valor de una mercancía supone como premisa el que otra mercancía cualquiera desempeñe la función de forma equivalente. A su vez, esa otra mercancía, en nuestro caso el abrigo, siendo la forma equivalente no puede desempeñar al mismo tiempo el papel de la forma relativa del valor, pues no es su propio valor lo que él expresa, se limita a ser el material para la expresión del valor del lino.

> [. . .] *para que las magnitudes de objetos distintos puedan ser cuantitativamente comparables entre sí, es necesario ante todo reducirlas a la misma unidad*. Sólo representándolas como *expresiones de la misma unidad* podremos ver en ellas *magnitudes de signo igual* y, por tanto, *conmensurables* [. . .] Es la expresión de equivalencia de diversas mercancías la que pone de manifiesto el carácter específico del trabajo como fuente de valor, al reducir a su nota común, la de trabajo humano puro y simple, los diversos trabajos contenidos en las diversas mercancías. (Marx 1982, p. 18)

En este esbozo de algunos de los principales términos marxistas que explican el sistema capitalista, puede apreciarse de manera clara cómo el ser humano trabajó con razones, proporciones, equivalencias surgidas directamente de su entorno social, empleándolas como un medio para entender su realidad. El desarrollo de una de las principales ramas de las matemáticas, el álgebra, está inminentemente ligado a la historia del desarrollo social del hombre.

CAPÍTULO III

EL PRELUDIO DE LA NUEVA CIENCIA: LAS EXPRESIONES CÚBICAS Y AQUELLAS SIN REFERENTE FÍSICO

Girolamo Cardano y las expresiones cúbicas del *Ars Magna*

Girolamo Cardano nació en 1501 en Pavia y murió en 1576 en Roma. Escribió sobre medicina, astronomía, astrología, filosofía, matemáticas y otras materias. Como escritor y maestro de matemáticas se interesó en exponer todas las ideas que llegaban a sus manos, fueran suyas o no. Esta actitud lo llevó al conflicto con Niccolò Tartaglia (1499 - 1557) por la solución general de la ecuación de tercer grado (Cardano 1968, p. xviii).

En su *De libris propiis* así como en su autobiografía titulada *De propria vita*, Cardano afirmó ser el autor de una gran cantidad de escritos, aunque no todos vieron la luz. La obra más significativa de todo su trabajo fue el *Artis magnae sive de regulis algebraicis liber unus* (1545), comunmente llamado *Ars Magna* que perteneció a una colección de 14 volúmenes que tituló *El trabajo perfecto*, donde trató problemas actualmente relacionados con aritmética, álgebra y geometría. Aunque Girolamo Cardano fue el autor del *Ars Magna*, reconoció que las ideas de tres hombres más aparecen en la obra: la solución de una de las trece posibles formas de la ecuación cúbica pertenece a Scipione del Ferro (1465 - 1526) y a Niccolò Tartaglia, y el método para resolver la ecuación bicuadrática es de Ludovico Ferrari (1522 - 1560), quien fue pupilo y sirviente de Cardano (Cardano 1968, pp. xvi - xvii).

Al comienzo del *Ars Magna* Girolamo Cardano mostró por medio de un pasaje histórico las influencias que lo llevaron a desarrollar sus ideas, sobresaliendo al-Khwarizmi (720 - 840 aprox.), Fibonacci (1170 - 1240) y Luca Pacioli (1445 - 1517) (Parshall 1988, p. 144). Refiriéndose al arte que ahora conocemos como álgebra, expresó:

> Puesto que este arte sobrepasa toda la agudeza humana y la claridad del talento mortal y es un obsequio verdaderamente celestial y una muy clara prueba de la capacidad de las mentes de los hombres, el que lo aplique creerá que nada hay que no pueda comprender. (Cardano 1968, p. 8)

Siendo un filósofo natural, Cardano buscó encontrar los medios que le permitieran conocer y comprender la naturaleza, lo que representó una de las principales limitantes en su trabajo. En el *Ars Magna* escribió: “Como *positio* [la primera potencia de la desconocida] se hace referencia a una línea, *quadratum* [el cuadrado de la desconocida] a una superficie, y *cubum* [el cubo de la desconocida] a un cuerpo sólido, sería muy tonto ir más allá de este punto. La naturaleza no lo permite”. Entonces restringió sus estudios sobre ecuaciones a las de tercer grado a lo más, pues desde su perspectiva son las únicas que tienen sentido y describen la naturaleza permitiendo modelar aspectos del espacio tridimensional (Parshall 1988, p. 144).

En el capítulo XI del *Ars Magna*, *Sobre el cubo y la primera potencia igual al número* (en notación moderna $x^3 + cx = d$), Cardano aclaró que Scipione del Ferro de Bologna descubrió la regla para resolver este problema casi 30 años antes de su publicación, y que Niccolò Tartaglia también la desarrolló de manera independiente. Cardano afirmó haber obteniendo de Tartaglia la solución sin demostración de este problema (Cardano 1968, p. 96). Procederemos a mostrar cómo Cardano resolvió esta ecuación cúbica; para ello nos auxilaremos de la figura 3.1.

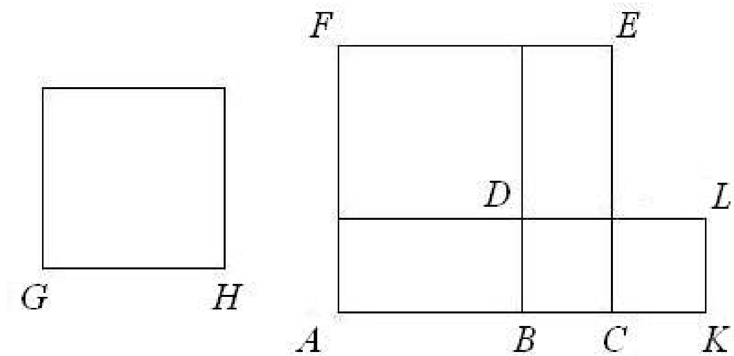

Figura 3.1. Esquema de Girolamo Cardano

> Por ejemplo, supongamos GH^3 más seis veces su lado GH es igual a 20, y supongamos que AE y CL son dos cubos cuya diferencia es 20 y tales que el producto de AC, el lado [de uno], y CK, el lado [del otro], es 2; a saber, un tercio del coeficiente de x [tomando $GH = x$]. Delimitando BC igual a CK, digo que si esto está hecho, la línea restante AB es igual a GH y es, por lo tanto el valor de x, para GH que ha sido dado como [igual a x]. (Cardano 1968, p. 96)

En lenguaje moderno, las hipótesis del problema son:

$$GH^3 + 6GH = 20, \quad AE - CL = 20 \text{ o } AC^3 - CK^3 = 20 \quad \text{y} \quad AC \times CK = 2.$$

Lo que Cardano desea demostrar es que si $BC = CK$ entonces $AB = GH$, siendo $GH = x$. Para resolver este problema Cardano utilizó dos proposiciones que aparecen en el capítulo VI del *Ars Magna*:

Proposición 1. *Si una cantidad es dividida en dos partes, el cubo del todo es igual a los cubos de las dos partes más tres veces los productos de uno y el cuadrado del otro* (Cardano 1968, p. 52).

En notación moderna, establece que si $N = a + b$ entonces $N^3 = a^3 + 3ab^2 + 3a^2b + b^3$. El autor aclaró que esto había sido probado previamente por Euclides y realizó una demostración fundamentada en las proposiciones II.4 y I.43 de los *Elementos*. En la figura 3.2 se ilustra la proposición.

Proposición 2. *[…]* $\mathrm{AB}^3 + 3(\mathrm{AB} \times \mathrm{BC}^2)$ *es mayor que* $\mathrm{BC}^3 + 3(\mathrm{BC} + \mathrm{AB}^2)$ *por el cubo de la diferencia entre* AB *y* BC (Cardano 1968, p. 53).

Esto es, $AB^3 + 3(AB \times BC^2) = BC^3 + 3(BC + AB^2) + (AB - BC)^3$. En la figura 3.2 se ilustra geométricamente este resultado. De aquí se desprende el siguiente corolario.

Corolario. *[…] si* BC *se asume como negativo [o se resta del segmento* AC*],* AB^3 *consistirá de* $\mathrm{AC}^3 + 3(\mathrm{AC} \times \mathrm{BC}^2) + (-\mathrm{BC}^3) + (-3\mathrm{BC} \times \mathrm{AC}^2)$ (Cardano 1968, p. 54).

(Las demostraciones de las proposiciones 1 y 2, así como del corolario pueden verse en el apéndice 3.1 de este capítulo.)

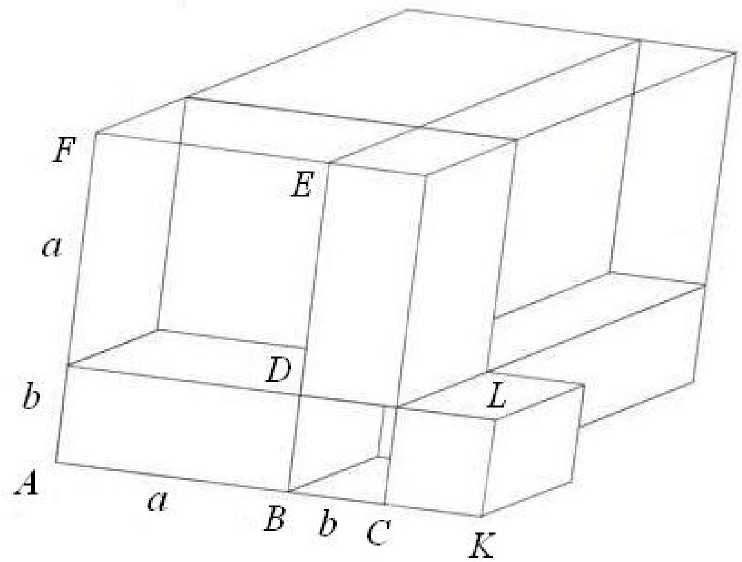

Figura 3.2. Representación geométrica de un binomio al cubo

Regresando al problema, se tienen como hipótesis que $AE - CL = 20$ y $AC \times CK = 2$ (véase la figura 3.1). Con base en la primera proposición del capítulo VI del *Ars Magna*, el cubo de lado $AB + BC$ está formado por los cuerpos BC^3, AB^3, $3(BC \times AB^2)$ y $3(AB \times BC^2)$; obsérvese la figura 3.2. Por hipótesis, $AC \times CK = 2$; luego, $AC \times 3CK = 6$, que es el coeficiente de x. De lo anterior se tiene que

$AB \times 3(AC \times CK) = 6AB$, y puesto que BC es igual a CKse tiene que

$3(AB \times BC \times AC) = 6AB$.

Supongamos ahora que se resta BC del segmento AC; el corolario de la proposición 2 establece que

$$AB^3 = AC^3 + 3(AC \times CB^2) - BC^3 - 3(BC \times AC^2).$$

Por otra parte, se tiene que

$$\begin{aligned} 3(BC \times AC^2) - 3(AC \times BC^2) &= 3(BC \times AC)(AC - BC) \\ &= 3(BC \times AC)(AB) \\ &= 3(AB \times BC \times AC) \\ &= 6AB, \end{aligned}$$

de lo cual se tiene que $3(AC \times BC^2) - 3(BC \times AC^2) + 6AB = 0$.

Como la segunda proposición establece que

$$AB^3 = AC^3 + 3(AC \times BC^2) - BC^3 - 3(BC \times AC^2),$$

al sumar $6AB$ a ambos miembros de la igualdad se obtiene

$$AB^3 + 6AB = AC^3 + 3(AC \times BC^2) - 3(BC \times AC^2) + 6AB - BC^3;$$

así, $AB^3 + 6AB = AC^3 - BC^3$, puesto que $3(AC \times BC^2) - 3(BC \times AC^2) + 6AB = 0$.

Además, una de las hipótesis del problema es que $AE - CL = 20$, es decir,

$AC^3 - CK^3 = 20$ o $AC^3 - BC^3 = 20$. Con esto se tiene que $AB^3 + 6AB = 20$ y, como $GH^3 + 6GH = 20$, de las proposiciones I.35 y XI.31 de los *Elementos* se tiene que $GH = AB$ (Cardano 1968, pp. 97 - 98).

Del procedimiento anterior el autor del *Ars Magna* obtuvo la siguiente regla para resolver este tipo de ecuaciones:

> Al cubo de un tercio del coeficiente de x se suma el cuadrado de la mitad de la constante de la ecuación y se toma la raíz cuadrada de todo. Se hará esto dos veces, a uno de los dos se le suma la mitad del número que ha sido elevado al cuadrado y al otro se le resta la mitad del mismo. Se tendrán un *binomium* y su *apotome* [por *binomium* se entiende una expresión de la forma $a + b$ y por *apotome* una expresión del tipo $a - b$]. Se resta la raíz cúbica del *apotome* de la raíz cúbica del *binomium*, el resto [o] lo que queda es el valor de x.
> (Cardano 1968, pp. 98 - 99)

Aplicando este razonamiento en la ecuación $x^3 + 6x = 20$ del ejemplo de Cardano, se obtiene que el valor de x es $\sqrt[3]{\sqrt{108} + 10} - \sqrt[3]{\sqrt{108} - 10}$. La regla para resolver una ecuación cúbica del tipo $x^3 + cx = d$ se obtiene al resolver el sistema de ecuaciones

$$\begin{aligned} u - v &= d \\ uv &= \left(\frac{c}{3}\right)^3, \end{aligned}$$

donde u y v son dos cubos cuya diferencia es la constante de la ecuación, y cuyos lados multiplicados entre sí producen un tercio del coeficiente de la cantidad desconocida. Así, Cardano dio como solución de la ecuación $x^3 = cx + d$ la expresión

$$x = \sqrt[3]{\sqrt{\left(\frac{d}{2}\right)^2 + \left(\frac{c}{3}\right)^3} + \frac{d}{2}} - \sqrt[3]{\sqrt{\left(\frac{d}{2}\right)^2 + \left(\frac{c}{3}\right)^3} - \frac{d}{2}}\,.$$

Cardano adoptó el esquema de demostración geométrica que sus antecesores siguieron, aunque lo combinó con argumentos que actualmente serían calificados como puramente algebraicos. Sin embargo, rechazó ese esquema para tratar con ecuaciones de grado superior. Nótese que para resolver la ecuación cúbica Cardano completó un cubo de manera similar a como al-Khwarizmi completó un cuadrado, pero no concibió una figura de cuatro dimensiones y, con mayor razón, le fue imposible completarla. Pensó disparatado usar geometría para resolver la ecuación irreducible de cuarto grado que su estudiante Ludovico Ferrari resolvió con éxito; entonces, al creer que la geometría le quedaba a deber, se inclinó por confiar sólo en los algoritmos algebraicos (Parshall 1988, p. 146).

La resolución de las ecuaciones de cuarto grado marcó un creciente interés por los algoritmos en este tipo de problemas; las cuestiones geométricas ya no fueron consideradas como método prioritario para resolverlos.

La *Arithmetica* de Diofanto y su influencia en el trabajo de Rafael Bombelli

La Arithmetica *de Diofanto de Alejandría*

Aproximadamente en 1463, Johannes Müller (1436 - 1476), mejor conocido como Regiomontano, decubrió seis de los libros de la colección titulada *Arithmetica* escrita por Diofanto de Alejandría alrededor del año 250. Regiomontano tradujo al latín los seis libros que sobrevivieron en griego (Parshall 1988, p. 149). Cuatro libros más fueron descubiertos en versiones árabes, pero el estilo de éstos difiere un poco de los libros griegos. Los libros árabes explican con más detalle los pasos que deben seguirse para encontrar las soluciones de los problemas que ahí se presentan. Es muy probable que estos libros no sean traducciones de los originales de Diofanto, sino de comentarios sobre la *Arithmetica* escritos por Hypatia (h. 370 - h. 415) en el año 400 aproximadamente (Katz 1993, p. 163). Diofanto introdujo en su trabajo letras para representar las cantidades que solucionaban sus problemas propuestos. Este hecho es comunmente ligado al surgimiento del simbolismo algebraico, pero faltaban muchos siglos aún para su nacimiento; sin embargo, el trabajo de Diofanto desempeñó un papel importante para que el álgebra adquiriera su forma actual.

Generalmente, la evaluación que Paul Tannery hizo de la obra de Diofanto y que se menciona en el trabajo de Jacob Klein, *Greek mathematical thought and the origin of algebra* (1992), es bien aceptada por los historiadores de las matemáticas. En ella, el trabajo de Diofanto es considerado como original o, en el peor de los escenarios, como basado en muy pocos predecesores. El punto más importante de esta originalidad, según Tannery, reside en la nueva concepción que Diofanto tuvo del objeto de estudio de la logística griega. Se incluye una breve descripción de la logística griega en el apéndice 3.2 de este capítulo. Klein analizó la obra de Diofanto desde otra perspectiva.

Se considera que Diofanto elevó la logística al rango de verdadera ciencia al introducir en ella la demostración; es decir, hizo de la logística una ciencia *apodeictica*. Además, los números con los que trabajó Diofanto fueron abstractos y no concretos, como los empleados por los estudiosos de la logística anteriores a él. En sus investigaciones Tannery llegó a la conclusión de que el carácter abstracto de los números de Diofanto tuvo un origen religioso. Según Tannery, Diofanto fue cristiano y su obra fue utilizada como libro de texto de la escuela ecuménica de Alejandría. De acuerdo con las costumbres cristianas, no se podían enseñar los problemas verbales de orden matemático en términos paganos o mitológicos. Por este motivo surgieron los términos abstractos empleados en toda la obra diofantina, con excepción del problema 13 del libro quinto, en el que se empleó una formulación verbal que no aludía a la mitología (Klein 1992, pp. 128 - 129). Sin embargo, Tannery resolvió el problema de explicar cómo surgió la abstracción en la logistica diofantina de manera muy superficial, encontró una solución de forma pero no de fondo. ¿Qué era un número para Diofanto y cómo logró expresarlo de manera general? El argumento de Tannery no contestó esta pregunta.

En la Definición I de la *Arithmetica*, Diofanto escribió: “Todos los números consisten de determinada multitud de mónadas”. En sus cálculos las cantidades conocidas son consideradas como números de mónadas puras, el signo que representa a las mónadas es el prefijo $\overset{\circ}{M}$, por ejemplo $\overset{\circ}{M}\overline{\delta}$ quiere decir 4 unidades. El uso del signo $\overset{\circ}{M}$ se relacionó con el significado de *arithmos* que, en el habla común de la época, significaba determinado número de determinadas cosas; de manera más elaborada diremos que significa un número de unidades de medida o de mónadas “puras” o “neutrales” (Klein 1992, p. 130).

Según Theon de Alejandría (fl. finales del s. IV d.C.), en su carácter de medida la mónada no puede sufrir partición, pues es invariable y siempre constante (Klein 1992, p. 137), lo que se divide es algo corpóreo o geométrico. Por otra parte, para Diofanto (como en toda la matemática griega) *arithmos* significaba “un número de”, y aplicó esta definición en su concepción de fracción. Diofanto entendió una fracción como un número de partes o una multitud de partes de una mónada; así, para él las mónadas fueron piezas de medida que admitían división (Klein 1992, p. 137). Lo anterior llevó a Diofanto a concluir que si se multiplicaba una parte fraccionaria de una unidad por el número de divisiones que sufrió dicha unidad, la unidad original estaría reconstruida. Una observación importante es que la parte fraccional de una mónada puede ser entendida como una nueva mónada.

Algo digno de distinguir es que desde el concepto de número de Diofanto es completamente imposible llegar al concepto de número negativo y número irracional. Así, una ecuación con solución negativa o irracional era una ecuación “imposible” o “no enunciable”, pues un *arithmos* irracional no era precisamente un *arithmos* del mismo modo que un *arithmos* negativo (Klein 1992, p. 138).

Por otra parte, Diofanto empleó en la *Arithmetica* números de orden superior (potencias) que fueron llamados con anterioridad por Arquímedes (287 - 212 a.C.) “números [formados] análogamente a aquéllos basados sobre mónadas” y por Apolonio (fl. finales del s. III - principios s. II a.C.) “números análogos” o “análogos”. Respecto a estos números, escribió:

> [. . .] Entre ellos están los *cuadrados*, que se forman cuando cualquier número es multiplicado por sí mismo; el número es llamado *lado del cuadrado*; *cubos*, que se forman cuando cuadrados se multiplican por sus lados; *cuadrado-cuadrados*, que se forman cuando cuadrados son multiplicados por ellos mismos; *cuadrado-cubos*, que se forman cuando cuadrados son multiplicados por cubos formados por el mismo lado; *cubo-cubos*, se forman cuando los cubos son multiplicados por sí mismos.
>
> (Tomado de: Katz 1993, p. 163)

A cada uno de los números anteriores Diofanto le asignó un nombre abreviado: el cuadrado de una cantidad desconocida fue llamado *dynamis* y se representó con Δ^Y, el cubo fue llamado *kubos* y se representó con K^Y, el cuadrado de un cuadrado recibió el nombre de *dynamo-dynamis* representado por $\Delta^Y\Delta$, el cuadrado multiplicado por el cubo fue *dynamo-kubos* y lo representaron las letras ΔK^Y, las letras $K^Y K$ representaron el cubo multiplicado por sí mismo, es decir el *kubo-kubos*. Diofanto afirmó que al sumar, restar, multiplicar o dividir estas magnitudes se generaban nuevas magnitudes (Katz 1993, p. 163). Las cantidades desconocidas y sus potencias también aparecieron como denominadores de las fracciones en el trabajo diofantino, éstas obtuvieron su nombre de la misma cantidad desconocida. Por ejemplo, la "fracción cúbica" es el recíproco de la desconocida cúbica (Klein 1992, p. 132). Diofanto utilizó la letra χ para los recíprocos; por ejemplo, $\Delta^{Y\chi}$ quiere decir $\frac{1}{x^2}$.

La obra de Diofanto es considerada parte de la logística teórica; se tiene que poner énfasis en que el dividir y manipular la unidad no implica que dicha logística pierda su caracter teórico o apodeictico. Por otra parte, el proceso de resolución de cada problema de la *Arithmetica* es estrictamente metódico en cada caso, lo que le imprime un carácter general; es decir, se puede emplear dicho procedimiento en otros problemas similares, aunque la dificultad de los cálculos varía según los números determinados con los que se trate.

Una de las principales características, y por la que podría considerarse al trabajo de Diofanto como "álgebra primitiva", es que planteó problemas de naturaleza indeterminada, en nuestros días llamados *ecuaciones diofantinas*, los cuales siempre son transformados por el autor en ecuaciones determinadas a través de suposiciones numéricas que permiten una solución única aunque no siempre entera. La debilidad del trabajo de Diofanto a la luz de nuestros días radica en que no se propuso llegar a soluciones generales (Klein 1992, pp. 133 - 134). De lo anterior se concluye que Diofanto conoció un tipo general de problemas pero no un objeto matemático general. El objetivo de Diofanto fue encontrar números determinados en cada uno de sus problemas, ya fueran números enteros o fracciones de la unidad. Se debe distinguir estrictamente entre el procedimiento y el objeto: el procedimiento es aplicado a las *eide* (ideas), que son como tales independientes de cada multitud de mónadas y por tanto dicho procedimiento se vuelve "general", pero el objeto entendido en cada caso es un número determinado de mónadas (Klein 1992, pp. 144 - 145).

Por otra parte, a Diofanto no le interesó organizar el conocimiento alrededor de los diferentes tipos de ecuaciones y sus métodos de resolución, como sucede actualmente al estudiar estas cuestiones: él buscó las posibles relaciones que los números cuadráticos y cúbicos tenían con sus raíces. Esto coloca a la *Arithmetica* como una obra no de carácter logístico sino aritmético, en la que la mónada aparece como unidad de cálculo “abstraida” de las unidades de medida (Klein 1992, p. 135).

La asimilación de los trabajos diofantinos fue importante para el surgimiento del álgebra de nuestros días y particularmente para la resolución de problemas en los que estaban implicadas lo que hoy conocemos como ecuaciones cuadráticas y de grado mayor que 2. La ruta histórica pasa por los árabes, Leonardo de Pisa, los cosistas, Cardano (1545), Tartaglia (1499 - 1560) y Rafael Bombelli (1526 - 1572), por mencionar a algunos. A lo largo de este camino ocurrieron restructuraciones respecto a las técnicas de cálculo, introduciendo los números negativos, los irracionales y los imaginarios. Sin embargo, se tuvo conciencia del caracter “científico” de este proceso hasta que estuvo en contacto con la *Arithmetica* de Diofanto a través de la obra de Rafael Bombelli.

Las aportaciones de Rafael Bombelli

Rafael Bombelli (1526 - 1572) fue un ingeniero y arquitecto sin instrucción formal en matemáticas. Consideró que el *Ars Magna* de Cardano era un libro con argumentos difíciles de seguir, y con el objetivo de elaborar un tratado sobre estas cuestiones más claro y profundo escribió el primer borrador de su obra *L’algebra parte maggiore dell’arithmetica divisa in tre libri* entre los años 1557 y 1560. Apareció impresa en 1572 y se convirtió en el punto máximo alcanzado por las matemáticas del Renacimiento (Parshall 1988, p. 152).

Con el trabajo de Rafael Bombelli se inició el camino hacia la creación de un verdadero simbolismo matemático. En esta obra el autor expuso el trabajo realizado por Diofanto aproximadamente en el año 250. Así, gracias a que Tartaglia y Bombelli emplearon las obras de Arquímedes (287 - 212 a.C.) y de Diofanto en la construcción de sus conocimientos, los europeos del siglo XVI puedieron apreciar las matemáticos griegas (Hadden 1994, p. 112).

La obra original de Bombelli sufrió modificaciones importantes. La primera versión del *Algebra* cambió cuando Antonio Maria Pazzi mostró a Bombelli una copia de la *Arithmetica* de Diofanto encontrada en la Biblioteca del Vaticano. En su forma original el Libro III contenía problemas prácticos sobre intercambio, asociaciones, intereses y tiempo laborable. Bombelli transformó este libro después de leer la *Arithmetica*; no consideró "acciones humanas", tal como las obras de autores contemporáneos de matemáticas; sustituyó los problemas prácticos por 143 problemas tomados de Diofanto. De este modo Rafael Bombelli pretendió enseñar aritmética superior a la manera de los antiguos. En su obra Bombelli reemplazó la *cosa* y el *censo* empleados hasta entonces en los trabajos matemáticos por el *tanto* y la *potenza* utilizados por Diofanto. Representó las potencias de la cantidad desconocida por medio de un semicírculo dentro del que se colocó el exponente; por ejemplo, $\underset{\smile}{1}$ representó x, $\underset{\smile}{2}$ representó x^2 y 7 $\underset{\smile}{1}$ fue $7x$ (Hadden 1994, pp. 112 - 113). Usó $R.q.$ para denotar raíz cuadrada, $R.c.$ para la raíz cúbica y otras expresiones similares para raíces superiores. Incluyó $\lfloor\ \rfloor$ como paréntesis para encerrar expresiones largas; por ejemplo, $R.c.\ \lfloor 2p\, R.q.\, 21 \rfloor$, y conservó los símbolos p y m para más y menos respectivamente (Katz 1993, p. 335).

