Desarrollo del Cálculo[1]

José Luis Abreu y Michael Barot

13 de septiembre de 2017

[1]Parte de una versión preliminar del libro *Desarrollo del pensamiento matemático*.

Índice general

1. Resumen histórico **5**
- 1.1. Introducción . . . 5
- 1.2. Los orígenes (3000 - 350 aec) . . . 6
- 1.3. Las matemáticas griegas, época de oro (350 - 100 aec) . . . 6
- 1.4. Alejandría, siglos I al IV . . . 7
- 1.5. Del siglo V al Renacimiento (siglo XVI) . . . 7
- 1.6. Siglo de oro, el XVII . . . 8
- 1.7. Expansión en el siglo XVIII . . . 9
- 1.8. Consolidación en el siglo XIX . . . 9
- 1.9. Siglos XX y XXI . . . 10
- 1.10. Presente y futuro de las Matemáticas . . . 10
- Bibliografía . . . 10

2. Los números **11**
- 2.1. Los sistemas de numeración . . . 11
- 2.2. Numerología y Teoría de números. . . . 19
- 2.3. La Teoría de Números . . . 20
- 2.4. El estudio de los números naturales en Euclides . . . 23
- 2.5. La irracionalidad de la raíz cuadrada de dos . . . 27

3. Geometría y Medición **31**
- 3.1. El uso de los números para cuantificar la realidad . . . 31
- 3.2. Longitud, área y volumen . . . 34
- 3.3. El método de exhaución . . . 44
- 3.4. La cuadratura de la parábola . . . 49
- 3.5. El número π a través de los siglos . . . 52
- 3.6. El volumen y la superficie de la esfera . . . 68
- 3.7. Cálculos de distancias celestes . . . 70
- 3.8. La trigonometría . . . 78

4. Cálculo, funciones y límites **95**
- 4.1. El Cálculo Diferencial e Integral . . . 95
- 4.2. El cálculo . . . 100
- 4.3. Las funciones en el cálculo . . . 104

4.4. Desarrollo del concepto de límite 112
4.5. Leyes del movimiento Planetario 121
4.6. Expansión del Cálculo 121
4.7. Ecuaciones diferenciales 121
4.8. Series de Potencias 121
4.9. Formalización del Cálculo 121
4.10. Métodos numéricos en el Cálculo 122

Capítulo 1

Resumen histórico

1.1. Introducción

La tradición nos dice que los primeros desarrollos matemáticos se hicieron en el antiguo Egipto, en Sumeria y Babilonia hace unos cuatro o cinco mil años. En realidad los primeros elementos de las matemáticas surgen mucho antes con el desarrollo del lenguaje. Por ejemplo las palabras que se usan para contar pequeños conjuntos de objetos, e incluso la manera de representar esos números, aunque fuese muy rudimentaria, por ejemplo haciendo marcas en un hueso, tantas como ovejas hay en un rebaño, surge mucho antes de que las culturas del oriente mediterráneo se establecieran. Surgen como parte integral del lenguaje humano. Keith Devlin (Ref. 1), un conocido investigador en el tema del pensamiento matemático piensa que las habilidades matemáticas son las mismas que las lingüísticas, ambas se ocupan de abstracciones, la única diferencia es que las de las matemáticas tienen un grado de abstracción mayor que las del lenguaje ordinario. Pero las relaciones entre los conceptos abstractos que manejan las matemáticas son, si no más simples, por lo menos más claras que las que ocurren entre las abstracciones más mundanas como podrían ser las personalidades de los miembros de una tribu o las plantas y sus productos o los animales y sus costumbres, asuntos con los que hay que lidiar en la vida diaria.