Cuando Bombelli trabajó con la expresión cúbica del tipo $x^3 = cx + d$ consideró cantidades negativas bajo un radical; así, es el primero que reconoció y operó los números complejos, muy a su pesar. En la primera parte de su *Algebra* se refirió a ellos como "otro tipo de raíces cúbicas muy diferentes de las anteriores, las cuales vienen del capítulo sobre el cubo igual a cosa y número; [...] esta clase de raíces tiene sus propios algoritmos para varias operaciones y un nuevo nombre" (Katz 1993, p. 335). A los números conocidos actualmente como complejos Bombelli les llamó *piú di meno* (más de menos) y *meno di meno* (menos de menos). De este modo escribió el número complejo $2 + 3i$ como $2\,p\,di\,m\,3$ y el número $2 - 3i$ como $2\,m\,di\,m\,3$. Bombelli presentó por primera vez las reglas para sumar, restar, multiplicar y dividir estos nuevos números, pero escribió: "el asunto entero parece descansar sobre un sofisma más bien que sobre la verdad" (Katz 1993, pp. 335 - 336).

Algo digno de considerar es el contexto en el que surge la obra de Rafael Bombelli. El *Algebra* fue escrita antes de que su autor leyera la *Arithmetica*. Este hecho constituye evidencia de que el trabajo de Bombelli no se basó en la obra de Diofanto. El *Algebra* nació en un contexto social y económico específico en una cultura comercial, lo que permitió a su autor hacer una lectura moderna del "simbolismo" diofantino. De este modo, los objetos matemáticos abstractos empleados en las matemáticas actuales tomaron forma en el contexto del cálculo comercial (Hadden 1994, p. 114).

Los orígenes de la nueva ciencia. Viète como proveedor de un nuevo lenguaje

Desde el punto de vista de los griegos de la antigüedad, el hombre puede conocer cuando se dedica a la contemplación de lo que le rodea en completa libertad y ocio, y es a través de esta actividad que puede lograr la felicidad. De aquí que la ciencia se colocó en franca oposición a las tareas de la vida cotidiana que esclavizan al individuo. Sin embargo, en ellas reconoció sus propias raíces (Klein 1992, p. 118), pues es en la realización de las mismas que surge la curiosidad y el deseo de comprender el porqué de lo que se ejecuta.

La ciencia griega tuvo su punto de partida en la naturaleza, aunque surgió como tal fuera de esos fundamentos cuando el hombre comenzó a razonar respecto a lo que le rodeaba, entregándose al mundo de las ideas. Así, la ciencia griega está definida en términos de su distinción y su origen en esos fundamentos naturales. La creación de sus conceptos dependió de lo natural, de las experiencias precientíficas desde las cuales el concepto científico fue abstraído. El carácter del concepto, y por tanto de la abstracción, estuvo determinado por el problema ontológico que apremió las investigaciones de la antigüedad (Klein 1992, p. 120).

La ciencia de nuestros tiempos, llamada *la nueva ciencia*, se cimentó durante los siglos XVI y XVII. En esa época los fundamentos naturales de la ciencia antigua fueron remplazados por una ciencia ya en existencia, la misma ciencia griega, pero sus principios y métodos fueron rechazados. Los fundadores de la nueva ciencia, hombres como Galileo Galilei (1564 - 1642), Simon Stevin (1548 - 1620), Johannes Kepler (1571 - 1630) y René Descartes (1596 - 1650), fueron impulsados por problemas de interés práctico relacionados con mecánica aplicada, arquitectura, construcción de máquinas, pintura y la construcción de instrumentos ópticos. Sin embargo, la estructura conceptual de sus ideas se derivó de los conceptos tradicionales (Klein 1992, p. 119).

La nueva ciencia regresó a las fuentes griegas que fueron negadas por la anterior ciencia escolástica basada en una filosofía cristiana; sin embargo estas fuentes fueron interpretadas desde una base ajena a la de la antigüedad. Esta nueva interpretación de la doctrina griega condujo a la transformación de sus conceptos generando los cimientos de la nueva ciencia así como de su intención. Sus conceptos fueron obtenidos en su polémico proceso contra la ciencia escolástica y no tuvieron un significado natural, ya no representaron una idea inmediata sino que cobraron sentido a través de las conexiones con otros conceptos, es decir, a través de sus relaciones mutuas y su subordinación total al nuevo edificio de la ciencia (Klein 1992, p. 120).

Uno de los principales rasgos de la nueva ciencia fue la relación que guardaban su carácter de arte y la generalización de sus métodos. Las expresiones más significativas de esa conexión son el formalismo simbólico y las técnicas de cálculo de las matemáticas modernas. Esa relación entre la generalización y el carácter de arte determinó la interpretación moderna de las matemáticas antiguas. Un ejemplo de esto es el término *álgebra geométrica*, con el que se hace referencia al modo antiguo de presentar los hechos matemáticos. Esta interpretación surge de una distinción insuficiente entre la generalidad del método y la generalidad del objeto de investigación (Klein 1992, p. 122). El concepto de *álgebra geométrica* se relaciona inmediatamente con el concepto de *magnitud general*; esta visión de las matemáticas antiguas choca con la verdadera concepción griega. Los antiguos centraban sus investigaciones en preguntas relacionadas con el ser de los objetos matemáticos, que podían ser figuras o curvas geométricas, proporciones de magnitudes conmensurables o inconmensurables, y razones, entre otros. Estos objetos marcaban la dirección de la investigación así como el punto de partida y el punto final de la misma. De este modo, el problema de la *aplicación general de un método* representaría para los antiguos el problema de la *generalidad de los objetos matemáticos*; este problema puede resolverse sólo sobre la base de una ontología de los objetos matemáticos (Klein 1992, pp. 122 - 123).

De manera contrastante las matemáticas modernas, y por tanto la interpretación moderna de las matemáticas antiguas, centran su atención de principio a fin en el método. Los objetos quedan determinados por la forma en que éstos se vuelven accesibles a través de un método general. Uno de los principales representantes de este nuevo movimiento fue François Viète (1540 - 1603). En 1591 apareció su obra más importante, *In artem analyticem isagoge seorsim excussa ah opere restitutae mathematicae analyceos seu algebra nova* mejor conocida como *Introducción al arte analítico* (Piaget y García 1982, p. 138). Las principales fuentes griegas del *arte analítico* de François Viète fueron el *libro séptimo* de Pappo (s. IV) y la *Arithmetica* de Diofanto.

De la obra de Pappo estudió sus ideas sobre el análisis. Viète escribió que en matemáticas existe un procedimiento especial para el descubrimiento, "una vía de inquisición de la verdad" que fue descubierta primero por Platón (h. 427 - h. 348 a.C.). De acuerdo con Viète, Theón de Alejandría llamó a ese procedimiento *análisis* y lo definió como un proceso que inicia con "la suposición de lo que es buscado como si fuera admitido y por medio de las consecuencias [se prosigue] a una verdad [la cual fue de hecho ya admitida]" (citado en: Klein 1992, p. 155). Por otra parte definió su procedimiento recíproco, la *síntesis*, como un proceso que comienza con "la suposición de lo que es admitido y por medio de las consecuencias [se prosigue a] la conclusión y comprensión de lo que es buscado" (Klein 1992, p. 155).

Estas definiciones se localizaron de manera modificada en el *libro séptimo* de Pappo. De acuerdo con la aplicación del análisis, es decir, si se empleaba para encontrar la demostración de un teorema o para resolver un problema mediante una construcción geométrica, Pappo disitinguió dos tipos de análisis: el análisis *zetético* o también llamado *teórico* que sirve para investigar la verdad de un teorema, y el análisis *porístico* o *problemático* que proporciona lo que es requerido para resolver un problema. Para Pappo la diferencia entre análisis *teórico* y análisis *problemático* es el tipo de objetos presentados en un *teorema* y en un *problema* (Klein 1992, p. 155).

Si se desea convertir el análisis *zetético* en *síntesis* el procedimiento es directo, basta con invertir el orden de los pasos ejecutados. En cambio, para transformar el análisis *porístico* en síntesis se requiere de una construcción geométrica previa para apoyarse. Sin importar la diferencia entre análisis *teórico* y *problemático*, en ambos casos Pappo llamó a las respectivas *síntesis* simplemente *apodeixis*, e inversamente consideró la *apodeixis* (o demostración) la conversa del *análisis* (Klein 1992, pp. 155 - 156).

De lo anteriormente expuesto, se tiene que el *análisis* concierne inmediatamente a la generalidad del procedimiento, y la *síntesis* está obligada, de acuerdo con la concepción griega de los objetos matemáticos, a realizar este procedimiento general en un objeto determinado. El *análisis* muestra la posibilidad de una demostración o construcción; un teorema puede considerarse probado sólo cuando es derivado de las relaciones *dadas* entre las magnitudes *dadas* en él, aunque la *dades* de la que hace uso el análisis es entendida sólo como una *posible dades*. Esta posibilidad aparece en el análisis geométrico en la construcción efectuada que no necesita usar las magnitudes que aparecen en él como determinadas sino sólo con carácter de ser posibles. Esta situación puede transferirse al análisis *aritmético* cuando los números dados en un problema son entendidos en su carácter de posibles y no justo como aquellos números determinados. En otras palabras, para comparar el análisis aritmético con el geométrico los números *dados* deben permitirse con algunas indeterminaciones limitadas sólo por la condición de posibilidad del problema (Klein 1992, pp. 163 - 164).

Así, en el análisis *aritmético* una solución indeterminada debe ser posible en cualquier caso y se podrán distinguir tres etapas en el proceso de solución:

1.- La construcción de la ecuación.

2.- La transformación a la que se sujeta hasta adquirir una forma canónica que proporciona inmediatamente la solución indeterminada.

3.- La explotación numérica de esta última, es decir, el cálculo de los números determinados que cumplen las condiciones puestas en el problema.

Sólo las primeras dos etapas representan el procedimiento propiamente *analítico*; la última etapa pertenece a la *síntesis*, pues se corresponde con la construcción geométrica en el análisis *geométrico*. En Diofanto la conversión de cada solución, que tiene la finalidad de mostrar que los números encontrados cumplen las condiciones del problema, fue llamada *apodeixis*.

Después de comparar el análisis problemático geométrico con el análisis aritmético de Diofanto en la manera antes mencionada, Viète concibió un modelo de cálculo que se realizaba en términos de *especies* de números llamado *logistice speciosa* o *logística especiosa*, en contraste con el cálculo sobre números llamado *logistice numerosa* o *logística numerosa*. De este modo la *logística especiosa* está estrechamente conectada con el procedimiento de Diofanto que a su vez es el análogo aritmético del análisis geométrico.

Expliquemos un poco a qué se refirió Viète con el concepto de *especies*. Existen algunas interpretaciones históricas que señalan a Diofanto como el creador del álgebra simbólica. Es verdad que Diofanto introdujo letras a modo de *"símbolos"* para representar las incógnitas de los problemas artiméticos que trató, pero dichos *"símbolos"* representaron una cantidad específica en un lugar específico, y no fueron factibles para las operaciones y mucho menos para la representación de un "objeto general" (Piaget y García 1982, p. 137). Sin embargo, la idea de fondo que utilizó Diofanto fue muy bien empleada por Viète.

En su trabajo, Diofanto representó a los números desconocidos aunque en relidad eran números "determinados" por su *eidos*, su idea, su especie. De este modo, cuando las *eide* de la desconocida y sus potencias aparecieron como nuevas unidades de cálculo, mientras el cálculo real se estableció en términos de números determinados, el cálculo que se ejecutó con las *especies* de los números cambió completamente de dominio, se realizó en el dominio de lo indeterminado. Este paso fue crucial para el desarrollo del trabajo de Viète, quien lo consideró como una práctica ampliamente usada entre los antiguos y encontrada especialmente en Diofanto, aunque no suficientemente aclarada por ellos, lo que le permitió reinterpretar la tradición.

Diofanto hizo uso instrumental del *arithmos* o *eidos* platónico, es decir de aquello que hace posible el ser unificado de cada número y que hace referencia inmediata a las cosas o unidades. Con el concepto de *especies*, Viète llegó a la realización simbólica del *eidos*, pues con el concepto de *especies* se refiere a la propiedad de "ser número" en general, propiedad que además de pertenecer a cada número pertenece a las cosas u objetos cuya "numerosidad" está representada por un número. De este modo surgió un concepto que se refiere a otro concepto y no a un ser. Con esta interpretación se puede representar lo que se busca y lo que se tiene para buscarlo, es decir, la fórmula se hace posible y con el surgimiento de la misma se crea una nueva manera de alcanzar el conocimiento.

A los dos tipos de análisis explicados por Pappo, el análisis *teórico* o *zetético* y el análisis *problemático* o *porístico*, Viète agregó un tercer tipo al que llamó *rético* o *exegético*. Para Viète el análisis *zetético* es un procedimiento a través del cual se construye la ecuación con las magnitudes dadas, el análisis *porístico* es el procedimiento por medio del que la verdad del teorema se investiga a través de la ecuación obtenida en la *zetética*, y finalmente por medio del análisis *rético* la magnitud buscada es producida por la ecuación establecida. La magnitud requerida es un número determinado y por tanto comunicable al habla, o una magnitud geométrica visible y medible. De aquí el doble nombre del tercer tipo de análisis: *rético*, porque los números pueden ser expresados a través de sus nombres ordinarios en nuestro lenguaje; y *exegético*, porque pueden representarse por medio de magnitudes geométricas disponibles a la vista (Klein 1992, p. 167).

Lo relevante y crucial del arte de Viète queda establecido en la siguiente cita, en la que se refirió a las *especies*:

> Ahora bien, lo que pertenece realmente al arte zetético es establecido por el arte de la lógica a través de silogismos o entimemas, cuya fundamentación está dada por las estipulaciones mismas (símbolos) por medio de las cuales se llega a las ecuaciones y proporciones, estipulaciones que a su vez se derivan de nociones comunes como también de teoremas que son demostrados por el poder del análisis mismo. En el arte zetético, sin embargo, la forma de proceder es peculiar al arte mismo, en tanto el arte zetético no emplea su lógica en los números – que fue el tedio de los analistas antiguos – sino que usa su lógica a través de una logística que tiene que ver, en una forma que es nueva, con *especies*. Esta logística es mucho más eficaz y poderosa para comparar magnitudes entre sí que la logística numérica, una vez que la ley de homogeneidad ha sido establecida. (Citado en: Piaget y García 1982, p. 139)

Viète dedicó la *logística especiosa* al servicio del álgebra pura, aplicada de forma indiferente a números y magnitudes geométricas; el concepto de *especies* sufrió una extensión, aunque preservó su vínculo con el reino de los números. Las *especies* o *forma de las cosas*, como también las llamó Viète, representaron simplemente magnitudes generales. Recordemos que para Pappo la diferencia entre análisis *teórico* y *problemático* reside en el tipo de objetos presentados en cada caso. Viète distinguió muy externamente entre estos dos tipos de análisis ya que sus reflexiones se centraron en los procedimientos; consideró todos los teoremas como problemas (Klein 1992, pp. 165 - 166). Por otra parte, Viète se interesó menos en las verdades encontradas, por ser verdades, que en la forma de encontrarlas correctamente; definió el "arte analítico" como la "teoría para encontrar lo que es buscado en matemáticas en general".

Para Viète, su *arte analítico* representó un álgebra tanto geométrica como aritmética y consideró que dejaba atrás todos los estudios sobre estas cuestiones realizados con anterioridad. Aunque para Viète el principal objetivo de su obra no fue encontrar construcciones geométricas o números, sino encontrar los números "posibles" que, según él, eran aquellos que podían interpretarse geométricamente (Klein 1992, p. 159). Viète concluyó el capítulo V de su obra, donde aborda estas reglas, con lo siguiente:

> Diofanto, en aquellos libros que tenían que ver con la aritmética, empleó la zetética en la forma más sutil. Pero él lo presentó como si fuera establecido por medio de números y no también por especies (las cuales fueron, sin embargo, usadas por él), a fin de que su sutileza y habilidad pudiera ser admirada, en tanto aquellas cosas que parecen más sutiles y ocultas para quien usa consideraciones acerca de números (*logistice numerosa*) son enteramente comunes e inmediatamente obvias para quien usa consideraciones acerca de especies (*logistice speciosa*). (Piaget y García 1982, pp. 140 - 141)

En este capítulo hemos observado cómo las expresiones cuadráticas y cúbicas cuya resolución está ligada a la percepción abrieron paso a expresiones que no fueron asociadas a ningún referente físico. Se ha esbozado cómo los conocimientos heredados de una tradición social comercial adquirireron carácter científico a la luz de los conocimientos griegos, particularmente, con el redescubrimiento de las obras diofantinas estudiadas con anterioridad por los árabes pero poco difundidas a causa de la caída de su imperio y la destrucción de las bibliotecas. Además, se muestran las principales raíces del simbolismo algebraico actual. El estudio de la resolución de las ecuaciones cuadráticas invita a conocer cómo es que el álgebra actual adquirió su estructura.

Apéndice 3.1. Dos proposiciones y un corolario del *Ars Magna*

Proposición 1. *Si una cantidad es dividida en dos partes, el cubo del todo es igual a los cubos de las dos partes más tres veces los productos de uno y el cuadrado del otro* (Cardano 1968, p. 52).

En notación moderna esta proposición dice que

$$\text{si } N = a + b \text{ entonces } N^3 = a^3 + 3ab^2 + 3a^2b + b^3.$$

Cardano afirmó que esta proposición había sido demostrada en el libro VII de los *Elementos* de Euclides, y decidió incluirla en su trabajo. Considérese la figura 3.1.1.

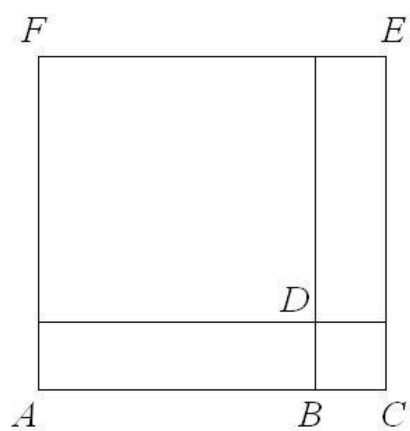

Figura 3.1.1. Cuadrado AE

Sea AC un segmento de recta dividida en dos partes por el punto B, y sea AE el cuadrado formado a partir de AC. Con base en la proposición 4 del libro II de los *Elementos* de Euclides, que establece que

> Si se corta al azar una línea recta, el cuadrado de la (recta) entera es igual a los cuadrados de los segmentos y dos veces el rectángulo comprendido por los segmentos. (Euclides 2000, p. 90)

AE estará integrado por los rectángulos DA y DE, que son el producto de AB y BC; por DF, el cuadrado de AB; y DC, el cuadrado de BC.

Pensemos ahora en el cubo de AC. Estará formado por AC veces el cuadrado AE; es decir, estará integrado por AC veces las superficies DA, DE, DF y DC. Puesto que $AC = AB + BC$, se tiene que AC^3 estará formado por 8 cuerpos, a saber: AB veces $DA,\ DE,\ DF,\ DC$; y BC veces $DA,\ DE,\ DF,\ DC$. Se tiene que $AB \times FD = AB^3$ y $BC \times CD = BC^3$. Estudiemos los cuerpos restantes;

$$AB \times DA,\ AB \times DC,\ AB \times DE,\ CB \times DA,\ CB \times DF \text{ y } CB \times DE.$$

Los paralelepípedos $AB \times DA,\ AB \times DE$ y $CB \times DA$ son iguales, pues en los tres casos las bases y las alturas son iguales; además, $AB \times DC,\ CB \times DA$ y $CB \times DE$ también son iguales, todo lo anterior fundamentado en la proposición 43 del libro I de los *Elementos*, que establece:

> En todo paralelogramo los complementos de los paralelogramos situados en torno a la diagonal son iguales entre sí. (Euclides 2000, p. 71)

De aquí que el cubo de AC está formado por el cubo de AB, el cubo de BC, tres veces $AB \times BC^2$ y tres veces $BC \times AB^2$.

Proposición 2. *Desde esta [primera proposición] viene una segunda, a saber, que* $\mathrm{AB}^3 + 3(\mathrm{AB} \times \mathrm{BC}^2)$ *es mayor que* $\mathrm{BC}^3 + 3(\mathrm{BC} + \mathrm{AB}^2)$ *por el cubo de la diferencia entre* AB *y* BC (Cardano 1968, p. 53).

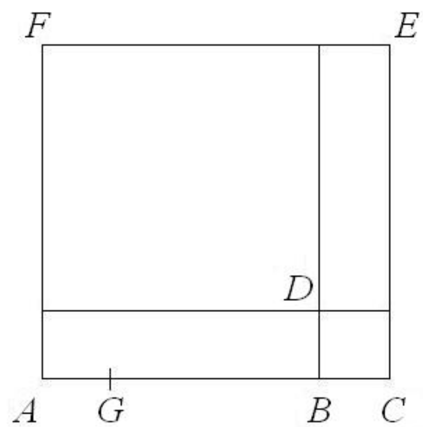

Figura 3.1.2. Esquema de Cardano

Se afirma que $AB^3 + 3(AB \times BC^2) = BC^3 + 3(BC \times AB^2) + (AB - BC)^3$. Para demostrar este resultado Cardano utilizó un esquema como el de la figura 3.1.2. En él se tiene que AG es igual a BC y por tanto $AB - BC = GB$. Procederemos entonces según el autor. De acuerdo con la proposición 1, se tiene que

$$AB^3 = AG^3 + GB^3 + 3(AG \times GB^2) + 3(AG^2 \times GB);$$

sumando en ambos miembros $3(AB \times BC^2)$ se llega a que

$$AB^3 + 3(AB \times BC^2) = AG^3 + GB^3 + 3(AG \times GB^2) + 3(AG^2 \times GB) + 3(AB \times BC^2);$$

sustituyendo AG por BC se obtiene que

$$AB^3 + 3(AB \times BC^2) = BC^3 + GB^3 + 3(BC \times GB^2) + 3(BC^2 \times GB) + 3(AB \times BC^2).$$

Por otra parte, $3(BC^2 \times GB)$ es igual a $3BC$ veces el rectángulo $BC \times GB$, es decir, $3BC(BC \times GB)$; y $3(AB \times BC^2)$ es igual a $3BC$ veces el rectángulo $AB \times BC$, es decir, $3BC(AB \times BC)$. Sustituyendo estos resultados, la expresión queda escrita del siguiente modo:

$$AB^3 + 3(AB \times BC^2) = BC^3 + GB^3 + 3BC(BC \times GB) + 3BC(AB \times BC) + 3BC(GB^2).$$

Obsérvese ahora que el rectángulo $AB \times BC$ es igual al rectángulo $GB \times BC$ más el cuadrado del lado AG, es decir, $AB \times BC = AG^2 + (GB \times BC)$. Teniendo en cuenta esta observación y la proposición 4 del libro II de los *Elementos*, se tiene que el rectángulo $AB \times BC$ más el rectángulo $GB \times BC$ más BG^2 es igual a AB^2, es decir,

$$(AB \times BC) + (GB \times BC) + BG^2 = AB^2.$$

De este modo,

$$\begin{aligned} AB^3 + 3(AB \times BC^2) &= BC^3 + GB^3 + 3BC(BC \times GB) + 3BC(BC \times AB) + 3BC(GB^2) \\ &= BC^3 + GB^3 + 3BC\left[(BC \times GB) + (BC \times AB) + (GB^2)\right] \\ &= BC^3 + GB^3 + 3BC(AB^2) \\ &= BC^3 + GB^3 + 3(BC \times AB^2). \end{aligned}$$

Como $GB^3 = (AB - BC)^3$, se concluye que $AB^2 + 3(AB \times BC^2)$ es mayor que $BC^3 + 3(BC \times AB^2)$ por el cubo de la diferencia entre AB y BC.

De la proposición anteriormente demostrada, Cardano desprendió el siguiente corolario:

Corolario. *[…] si* BC *se asume como negativo [o se resta del segmento* AC*],* AB^3 *consistirá de* $\mathrm{AC}^3 + 3(\mathrm{AC} \times \mathrm{BC}^2) + (-\mathrm{BC}^3) + (-3\mathrm{BC} \times \mathrm{AC}^2)$ (Cardano 1968, p. 54).

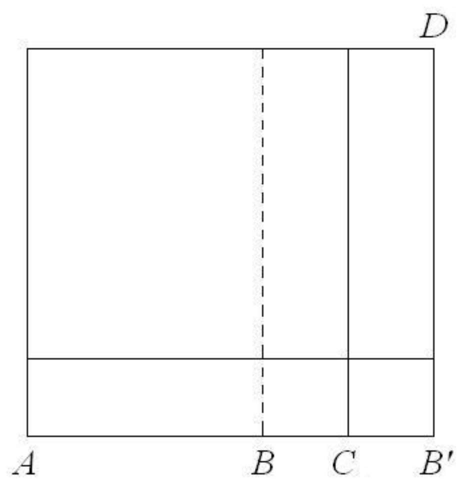

Figura 3.1.3. Cuadrado AD.

Sea AD el cuadrado cuyo lado es el segmento AB' (véase la figura 3.1.3), siendo $BC = B'C$. De acuerdo con la proposición 2 anterior,

$$AC^3 + 3(AC \times CB'^2) = (CB')^3 + 3(AC^2 \times CB') + (AC - CB')^3.$$

Como $AC - CB' = AB$, se tiene que

$$AC^3 + 3(AC \times CB'^2) = (CB')^3 + 3(AC^2 \times CB') + AB^3.$$

Al sustituir a CB' por BC, obtenemos que

$$AC^3 + 3(AC \times BC^2) = (BC)^3 + 3(AC^2 \times BC) + AB^3.$$

De este modo,

$$AB^3 = AC^3 + 3(AC \times BC^2) - (BC)^3 - 3(AC^2 \times BC).$$

Apéndice 3.2. La filosofía de Platón: Las artes de la aritmética y la logística

El estudio de los números en la antigua Grecia perteneció a dos disciplinas que algunos creyeron opuestas: la *aritmética* y la *logística*. La aritmética fue considerada como una disciplina teórica y la logística como el arte práctico del cálculo.

Para Platón (h. 427 - 348 a.C.) no fueron materias de niveles diferentes, él trató con una sola ciencia de la que se desprendieron dos ramas. Según Platón, adquirimos esa ciencia cuando tratamos con los objetos de la vida diaria y posteriormente desarrollamos nuestras experiencias para convertirlas en conocimientos expertos o especiales. Expliquemos un poco: cuando nos encontramos frente a una cantidad de objetos podemos determinar exactamente el número de ellos, es decir, podemos contarlos. Esta acción supone una familiaridad con los números, una capacidad de distinguir cada número de manera individual, por ejemplo distinguir el dos del tres. Platón llamó a la ciencia de todos los números posibles el *arte del número* o *aritmética*. Nótese que la aritmética de Platón no es la aritmética de nuestros tiempos.

Por otra parte, una vez contados los elementos de determinado conjunto de cosas no nos contentamos con ello, podemos incrementar la cantidad determinado número de veces, por ejemplo cinco veces más, o podemos separar elementos como la mitad o la tercera parte de ellos. En otras palabras, es habitual multiplicar o dividir. En esas multiplicaciones y divisiones, así como en todos los cálculos que se pueden realizar con las cantidades, está implícito el conocimiento de los diferentes números y el de las relaciones que guardan unos con otros. La ciencia de las relaciones mutuas entre los números que nos permiten calcular con ellos fue llamada el *arte del cálculo* o *logística*.

Para los pensadores platónicos la *aritmética* no fue una teoría de números sino el arte de contar correctamente, y sólo sobre sus bases era posible adquirir nuevos conocimientos en el terreno de los números. Sobre la misma línea, la *logística* no fue únicamente una disciplina que mostraba los posibles procedimientos de operación con los números, las operaciones significativas requieren un verdadero conocimiento de las relaciones que conectan un número con otros y, por ende, se requiere de un conocimiento firme de los números (Klein 1992, p. 19). De este modo, la aritmética y la logística no se oponen.

Como se dijo con anterioridad, podemos adquirir estos conocimientos mediante las operaciones elementales que aplicamos a cosas contables. Sin embargo, en este primer nivel resulta muy complicado distinguir la *aritmética* de la *logística*; los conocimientos que adquirimos son netamente prácticos y los usamos para satisfacer necesidades de nuestra vida cotidiana, ambas disciplinas se tratan como una sola y esta unidad tiene sus raíces en el objeto con que ambas trabajan: el número.

Por otra parte, la aritmética y la logística fueron divididas a su vez en aritmética teórica y aritmética práctica y en logística teórica y logística práctica. De acuerdo con Jacob Klein, Platón contrastó en *la República* y en el *Filebo* la *aritmética práctica* y la *logística práctica* con sus contrapartes teóricas. Platón hizo notar que la gente común cuenta de manera diferente a los "amantes de la sabiduría", la principal diferencia son las unidades con las que unos y otros trabajan (Klein 1992, p. 22). Respecto a esto, hablemos de la logística: la diferencia entre *logística práctica* y *logística teórica* es el tipo de multitudes con las que cada una trata. La *logística práctica* trabaja con multitudes de objetos distintos y obviamente son objetos de sentido; sin embargo, a pesar de sus diferencias se consideran iguales. Por ejemplo, dos campamentos armados, dos cabezas de ganado, o dos de las más pequeñas o más grandes cosas. La *logística teórica* trabaja con multitudes de unidades iguales, a saber, aquellas que no ocurren entre los objetos de sentido sino sólo en el pensamiento (Klein 1992, pp. 22 - 23).

Así, la *logística teórica* surge de la *logística práctica* cuando sus aplicaciones prácticas son rechazadas y sus suposiciones son investigadas en beneficio del conocimiento mismo. En este proceso los números son comprendidos en su propia naturaleza sólo por el pensamiento y no como se entienden cuando se tornan hacia objetos de sentido donde toman cuerpos visibles o tangibles.

Platón postuló una ciencia completamente libre del sentido de la percepción, una ciencia cuyo objeto final fue el orden cósmico, invisible e inaudible sobre el que se fundamenta nuestro mundo de los sentidos. La exigencia de Platón sobre la formación de la *logística teórica* se fundamentó en que dentro de la estructura de las ciencias noéticas, es decir, las ciencias que son sólo para el pensamiento, tenía que existir una dirigida a las relaciones puras de los números que se corresponden con el arte común del cálculo para proporcionarles los cimientos necesarios. El obstáculo crucial para la *logística teórica*, considerando su relación con el cálculo, surge con las fracciones, de manera precisa, con la división de la unidad de cálculo. La *logística teórica* trata con unidades que pueden ser entendidas sólo en el pensamiento, que son iguales entre sí y sobre todo que rechazan cualquier partición (Klein 1992, p. 39).

Un objeto que podemos percibir a través de los sentidos puede ser dividido en varias partes. Dicho objeto disminuye respecto a su tamaño original, pero se multiplica en el proceso respecto a la unidad anterior; en otras palabras, el objeto es dividido pero no así la unidad respecto a la cual cada parte adquiere su carácter individual. En todo cálculo o proceso de conteo se trabaja con unidades, y si en el proceso se "divide" alguna unidad lo que sucede en realidad es que se sustituye la unidad original por una nueva que también está sujeta a "división", en lenguaje correcto diremos que la unidad no se divide sino que se multiplica. Este proceso puede repetirse una y otra vez sin tocar la naturaleza esencial de las unidades: la indivisibilidad. De este modo surge la necesidad de hacer una distinción entre el "uno" como objeto de sentido y que está sujeto a conteo y cálculo, y la unidad como tal, que mantiene estrictamente separadas una cosa individual de todas las demás (Klein 1992, p. 40). Cada cosa individual puede ser dividida infinitamente gracias a su naturaleza corpórea, la unidad que sólo puede ser entendida en el pensamiento es indivisible precisamente por ese carácter noético que posee. Se debe considerar que la necesidad de introducir partes fraccionarias de la unidad surge precisamente en el cálculo mismo. En este punto emerge un desajuste muy importante e interesante entre el material sobre el que se ejecutan los cálculos, por una parte, y los números puros cuyo carácter noético se expresa precisamente en la indivisibilidad de las unidades (Klein 1992, p. 43).

A lo largo de esta breve descripción de la *aritmética* y la *logística* platónicas se ha trabajado alrededor del concepto de *arithmos* ($\alpha\rho\iota\theta\mu o\varsigma$). El significado de *arithmos* es contar o, de manera más precisa, *separar de* determinado número de cosas que pueden ser diferentes pero que son consideradas uniformes cuando son contadas (Klein 1992, p. 46).

La palabra pronunciada al final de la separación o numeración da el "número contado", el *arithmos* de las cosas implicadas. El 7, por ejemplo, es siempre un número definido de cosas definidas: manzanas, perros, gatos y, en el caso extremo, unidades puras. El *arithmos* está relacionado indisolublemente a eso que se enumera. Los griegos determinaron que la unidad como tal no es *arithmos*, puesto que el *arithmos* se refiere a una multitud de cosas contadas y ni un objeto de sentido ni una unidad pura son un número de cosas o unidades (Klein 1992, p. 49).

Algo digno de estudios profundos pero que no será tratado en este escrito es el surgimiento de los los números "puros" como opuestos a números "visibles o tangibles" fuera del fenómeno natural de conteo. De acuerdo con Platón, la práctica continua de conteo y cálculo crea en nosotros familiaridad con los números y sus relaciones o, en términos platónicos, con la *aritmética* y la *logística*; pero aquellos números que tenemos a nuestra disposición antes de comenzar un conteo o cálculo y los que deben ser claramente independientes de las cosas que experimentan dicho conteo, ¿a qué están asociados?, ¿qué numeran? (Klein 1992, p. 49). En este nuevo contexto evidentemente ya no hay interés alguno en los requirimientos de la vida diaria. De este modo se pone especial atención a la naturaleza del objeto de la *aritmética* y la *logística*, dicho objeto requiere un carácter puramente noético y que al mismo tiempo exhiba las características esenciales de lo contable. Estos requisitos son cubiertos completamente por las unidades puras que son no sensibles, accesibles solo al pensamiento, indistinguibles una de otra y que no admiten partición (los antiguos las llamaron mónadas).

La filosofía de Platón fue influida de manera decisiva por la ciencia pitagórica; sin embargo, Platón, en oposición a los pitagóricos, separó a los "números" de los objetos de sentido de tal forma que aparecieron junto a las cosas perceptibles con un ser aparte. Este nuevo reino estuvo representado por las unidades "puras" o mónadas. Nuestra capacidad de contar una multitud definida de objetos de sentido, de acuerdo con Platón, está basada en la existencia de mónadas que pueden ser unidas para formar el número en cuestión. Aquí se pone de manifiesto uno de los principales rasgos de la filosofía platónica: que lo que proporciona el fundamento de otra cosa, lo que hace posible el ser de otro, es más significativo y poderoso en su propio ser que eso otro a lo que da sentido. El segundo no puede existir sin el primero, pero el primero puede ser sin el segundo. En nuestro caso el modo especial de ser al que pertenecen los números de mónadas "puras" se vuelve perceptible a la luz de su papel de fundamento en la actividad de contar (Klein 1992, pp. 70 - 71).

En *la República* Platón divide en dos el universo de todo aquello que es accesible al pensamiento. Una de esas divisiones está caracterizada por los objetos de sentido que comprendemos como imagen de algún otro objeto; en otras palabras, examinamos algunos objetos por medio de los sentidos, aunque no nos referimos exactamente a ellos sino a lo que los fundamenta y de lo cual son imagen, es decir a un objeto de pensamiento (Klein 1992, p. 72). Como ejemplo digamos que algunos logísticos reflejan sus ideas sobre los números pares e impares en algunos objetos contables, pero estas reflexiones surgidas en el pensamiento no apuntan a aquellos objetos particulares sino a los números "puros" o sus *eide* (ideas) que son supuestos en el pensamiento y reflejados en los objetos. (Este esbozo sobre la *aritmética* y la *logística*, así como sobre lo que es accesible al pensamiento desde el punto de vista de Platón es de utilidad para comprender algunos sucesos ocurridos en siglos posteriores y que influyeron directa o indirectamente en las matemáticas de nuestros tiempos.)

CAPÍTULO IV

EL NACIMIENTO DE LA GEOMETRÍA ANALÍTICA EN EL SIGLO XVII

El aporte de René Descartes

René Descartes (1596 - 1650) y Pierre de Fermat (1601 - 1665) de manera independiente crearon la geometría analítica, una nueva forma de reinterpretar y conjuntar el análisis griego y el álgebra elaborada por sus antecesores. Según algunos autores, Descartes fue influenciado en un principio por los textos de Cristóforo Clavio (1538 - 1612). Javier de Lorenzo señaló: "[. . .] a Descartes no le interesa la Matemática por la Matemática misma. Busca en ella, como antes PLATÓN, como en su misma época PASCAL, como después LEIBNIZ, como después la lógica simbólica, un modelo de razonar" (Lorenzo 1989, p. 81). En esta búsqueda creó la geometría analítica como una nueva unificación del método de la aritmética y la geometría clásicas; este descubrimiento apareció en el *Discurso del método* publicado en 1637 como un apéndice titulado *La Geometría*. Otras obras de Descartes muestran su filosofía y modo de proceder en la construcción de su matemática, entre ellas su *Regulae ad directionem ingenii*, escrita hacia 1628, pero publicada de manera póstuma e inacabada.

El trabajo de Descartes partió de una crítica a las materias clásicas y a los estudiosos de las mismas. En su *Regulae ad directionem ingenii*, en la Regla IV, citada por Javier de Lorenzo, afirmó:

> Cuando comencé a aplicarme al estudio de las disciplinas matemáticas, leí uno tras otro a la mayor parte de lo que la tradición corriente nos ha legado procedente de quienes son autoridad en estas materias, y cultivé, sobre todo la aritmética y la geometría... Pero ni para la una ni para la otra conseguí dar con ningún autor capaz de satisfacerme plenamente... Es que en verdad nada es más vano que ocuparme de números abstractos y de figuras imaginarias, hasta el extremo de querer contentarse en conocer parecidas bagatelas; nada es más vano que aplicarse a esas demostraciones superficiales, que se encuentran más frecuentemente por azar que por *savoir faire*, y que son fuente de los ojos y de la imaginación más que del entendimiento, hasta el extremo de desacostumbrarse de alguna manera en el uso de la razón; y al mismo tiempo, nada es más complicado que poner en claro, por tal método de demostración las dificultades nuevas que se esconden bajo la alineación confusa de los números. (Citado en: Lorenzo 1989, pp. 81 - 82)

En el *Discurso del método* Descartes manifestó:

> En cuanto al análisis de los antiguos y el álgebra de los modernos, además de que no se extienden más que a materias muy abstractas y que no parecen de utilidad alguna, la primera está siempre tan restringida a la consideración de

> las figuras que no puede ejercitar el entendimiento sin fatigar mucho la imaginación; y se está de tal forma sometido, en la última, a ciertas reglas y ciertas cifras, que se ha hecho un arte confuso y oscuro que embaraza el espíritu, en lugar de una ciencia que lo cultive. (Citado en: Lorenzo 1989, p. 82)

Viète fue el primero en relacionar el análisis griego con el proceso matemático que se realizó hasta poco antes de 1591 (fecha de publicación del *Isagoge*); por otra parte, los trabajos de álgebra de Viète aparecieron con una década de anticipación al *Algebra* de Clavio. Sin embargo, Descartes aceptó haber leído el trabajo de Viète sólo después de que publicó su *Geometría*. En una carta que dirigió a Mersenne (1588 - 1648) con fecha 20 de febrero de 1639 escribió: “No tengo ningún conocimiento de ese geómetra del que me escribe [Beaugrand?], y me asombro de eso que él dice, que estudiamos juntos a Viète en París; ya que es un libro del que no recuerdo haber visto nunca la cubierta, mientras estuve en Francia” (Citado en: Mahoney 1994, pp. 27 - 28). Sin lugar a dudas, la obra de Viète fue una de las fuentes más importantes para dar un paso adelante en la unificación algebra – geometría que él mismo inició. Sin embargo, algunos matemáticos como Descartes, Mersenne y otros parisinos conocieron el trabajo de Viète y obtuvieron provecho de él, pero no adoptaron ni su simbolismo ni su espíritu.

Dentro de las críticas realizadas a los trabajos de sus antecesores, Descartes escribió sobre los términos empleados para referirse a las cantidades desconocidas: cosa, raíz, cuadrado, cubo, bicuadrado, entre otros. En la Regla XVII rechazó la representación espacial de la “desconocida” argumentando que

> [...] acabé por darme cuenta de que con esta manera de representar las cosas, no había descubierto absolutamente nada que no hubiese podido reconocer sin ella con mucha mayor facilidad y distinción, y que es necesario rechazar completamente los términos de este género bajo pena de hacer confusa la representación [...] (Citado en: Lorenzo 1989, p. 83)

Es así como considera indispensable la creación de un lenguaje unívoco, que permita el desarrollo de una interpretación más abstracta del significado de las incógnitas y que evite cualquier equivocación; del lenguaje geométrico sólo conservó la línea recta. El número o cantidad discontinua fue representado por la línea recta o cantidad continua, de este modo se suma, resta, divide, multiplica y se extren raíces por medio de líneas rectas.

Entonces el producto de dos líneas genera otra línea y no un rectángulo, y elevar al cuadrado no generó una magnitud plana (área de un cuadrado), el resultado fue la segunda potencia de una cantidad, y asignando un número a cada línea las operaciones estuvieron representadas en terreno algebraico solamente (Lorenzo 1989, p. 85).

En la Regla XVI Descartes expresó que las magnitudes conocidas las denotaría con las letras $a, b, c, \ldots$ y las desconocidas con las letras $A, B, C, \ldots$ precedidas frecuentemente por los números $1, 2, 3, \ldots$ para indicar su multiplicidad. Los números también indicarían el número de relaciones de la magnitud, por ejemplo $2a^3$ significa el doble de la magnitud designada por la letra a que contiene tres relaciones. El punto clave en los trabajos posteriores de Descartes fue convenir en una notación diferente a la empleada por Viète: dejó de lado las vocales mayúsculas que representaban a las incógnitas para introducir los signos $x, y, z, \ldots$ para el mismo fin, pues consideró, como

Javier de Lorenzo señaló, que el signo de la incógnita es un signo artificial enteramente arbitrario y no abreviador.

La nueva notación cartesiana permitió representar las potencias de manera exponencial ($x \cdot x \cdot x \cdot x \cdot x = x^5$) y operar con potencias de exponente generalizado x^n, así como efectuar la división de polinomios. Por otra parte, se atribuye a Descartes la transposición de términos de una ecuación; por ejemplo, $ax^3 + bx^2 + cx + d = 0$, lo que permitió eliminar del estudio de una ecuación de determinado grado la multitud de casos que debían considerarse.

En un pasaje del Libro Tercero de *La Geometría*, citado por Javier de Lorenzo, Descartes se refirió a las raíces de una ecuación, a su naturaleza y a su relación con el grado de la misma:

> Cuántas raíces puede haber en cada ecuación.
>
> Sépase, pues, que en cada ecuación, según cuantas dimensiones tenga la cantidad desconocida, otras tantas serán las diversas raíces que puede haber, es decir valores de esa cantidad; pues, por ejemplo, si se supone $x = 3$, o bien $x - 3 = 0$ y multiplicamos estas dos ecuaciones $x - 2 = 0$ y $x - 3 = 0$ una por otra, se tendrá
>
> $$xx - 5x + 6 = 0, \quad \text{o bien} \quad xx = 5x - 6$$
>
> que es una ecuación en la cual la cantidad x vale 2 y al mismo tiempo vale 3. Que si luego se hace $x - 4 = 0$ y se multiplica esta suma por $xx - 5x + 6$, se

> tendrá
>
> $$x^3 - 9xx + 26x - 24 = 0$$
>
> que es otra ecuación, en la cual x, teniendo tres dimensiones, tendrá tres valores, que son, 2, 3 y 4. (Citado en: Lorenzo 1989, p. 89)

También considera las "raíces falsas" asumiendo el caso en que x sea igual a un defecto (número negativo), por ejemplo x igual al defecto de 5, entonces $x + 5 = 0$. Si se multiplica $x + 5 = 0$ por $x^3 + 9xx + 26x - 24 = 0$ se obtiene $x^4 - 4x^3 - 19xx + 106x - 120 = 0$ cuyas raíces son 2, 3, 4 y una falsa que es 5. Además escribió sobre cómo saber si una cantidad dada es el valor de una raíz:

> Y recíprocamente, si la suma de una ecuación no puede ser dividida por un binomio, compuesto de la cantidad conocida $+$ o $-$ alguna otra cantidad, esto prueba que esa otra cantidad no es el valor de ninguna de sus raíces; así la última
>
> $$x^4 - 4x^3 - 19xx + 106x - 120 = 0$$
>
> puede ser dividida por $x - 2$ y por $x - 3$ y por $x - 4$ y por $x - 5$; pero no por x $+$ o $-$ ninguna otra cantidad: lo que muestra que ella no puede tener más que las cuatro raíces 2, 3, 4 y -5. (Citado en: Lorenzo 1989, p. 90)

Es indudable que Descartes realizó una de las más grandes aportaciones a la matemática con su convenio notacional, pero de manera conciente o inconciente, en 1628 siguió los pasos de Viète, pues el álgebra simbólica fue fundamento de sus "matemáticas genuinas", y en particular de la teoría de ecuaciones del Libro III de *La Geométría*, que en esencia no es más que el arte del análisis de Viète (Mahoney 1994, p. 32).

Pierre de Fermat y el proyecto de reconstrucción de las obras griegas

Durante el siglo XVII las matemáticas sufrieron una reorientación. Ya no se trataba de resolver problemas específicos ni de clasificarlos, los esfuerzos de los estudiosos se dirigieron hacia los métodos de solución. En el siglo anterior, Viète retomó el álgebra de los cosistas y la *Arithmetica* de Diofanto para analizarlas y sistematizarlas por medio de los trabajos citados por Pappo en el Libro VII de la *Colección matemática*. Para incorporar este material al arte analítico Viète se enfrentó a dos problemas que se convirtieron en los principales objetivos de la nueva escuela del análisis. El primer objetivo fue claro: reconstruir tan fielmente como fuera posible el contenido de los trabajos perdidos del análisis geométrico griego citados por Pappo. El segundo objetivo fue traducir ese contenido geométrico al lenguaje del arte analítico, es decir, convertirlo en logística especiosa o álgebra simbólica. Viète inició este proyecto de restauración y traducción con el *Apollonius Gallus, seu exsuscitata Apollonii Pergaei peri epaphon geometria* (Mahoney 1994, p. 39).

Influido por los trabajos de Viète y el ambiente de reconstrucción y restauración de aquella época, Pierre de Fermat (1601 - 1665) comenzó la restauración del *Plane Loci* de Apolonio (fines s. III - principios s. II a.C.). Por medio de la geometría analítica que surgió gracias a esa labor, Fermat encontró el camino para reconstruir el *Solido Loci* de Aristeus, y el *Superficie Loci* y los *Porismas* de Euclides.

Algunos trabajos fueron más fáciles de traducir al arte analítico, como las *Secciones en una razón dada*, pues contenían problemas de métrica que tenían un punto por solución, por ejemplo dividir un segmento de recta en una razón dada. Este tipo de problemas se representó con ecuaciones con una desconocida. Sin embargo, trabajos que implicaban lugares geométricos, como las *Cónicas* de Apolonio, no contaban con un modelo para su traducción. Así, quedó como tarea para Fermat y Descartes encontrar una técnica adecuada que permitiera expresar los *loci* en ecuaciones (Mahoney 1994, pp. 40 - 41).

Un nuevo obstáculo en la traducción de las antiguas obras fue el problema de la homogeneidad ligado a la dimensionalidad. Empleando las combinaciones de las cuatro operaciones básicas de los griegos en geometría (suma, multiplicación, división, obtención de raíces) Viète trabajó con un elemento totalmente ajeno al álgebra numérica: la dimensión, dejando en su *Arte analítico* un problema conceptual que Fermat y Descartes tuvieron que librar para lograr sus objetivos.

Para los griegos los números no tenían dimensión, cualquier combinación de operaciones aplicada a un par de números daba como resultado otro número. Un punto tampoco tenía dimensión, una línea era unidimensional y un rectángulo era bi-dimensional. Podían sumarse únicamente elementos con la misma dimensión y el producto de elementos generaba un elemento de dimensión mayor; por ejemplo, el producto de dos líneas era un rectángulo y el producto de un rectángulo y una línea generaba un paralelepípedo.

Viète respetó estos aspectos de las obras griegas, pues dichas obras fueron su motivación para la realización de su proyecto, y con ellos estableció la *Ley de Homogeneidad* en su *Introducción al arte analítico* de 1646, cuya esencia fue: sólo pueden sumarse magnitudes homogéneas obteniéndose también una magnitud homogénea, el producto de cualesquiera dos magnitudes es heterogéneo con ellas (Mahoney 1994, p. 42). Esto dio como resultado que todas las ecuaciones del *Arte analítico* de Viète tuvieran una dimensión relacionada directamente al grado de la ecuación.

De acuerdo con lo anterior, la ley de homogeneidad y su vinculación a la geometría clásica griega no dejaban espacio para las ecuaciones de grado mayor que 3 en el nuevo arte de Viète, pues bajo esta visión carecían de significado. Pero el autor rechazó esto tajantemente e introdujo procedimientos mecánicos como el *Pseudo-Mesolabe* cuyos resultados podían ser descritos abstractamente mediante el arte analítico revelando su relación con los problemas de grado superior. Tal fue el caso de las construcciones de las cuarta, quinta, sexta, etc. proporcionales a dos segmentos de líneas dados.

Fermat y Descartes encontraron dificultades con la dimensionalidad al traducir al lenguaje algebraico los problemas geométricos relacionados con el *loci*. Aunque Fermat pareció estar de acuerdo con la *Ley de homogeneidad* de Viète, no relacionó la dimensión con el grado de una ecuación cuando empleó el álgebra en problemas geométricos. Nunca justificó su desarrollo, aunque al parcer asumió la misma línea de razonamiento que Descartes hizo explícita al principio de su *Geometría* (Mahoney 1994, p. 42).

Viète y después Descartes observaron que las potencias sucesivas de la incógnita formaban una proporción continua, es decir, $x : x^2 :: x^2 : x^3 :: x^3 : x^4 :: x^4 : x^5 \ldots$ Pero Descartes dio un paso adelante comenzando la proporción continua con la unidad de la magnitud, es decir, $1 : x :: x : x^2 :: x^2 : x^3 :: x^3 : x^4 :: x^4 : x^5 \ldots$ y argumentó con base en Euclides que las proporciones sólo pueden obtenerse entre cantidades homogéneas (Mahoney 1994, p. 44).

Los trabajos de Fermat, por otra parte, no mostraron explícitamente la solución al problema de la homogeneidad en la combinación de operaciones de la geometría. Sin embargo, el proceder de Fermat hizo evidente que su conclusión sobre el problema era la misma que la de Descartes. De este modo, Fermat y Descartes rompieron la barrera conceptual heredada por Viète e introdujeron los trabajos analíticos griegos sobre *loci* en la teoría de ecuaciones, preparando el terreno para el inminente surgimiento del concepto de función (Mahoney 1994, p. 44).

Fermat heredó de Viète todo un sistema algebraico con una teoría de ecuaciones definida así como un programa de investigación con un enorme potencial de aplicación a distintos campos de la matemática que explotó fructíferamente, es decir, tomó el *Arte analítico* como modelo. Adoptando la notación de esta obra se colocó en desventaja ante la publicación de *La geometría* de Descartes con su reluciente "nueva" teoría de ecuaciones, pues aunque no fue una obra tan completa como el sistema creado por Viète, muchos aspectos importantes de la estructura de las ecuaciones se hicieron claros con la notación cartesiana a pesar de que fueron considerados con mucha anterioridad en el *Arte analítico*. Así, aunque la teoría de ecuaciones de Fermat igualó a la de Descartes e incluso la superó, no fue accesible a estudiosos educados bajo el modelo cartesiano (Mahoney 1994, p. 45).