A lo largo de este libro nos adherimos a la idea de que las matemáticas constituyen un lenguaje especializado en el manejo de conceptos abstractos entre los que hay unas relaciones bien definidas. De hecho el nombre mismo de las Matemáticas proviene del griego MATHEMATA que significa aquello que se puede entender racionalmente. Y entender algo significa tener claras las relaciones que operan entre los elementos de un sistema y poseer la capacidad de deducir que otras relaciones también se cumplen a partir de las primeras mediante un proceso de deducción lógica. Esto es, a grandes rasgos, lo que se entiende por un sistema axiomático. Sin embargo no deseamos reducir las matemáticas a sistemas axiomáticos ya que éstos suponen una claridad ab-initio de las relaciones entre los conceptos abstractos de que se ocupan, mientras que en el desarrollo de las matemáticas la formación de conceptos y la clarificación de las relaciones entre ellos y la búsqueda de las más básicas a partir de las cuales las demás pueden deducirse, suele ser un proceso complejo que lleva tiempo y esfuerzo. Consideramos que estos procesos son parte integrante del quehacer matemático y que describir a las Matemáticas como un conjunto de sistemas axiomáticos deja de lado lo que

puede ser, quizás, su aspecto más importante.

Las matemáticas constituyen una de las actividades humanas más refinadas. Fueron desarrolladas en muchas culturas diferentes por gente que observaba el mundo en que vivía, tratando de reconocer orden y coherencia, con la esperanza de predecir y tal vez controlar eventos futuros.

1.2. Los orígenes (3000 - 350 aec)

Las matemáticas nacieron cuando el hombre primitivo tuvo necesidad de contar y evolucionaron como un lenguaje escrito hace unos cinco mil años cuando los creadores de calendarios calcularon el inicio de las estaciones e hicieron marcas en tabletas o papiros para registrar esos datos. Siglos después, el escolar renacentista Galileo Galilei escribió que el gran libro de la naturaleza está abierto ante nuestros ojos y la filosofía verdadera está escrita en él, pero no podemos leerlo sin antes aprender el lenguaje y los símbolos con los que está escrito; está escrito en lenguaje atemático y los símbolos son triángulos, círculos y otras figuras geométricas.

Pero los símbolos matemáticos más antiguos no fueron figuras geométricas sino marcas en tabletas de barro babilonias que se usaban como calendarios o jeroglíficos cavados en piedra en los templos mayas. Con el desarrollo del comercio, la necesidad de medir y contar acompañó a las caravanas de viajeros comerciantes.

La medición de la tierra, de donde la Geometría toma su nombre, tuvo que satisfacer las necesidades de la irrigación y el cálculo de impuestos en el antiguo Egipto y en Mesopotamia. Al construir las pirámides y tener que orientarlas debidamente exactamente al norte o al oriente, los arquitectos egipcios descubrieron relaciones entre triángulos, rectángulos y círculos. En la antigua Alejandría, la Geometría y la Trigonometría se usaron para medir el tamaño de la Tierra y las distancias al Sol y a la Luna.

Grandes desarrollos ocurrieron cuando los antiguos griegos como Tales de Mileto (620 - 546 aec) o Pitágoras (569 - 475 aec) y sus discípulos llamados "los pitagóricos" trataron a los números y las figuras geométricas como abstracciones puras. Estos juguetes mentales generaron las bases para toda la ciencia y tecnología con que contamos en la actualidad. ¿Cómo ocurrió esto?

1.3. Las matemáticas griegas, época de oro (350 - 100 aec)

El desarrollo de los conceptos de área y volumen como entes cuantificables por medio de los números, y del conceptos de razón, también expresable mediante números llevan al inicio de la época de oro de las Matemáticas griegas en las que se descubren el teorema de los triángulos semejantes (los triángulos semejantes tienen sus lados proporcionales) y el Teorema de Pitágoras. También se formaliza el número π como la razón del perímetro de un círculo a su diámetro y se obtienen buenas aproximaciones de su valor. Éstas y otras ideas importantes que siguieron desarrollándose a lo largo de estas líneas culminaron en el siglo XVII en la creación del Cálculo dando un impulso sin precedentes al desarrollo de la Ciencia y la Tecnología.