CAPÍTULO V

GAUSS Y LA PRIMERA DEMOSTRACIÓN DEL TEOREMA FUNDAMENTAL DEL ÁLGEBRA

La tesis de doctorado de Carl Friedrich Gauss

La primera demostración válida del teorema fundamental del álgebra fue elaborada por Carl Friedrich Gauss (1777 - 1855) y con ella obtuvo su doctorado en filosofía en 1799. A continuación presentaremos esta demostración.

Teorema Fundamental del Álgebra. *Toda ecuación algebraica de la forma* $\mathrm{x^m + Ax^{m-1} + Bx^{m-2} + \cdots + M = 0}$ *se puede resolver en* m *factores reales de primer o segundo grado.*

Convengamos en llamar X a $x^m + Ax^{m-1} + Bx^{m-2} + \cdots + M$. Para comenzar, Gauss planteó las condiciones necesarias para la demostración del teorema a través de dos lemas:

Lema 1. *Siendo* m *cualquier número entero positivo, la función* $\mathrm{x^m \operatorname{sen}\varphi - r^{m-1}x \operatorname{sen} m\varphi + r^m \operatorname{sen}(m-1)\varphi}$ *es divisible entre* $\mathrm{x^2 - 2rx\cos\varphi + r^2}$.

Demostración

Si $m = 1$, tenemos que $(x \operatorname{sen} \varphi - x \operatorname{sen} \varphi + r \operatorname{sen} 0\varphi) \div (x^2 - 2rx \cos \varphi + r^2) = 0$.

Para $m = 2$, se tiene que

$$(x^2 \operatorname{sen} \varphi - rx \operatorname{sen} 2\varphi + r^2 \operatorname{sen} \varphi) \div (x^2 - 2rx \cos \varphi + r^2) = \operatorname{sen} \varphi.$$

Se demostrará que para $m \geq 3$ el cociente es

$$x^{m-2} \operatorname{sen} \varphi + rx^{m-3} \operatorname{sen} 2\varphi + r^2 x^{m-4} \operatorname{sen} 3\varphi + \cdots + r^{m-2} \operatorname{sen}(m-1)\varphi.$$

Procederemos por inducción probando que el producto de esta función por $x^2 - 2rx \cos \varphi + r^2$ da como resultado $x^m \operatorname{sen} \varphi - r^{m-1} x \operatorname{sen} m\varphi + r^m \operatorname{sen}(m-1)\varphi$.

Sea $m = 3$. Se tiene que

$(x \operatorname{sen} \varphi + r \operatorname{sen} 2\varphi)(x^2 - 2rx \cos \varphi + r^2) =$

$x^3 \operatorname{sen} \varphi + rx^2 \operatorname{sen} 2\varphi - 2rx^2 \operatorname{sen} \varphi \cos \varphi - 2r^2 x \operatorname{sen} 2\varphi \cos \varphi + r^2 x \operatorname{sen} \varphi + r^3 \operatorname{sen} 2\varphi$

$= x^3 \operatorname{sen} \varphi + 2rx^2 \operatorname{sen} \varphi \cos \varphi - 2rx^2 \operatorname{sen} \varphi \cos \varphi - 4r^2 x \operatorname{sen} \varphi \cos^2 \varphi + r^2 x \operatorname{sen} \varphi + r^3 \operatorname{sen} 2\varphi$

$= x^3 \operatorname{sen} \varphi - 4r^2 x \operatorname{sen} \varphi \cos^2 \varphi + r^2 x \operatorname{sen} \varphi + r^3 \operatorname{sen} 2\varphi$.

Como

$$\begin{aligned}
-4r^2 x \operatorname{sen} \varphi \cos^2 \varphi + r^2 x \operatorname{sen} \varphi &= -3r^2 x \operatorname{sen} \varphi \cos^2 \varphi - r^2 x \operatorname{sen} \varphi \cos^2 \varphi + r^2 x \operatorname{sen} \varphi \\
&= -3r^2 x \operatorname{sen} \varphi \cos^2 \varphi + r^2 x \operatorname{sen} \varphi(-\cos^2 \varphi + 1) \\
&= -3r^2 x \operatorname{sen} \varphi \cos^2 \varphi + r^2 x \operatorname{sen} \varphi(\operatorname{sen}^2 \varphi)
\end{aligned}$$

$$= -3r^2x \operatorname{sen}\varphi \cos^2\varphi + r^2x \operatorname{sen}^3\varphi$$

$$= -2r^2x \operatorname{sen}\varphi \cos^2\varphi - r^2x \operatorname{sen}\varphi \cos^2\varphi + r^2x \operatorname{sen}^3\varphi$$

$$= -r^2x(2 \operatorname{sen}\varphi \cos^2\varphi + \operatorname{sen}\varphi \cos^2\varphi - \operatorname{sen}^3\varphi)$$

$$= -r^2x[2 \operatorname{sen}\varphi \cos^2\varphi + \operatorname{sen}\varphi(\cos^2\varphi - \operatorname{sen}^2\varphi)]$$

$$= -r^2x(\operatorname{sen} 2\varphi \cos\varphi + \operatorname{sen}\varphi \cos 2\varphi)$$

$$= -r^2x \operatorname{sen} 3\varphi,$$

resulta que

$$(x \operatorname{sen}\varphi + r \operatorname{sen} 2\varphi)(x^2 - 2rx\cos\varphi + r^2) = x^3 \operatorname{sen}\varphi - r^2x \operatorname{sen} 3\varphi + r^3 \operatorname{sen} 2\varphi.$$

Supongamos ahora que el resultado se cumple para $m = k$, es decir,

$$[x^{k-2} \operatorname{sen}\varphi + rx^{k-3} \operatorname{sen} 2\varphi + r^2x^{k-4} \operatorname{sen} 3\varphi + \cdots + r^{k-3}x \operatorname{sen}(k-2)\varphi +$$
$$r^{k-2} \operatorname{sen}(k-1)\varphi](x^2 - 2rx\cos\varphi + r^2) = x^k \operatorname{sen}\varphi - r^{k-1}x \operatorname{sen} k\varphi + r^k \operatorname{sen}(k-1)\varphi. \quad \textit{(1)}$$

Demostremos que la afirmación es válida para $m = k + 1$; es decir,

$$[x^{k-1} \operatorname{sen}\varphi + rx^{k-2} \operatorname{sen} 2\varphi + r^2x^{k-3} \operatorname{sen} 3\varphi + \cdots + r^{k-2}x \operatorname{sen}(k-1)\varphi +$$
$$r^{k-1} \operatorname{sen} k\varphi](x^2 - 2rx\cos\varphi + r^2) = x^{k+1} \operatorname{sen}\varphi - r^k x \operatorname{sen}(k+1)\varphi + r^{k+1} \operatorname{sen} k\varphi. \quad \textit{(2)}$$

Notemos que el primer miembro de *(2)* puede rescribirse en la forma

$$[x^{k-1} \operatorname{sen}\varphi + rx^{k-2} \operatorname{sen} 2\varphi + \cdots + r^{k-2} \operatorname{sen}(k-1)\varphi](x^2 - 2rx\cos\varphi + r^2) +$$
$$(r^{k-1} \operatorname{sen} k\varphi)(x^2 - 2rx\cos\varphi + r^2). \quad \textit{(3)}$$

Si se multiplican ambos miembros de la ecuación *(1)* por x, se obtiene

$$(x^{k-1} \operatorname{sen}\varphi + rx^{k-2} \operatorname{sen} 2\varphi + \cdots + r^{k-2}x \operatorname{sen}(k-1)\varphi)(x^2 - 2rx\cos\varphi + r^2) =$$

$= x^{k+1}\operatorname{sen}\varphi - r^{k-1}x^2\operatorname{sen}k\varphi + r^k x\operatorname{sen}(k-1)\varphi.$

El primer miembro de esta nueva expresión es el primer sumando de *(3)*; sustituyendo, se tiene que

$(x^{k+1}\operatorname{sen}\varphi - r^{k-1}x^2\operatorname{sen}k\varphi + r^k x\operatorname{sen}(k-1)\varphi) + (r^{k-1}\operatorname{sen}k\varphi)(x^2 - 2rx\cos\varphi + r^2) =$

$= (x^{k+1}\operatorname{sen}\varphi - r^{k-1}x^2\operatorname{sen}k\varphi + r^k x\operatorname{sen}(k-1)\varphi) + (r^{k-1}x^2\operatorname{sen}k\varphi - 2r^k x\operatorname{sen}k\varphi\ \cos\varphi +$
$r^{k+1}\operatorname{sen}k\varphi) = x^{k+1}\operatorname{sen}\varphi + r^k x\operatorname{sen}(k-1)\varphi - 2r^k x\operatorname{sen}k\varphi\cos\varphi + r^{k+1}\operatorname{sen}k\varphi.$

De esta última igualdad analicemos $r^k x\operatorname{sen}(k-1)\varphi - 2r^k x\operatorname{sen}k\varphi\cos\varphi$. Se tiene que

$$r^k x\operatorname{sen}(k-1)\varphi - 2r^k x\operatorname{sen}k\varphi\cos\varphi = r^k x\big[\operatorname{sen}(k-1)\varphi - 2\operatorname{sen}k\varphi\ \cos\varphi\big].$$

Utilizando la identidad trigonométrica $\operatorname{sen}\theta\ \cos\beta = \frac{1}{2}\big[\operatorname{sen}(\theta-\beta) + \operatorname{sen}(\theta+\beta)\big]$, se obtiene que

$$\begin{aligned}
r^k x\big[\operatorname{sen}(k-1)\varphi - 2\ \operatorname{sen}k\varphi\ \cos\varphi\big] &= r^k x\Big[\operatorname{sen}(k-1)\varphi - 2\Big(\tfrac{1}{2}[\operatorname{sen}(k\varphi-\varphi) + \operatorname{sen}(k\varphi+\varphi)]\Big)\Big]\\
&= r^k x\Big[\operatorname{sen}(k-1)\varphi - \Big(\operatorname{sen}(k-1)\varphi + \operatorname{sen}(k+1)\varphi\Big)\Big]\\
&= r^k x\Big[-\operatorname{sen}(k+1)\varphi\Big]\\
&= -r^k x\operatorname{sen}(k+1)\varphi.
\end{aligned}$$

Así, después de estas consideraciones, se obtiene que

$[x^{k-1}\operatorname{sen}\varphi + rx^{k-2}\operatorname{sen}2\varphi + r^2x^{k-3}\operatorname{sen}3\varphi + \cdots + r^{k-2}x\operatorname{sen}(k-1)\varphi +$
$r^{k-1}\operatorname{sen}k\varphi](x^2 - 2rx\cos\varphi + r^2) =$

$= (x^{k+1}\operatorname{sen}\varphi - r^{k-1}x^2\operatorname{sen}k\varphi + r^k x\operatorname{sen}(k-1)\varphi) + (r^{k-1}\operatorname{sen}k\varphi)(x^2 - 2rx\cos\varphi + r^2)$

$= x^{k+1}\operatorname{sen}\varphi - r^k x\operatorname{sen}(k+1)\varphi + r^{k+1}\operatorname{sen}k\varphi.$

Es decir, *(2)* se cumple y por lo tanto se cumple el lema.

Lema 2. *Si la cantidad* r *y el ángulo* φ *están determinados de tal manera que se cumplan las ecuaciones*

$\mathrm{r^m}\cos\mathrm{m}\varphi + \mathrm{Ar^{m-1}}\cos(\mathrm{m}-1)\varphi + \mathrm{Br^{m-2}}\cos(\mathrm{m}-2)\varphi + \cdots + \mathrm{Kr^2}\cos 2\varphi + \mathrm{Lr}\cos\varphi$

$+\mathrm{M} = 0$ (4)

y

$\mathrm{r^m}\operatorname{sen}\mathrm{m}\varphi + \mathrm{Ar^{m-1}}\operatorname{sen}(\mathrm{m}-1)\varphi + \mathrm{Br^{m-2}}\operatorname{sen}(\mathrm{m}-2)\varphi + \cdots + \mathrm{Kr^2}\operatorname{sen}\ 2\varphi +$

$\mathrm{Lr}\operatorname{sen}\varphi = 0$ (5)

entonces la función

$$\mathrm{x^m + Ax^{m-1} + Bx^{m-2} + \cdots + Kx^2 + Lx + M = X}$$

será divisible entre el factor de segundo grado $\mathrm{x^2 - 2rx}\cos\varphi + \mathrm{r^2}$ *solamente si* $\mathrm{r}\operatorname{sen}\varphi$ *no es igual a* 0*. Pero si* $\mathrm{r}\operatorname{sen}\varphi = 0$ *entonces esa función será divisible entre el factor simple* $\mathrm{x - r}\cos\varphi$*.*

Demostración

De acuerdo con la demostración del lema 1, las siguientes cantidades serán divisibles entre $x^2 - 2rx\cos\varphi + r^2$:

$$\left.\begin{aligned}
&rx^m \operatorname{sen}\varphi - r^m x \operatorname{sen}\ m\varphi + r^{m+1}\operatorname{sen}(m-1)\varphi,\\
&Arx^{m-1}\operatorname{sen}\varphi - Ar^{m-1}x\operatorname{sen}(m-1)\varphi + Ar^m\operatorname{sen}(m-2)\varphi,\\
&Brx^{m-2}\operatorname{sen}\varphi - Br^{m-2}x\operatorname{sen}(m-2)\varphi + Br^{m-1}\operatorname{sen}(m-3)\varphi,\\
&\vdots\\
&Krx^2\operatorname{sen}\varphi - Kr^2x\operatorname{sen}2\varphi + Kr^3\operatorname{sen}\varphi,\\
&Lrx\operatorname{sen}\varphi - Lrx\operatorname{sen}\varphi + Lr^2\operatorname{sen}(0)\varphi = Lrx\operatorname{sen}\varphi - Lrx\operatorname{sen}\varphi,\ \text{y}\\
&Mr\operatorname{sen}\varphi - Mx\operatorname{sen}(0)\varphi + Mr\operatorname{sen}(-\varphi) = Mr\operatorname{sen}\varphi + Mr\operatorname{sen}(-\varphi).
\end{aligned}\right\}\textit{(6)}$$

Observemos que para la primera cantidad se tiene que

$$rx^m \operatorname{sen}\varphi - r^m x \operatorname{sen}\ m\varphi + r^{m+1}\operatorname{sen}(m-1)\varphi = r(x^m\operatorname{sen}\varphi - r^{m-1}x\operatorname{sen}\ m\varphi + r^m\operatorname{sen}(m-1)\varphi).$$

El lema 1 afirma que $x^k \operatorname{sen}\varphi - r^{k-1}x\operatorname{sen}k\varphi + r^k\operatorname{sen}(k-1)\varphi$ es divisible entre $x^2 - 2rx\cos\varphi + r^2$ para cualquier k entero; luego, $rx^m \operatorname{sen}\varphi - r^m x \operatorname{sen}\ m\varphi + r^{m+1}\operatorname{sen}(m-1)\varphi$ es divisible entre $x^2 - 2rx\cos\varphi + r^2$.

Veamos ahora que

$Arx^{m-1}\operatorname{sen}\varphi - Ar^{m-1}x\operatorname{sen}(m-1)\varphi + Ar^m\operatorname{sen}(m-2)\varphi =$

$Ar(x^{m-1}\operatorname{sen}\varphi - r^{m-2}x\operatorname{sen}(m-1)\varphi + r^{m-1}\operatorname{sen}(m-2)\varphi)$;

$Brx^{m-2}\operatorname{sen}\varphi - Br^{m-2}x\operatorname{sen}(m-2)\varphi + Br^{m-1}\operatorname{sen}(m-3)\varphi =$

$Br(x^{m-2}\operatorname{sen}\varphi - r^{m-3}x\operatorname{sen}(m-2)\varphi + r^{m-2}\operatorname{sen}(m-3)\varphi)$;

$\vdots$

$Krx^2\operatorname{sen}\varphi - Kr^2x\operatorname{sen}2\varphi + Kr^3\operatorname{sen}\varphi = Kr(x^2\operatorname{sen}\varphi - rx\ \operatorname{sen}2\varphi + r^2\operatorname{sen}\varphi)$.

Por medio de estas igualdades se puede ver que todas las expresiones son divisibles entre $x^2 - 2rx\cos\varphi + r^2$, de acuerdo con el lema 1. Obsérvese que

$$Lrx\ \operatorname{sen}\varphi - Lrx\ \operatorname{sen}\varphi + Lr^2\operatorname{sen}(0)\varphi = Lrx\ \operatorname{sen}\varphi - Lrx\ \operatorname{sen}\varphi = 0$$

y

$$\begin{aligned} Mr\operatorname{sen}\varphi - Mx\operatorname{sen}(0)\varphi + Mr\operatorname{sen}(-1)\varphi &= Mr\operatorname{sen}\varphi + Mr\operatorname{sen}(-\varphi) \\ &= Mr\operatorname{sen}\varphi - Mr\operatorname{sen}\varphi \\ &= 0. \end{aligned}$$

Así, las dos últimas expresiones de *(6)* también son divisibles entre $x^2 - 2rx\cos\varphi + r^2$.

De lo anterior se tiene que la suma de todas las cantidades especificadas en *(6)* es divisible entre $x^2 - 2rx\cos\varphi + r^2$. La suma del primer término de cada una de las expresiones *(6)* produce

$$rx^m\operatorname{sen}\varphi + Arx^{m-1}\operatorname{sen}\varphi + Brx^{m-2}\operatorname{sen}\varphi + \cdots + Krx^2\operatorname{sen}\varphi + Lrx\operatorname{sen}\varphi + Mr\operatorname{sen}\varphi$$

$= r\,\text{sen}\,\varphi(x^m + Ax^{m-1} + Bx^{m-2} + \cdots + Kx^2 + Lx + M)$

$= rX\;\text{sen}\,\varphi.$

Analicemos la suma de los segundos términos:

$-r^m x\,\text{sen}\;m\varphi - Ar^{m-1}x\,\text{sen}(m-1)\varphi - Br^{m-2}x\,\text{sen}(m-2)\varphi - \cdots - Lrx\,\text{sen}\,\varphi$

$= -(r^m x\,\text{sen}\;m\varphi + Ar^{m-1}x\,\text{sen}(m-1)\varphi + Br^{m-2}x\,\text{sen}(m-2)\varphi + \cdots + Lrx\,\text{sen}\,\varphi)$

$= -x\,[r^m\,\text{sen}\;m\varphi + Ar^{m-1}\,\text{sen}(m-1)\varphi + Br^{m-2}\,\text{sen}(m-2)\varphi + \cdots + Lr\,\text{sen}\,\varphi].$

El segundo factor de esta última expresión es igual a 0, por la hipótesis *(5)*; luego,

$-r^m x\,\text{sen}\;m\varphi - Ar^{m-1}x\,\text{sen}(m-1)\varphi - Br^{m-2}x\,\text{sen}(m-2)\varphi - \cdots - Kr^2x\,\text{sen}\,\varphi -$

$Lrx\,\text{sen}\,\varphi = 0.$

Con la finalidad de analizar la suma de los terceros términos, multipliquemos la igualdad *(4)* por $\text{sen}\,\varphi$ y la *(5)* por $\cos\varphi$ (después se restará un producto del otro):

$\text{sen}\,\varphi\big(r^m\cos m\varphi + Ar^{m-1}\cos(m-1)\varphi + Br^{m-2}\cos(m-2)\varphi + \cdots + Kr^2\cos 2\varphi +$

$Lr\cos\varphi + M\big) = 0.$

$\cos\varphi\big(r^m\,\text{sen}\,m\varphi + Ar^{m-1}\,\text{sen}(m-1)\varphi + Br^{m-2}\,\text{sen}(m-2)\varphi + \cdots + Kr^2\,\text{sen}\,2\varphi +$

$Lr\,\text{sen}\,\varphi\big) = 0.$

Desarrollando ambas ecuaciones obtenemos respectivamente:

$r^m\cos m\varphi\;\text{sen}\,\varphi + Ar^{m-1}\cos(m-1)\varphi\;\text{sen}\,\varphi + Br^{m-2}\cos(m-2)\varphi\;\text{sen}\,\varphi + \cdots +$

$Kr^2\cos 2\varphi\;\text{sen}\,\varphi + Lr\cos\varphi\;\text{sen}\,\varphi + M\,\text{sen}\,\varphi = 0.$ *(7)*

$r^m \operatorname{sen} m\varphi \, \cos\varphi + Ar^{m-1} \operatorname{sen}(m-1)\varphi \, \cos\varphi + Br^{m-2} \operatorname{sen}(m-2)\varphi \, \cos\varphi + \cdots +$

$Kr^2 \operatorname{sen} 2\varphi \, \cos\varphi + Lr \operatorname{sen}\varphi \, \cos\varphi = 0.$ *(8)*

Mediante la identidad $\cos\alpha \, \operatorname{sen}\beta = \frac{1}{2}\big[\operatorname{sen}(\alpha+\beta) - \operatorname{sen}(\alpha-\beta)\big]$ transformaremos cada uno de los sumandos de las expresiones anteriores, para restar fácilmente una de la otra; así, tenemos que

$$
\begin{aligned}
r^m \cos\, m\varphi \, \operatorname{sen}\varphi &= \tfrac{1}{2} r^m \operatorname{sen}(m+1)\varphi - \tfrac{1}{2} r^m \operatorname{sen}(m-1)\varphi,\\
r^m \operatorname{sen}\, m\varphi \, \cos\varphi &= \tfrac{1}{2} r^m \operatorname{sen}(m+1)\varphi - \tfrac{1}{2} r^m \operatorname{sen}(1-m)\varphi\\
&= \tfrac{1}{2} r^m \operatorname{sen}(m+1)\varphi + \tfrac{1}{2} r^m \operatorname{sen}(m-1)\varphi,\\
Ar^{m-1} \cos(m-1)\varphi \, \operatorname{sen}\varphi &= \tfrac{1}{2} Ar^{m-1} \operatorname{sen}\, m\varphi - \tfrac{1}{2} Ar^{m-1} \operatorname{sen}(m-2)\varphi,\\
Ar^{m-1} \operatorname{sen}(m-1)\varphi \, \cos\varphi &= \tfrac{1}{2} Ar^{m-1} \operatorname{sen}\, m\varphi - \tfrac{1}{2} Ar^{m-1} \operatorname{sen}(-m+2)\varphi\\
&= \tfrac{1}{2} Ar^{m-1} \operatorname{sen}\, m\varphi + \tfrac{1}{2} Ar^{m-1} \operatorname{sen}(m-2)\varphi,\\
Br^{m-2} \cos(m-2)\varphi \, \operatorname{sen}\varphi &= \tfrac{1}{2} Br^{m-2} \operatorname{sen}(m-1)\varphi - \tfrac{1}{2} Br^{m-2} \operatorname{sen}(m-3)\varphi,\\
Br^{m-2} \operatorname{sen}(m-2)\varphi \, \cos\varphi &= \tfrac{1}{2} Br^{m-2} \operatorname{sen}(m-1)\varphi - \tfrac{1}{2} Br^{m-2} \operatorname{sen}(-m+3)\varphi\\
&= \tfrac{1}{2} Br^{m-2} \operatorname{sen}(m-1)\varphi + \tfrac{1}{2} Br^{m-2} \operatorname{sen}(m-3)\varphi,\\
&\vdots\\
Kr^2 \cos 2\varphi \, \operatorname{sen}\varphi &= \tfrac{1}{2} Kr^2 \operatorname{sen} 3\varphi - \tfrac{1}{2} Kr^2 \operatorname{sen}\varphi,\\
Kr^2 \operatorname{sen} 2\varphi \, \cos\varphi &= \tfrac{1}{2} Kr^2 \operatorname{sen} 3\varphi - \tfrac{1}{2} Kr^2 \operatorname{sen}(-\varphi)\\
&= \tfrac{1}{2} Kr^2 \operatorname{sen} 3\varphi + \tfrac{1}{2} Kr^2 \operatorname{sen}\varphi,\\
Lr \cos\varphi \, \operatorname{sen}\varphi &= \tfrac{1}{2} \operatorname{sen} 2\varphi - \tfrac{1}{2} \operatorname{sen}(0), \quad \text{y}\\
Lr \operatorname{sen}\varphi \, \cos\varphi &= \tfrac{1}{2} \operatorname{sen} 2\varphi - \tfrac{1}{2} \operatorname{sen}(0)\\
&= \tfrac{1}{2} \operatorname{sen} 2\varphi + \tfrac{1}{2} \operatorname{sen}(0).
\end{aligned}
$$

Restando *(8)* de *(7)*, se obtiene:

$$-r^m \operatorname{sen}(m-1)\varphi - Ar^{m-1}\operatorname{sen}(m-2)\varphi - \cdots - Kr^2 \operatorname{sen}\varphi + M\operatorname{sen}\varphi = 0,$$

es decir,

$$r^m \operatorname{sen}(m-1)\varphi + Ar^{m-1}\operatorname{sen}(m-2)\varphi + \cdots + Kr^2 \operatorname{sen}\varphi - M\operatorname{sen}\varphi = 0.$$

Multiplicando la expresión anterior por r, tenemos que

$$r^{m+1} \operatorname{sen}(m-1)\varphi + Ar^{m}\operatorname{sen}(m-2)\varphi + \cdots + Kr^3 \operatorname{sen}\varphi - Mr\operatorname{sen}\varphi = 0,$$

que es lo mismo que

$$r^{m+1} \operatorname{sen}(m-1)\varphi + Ar^{m}\operatorname{sen}(m-2)\varphi + \cdots + Kr^3 \operatorname{sen}\varphi + Lr^2\operatorname{sen}(0)\varphi + Mr\operatorname{sen}(-\varphi) = 0.$$

Con esto concluimos que la suma de los terceros términos de *(6)* es nula. Se sigue que la función $rX\operatorname{sen}\varphi$ es divisible entre $x^2 - 2rx\cos\varphi + r^2$ y por lo tanto la función X lo es siempre que $r\operatorname{sen}\varphi$ no sea igual a 0.

Si $r\operatorname{sen}\varphi = 0$ entonces $r = 0$ o $\operatorname{sen}\varphi = 0$. En el primer caso $M = 0$, por la hipótesis *(4)*; así,

$$X = x^m + Ax^{m-1} + Bx^{m-2} + \cdots + Kx^2 + Lx$$

$$= x(x^{m-1} + Ax^{m-2} + Bx^{m-3} + \cdots + Kx + L),$$

y por lo tanto X será divisible entre x.

Si $\operatorname{sen}\varphi = 0$ entonces $\varphi = k\pi$ con k en los enteros. Si k es par se tendrá que $\operatorname{sen}k\pi = 0$ y $\cos k\pi = 1$; si k es impar se tendrá que $\operatorname{sen}k\pi = 0$ y $\cos k\pi = -1$.

Observemos lo siguiente:

Si $\varphi = k\pi$ con k par entonces $\cos 2\varphi = \cos 2k\pi$; como $2k$ es par se tiene que $\cos 2\varphi = 1$.

Si $\varphi = k\pi$ con k par entonces $\cos 3\varphi = \cos 3k\pi$; como $3k$ es par se tiene que $\cos 3\varphi = 1$.

Si $\varphi = k\pi$ con k impar entonces $\cos 2\varphi = \cos 2k\pi$; como $2k$ es par se tiene que $\cos 2\varphi = 1$.

Si $\varphi = k\pi$ con k impar entonces $\cos 3\varphi = \cos 3k\pi$; como $3k$ es impar se tiene que $\cos 3\varphi = -1$.

Esto implica que $\cos 2\varphi = 1$, $\cos 4\varphi = 1$, $\cos 6\varphi = 1$ y $\cos t\varphi = 1$, siendo $\varphi = k\pi$ con k par o impar y t par. De manera similar, $\cos 3\varphi = 1$, $\cos 5\varphi = 1$, $\cos 7\varphi = 1$ y $\cos t\varphi = 1$, siendo $\varphi = k\pi$ con k par y t impar. Asimismo, $\cos 3\varphi = -1$, $\cos 5\varphi = -1$, $\cos 7\varphi = -1$ y $\cos t\varphi = -1$, siendo $\varphi = k\pi$ con k impar y t impar. Así, de manera general para estos casos se tiene que $\cos k\varphi = (\cos\varphi)^k$. Por lo que si $x = r\cos\varphi$, se tendrá que $X = 0$ por la hipótesis *(4)*, pues haciendo la sustitución convenida en $X = x^m + Ax^{m-1} + Bx^{m-2} + \cdots + M$, se tiene que

$$
\begin{aligned}
X &= (r\cos\varphi)^m + A(r\cos\varphi)^{m-1} + B(r\cos\varphi)^{m-2} + \cdots + K(r\cos\varphi)^2 + L(r\cos\varphi) + M \\
&= r^m\cos^m\varphi + Ar^{m-1}\cos^{m-1}\varphi + Br^{m-2}\cos^{m-2}\varphi + \cdots + Kr^2\cos^2\varphi + Lr\cos\varphi + M \\
&= r^m\cos\ m\varphi + Ar^{m-1}\cos\ m-1\varphi + Br^{m-2}\cos\ m-2\varphi + \cdots + Kr^2\cos\ 2\varphi + Lr\cos\varphi + M \\
&= 0.
\end{aligned}
$$

En consecuencia, la función $X = x^m + Ax^{m-1} + Bx^{m-2} + \cdots + M$ es divisible entre $x - r\cos\varphi$, y por lo tanto $x^m + Ax^{m-1} + Bx^{m-2} + \cdots + M$ es divisible entre

$x^2 - 2rx\cos\varphi + r^2$ o entre $x - r\cos\varphi$ siempre que se satisfaga el sistema de ecuaciones simultáneas

$$\begin{cases} r^m \cos m\varphi + Ar^{m-1}\cos(m-1)\varphi + \cdots + Kr^2\cos 2\varphi + Lr\cos\varphi + M = 0 \\ \\ r^m \operatorname{sen} m\varphi + Ar^{m-1}\operatorname{sen}(m-1)\varphi + \cdots + Kr^2\operatorname{sen} 2\varphi + Lr\operatorname{sen}\varphi = 0. \end{cases}$$

Respecto a los lemas anteriormente demostrados, Gauss escribió en el apartado 15 de su disertación de doctorado (Gauss, 1799):

> Manifestum iam est, ad demonstrationem theorematis nostri nihil aliud requiri quam at ostendatur: *Proposita functione quacunque X formae $x^m + Ax^{m-1} + Bx^{m-2} + etc. + Lx + M = 0$, r et φ ita determinari posse, ut aequationes* [1] *et* [2] *locum habeant* . Hinc enim sequetur, X habere factorem realem primi vel secundi gradus; divisio autem necessario producet quotientem realem inferioris gradus, qui ex eadem ratione quoque factorem primi vel secundi gradus habebit. Per continuationem huius operationis X tandem in factores reales simplices vel duplices resolvetur.

$\Big[$ [. . .] Ya es evidente que para la demostración de nuestro teorema no se requiere otra cosa sino mostrar que:

Propuesta cualquier función X de la forma

$$x^m + Ax^{m-1} + Bx^{m-2} + etc. + Lx + M = 0$$

r y φ se pueden determinar de tal manera que se cumplan las ecuaciones [1] *y* [2]. [(4) y (5) en este capítulo].

De aquí en efecto se sigue que X tiene un factor real de primer o segundo grado; por otra parte, la división necesariamente produce un cociente real de menor grado, el cual por la misma razón también tendrá un factor de primer o segundo grado. Mediante la continuación de esta operación X finalmente se descompondrá en factores reales simples o de segundo grado. [. . .] $\Big]$

Pero, ¿cuál es el significado que Gauss da a las ecuaciones

$$r^m \cos m\varphi + Ar^{m-1}\cos(m-1)\varphi + \cdots + Kr^2\cos 2\varphi + Lr\cos\varphi + M = 0$$

y

$$r^m \operatorname{sen} m\varphi + Ar^{m-1}\operatorname{sen}(m-1)\varphi + \cdots + Kr^2\operatorname{sen} 2\varphi + Lr\operatorname{sen}\varphi = 0?$$

¿Cómo podemos comprender con mayor claridad la manera en que estas ecuaciones generan la descomposición de una ecuación polinomial en factores de primer y segundo grado?

Consideremos un plano infinito, y en él una línea recta o eje GG' en la que seleccionamos un punto C y determinemos un segmento rectilíneo como unidad de medida de longitud (véase la figura 5.1).

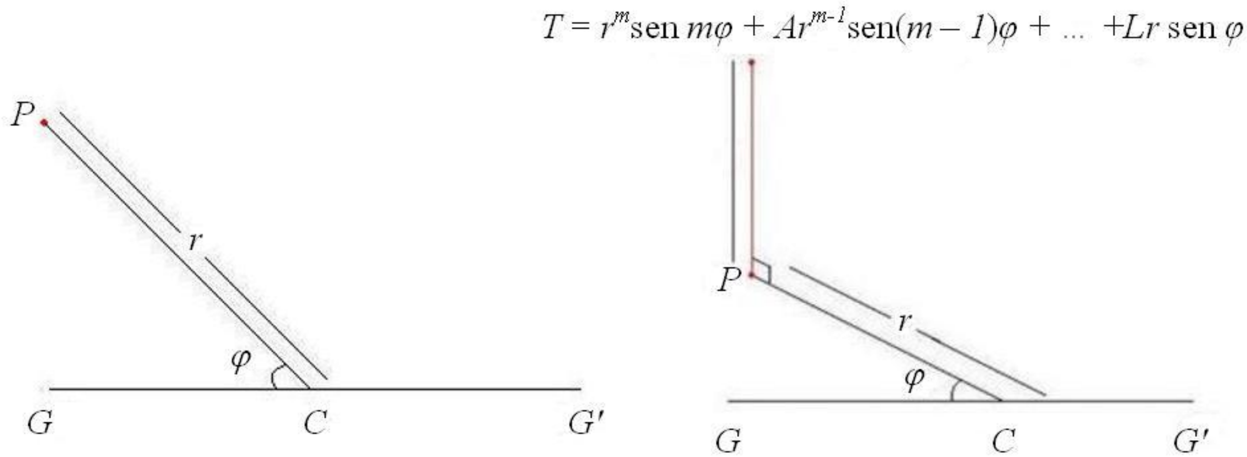

Figura 5.1. Construcción de la perpendicular con longitud T sobre el punto P

Ubiquemos un punto cualquiera P en el plano cuya distancia a C sea r. Se tiene que $\overline{PC}$ y $\overline{GC}$ forman el $\measuredangle PCG = \measuredangle\varphi$. Sobre P levantemos un segmento rectilíneo perpendicular al plano cuya longitud sea

$$T = r^m \operatorname{sen} m\varphi + Ar^{m-1} \operatorname{sen}(m-1)\varphi + \cdots + Lr \operatorname{sen}\varphi,$$

como en la figura 5.1. Consideremos que r siempre será un valor positivo y que para los puntos P que caen al otro lado del eje GG', φ se incrementará en π o será tomado como negativo.

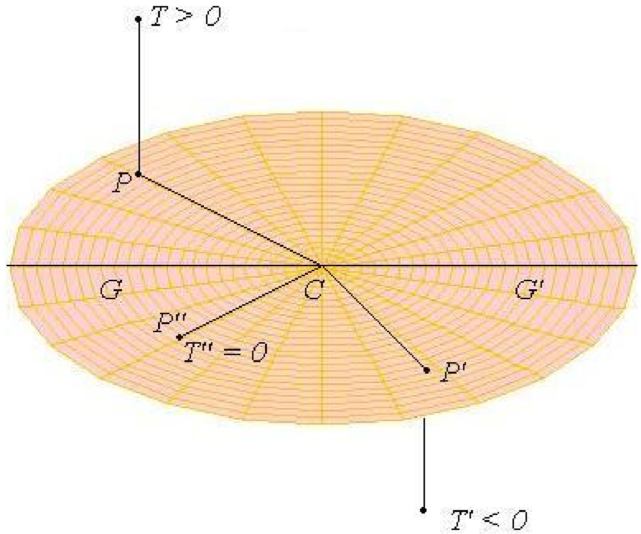

Figura 5.2. $T > 0,\ T' < 0$ y $T'' = 0$

Si "levantamos" segmentos rectilíneos perpendiculares de longitud correspondiente T en todos los puntos del plano, los extremos de los mismos se ubicarán arriba del plano si

$$T = r^m \operatorname{sen} m\varphi + Ar^{m-1} \operatorname{sen}(m-1)\varphi + \cdots + Lr \operatorname{sen} \varphi$$

es positiva, bajo el plano si T es negativa y en el plano si $T = 0$, como se muestra en la figura 5.2. Respecto a los extremos de las perpendiciulares, Gauss escribió:

> Extremitates horum perpendiculorum (quae pro valore positivo ipsius T supra planum accipiendae sunt, pro negativo infra, pro evanescente in plano ipso) erunt ad superficiem curvam continuam quaquaversum infinitam, quam brevitatis gratia in sequentibus *superficiem primam* vocabo. (Gauss, 1799, §16)

> [Los extremos de estas perpendiculares –los cuales para valores positivos de T se encuentran por arriba del plano; para negativos, por abajo; y cuando T desaparece, en el plano mismo– estarán en una superficie curva continua e infinita en todas direcciones, a la cual por brevedad llamaré en lo que sigue *primera superficie*.]

De manera similar construiremos una *segunda superficie* relacionada con los puntos del plano tomando como referencia el mismo eje GG' y el centro C. Las alturas de los puntos de la superficie serán

$$U = r^m \cos m\varphi + Ar^{m-1}\cos(m-1)\varphi + \cdots + Lr\cos\varphi + M.$$

Esta nueva superficie también será continua e infinita en todas direcciones, y se localizará arriba del plano para los puntos en que U sea positiva, bajo el plano para los puntos en que U sea negativa y en el plano cuando U sea igual a 0. Un ejemplo de estas superficies se muestra en la figura 5.3.

Entonces nuestro problema para demostrar el teorema fundamental del álgebra se reduce a mostrar que existe al menos un punto ubicado en el plano simultáneamente sobre la primera superficie *y sobre la* segunda superficie, como se ilustra en la figura 5.4.

Analicemos qué ocurre cuando la distancia r del centro al punto P es muy grande. En este caso, en la expresión $T = r^m \operatorname{sen} m\varphi + Ar^{m-1}\operatorname{sen}(m-1)\varphi + \cdots + Lr\operatorname{sen}\varphi$ los términos $Ar^{m-1}\operatorname{sen}(m-1)\varphi + \cdots + Lr\operatorname{sen}\varphi$ no contribuyen en nada a la suma

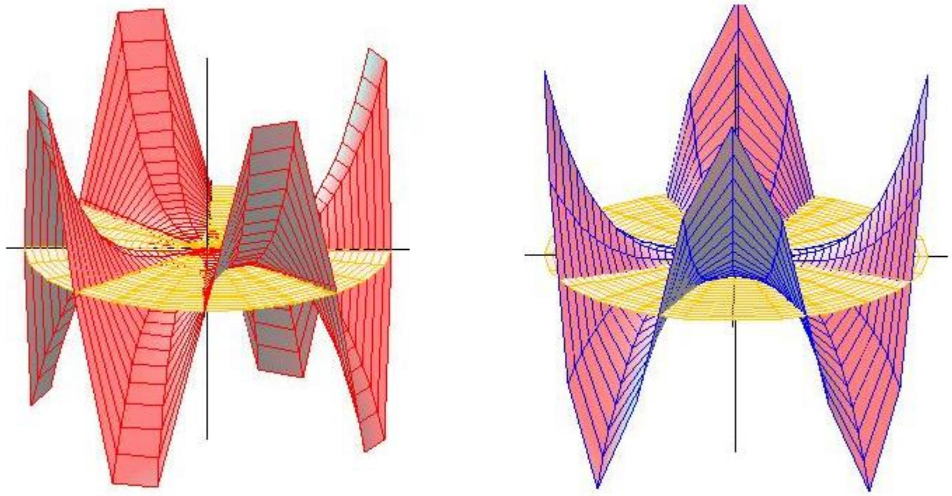

Figura 5.3 Superficies $T = r^4 \operatorname{sen} 4\varphi - 3r^2 \operatorname{sen} 2\varphi + 6r \operatorname{sen} \varphi$

y $U = r^4 \cos 4\varphi - 3r^2 \cos 2\varphi + 6r \cos \varphi - 2$

asociadas a la ecuación $x^4 - 3x^2 + 6x - 2 = 0$

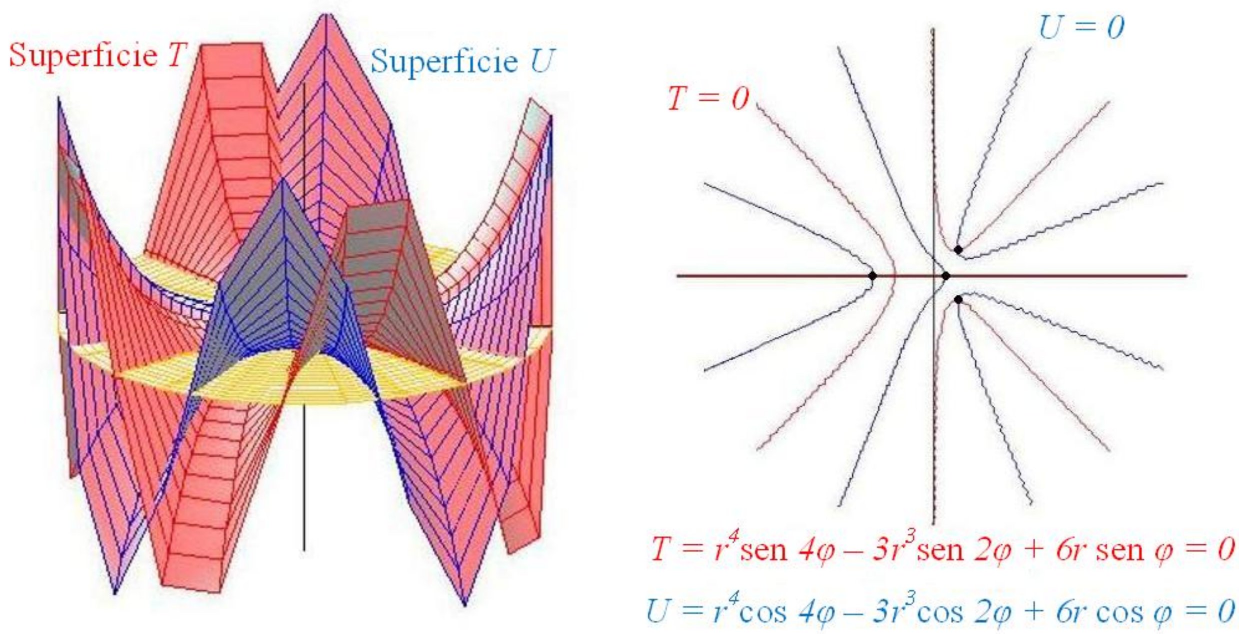

Figura 5.4. Intersecciones de las curvas $T = 0$ y $U = 0$

comparados con $r^m \operatorname{sen} m\varphi$, por lo que para r muy grande

$$T = r^m \operatorname{sen} m\varphi + Ar^{m-1} \operatorname{sen}(m-1)\varphi + \cdots + Lr \operatorname{sen} \varphi$$

se vuelve prácticamente en $T = r^m \operatorname{sen} m\varphi$; es fácil notar que T puede tomar valores positivos o negativos, según varíe el ángulo φ. De este modo, la primera superficie se intersecará necesariamente con el plano; a dicha intersección le llamaremos *primera curva*.

De igual manera, para r muy grande la expresión $U = r^m \cos m\varphi + Ar^{m-1} \cos(m-1)\varphi + \cdots + Lr \cos \varphi + M$ estará dominada por el primer término, $r^m \cos m\varphi$. Así, por la razón expuesta anteriormente, la segunda superficie se intersecará tembién con el plano; dicha intersección será llamada *segunda curva*. Nuestra tarea se ha simplificado aún más: basta mostrar que existe por lo menos un punto común de la *primera curva* y de la *segunda curva*.

Analicemos la curva T,

$$r^m \operatorname{sen} m\varphi + Ar^{m-1} \operatorname{sen}(m-1)\varphi + \cdots + Lr \operatorname{sen} \varphi = 0.$$

Dividiendo ambos miembros de esta ecuación entre r^m, tenemos que

$$\operatorname{sen} m\varphi + \frac{A}{r} \operatorname{sen}(m-1)\varphi + \frac{B}{r^2} \operatorname{sen}(m-2)\varphi + \cdots + \frac{L}{r^{m-1}} \operatorname{sen} \varphi = 0.$$

A una distancia infinita del punto C, la curva $T = 0$ coincidirá con la curva $\operatorname{sen} m\varphi = 0$; pero $\operatorname{sen} m\varphi = 0$ si y sólo si $m\varphi = k\pi$, siendo k un número entero; es decir, si y sólo si

$\varphi = \frac{k}{m}\pi$. Luego, la curva $\operatorname{sen} m\varphi = 0$ está formada por m líneas rectas que se intersecan en el punto C y cuyos ángulos de inclinación respecto al eje GG' son

$0,\ \frac{1}{m}\pi,\ \frac{2}{m}\pi,\ \ldots,\ \frac{m-1}{m}\pi$. Estas rectas dividen al plano en $2m$ regiones iguales, m por encima del eje GG' y las otras m por debajo del mismo. De aquí se concluye que la primera curva tiene $2m$ ramas infinitas que dividen a la circunferencia de radio infinito y centro C en $2m$ partes iguales. Además, el eje mismo forma dos ramas infinitas de esta *primera curva*, la primera y la $(m+1)$-ésima (obsérvese la figura 5.5).

Analicemos ahora la segunda curva, U;

$$r^m \cos m\varphi + Ar^{m-1}\cos(m-1)\varphi + \cdots + Lr\cos\varphi + M = 0.$$

Dividiendo ambos miembros de esta ecuación entre r^m, tenemos

$$\cos m\varphi + \frac{A}{r}\cos(m-1)\varphi + \frac{B}{r^2}\cos(m-2)\varphi + \cdots + \frac{L}{r^{m-1}}\cos\varphi + M = 0.$$

A una distancia infinita del punto C, la curva $U = 0$ coincidirá con la curva $\cos m\varphi = 0$; pero $\cos m\varphi = 0$ si y sólo si $m\varphi = k\frac{\pi}{2}$, siendo k entero impar; es decir, si y sólo si

$\varphi = \frac{k}{m}\left(\frac{\pi}{2}\right)$. Así, la segunda curva está compuesta de m rectas que se intersecan en C, cuyas inclinaciones respecto al eje GG' son $\frac{1}{2m}\pi,\ \frac{3}{2m}\pi,\ \frac{5}{2m}\pi,\ \ldots,\ \frac{(2m-1)}{2m}\pi$.

Sean $\frac{k}{m}\pi$ y $\frac{k+1}{m}\pi$, con $k \in \{0, 1, \ldots, m-1\}$, los ángulos de inclinación de dos rectas consecutivas que forman parte de la curva $T = 0$; éstos son iguales a $\left(\frac{2k}{m}\right)\frac{\pi}{2}$ y $\left(\frac{2k+2}{m}\right)\frac{\pi}{2}$ respectivamente. Entre ellos se encuentra el ángulo $\left(\frac{2k+1}{m}\right)\frac{\pi}{2}$, que es el ángulo de inclinación de una recta perteneciente a $U = 0$. De lo anterior se concluye que cada rama de la *segunda curva* se localiza entre dos ramas consecutivas de la *primera curva*, como se muestra en la figura 5.5.

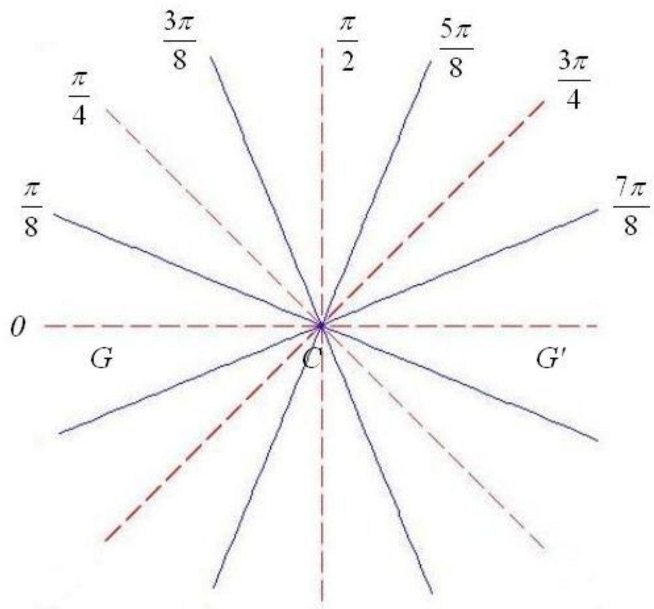

Figura 5.5. Primera curva y segunda curva para r infinita y $m = 4$

Veamos qué ocurre con las superficies y las curvas cuando la distancia de P a C es finita. Gauss demostró el siguiente teorema:

THEOREMA. *Manentibus cunctis ut supra, ex centro C describi poterit circulus, in cuius peripheria sint 2m puncta, in quibus T=0, totidemque, in quibus U=0, et quidem ita, ut singula posteriora inter bina priorum iaceant*. (Gauss 1799, §19)

[TEOREMA. *Manteniendo todo como antes, se puede describir un círculo de centro C en cuya circunferencia están 2m puntos en los cuales T = 0, y la misma cantidad en los cuales U = 0; y de tal manera que cada uno de estos últimos yace entre dos de los primeros*.]

Demostración

Sea un círculo de radio r (más adelante determinaremos r). Designemos con (1) el punto sobre la circunferencia del mismo que se localiza a $\left(\frac{1}{m}\right)\frac{\pi}{4}$ de la intersección de la circunferencia con la parte izquierda del eje GG', es decir, para el punto (1) el ángulo φ es igual a $\left(\frac{1}{m}\right)\frac{\pi}{4}$, como se muestra en la figura 5.6. Denótese con (3) el punto sobre la circunferencia para el cual $\varphi=\left(\frac{3}{m}\right)\frac{\pi}{4}$, con (5) aquel para el que $\varphi=\left(\frac{5}{m}\right)\frac{\pi}{4}$, y así sucesivamente hasta el punto $(8m-1)$ para el cual $\varphi=\left(\frac{8m-1}{m}\right)\frac{\pi}{4}$.

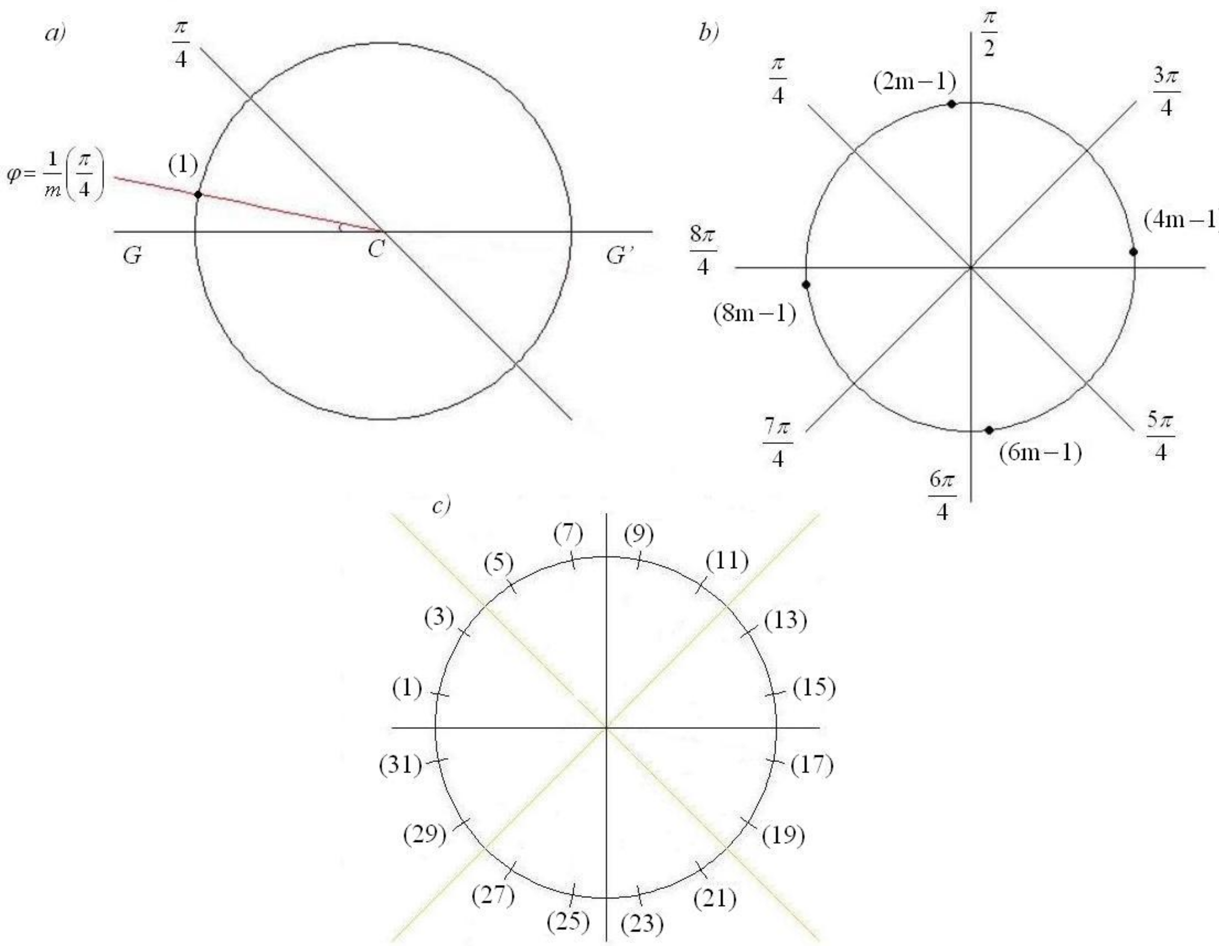

Figura 5.6. *a)* y *b)*: Proceso de división de la circunferencia en $4m$ puntos;

c): Una circunferencia dividida para $m=4$

De este modo se tienen $4m$ puntos sobre la circunferencia espaciados a intervalos iguales (véase la figura 5.6). Veamos que se cumplirá que $T = 0$ en algún punto entre $(8m-1)$ y (1), así como en alguno entre (3) y (5), (7) y (9), (11) y (13), etc. Además, se cumplirá que $U = 0$ para algún punto entre (1) y (3), (5) y (7), (9) y (11), etcétera.

Notemos que con los $4m$ puntos que localizamos en la circunferencia podemos formar $4m$ parejas de puntos consecutivos. Entonces se tendrán $2m$ puntos en los que $T = 0$ y, de manera similar, $2m$ puntos donde $U = 0$. Demostremos las afirmaciones anteriores. Primero, analicemos en qué intervalos de la circunferencia de radio r la función $T = r^m \operatorname{sen} m\varphi + Ar^{m-1}\operatorname{sen}(m-1)\varphi + Br^{m-2}\operatorname{sen}(m-2)\varphi + \cdots + Kr^2 \operatorname{sen} 2\varphi + Lr \operatorname{sen}\varphi$ tiene valores positivos y en cuáles la función tiene valores negativos; es decir, encontremos los intervalos de la circunferencia en los cuales la superficie T se encuentra sobre ella, y por tanto sobre el plano; y los intervalos en los que la superficie está debajo de ella, y en consecuencia debajo del plano. Consideremos que podemos escribir la función

$$T = r^m \operatorname{sen} m\varphi + Ar^{m-1}\operatorname{sen}(m-1)\varphi + Br^{m-2}\operatorname{sen}(m-2)\varphi + \cdots + Lr \operatorname{sen}\varphi$$

como

$$T = r^m \left(\operatorname{sen} m\varphi + \frac{A}{r}\operatorname{sen}(m-1)\varphi + \frac{B}{r^2}\operatorname{sen}(m-2)\varphi + \cdots + \frac{L}{r^{m-1}}\operatorname{sen}\varphi\right).$$

Sea $D = \frac{A}{r}\operatorname{sen}(m-1)\varphi + \frac{B}{r^2}\operatorname{sen}(m-2)\varphi + \cdots + \frac{L}{r^{m-1}}\operatorname{sen}\varphi$; con esta nueva notación se tiene que $T = r^m(\operatorname{sen} m\varphi + D)$. Recordemos que en la circunferencia de radio r el ángulo de inclinación de (1) respecto a GG' es $\varphi = \left(\frac{1}{m}\right)\frac{\pi}{4}$; luego, en ese punto $m\varphi = \frac{\pi}{4}$. Evaluando T en el punto (1) de la circunferencia se tiene que $T = r^m(\operatorname{sen}\frac{\pi}{4} + D)$; es decir,

$T = r^m\left(\sqrt{\frac{1}{2}} + D\right)$. Para que T sea positivo para cualesquier valores de $A, B, \ldots, L$, como sucede en el caso de r infinito cuando $\varphi = \frac{\pi}{4}$, se necesita que $|D| < \sqrt{\frac{1}{2}}$.

Analicemos cuidadosamente a $|D|$:

$$\begin{aligned}
|D| &= \left|\tfrac{A}{r}\operatorname{sen}(m-1)\varphi + \tfrac{B}{r^2}\operatorname{sen}(m-2)\varphi + \cdots + \tfrac{L}{r^{m-1}}\operatorname{sen}\varphi\right| \\
&\leq \left|\tfrac{A}{r}\operatorname{sen}(m-1)\varphi\right| + \left|\tfrac{B}{r^2}\operatorname{sen}(m-2)\varphi\right| + \cdots + \left|\tfrac{L}{r^{m-1}}\operatorname{sen}\varphi\right| \\
&= \left|\tfrac{A}{r}\right|\left|\operatorname{sen}(m-1)\varphi\right| + \left|\tfrac{B}{r^2}\right|\left|\operatorname{sen}(m-2)\varphi\right| + \cdots + \left|\tfrac{L}{r^{m-1}}\right|\left|\operatorname{sen}\varphi\right| \\
&< \left|\tfrac{A}{r}\right| + \left|\tfrac{B}{r^2}\right| + \cdots + \left|\tfrac{L}{r^{m-1}}\right| \quad \text{(pues } 0 \lneq |\operatorname{sen} k\varphi| \leq 1 \text{ con } k \in \{1, 2, \ldots, m-1\}) \\
&= \tfrac{|A|}{r} + \tfrac{|B|}{r^2} + \cdots + \tfrac{|L|}{r^{m-1}}.
\end{aligned}$$

Siendo g el máximo del conjunto $\{|A|, |B|, |C|, \ldots, |L|\}$,

$$\begin{aligned}
|D| &< \tfrac{g}{r} + \tfrac{g}{r^2} + \cdots + \tfrac{g}{r^{m-1}} \\
&= g\left(\tfrac{1}{r} + \tfrac{1}{r^2} + \cdots + \tfrac{1}{r^{m-1}}\right).
\end{aligned}$$

Consideremos que

$$\frac{1}{r} + \frac{1}{r^2} + \cdots + \frac{1}{r^{m-1}} < \frac{1}{r} + \frac{1}{r^2} + \cdots + \frac{1}{r^{m-1}} + \ldots .$$

Sabemos que $r > 0$; si se toma $r > 1$, se tendrá que $\frac{1}{r} < 1$. Luego,

$$|D| < g\left(\frac{\frac{1}{r}}{1 - \frac{1}{r}}\right) = g\left(\frac{1}{r-1}\right).$$

Obtuvimos una relación más clara entre D y r. Si $\frac{g}{r-1} < \sqrt{\frac{1}{2}}$, se cumplirá que $|D| < \sqrt{\frac{1}{2}}$; por lo que $r > g\sqrt{2} + 1$. Así, en el punto (1) de cualquier circunferencia cuyo radio r cumpla la condición $r > g\sqrt{2} + 1$ se tendrá que $|D| < \sqrt{\frac{1}{2}}$; es decir, la función $T = r^m\left(\operatorname{sen}\frac{\pi}{4} + D\right)$ conservará el mismo signo que $\operatorname{sen}\frac{\pi}{4}$. Consecuentemente, para todos los ángulos φ para los cuales $|\operatorname{sen} m\varphi| > \sqrt{\frac{1}{2}}$ se tendrá que $T = r^m(\operatorname{sen} m\varphi + D)$ conservará el mismo signo que $\operatorname{sen} m\varphi$. Analicemos los valores de φ y de $m\varphi$ para cada uno de los $4m$ puntos de la circunferencia:

en (1), $\varphi = \left(\frac{1}{m}\right)\frac{\pi}{4} \;\Rightarrow\; m\varphi = \frac{\pi}{4}$,

en (3), $\varphi = \left(\frac{3}{m}\right)\frac{\pi}{4} \;\Rightarrow\; m\varphi = \frac{3\pi}{4}$,

en (5), $\varphi = \left(\frac{5}{m}\right)\frac{\pi}{4} \;\Rightarrow\; m\varphi = \frac{5\pi}{4}$,

en (7), $\varphi = \left(\frac{7}{m}\right)\frac{\pi}{4} \;\Rightarrow\; m\varphi = \frac{7\pi}{4}$,

en (9), $\varphi = \left(\frac{9}{m}\right)\frac{\pi}{4} \;\Rightarrow\; m\varphi = \frac{9\pi}{4}$,

$\vdots$

en (8m-3), $\varphi = \left(\frac{8m-3}{m}\right)\frac{\pi}{4} \;\Rightarrow\; m\varphi = \frac{(8m-3)\pi}{4}$,

en (8m-1), $\varphi = \left(\frac{8m-1}{m}\right)\frac{\pi}{4} \;\Rightarrow\; m\varphi = \frac{(8m-1)\pi}{4}$.

Podemos deducir que T tendrá un valor positivo siempre que $m\varphi$ caiga entre $\frac{\pi}{4}$ y $\frac{3\pi}{4}$, es decir, del punto (1) al punto (3). De igual manera, la primera superficie tendrá un valor positivo cuando $m\varphi$ caiga entre $\frac{9\pi}{4}$ y $\frac{11\pi}{4}$, es decir, del punto (9) al punto (11), y de manera general siempre que $m\varphi$ caiga entre $\frac{(8k+1)\pi}{4}$ y $\frac{(8k+3)\pi}{4}$ con $k \in \{0, 1, 2, \ldots, m-1\}$; dicho de otro modo, entre los puntos $(8k+1)$ y $(8k+3)$, como se ilustra en la figura 5.7. Por otra parte, T siempre tendrá un valor negativo cuando $m\varphi$ caiga entre $\frac{5\pi}{4}$ y $\frac{7\pi}{4}$ o entre los puntos (5) y (7) de la circunferencia de radio r; T también será negativa entre los puntos (13) y (15), y de manera general entre $(8k+5)$ y $(8k+7)$ con $k \in \{0, 1, 2, \ldots, m-1\}$, como se muestra en la figura 5.7. En ninguno de los intervalos mencionados $T = 0$, pues $\operatorname{sen} m\varphi$ nunca es nulo.

Puesto que en (3) el valor de T es positivo y en (5) es negativo, y como T es una función continua, se tendrá que en alguna parte entre (3) y (5) T será igual a 0. Por la misma razón, también será igual a 0 entre (7) y (9), entre (11) y (13) y de manera general entre $(4k+3)$ y $(4k+5)$, con k un número entero. En total se tendrán $2m$ puntos donde $T=0$, puesto que de los $4m$ intervalos consecutivos de la circunferencia, en la mitad de ellos la función T tiene valores positivos o negativos.

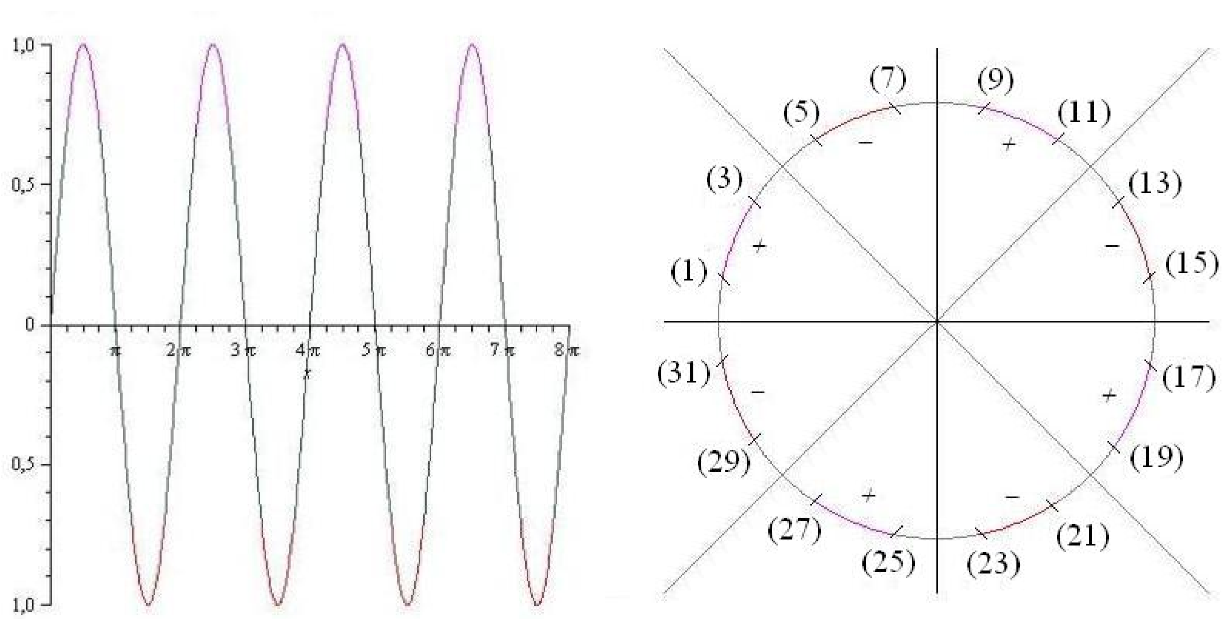

Figura 5.7. Valores positivos y negativos para la superfice T, según los valores del ángulo $m\varphi$ para la ecuación $x^4-3x^2+6x-2=0$

Pero, ¿puede existir más de un punto en algún intervalo $[4k+3, 4k+5]$ de la circunferencia donde $T=0$? Supongamos sin pérdida de generalidad que entre (3) y (5) hay al menos dos puntos donde $T=0$; así, T tiene un máximo o un mínimo en ese intervalo; es decir, $\frac{dT}{d\varphi}=0$ en algún punto de dicho intervalo. Analicemos la derivada de T:

$$\frac{dT}{d\varphi} = mr^m \cos m\varphi + (m-1)Ar^{m-1}\cos(m-1)\varphi + \cdots + Lr\cos\varphi$$
$$= mr^m\left(\cos m\varphi + \frac{(m-1)A}{mr}\cos(m-1)\varphi + \cdots + \frac{L}{mr^{m-1}}\cos\varphi\right).$$

Tomando $E = \frac{(m-1)A}{mr}\cos(m-1)\varphi + \frac{(m-2)B}{mr^2}\cos(m-2)\varphi + \cdots + \frac{L}{mr^{m-1}}\cos\varphi$, $\frac{dT}{d\varphi} = mr^m(\cos m\varphi + E)$. Veamos bajo qué circunstancias $\frac{dT}{d\varphi}$ conserva el mismo signo que $\cos m\varphi$. En (3) $m\varphi = \frac{3\pi}{4}$ y $\cos\frac{3\pi}{4} = -\sqrt{\frac{1}{2}}$. Para que $\frac{dT}{d\varphi} = mr^m(\cos m\varphi + E)$ tenga el mismo signo que $\cos m\varphi$ para cualesquier valores de $A, B, \ldots, L$, se necesita que $|E| < \sqrt{\frac{1}{2}}$. Tenemos que

$$|E| = \left|\tfrac{(m-1)A}{mr}\cos(m-1)\varphi + \tfrac{(m-2)B}{mr^2}\cos(m-2)\varphi + \cdots + \tfrac{L}{mr^{m-1}}\cos\varphi\right|$$
$$|E| = \left|\tfrac{(m-1)A}{mr}\cos(m-1)\varphi + \tfrac{(m-2)B}{mr^2}\cos(m-2)\varphi + \cdots + \tfrac{L}{mr^{m-1}}\cos\varphi\right|$$
$$= \left|\tfrac{(m-1)A}{mr}\right||\cos(m-1)\varphi| + \left|\tfrac{(m-2)B}{mr^2}\right||\cos(m-2)\varphi| + \cdots + \left|\tfrac{L}{mr^{m-1}}\right||\cos\varphi|$$
$$\leq \left|\tfrac{(m-1)A}{mr}\right| + \left|\tfrac{(m-2)B}{mr^2}\right| + \cdots + \left|\tfrac{L}{mr^{m-1}}\right| \text{ (pues } 0 \lneq |\cos k\varphi| \leq 1)$$
$$= \tfrac{(m-1)}{m}\left|\tfrac{A}{r}\right| + \tfrac{(m-2)}{m}\left|\tfrac{B}{r^2}\right| + \cdots + \tfrac{1}{m}\left|\tfrac{L}{r^{m-1}}\right|$$
$$< \left|\tfrac{A}{r}\right| + \left|\tfrac{B}{r^2}\right| + \cdots + \left|\tfrac{L}{r^{m-1}}\right|$$
$$< \tfrac{g}{r} + \tfrac{g}{r^2} + \cdots + \tfrac{g}{r^{m-1}} \text{ (donde } g \text{ es el máximo de } \{|A|, |B|, \ldots, |L|\})$$
$$< \tfrac{g}{r} + \tfrac{g}{r^2} + \cdots + \tfrac{g}{r^{m-1}} \text{ (donde } g \text{ es el máximo de } \{|A|, |B|, \ldots, |L|\})$$

Entonces, $|E| < g\left(\frac{1}{r-1}\right)$; pero si $\frac{g}{r-1} < \sqrt{\frac{1}{2}}$, tendremos que $|E| < \sqrt{\frac{1}{2}}$. Esta condición nos lleva a que $r > g\sqrt{2} + 1$, que es la misma condición que empleamos para que la superficie $T = r^m(\operatorname{sen} m\varphi + D)$ conservara el mismo signo que $\operatorname{sen} m\varphi$. Se deduce entonces que en la misma circunferencia de radio $r > g\sqrt{2} + 1$ la derivada $\frac{dT}{d\varphi} = mr^m(\cos m\varphi + E)$ mantiene el mismo signo que $\cos m\varphi$. De aquí que para todos los ángulos φ tales que $|\cos m\varphi| > \sqrt{\frac{1}{2}}$ se tendrá que $\frac{dT}{d\varphi}$ consevará el mismo signo que $\cos m\varphi$. De lo anterior se tiene que $\frac{dT}{d\varphi}$ es negativa cuando $m\varphi$ cae entre $\frac{3\pi}{4}$ y $\frac{5\pi}{4}$, es decir entre los puntos (3) y (5) de la circunferencia, como se muestra en la figura 5.8, por lo que $\frac{dT}{d\varphi} \neq 0$ en ese intervalo de la circunferencia. Luego $T = 0$ en un solo punto entre (3) y (5).

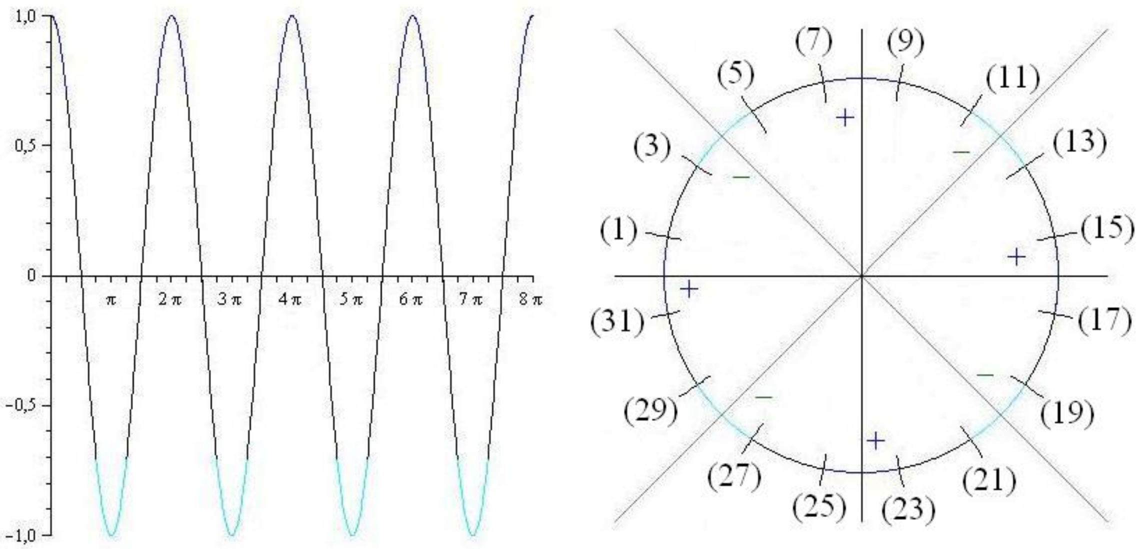

Figura 5.8. Derivada de T

De manera similar se demuestra que U tiene un valor negativo en cualquier punto desde (3) hasta (5), de (11) a (13) y de manera general de $(8k + 3)$ a $(8k + 5)$. La supeficie U tendrá un valor positivo de (7) a (9), de (15) a (17) y de forma general de $(8k + 7)$ a $(8k + 9)$. Entonces $U = 0$ en algún punto entre (1) y (3), entre (5) y (7), entre (9) y (11) y de forma general entre $(4k + 1)$ y $(4k + 3)$ con $k \in \{0, 1, 2, ..., 2m - 1\}$; es decir, en $2m$ puntos. Puede comprobarse que en esos intervalos la derivada $\frac{dU}{d\varphi} = 0$, por lo cual no existen más de $2m$ puntos sobre la circunferencia en los que $U = 0$.

En el apartado 20 de su disertación doctoral, Gauss (1799) explicó que si trazamos otro círculo con un radio mayor que r y lo dividimos de la misma manera, en ese círculo también caerá un punto entre (3) y (5), entre (7) y (9) y en general entre $(4k+3)$ y $(4k+5)$ donde $T=0$. Lo mismo pasará si trazamos un cículo de radio menor que r pero mayor que $g\sqrt{2}+1$. Entre más próximos se encuentren los círculos, más próximos se encontrarán aquellos puntos donde $T=0$ (por ejemplo, entre (3) y (5)); de tal forma que será más claro que la circunferencia de radio r se interseca con alguna rama de la primera curva precisamente en ese punto, como se ilustra en la figura 5.9. Lo mismo ocurrirá para los otros puntos donde $T=0$.

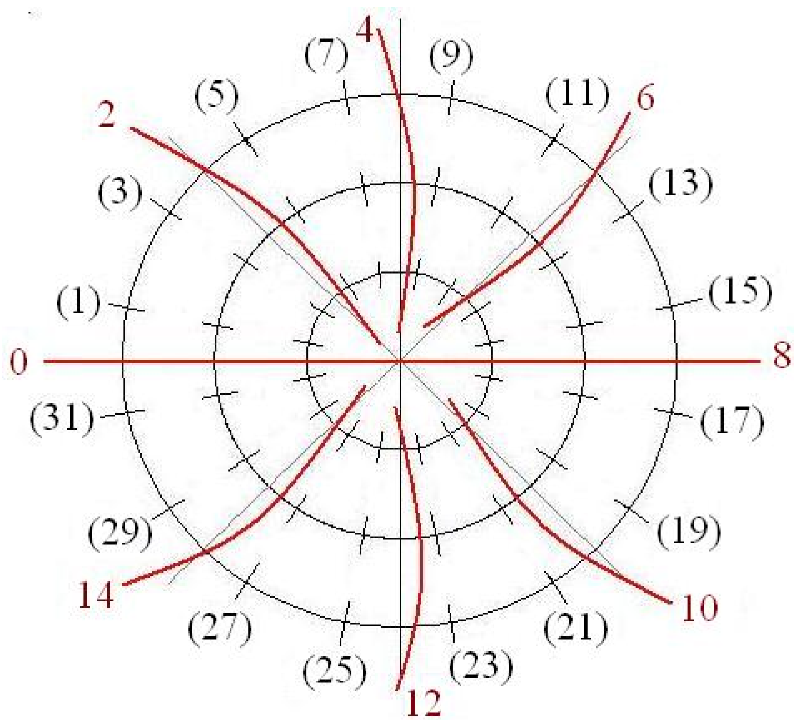

Figura 5.9. Ramas de $T=0$ entrando en circunferencias de radio mayor que $g\sqrt{2}+1$

De igual manera, el círculo de radio r se intersecará con alguna rama de la segunda curva en los $2m$ puntos donde $U = 0$. En resumen, cuando se traza un círculo de radio $r > g\sqrt{2} + 1$, $2m$ ramas de la primera curva y $2m$ ramas de la segunda entrarán en él, de tal modo que entre dos ramas de la primera curva se encontrará una de la segunda. En el apartado 21 de su disertación, Gauss escribió:

> Iam ex hoc situ relativo ramorum in circulum intrantium tot modis diversis deduci potest, intersectionem alicuius rami lineae primae cum ramo lineae secundae intra circulum necessario dari, [...] (Gauss, 1799, §21)
>
> [Ya se puede deducir de esta posición relativa de las ramas entrando al círculo de tantas maneras diversas que necesariamente debe ocurrir dentro del círculo una intersección de alguna rama de la primera línea con una rama de la segunda línea, [...]]

Designemos con el número 0 la parte izquierda del eje *GG'*, que es la primera rama de la primera curva. Asignemos con 1 a la rama inmediata que entra al círculo, la cual pertenece a la segunda curva. Con el número 2 marquemos la segunda rama de la primera curva, con el número 3 la segunda rama de la segunda curva y así sucesivamente hasta la rama $4m - 1$. Con los números pares marcamos las ramas de la primera curva y con los números impares las ramas de la segunda curva, como se muestra en la figura 5.10.

Mostraremos ahora que cualquier rama de cualquiera de las dos curvas que entra al círculo necesariamente saldrá de él. Observese la figura 5.11; supongamos que la rama 2, por ejemplo, no sale. Podemos entrar al círculo por cualquier parte permitida del espacio

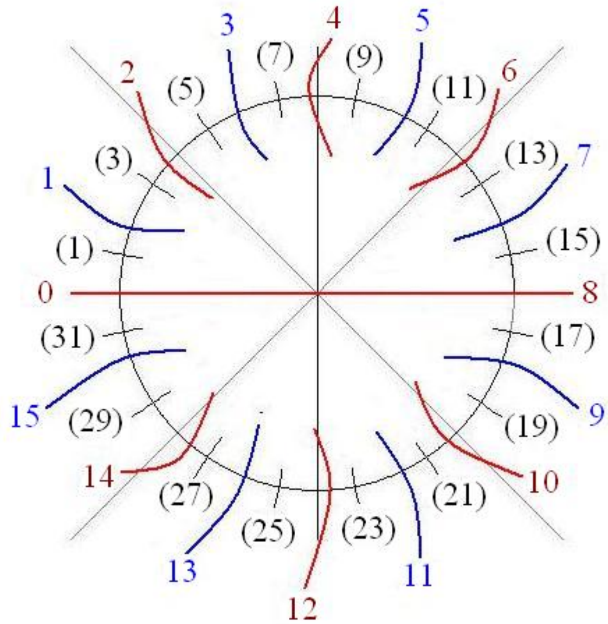

Figura 5.10. Ramas numeradas de las curvas $T = 0$ y $U = 0$ asociadas

a una ecuación de cuarto grado

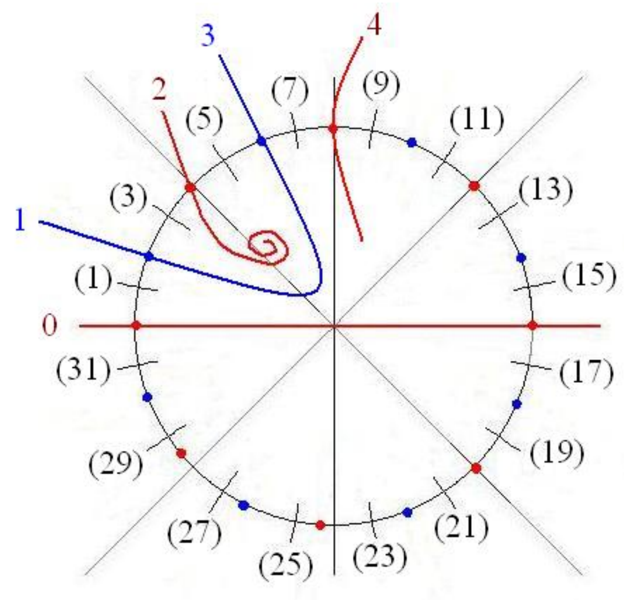

Figura 5.11. Supóngase que la rama 2 no sale del círculo de radio mayor

que $g\sqrt{2} + 1$

entre la rama 0 y la rama 2; esta última se ha perdido dentro del círculo de tal suerte que podemos rodearla y finalmente salir por algún lugar entre 2 y 4 sin intersecar la primera curva; veamos que esto es absurdo.

Al entrar entre las ramas 0 y 2 la *primera superficie* o superficie T tiene valores positivos; es decir, está sobre o por encima del círculo. Al salir entre las ramas 2 y 4, la superficie T tiene valores negativos; es decir, está bajo el círculo. Entonces en algún momento estuvimos sobre la primera superficie; es decir, sobre la primera curva y por tanto tocamos la rama 2. Luego, toda rama que entra en la circunferencia tiene que salir de ella.

De lo anterior se tiene que toda rama señalada con un número par está conectada con otra marcada con un número par dentro del círculo y, de igual forma, toda rama etiquetada con un número impar está conectada con otra impar en el interior del círculo. Veamos que independientemente de la naturaleza de la conexión entre números pares e impares, siempre ocurrirá una intersección de la primera curva con la segunda.

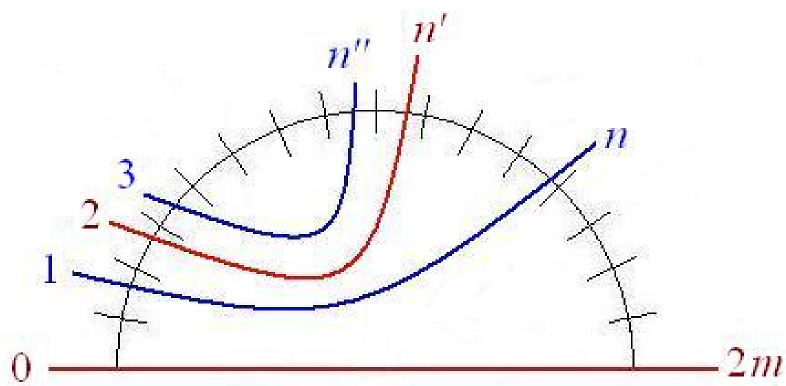

Figura 5.12. Conexiones entre ramas

Supongamos que se pueden unir dos puntos pares y dos puntos impares cualesquiera de tal manera que no resulta una intersección de la primera curva con la segunda. Puesto que el eje GG' forma parte de la primera curva, se tendrá que el punto 0 está conectado con el punto $2m$. El punto 1 puede conectarse entonces con cualquier otro punto impar n tal que $n < 2m$, pues de otro modo una rama de la primera curva se intersecaría con una rama de la segunda; esto se ilustra en la figura 5.12. Unamos ahora el punto 2 con un punto n', donde $n' < n$, para evitar que esta nueva rama de la primera curva se interseque con la rama de la segunda curva que hemos construido. El punto 3 tendrá que unirse entonces con un punto n'' tal que $n'' < n'$. Notemos que la cantidad de puntos que hay entre cualesquiera dos puntos impares es un número impar, y la cantidad de puntos entre dos puntos pares también es impar. Lo anterior nos conduce a que en algún momento sólo reste conectar el punto h con el punto $h + 2$, como en la figura 5.13, por lo que la rama que entra al círculo a través del punto $h + 1$ necesariamante se intersecará con la rama que une h con $h + 2$. Pero alguna de ellas pertenece a la primera curva mientras la otra pertenece a la segunda curva, lo que contradice la suposición inicial, y por lo tanto existe una intersección de la primera curva con la segunda curva.

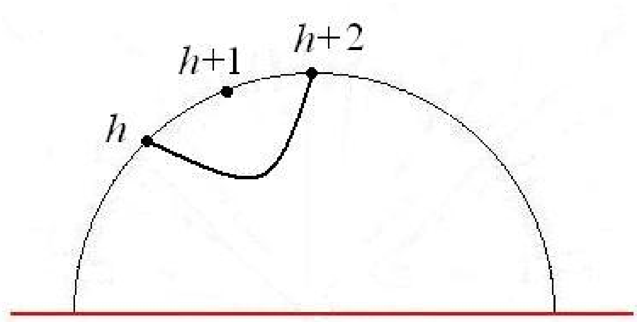

Figura 5.13. Conexiones entre ramas

En el apartado 23 de su demostración, Gauss explicó que no sólo habrá una sino al menos m intersecciones de la primera curva con la segunda, pudiendo suceder que la primera curva se interseque con varias ramas de la segunda en un mismo punto generando que X tenga varios factores iguales. En ese mismo apartado Gauss escribió:

> Denique observo, minime impossibile esse, ut demonstratio praecedens, quam hic principiis geometricis superstruxi, etiam in forma mere analytica exhibeatur: sed eam repraesentationem, quam hic explicavi, minus abstractam evadere credidi, verumque nervum probandi hic multo clarius ob oculos poni, quam a demonstratione analytica exspectari possit. (Gauss, 1799, §23)
>
> [En seguida advierto que de ninguna manera es imposible presentar la demostración precedente, la cual aquí construí sobre principios geométricos, también de manera puramente analítica: mas mediante esta representación, la cual expuse aquí, pienso que resulta menos abastracta, y realmente aquí se coloca ante los ojos el centro neurálgico de lo que se demuestra mucho más claramente de lo que se podría esperar de una demostración analítica.]

Las magnitudes complejas de Gauss

En una carta de Carl Friedrich Gauss (1777 - 1855), dirigida a Wolfang Bolyai (1775 - 1856), le dio a conocer que había elaborado un ensayo en el que demostraba el teorema fundamental del álgebra y por medio del cual pretendía obtener su grado de doctor. En este documento Gauss hizo una fuerte crítica a los intentos de demostración del teorema desarrollados hasta esos momentos por algunos estudiosos de la época (Director 2002, p. 22).

Para su demostración Gauss utilizó el lenguaje de la geometría de posición, pues aunque las ideas plasmadas en la misma no pertenecían a tal dominio, su lenguaje permitía mostrarlas con mayor facilidad. Los fundamentos de la demostración se remontan a la concepción platónica de magnitud en la antigua Grecia. Los estudiosos griegos distinguieron tres diferentes tipos de magnitudes, cada una generada por una acción distinta; así, estas magnitudes poseen potencias diferentes. En primer lugar se encuentra la magnitud lineal o magnitud simplemente extendida que se puede generar mediante la "extensión" de un segmento de recta; en segungo lugar se tiene la magnitud doblemente extendida que se genera por una acción circular o rotación de un segmento rectilíneo (radio) alrededor de un punto, siendo ésta la magnitud propia de las superficies; y finalmente está la acción circular extendida o triplemente extendida, acción con la que se generan por ejemplo un cono, un cilindro o un toro (Director 2002, p. 22).

Se debe aclarar que una magnitud de potencia inferior no puede generar una magnitud de potencia superior; es decir, una magnitud lineal no puede generar una magnitud doblemente extendida o de superficie, del mismo modo que esta magnitud no puede generar una magnitud de volumen o triplemente extendida. Como ejemplo, considérese la duplicación del área de un cuadrado.

Sea el cuadrado $ABCD$ de área 1 (véase la figura figura 5.14). Para construir un cuadrado con el doble de área partiendo de éste, es necesario construirlo sobre su diagonal

BD, cuya longitud es $\sqrt{2}$. Obsérvese que la longitud $\sqrt{2}$ es inconmensurable con la longitud del lado del cuadrado original. De aquí que la diagonal generadora del nuevo cuadrado no tiene el mismo poder que los lados originales; es decir, no se puede obtener como una extensión lineal de ellos.

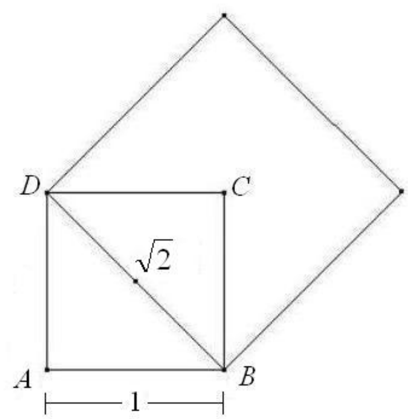

Figura 5.14. Construcción del cuadrado de área 2

En las investigaciones de la resolución de ecuaciones aparecieron raíces cuadradas de números negativos sin que alguien anterior a Gauss hubiera podido explicar qué clase de acción genaraba estas magnitudes. Tampoco existía un referente físico para las ecuaciones de grado mayor que 3, a pesar de la importancia formal que cobraron.

Desde tiempos muy anteriores a Gauss algunos estudiosos manipularon raíces cuadradas de números negativos, pero la interrogante persistía: si se puede representar el significado de $x^2 = 1$ como un cuadrado de área igual a 1, ¿qué referente físico había para $x^2 = -1$? Otras estudiosos como Euler (1707 - 1783) y Lagrange (1736 - 1813) afirmaron que ecuaciones del tipo $x^2 = -1$ tenían solución pero que tal solución era “imposible”.

Para elaborar la demostración del teorema fundamental del álgebra, Gauss desarrolló la concepción de las *raíces cuadradas de números negativos*, denominándolas *números complejos*; los consideró como la sombra de una acción sobre otra. Gauss, al igual que Leibniz (1646 - 1716), rechazó la idea de trabajar deductivamente en su investigación sobre los números complejos; eligió ver “hacia atrás” y desprender el significado físico de $\sqrt{-1}$ del concepto de magnitud de los antiguos griegos; específicamente, recurrió al principio de *cuadrar*.

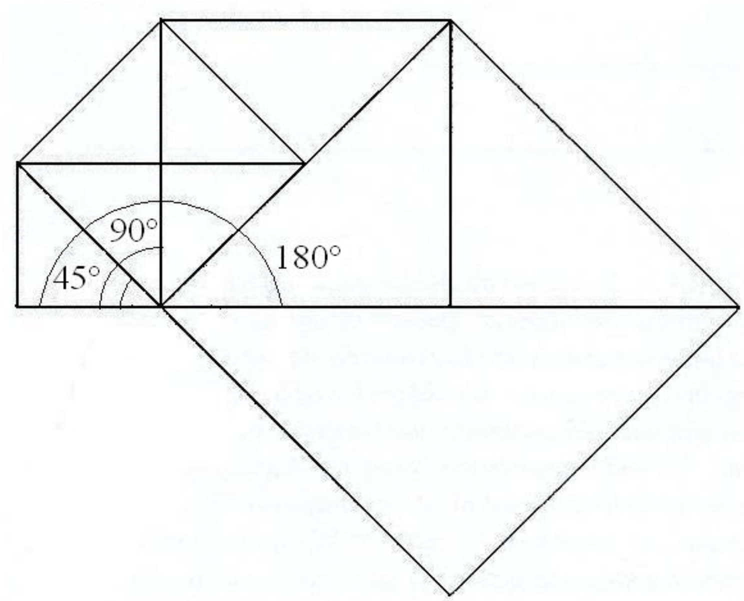

Figura 5.15. La duplicación del área de un cuadrado

Pero, ¿qué significa *cuadrar* según los estudiosos de Platón (h. 427 - h. 348)? Obsérvese que dado un cuadrado cuya área es 1, para obtener un nuevo cuadrado cuya área sea el doble basta con trazar su diagonal y sobre ésta construir el nuevo cuadrado. Tomando como referencia el lado horizontal inferior del cuadrado inicial, la acción que generó el cuadrado cuya área es 2 fue una rotación de $45\,°$ y una extensión de la longitud de 1 a $\sqrt{2}$. Para obtener un cuadrado de área 4 a partir del cuadrado de área 2 se necesita construirlo sobre la diagonal de éste, es decir, se hace una nueva rotación de 45° (con la que se tiene una rotación de 90° a partir del lado horizontal inferior del cuadrado inicial) y se hace una extensión de $\sqrt{2}$ a 2, como se ilustra en la figura 5.15. Podemos continuar este proceso indefinidamente.

En resumen, para cuadrar la magnitud $\sqrt{2}$ representada por la diagonal del cuadrado de área 1 y considerando como referencia el lado horizontal inferior de éste, el ángulo de 45° se duplicó a 90° y se hizo una extensión equivalente al cuadarado de $\sqrt{2}$ (véase la figura 5.16); para cuadrar la magnitud igual a 2 se duplicó el ángulo de rotación de 90° a 180° y se hizo una extensión de 2 a 4; para cuadrar la magnitud igual a 4 se duplicó el ángulo de 180° a 360° y se hizo una extensión de 4 a 16 (Director 2002, p. 24).

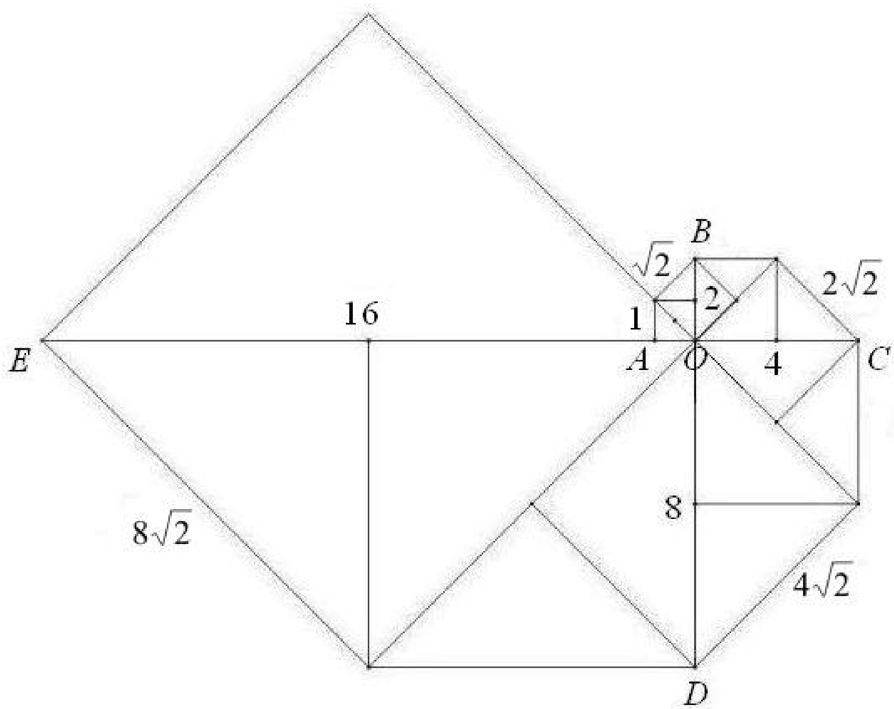

Figura 5.16. Generación de magnitudes cuadradas

De acuerdo con lo anterior, para obtener la raíz cuadrada se sigue un procedimiento inverso: el ángulo se reduce a la mitad y la longitud cambia por su raíz cuadrada. De este modo Gauss obtuvo los elementos para comprender el significado de $\sqrt{-1}$: asoció sus números complejos a las acciones combinadas de rotar y extender sobre el plano; así, cada punto del plano estaría representado por un número complejo.

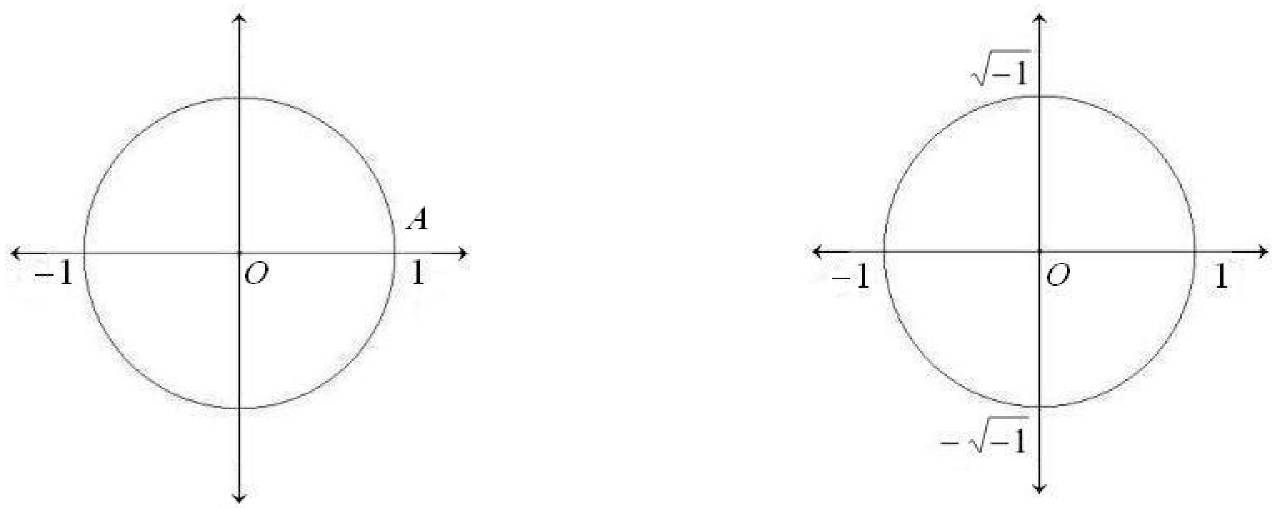

Figura 5.17. La unidad en el dominio complejo de Gauss

Consideremos el círculo unitario: el centro será el origen y los extremos de uno de sus diámetros corresponderán a 1 y -1, como se muestra en la figura 5.17. Tomando como eje de referencia este diámetro, la raíz cuadrada de -1 se obtendra al reducir el ángulo de rotación entre 1 y -1 a la mitad y la longitud del diámetro a $\sqrt{1}$. De este modo, $\sqrt{-1}$ se ubicará sobre la circunferencia a $90°$. Puesto que el ángulo de $180°$ es equivalente al ángulo de $180° + 360° = 540°$, tendremos que una segunda raíz cuadrada de $\sqrt{-1}$ se ubicara a $\frac{540°}{2} = 270°$ y a una distancia 1 del origen, es decir, es el punto antípoda de la primera raíz. Así, la unidad en el dominio complejo de Gauss es la circunferencia unitaria, donde 1 representa una rotación completa, -1 media rotación, $\sqrt{-1}$ un cuarto de rotación y $-\sqrt{-1}$ tres cuartos de rotación (Director 2002, p. 28).

Por medio de las acciones de rotar y extender (o acortar; en el caso de $\sqrt{1} = 1$ obviamente queda igual) cualquier poder algebraico o potencia algebraica puede observarse en el dominio complejo. Por ejemplo, para *cubicar* un número complejo, el ángulo de rotación se triplica y la longitud se multiplica tres veces por sí misma. Comprender este proceso fue necesario para Gauss porque para él una magnitud tenía sentido sólo si se conocía su principio generador (Director 2002, p. 28)

En su disertación de doctorado Gauss demostró que el principio general que produce cualquier ecuación algebraica podía hacerse visible por medio de superficies; es decir, se preocupó por la representación de aquello que genera a la ecuación y no por la representación de lo que la ecuación genera, como en otros tiempos sucedió al referirse a las líneas, cuadrados y cubos. Pensemos por ejemplo en un número complejo z; para cuadrarlo hay que duplicar su ángulo de rotación α para obtener 2α y su longitud cambia de A a A^2. Si se repitiera este proceso con todos los números complejos de la circunferencia en donde se ubica z, se tendría que el círculo de z se proyectaría dos veces en el círculo de z^2. Si z se cubicara, el ángulo α se convertiría en 3α y la longitd A cambiaría por A^3; entonces el círculo de z se proyectaría tres veces en el círculo de z^3.

Gauss utilizó esto para crear las superficies que son sólo las *sombras de sombras* del principio generador de cualquier ecuación algebraica. Como primer paso, ubiquemos en el círculo unitario el triángulo que producen las rotaciones y extensiones de los números complejos: el cateto horizontal del triángulo es el coseno del ángulo de rotación y el cateto vertical es el seno, como se muestra en la figura 5.18.

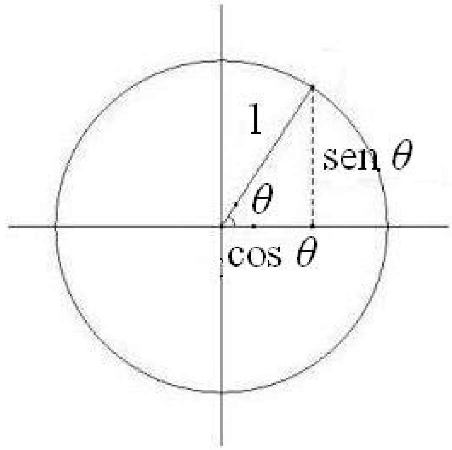

Figura 5.18. La unidad en el dominio complejo de Gauss

Consideremos las relaciones que guardan el seno y el coseno del ángulo mientras éste avanza en sentido levógiro y la longitud de z es 1. Si $\theta = 0°$, $\operatorname{sen}\theta = 0$ y $\cos\theta = 1$; si $\theta = 90\,°$, $\operatorname{sen}\theta = 1$ y $\cos\theta = 0$. Si $\theta = 180\,°$, $\operatorname{sen}\theta = 0$ y $\cos\theta = -1$; si $\theta = 270\,°$, $\operatorname{sen}\theta = -1$ y $\cos\theta = 0$; finalmente, cuando $\theta = 360\,°$, $\operatorname{sen}\theta = 0$ y $\cos\theta = 1$.

Observemos la relación entre el $\operatorname{sen}\theta$ correspondiente a z y $\operatorname{sen}\alpha$ que corresponde a z^2. Cuando el ángulo de rotación de z va de 0 a $90\,°$, el valor de $\operatorname{sen}\theta$ va de 0 a 1, mientras que el ángulo de rotación de z^2, α, va de 0 a $180\,°$, variando $\operatorname{sen}\alpha$ de 0 a 1 y nuevamente a 0 (obsérvese la figura 5.19).

Cuando el ángulo θ de z va de $90\,°$ a $180\,°$, $\operatorname{sen}\theta$ varía de 1 a 0 mientras que el ángulo α de z^2 va de $180\,°$ a $360\,°$ y $\operatorname{sen}\alpha$ de 0 a -1 y a 0, como se muestra en la figura 5.20. Entonces, $\operatorname{sen}\theta$ cambió de 0 a 1 y regresó a 0, mientras que $\operatorname{sen}\alpha$ se movió de 0 a 1 a 0 a -1, y finalmente a 0; es decir, en media rotación de z, $\operatorname{sen}\alpha$ cambió de positivo a negativo y regresó a 0. Entonces, en la rotación completa de z $\operatorname{sen}\alpha$ cambiaría de positivo a negativo 2 veces; es decir, se movería de 0 a 1 a 0 a -1 a 0 a 1 a 0 a -1 a 0.

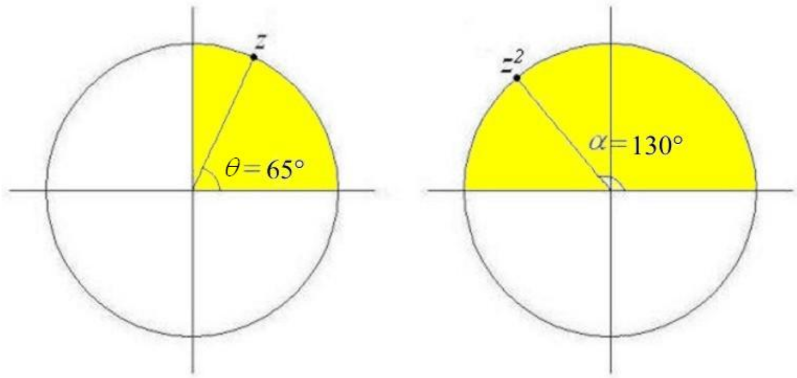

Figura 5.19. Ángulo de z que varía entre $0°$ y $90°$

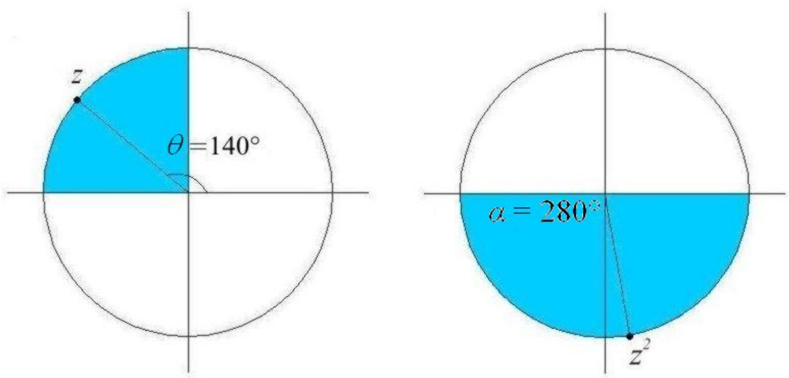

Figura 5.20. Ángulo de z que varía entre $90°$ y $180°$

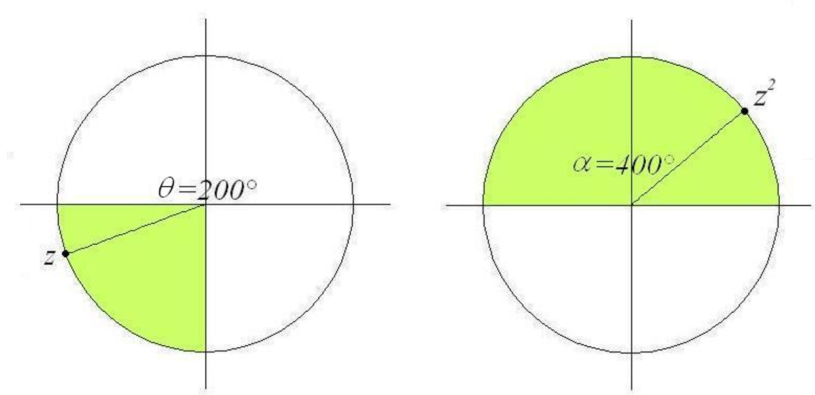

Figura 5.21. Ángulo de z que varía entre $180°$ y $270°$

Para cada ecuación Gauss construyó una superficie relacionada con el seno del ángulo de rotación de cada número complejo y otra relacionada con el coseno. Cada punto de la superficie correspondiente al seno se obtiene multiplicando la magnitud del número complejo por el seno de su ángulo de rotación. Así, la potencia de un punto en el plano está representado por la altura de la superficie sobre ese punto. Entonces, si un número complejo está más cerca del origen, su altura en la superficie disminuirá respecto a aquellos que se encuentran más alejados del origen.

Tomando en cuenta que el seno del ángulo de un número complejo varía de positivo a negativo, según el poder que multiplique dicho ángulo, la superficie subirá y bajará. Así, una superficie de una ecuación cuadrática tendrá dos "jorobas" sobre el plano y dos bajo él; una superficie de una ecuación cúbica tendrá tres "jorobas" sobre el plano y tres bajo él. Todo lo anteriormente descrito se aplica a la superficie relacionada con el coseno del ángulo de rotación de los números complejos (Director 2002, pp. 29 - 31).

Gauss observó que las intersecciones de estas dos curvas con el plano complejo dependían del número de "jorobas" y por tanto de la potencia más elevada de la ecuación, además de que estas curvas se intersecarían, puesto que el seno y el coseno se desfasan $90°$. Con ello Gauss demostró que una ecuación tiene tantas raíces como su potencia más alta.

Las críticas a la primera demostración de Gauss

El 16 de julio de 1799 la facultad de filosofía de Helmsted confirió el título de doctor a Carl Friedrich Gaus antes de que su tesis fuera publicada y sin presentar defensa oral

(Dunnington 2004, p. 35). En su disertación de doctorado Gauss presentó la primera demostración del ahora llamado *teorema fundamental del álgebra*. Desde su aparición hasta nuestros días, múltiples críticas se han suscitado en torno a esa demostración.

La demostración de Gauss ha sido comparada en varios escritos con los intentos de demostración de Jean Le Rond d'Alembert (1717 - 1783), Leonhard Euler (1707 - 1783) y Joseph Louis Lagrange (1736 - 1813), entre otros. Gauss elaboró otras tres demostraciones del teorema fundamental del álgebra. Algunos de sus más tempranos descubrimientos aparecieron en ese escrito, así como en sus *Disquisitiones Arithmeticae* de 1801 (Struik 1987, p. 142). La segunda demostración fue dada a conocer en 1815 (Gauss, 1815), la tercera en 1816 (Gauss, 1816) y la cuarta fue realizada en 1849 (Gauss, 1849). Esta última demostración fue una nueva versión de su disertación de doctorado con motivo de la celebración del 50 aniversario de la misma ($1849 - 1799 = 50$). Los cambios de la demostración no fueron muy profundos; la diferencia más notable fue el uso explícito del dominio complejo, el cual fue evitado en su disertación de doctorado, aunque Gauss hizo uso implícito de él (Bühler 1987, p. 150). En la introducción de su cuarta demostración, publicada en 1850, Gauss escribió sobre su disertación de doctorado:

> La memoria aparecida en 1799, *Demonstratio [. . .]*, tenía un doble objetivo, en primer lugar mostrar que todas las demostraciones intentadas hasta el momento del más importante teorema de la teoría de las ecuaciones algebraicas eran insuficientes e ilusorias, y después dar una prueba perfectamente rigurosa. (Citado en: Gilain 1991, p. 118)

Según Gilain, afirmaciones como ésta fueron repetidas por historiadores de las matemáticas, como Struik y Boyer, con el fin de desvalorizar los intentos de demostración anteriores por no otorgale al teorema estatus de teorema de existencia, aunado a la estima excesiva por la modernidad y el rigor de los trabajos de Gauss (Gilain 1991, p. 118). Al respecto consideremos la crítica de Bernard Bolzano (1781 - 1848). En el prefacio de la memoria de 1817 escrita por él, cuyo título comienza con *Rein Analytisher Beweis* (*Demostración puramente analítica*), explicó por qué las demostraciones existentes de lo que ahora llamamos teorema fundamental del álgebra no eran satisfactorias (Barbin 1994, p. 46). Las demostraciones de d'Alembert, Euler, Lagrange, Pierre Simon de Laplace (1749 - 1827) y Gauss existentes en aquellos tiempos no cubrieron los requisitos de Bolzano para ser consideradas demostraciones válidas. Particularmente, Bolzano criticó las demostraciones que recurrían a la intuición geométrica, pues según él las demostraciones no sólo deben apoyarse en la evidencia sino que deben exponer el fundamento objetivo de la verdad que se está demostrando (Barbin 1994, p. 47).

De acuerdo con la filosofía de Bolzano, un teorema se vuelve fundamentalmente verdadero cuando se relaciona con otros teoremas o proposiciones. En este contexto, para la existencia de las raíces de un polinomio la idea fundamental es la de *función continua*, noción que el mismo Bolzano definió (Barbin 1994, p. 47). Para él la verdad del teorema se establece de acuerdo con su utilidad como fundamento en la estructura completa de las matemáticas; el problema ontológico concerniente a la existencia y la verdad del objeto que el teorema trata ya no es relevante. Bolzano aplicó su definición al teorema fundamental del álgebra y no consideró como dificultad la existencia de la raíz ni el tipo de ser que representó; su objeto de investigación fue una proposición que encajara en una estructura y no contempló investigar como objetos los entes a los que dicha proposición hace referencia. Bolzano resolvió su problema filosófico de forma pero no de fondo. Con relación a lo anterior se afirma que Bolzano creía que las matemáticas (y por tanto las proposiciones matemáticas) tenían una existencia ontológica y lo que constituía la existencia de una proposición era su relación con otras proposiciones (Barbin 1994, p. 47); es decir, para Bolzano las matemáticas tenían existencia ontológica pero no directamente sobre el ser de los objetos que intervienen en sus proposiciones sino que consideró de mayor importancia el ser de la proposición dentro de la estructura matemática.

Bolzano rechazó las consideraciones geométricas sobre las que sus predecesores se apoyaron y esa actitud se extendió cada vez más (Barbin 1994, p. 47). En el caso de Gauss, la geometría fue un modelo indispensable para explicar el ser subyacente de lo que fundamenta su demostración, no las proposiciones mismas, sino aquello de lo que las proposiciones hablan.

Respecto a la continuidad, Stillwell (1989) indicó que Gauss no tuvo claro este concepto. Se valió de la afirmación de Gauss usada en su disertación de doctorado de que una rama de cualquier curva algebraica que entra en un espacio limitado debe necesariamente salir de este espacio en alguna parte. Así, Stillwell opinó que la demostración de Gauss podría ser juzgada de manera diferente en la actualidad. Esto es inobjetable: la definición de función continua surgió después de la disertación de Gauss y con un espíritu diferente al de ésta, una demostración completamente rigurosa de acuerdo con los estándares de nuestro siglo resultaría imposible en la época de Gauss. Sin embargo, la rigurosidad de Gauss no va encaminada al formalismo puro: su rigor tiene como punto de partida el principio generador de los elementos que intervienen en su demostración, y este hecho deja muy por detrás los intentos de sus contemporáneos.

Sobre el tema del rigor en la primera demostración del teorema fundamental del álgebra, Bühler escribió:

> La idea de la demostración es intuitivamente accesible; la demostración misma no es rigurosa de acuerdo con nuestros estándares presentes, pero es fácilmente superior a cualquier esfuerzo anterior, específicamente al de d'Alembert, al cual es relacionada. (Bühler 1987, p. 41)

Por otra parte, hay quienes opinan que Gauss intentó convencer a sus lectores de que su demostración era la primera válida mostrando las debilidades de los anteriores intentos, pero dejando de lado que la suya contenía suposiciones no probadas (Stillwell 1989, p. 196). Stillwell señaló dos puntos débiles en la demostración de Gauss que, según su criterio, impiden considerarla una demostración completa: no se emplearon en ella las propiedades geométricas de los números complejos y no se utilizó adecuadamente el concepto de continuidad. Según Stillwell, la identificación del número complejo $x + iy$ con el punto del plano (x, y) no fue reconocida por los matemáticos sino hasta fines del siglo XVIII, siendo hasta 1806 cuando Argand desarrolló completamente esta concepción. Stillwell agregó además que probablemente Gauss tuvo la misma idea que Argand sobre los números complejos, pero tal idea permaneció oculta en su demostración (Stillwell 1989, p. 196).

Sin embargo, lo que Stillwell consideró “oculto” en la demostración de Gauss es en realidad el sustento de la misma. Gauss desarrolló previamente el concepto de número complejo encontrando su principio generador como magnitud, lo cual volvió evidente en su demostración con la creación de dos superficies asociadas a cada ecuación polinomial. Asociado con esto, en 1825 y en 1831 aperecieron los trabajos de Gauss sobre residiuos bicuadráticos. El tratado de 1831 trata del álgebra de los números complejos. La nueva teoría sobre estos números expuesta por Gauss disipó todas las dudas relacionadas con su representación por puntos en el plano (Struik 1987, pp. 144-145).

Para otros autores la demostración de Gauss extendió exitosamente las ecuaciones binomiales a los terrenos de las magnitudes complejas. Dunnington considera como uno de los principales logros de Gauss la colocación de lo imaginario sobre bases firmes, pues Gauss interpretó $\sqrt{-1}$ como la media geométrica entre 1 y -1; además, fue el primero en utilizar el símbolo i para $\sqrt{-1}$. Desde la perspectiva de Dunnington, el hecho de materializar la solución completa de las ecuaciones binomiales fue suficiente para que la facultad de Helmstedt otorgara el doctorado a Gauss (Dunnington 2004, p. 40).

En nuestros días se concede privilegio a la creación de nuevas estructuras matemáticas que surgen de las ya establecidas. Gilain, por ejemplo, se refiere a una observación de Bachmacova sobre la posibilidad de construir un cuerpo de descomposición para una ecuación algebraica sin recurrir a la existencia del cuerpo de los números complejos (Gilain 1991, p. 119).

Es posible que se considere que el teorema fundamental del álgebra ha perdido ya su estatus de fundamental a causa de que el álgebra ya no se relaciona únicamente con la teoría de ecuaciones algebraicas con coeficientes reales o complejos. Sin embargo, es a partir del estudio de las raíces de las ecuaciones algebraicas que surgen estructuras más abstractas en el álgebra.

REFLEXIONES

Comprender por qué las ecuaciones de segundo grado tienen 2 raíces nos ha llevado por caminos que la memorización de un algoritmo no nos puede mostrar. Hemos observado cómo los árabes dieron una interpretación geométrica sobre lo que una expresión cuadrática representa y cómo sus ideas se difundieron por el continente europero gracias al intercambio comercial que prevaleció en la época en que dominaron España. Estas ideas llegaron hasta Girolamo Cardano, quien interpretó geométricamente lo que una expresión cúbica representaba, pero rechazó las expresiones de grado mayor que 3 porque no tenían sentido en la naturaleza, no podían modelar nada real; sin embargo, otros estudiosos de su época trabajaron con expresiones bicuadráticas resolviéndolas satisfactoriamente. El interés por los algoritmos creció. Las matemáticas surgidas en un contexto social comercial se convirtieron en base para el surgimiento de lo que hoy llamamos álgebra. Se tuvo conciencia del carácter científico de todas estas ideas cuando fueron confrontadas por Rafael Bombelli con los trabajos diofantinos.

Por otra parte, fue inminente la necesidad de un simbolismo que permitiera realizar cálculos con mayor facilidad. El surgimiento del simbolismo va encadenado directamente a la noción de número que los estudiosos tenían; finalmente, representar un número de forma general implicó comprender las propiedades que éste posee.

En cuanto al problema planteado inicialmente en este trabajo, se tiene que para resolver satisfactoriamente una ecuación de segundo grado necesitamos comprender cómo se genera un número complejo, qué es lo que lo hace posible y qué lo representa. En este contexto, la primera *demostración* válida del teorema fundamental del álgebra, además de “demostrar” *expone* de manera clara el significado de las ideas que lo componen, mostrando así que las matemáticas no son únicamente un lenguaje creado con una visión formalista.

La demostración de Gauss representa el desenlace de una importante etapa de las matemáticas, en la que se desarrollaron la concepción de los objetos de estudio de la misma, sus relaciones y sus transformaciones. La primera demostración de Gauss de dicho teorema promueve la creación y el descubrimiento e invita a la apropiación de las ideas: no sólo convence, sino que ilustra dejando en segundo plano los argumentos deductivos y el rigor del formalismo al que los estudiosos de su época y los actuales se aferran.

Gauss regresó al pensamiento de los antiguos, retomó su filosofía y planteó el teorema fundamental del álgebra recurriendo a la investigación de la existencia de las raíces de una ecuación y a su forma. Los argumentos de la demostración están íntimamente relacionados con el objeto de estudio del teorema: las magnitudes complejas.

El estudio de las matemáticas desde una perspectiva histórica enriquece el pensamiento gracias a que motiva y permite el planteamiento de muchas otras preguntas, situación que el estudiante no enfrenta en el ámbito escolar. En nuestros días pareciera que ya no se buscan principios generadores de objetos matemático. Esta posición ha quedado relegada por la creación de nuevas estructuras a partir de las ya establecidas. Es por ello que, a pesar de las críticas hacia la primera demostración de Gauss del teorema fundamental del álgebra, debe ser considerada vigente, apropiada y accesible para el entendimiento y la apropiación de los conceptos implicados.

El surgimiento del álgebra puede considerarse paralelo al desarrollo de las ecuaciones en un primer nivel, pues finalmente podemos considerar que una ecuación existe como tal hasta que puede expresarse un número cualquiera mediante un simbolismo universal. Aunque los griegos, los árabes y otros estudiosos del medievo tenían los conocimientos necesarios para resolver problemas que implicaban, desde nuestra actualidad, expresiones cuadráticas, no podemos afirmar que usaban álgebra, pues no tenían el simbolismo requerido ni los métodos generales para una expresión general en la que interviniera un objeto matemático general. Sin embargo, sus matemáticas son tan valiosas como las nuestras; sobresale que las ideas subyacentes de los procedimientos actuales de resolución de ecuaciones y la demostración del teorema fundamental del álgebra descansan en ellas.

APÉNDICE

UNA DEMOSTRACIÓN MODERNA DEL TEOREMA FUNDAMENTAL DEL ÁLGEBRA

En algunos textos de variable compleja se demuestra el teorema fundamental del álgebra con base en teoremas desarrollados después del siglo XVIII. Un ejemplo de estas demostraciones es la que se realiza por medio del teorema de Liouville (Joseph Liouville, 1809 - 1882). Este teorema dice:

Teorema de Liouville. *Si* f *es una función entera y acotada entonces* f *es constante*. (Conway 1995, p. 77)

Conway (1995) enuncia el teorema fundamental del álgebra de la siguiente manera:

Teorema fundamental del álgebra. *Si* $\mathrm{p(z)}$ *es un polinomio no constante entonces existe un número complejo* a *tal que* $\mathrm{p(a) = 0}$. (Conway 1995, p. 77)

La demostración del teorema fundamental del álgebra se realiza de manera indirecta, por reducción al absurdo.

Sea $p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$, con $a_n \neq 0$ y $n \geq 1$, una ecuación polinomial con coeficientes reales. El teorema fundamental del álgebra afirma que el segundo miembro de esta ecuación puede descomponerse en factores de primer o segundo grado; es decir, tiene raíces en los complejos. Concretamente, $p(z) = 0$ tendrá n raíces, algunas de las cuales pueden ser repetidas.

Supongamos que $p(z) = 0$ no tiene solución. Tomando como modelo del campo de los números complejos al plano, lo dividiremos en dos regiones: la primera será $R_1 = \{z \mid |z| \leq r\}$; la región R_2 será el resto del plano, es decir, $R_2 = \{z \mid |z| > r\}$.

Consideremos la función $\frac{1}{p(z)}$; como $p(z)$ nunca se anula, tenemos que $\frac{1}{p(z)}$ es una función continua en todo el plano, en particular en R_1. De acuerdo con la teoría actual de la variable compleja, si $f(z)$ es una función continua en alguna región R del plano complejo, $|f(z)|$ también lo será; además, si R es acotada y cerrada, es decir, si contiene a su frontera, entonces existirá un real positivo M tal que $|f(z)| \leq M$ para todo z en R. Este resultado nos conduce a que $\left|\frac{1}{p(z)}\right| \leq M$ para alguna M real, siempre que z se encuentre en R_1, que es una región cerrada y acotada.

Analicemos ahora el comportamiento de $\frac{1}{p(z)}$ en la región R_2. Para ello utilizaremos la desigualdad

$$|f(z) + g(z)| \geq |f(z)| - |g(z)|, \text{ en la que } |f(z)| \geq |g(z)|.$$

Expresemos $p(z)$ como

$$p(z) = f(z) + g(z) \text{ siendo } f(z) = a_n z^n \text{ y } g(z) = a_{n-1} z^{n-1} + a_{n-2} z^{n-2} + \cdots + a_1 z + a_0.$$

Notemos que r puede ser lo suficientemente grande para cumplir que $|f(z) \geq |g(z)|$.

Tendremos así que

$$\begin{aligned}|p(z)| = |f(z) + g(z)| &= |a_n z^n + a_{n-1}z^{n-1} + \cdots + a_0| \\ &\geq |a_n z^n| - |a_{n-1}z^{n-1} + \cdots + a_0| \\ &= |f(z)| - |g(z)|.\end{aligned}$$

Por otra parte,

$$|a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \cdots + a_0| \leq |a_{n-1}||z|^{n-1} + |a_{n-2}||z|^{n-2} + \cdots + |a_0|.$$

De este modo,

$$\begin{aligned}|a_n z^n + a_{n-1}z^{n-1} + \cdots + a_0| &\geq |a_n z^n| - |a_{n-1}z^{n-1} + \cdots + a_0| \\ &\geq |a_n||z|^n - (|a_{n-1}||z|^{n-1} + \cdots + |a_0|),\end{aligned}$$

siempre que r sea lo suficientemente grande para que el último miembro de esta doble desigualdad sea positivo. Tenemos así que

$$\begin{aligned}|p(z)| &\geq |a_n||z|^n - (|a_{n-1}||z|^{n-1} + \cdots + |a_0|) \\ &= |a_n||z|^n - |z|^{n-1}\left(|a_{n-1}| + \tfrac{|a_{n-2}|}{|z|} + \tfrac{|a_{n-3}|}{|z|^2} + \cdots + \tfrac{|a_1|}{|z|^{n-2}} + \tfrac{|a_0|}{|z|^{n-1}}\right).\end{aligned}$$

Sea A el número mayor del conjunto $\{|a_{n-1}|, |a_{n-2}|, |a_{n-3}|, \ldots, |a_0|\}$; se tendrá que

$$|a_{n-1}| + \frac{|a_{n-2}|}{|z|} + \frac{|a_{n-3}|}{|z|^2} + \cdots + \frac{|a_0|}{|z|^{n-1}} \leq A + \frac{A}{|z|} + \frac{A}{|z|^2} + \cdots + \frac{A}{|z|^{n-1}} \leq nA.$$

Como consecuencia, obtendremos que

$$|a_n||z|^n - |z|^{n-1}\left(|a_{n-1}| + \frac{|a_{n-2}|}{|z|} + \cdots + \frac{|a_0|}{|z|^{n-1}}\right) \geq |a_n||z|^n - |z|^{n-1}\left(A + \frac{A}{|z|} + \cdots + \frac{A}{|z|^{n-1}}\right)$$

$$\geq |a_n||z|^n - |z|^{n-1}(nA);$$

es decir,

$$|p(z)| \geq |a_n||z|^n - |z|^{n-1}\left(A + \frac{A}{|z|} + \cdots + \frac{A}{|z|^{n-1}}\right) \geq |a_n||z|^n - |z|^{n-1}(nA) = |z|^n\left(|a_n| - \frac{nA}{|z|}\right).$$

Lo anterior nos lleva a que

$$\frac{1}{|p(z)|} \leq \frac{1}{z^n\left(|a_n| - \frac{nA}{|z|}\right)};$$

el segundo miembro de la desigualdad alcanzará su valor máximo respecto a los puntos de R_2 cuando $|z| = r$, pues cada punto z de R_2 cumple que $|z| > r$. Así, en la segunda región

$$\frac{1}{|p(z)|} \leq \frac{1}{r^n\left(|a_n| - \frac{nA}{r}\right)}.$$

En R_1 se obtuvo que $\frac{1}{|p(z)|} \leq M$ y en R_2 se concluyó que $\frac{1}{|p(z)|} \leq \frac{1}{r^n\left(|a_n| - \frac{nA}{r}\right)}$. Sea M' la mayor de las dos cotas. Se tendrá que $\frac{1}{|p(z)|} \leq M'$ para cualquier z en el plano complejo. De este modo, la función $\frac{1}{p(z)}$ es acotada. Bajo estas condiciones puede aplicarse el teroema de Liouville obteniendose que $\frac{1}{p(z)}$ es una función constante, es decir, $\frac{1}{p(z)} = k$. Lo anterior lleva a que $p(z) = \frac{1}{k}$ y esto es absurdo ya que $p(z)$ es una función polinomial.

Por lo tanto la ecuación $p(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 = 0$ sí tiene solución en los complejos.

REFERENCIAS

Barbin, Evelyn. 1994. *The meanings of mathematical proof: Relations between history and mathematical education*. En: Joby Milo Anthony (comp.), *In Eves'circles* [MAA Notes, No. 34], pp. 41 - 52. Washington, DC: The Mathematical Association of America.

Bühler,Walter K. 1987. *Gauss. A biographical study*. Nueva York: Springer-Verlag.

Cardano, Girolamo. 1968. *The great art or the rules of algebra*. Massachusets: The MIT Press.

Conway, John B. 1995. *Functions of one complex variable* I. *Second Edition.* Nueva York: Springer.

Director, Bruce. 2002. The fundamental theorem of algebra. Bringing the invisible to the surface. *Fidelio. Journal of poetry, science an statecraft.* XI (3 - 4) pp. 17 - 29.

Dunnington, Waldo G. 2004. *Carl Friedrich Gauss: Titan of science*. The Mathemathical Association of America.

Euclides. 2000. *Elementos. Libros* I-IV. [Traducción: María Luisa Puertas Castaños. Introducción: Luis Vega.] Madrid: Gredos.

Fibonacci, Leonardo. 2002. *Fibonacci's Liber abaci: A translation into modern english of Leonardo Pisano's Book calculation*. Nueva York: Springer.

Gauss, Carl F. 1849. Beiträge zur theorie der algebraischen gleichungen. *Werke* III; pp. 71 - 103.

Gauss, Carolo F. 1799. Demonstratio nova theorematis omnem functionem algebraicam rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse. *Werke* III; pp. 1 - 30.

Gauss, Carolo F. 1815. Demonstratio nova altera theorematis omnem functionem algebraicam rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse. *Werke* III; pp. 32 - 56.

Gauss, Carolo F. 1816. Theorematis de resolubilitate functionum algebraicarum integrarum in factores reales. Demonstratio tertia. *Werke* III; pp. 57 - 64.

Gilain, Christian. 1991. Sur l'histoire du théorème fondamental de l'algebre: théorie des équations et calcul intégral. *Archieve for History of Exact Sciences* 42, pp. 91 - 136.

Hadden, Richard W. 1994. *On the shoulders of merchants*. Albany, Nueva York: State University of New York Press.

Katz, Victor J. 1993. *A history of mathematics. An introduction*. Nueva York: Harper Collins College Publishers.

Klein, Jacob. 1992. *Greek mathematical thought and the origin of algebra*. Nueva York: Dover.

Lorenzo, Javier de. 1989. *Introducción al estilo matemático*. Madrid: Tecnos.

Mahoney, Michael S. 1994. *The mathematical career of Pierre de Fermat, 1601 - 1665 (2a. ed.)*. Princeton, Nueva Jersey: Princeton University Press.

Marx, Carlos. 1847. *Miseria de la filosofía. Respuesta a la "Filosofía de la miseria" del señor Proudhon*. Moscú: Progreso.

Marx, Carlos. 1982. *El capital. Crítica de la economía política.* Volumen I. México: Fondo de Cultura Económica.

Parshall, Karen H. 1988. The art of algebra from al-Khwarizmi to Viète: A study in natural selection of ideas. *History of Science* 26 (72) pp. 129 - 164.

Piaget, J. y García, R. 1982. *Psicogénesis e historia de la ciencia*. México: Siglo XXI.

Pisa, Leonardo de. 1973. *El libro de los números cuadrados*. Buenos Aires: Eudeba.

Platón. 2007. *Diálogos*. México: Porrúa.

Programa de la asignatura de álgebra superior II. Plan de estudios 2007. Universidad Nacional Autónoma de México, Facultad de Ciencias, carrera de matemático.

Obtenido el 25 de mayo de 2008 de la internet:
http://www.matematicas.unam.mx/programas/obligatorias/algebra__superior – II.pdf

Programa de la asignatura de Variable compleja I. Plan de estudios 2007. Universidad Nacional Autónoma de México, Facultad de Ciencias, carrera de matemático. Obtenido el 25 de mayo de 2008 de la internet:
http://www.matematicas.unam.mx/programas/obligatorias/variable__compleja – I.pdf

Serrés, Michel. 1996. *Los orígenes de la geometría*. México: Siglo XXI.

Stillwell, John. 1989. *Mathematics and its history*. Nueva York: Springer-Verlag.

Struik, Dirk J. 1987. *A concise history of mathematics*. Nueva York: Dover.