Sin duda el libro más importante en toda la historia de las Matemáticas y uno de los más importantes en la historia de la humanidad es "Los elementos" de Euclides de Alejandría (325 - 265 aec), escrito alrededor del año 300 aec. La aparición de esta obra marca un momento decisivo en la

historia de la humanidad. Es una admirable colección de trece libros que contienen casi todo el conocimiento matemático de su tiempo, incluyendo los trabajos de Eudoxo de Cnido (390 - 337 aec), uno de los tres matemáticos más importantes de la antigüedad griega, junto con Arquímedes de Siracusa (287 - 212 aec) y Apolonio de Perga (262 - 190 aec). Bien presentado y organizado en orden lógico, es un logro cultural extraordinario que ha influido en el pensamiento científico de la humanidad durante más de dos mil años. Asentó las bases del método deductivo, mostrando que el intelecto humano puede descubrir verdades matemáticas utilizando únicamente la razón y la lógica. "Los Elementos" se escribió originalmente en griego sobre rollos de papiro, pero fue traducido a muchos idiomas y publicado en incontables ediciones.

Entre los logros más importantes que se atribuyen a Eudoxo están la teoría de las proporciones (Libro V de Los elementos), la fórmula para el volumen de una pirámide y el método de exhaución para calcular o estimar áreas y volúmenes de regiones acotadas por curvas y superficies que no son planos.

Entre los principales logros de Arquímedes se consideran su estimación de π, la cuadratura de la parábola (utilizando en ambos casos el método de exhaución) y los cálculos del volumen y la superficie de la esfera. El método de exhaución también lo utilizó en China, de manera independiente, Liu Hui en el siglo III aec para encontrar el área del círculo, logrando obtener un valor de π más preciso que el de Arquímedes.

A Apolonio se le conoce principalmente por su exhaustivo y elegante trabajo sobre las secciones cónicas, que solo podrá será superado casi dos milenios más tarde en el siglo XVII con las geometría analítica de Fermat y Descartes.

1.4. Alejandría, siglos I al IV

Durante el apogeo del Imperio romano que controló la cuenca del mar Mediterráneo desde unos años antes de Cristo hasta el siglo IV, el desarrollo de las matemáticas decayó. Esta época está dominada por personajes que más que desarrollar conceptos, ideas y métodos, lo que hacían era comentar o recopilar los conocimientos de sus predecesores. Sin embargo la ciudad griega de Alejandría, fundada por Alejandro Magno (356 - 323 aec) en la desembocadura del río Nilo, en Egipto, se convirtió en la capital y refugio cultural de la época y especialmente durante los siglos II al V hubo importantes desarrollo en matemáticas. Entre los personajes más notables de esta ciudad y en esta etapa (no olvidemos que Euclides vivió en Alejandría, pero mucho antes, alrededor del 300 aec) que hicieron contribuciones a las matemáticas destacan Menelao de Alejandría Claudio Ptolomeo (100 - 170), Diofanto de Alejandría (210 - 290 aprox), Pappus (siglos III y IV), Teón de Alejandría (335 - 405) y su hija Hipatia (370 - 415). Entre quienes hicieron principalmente comentarios a las obras de sus predecesores, sobre todo a las de Euclides y Arquímedes, destaca sobre todos Proclo (412 - 485).

1.5. Del siglo V al Renacimiento (siglo XVI)

De manera independiente, en Asia los matemáticos indios y chinos crearon sus propios símbolos matemáticos, incluidos los números indo-arábigos que usamos en la actualidad. De las manos de matemáticos como Al-Juarismi (780 - 850) y Omar Jayam (1048 - 1131), surgieron nuevos

métodos matemáticos que hoy conocemos como álgebra, de acuerdo con su nombre árabe que significa reintegración o recomposición.

En China, el siglo V, Zu Chongzhi usó lo que más tarde sería llamado la teoría de los indivisibles por el matemático italiano Bonaventura Cavalieri para encontrar el volumen de una esfera.

Cerca del año 1000, el matemático islámico Alhacén fue el primero en derivar la fórmula para la suma de la cuarta potencia de una progresión aritmética, usando un método a partir del cual es fácil encontrar la fórmula para la suma de cualquier potencia integral de mayor orden. Esto sentó las bases para obtener más adelante el área bajo las gráficas de las funciones polinomiales.

En el siglo XI, el polímata chino Shen Kuo desarrolló ecuaciones que se encargaban de integrar. En el siglo XII, el matemático indio, Bhaskara II, desarrolló una derivada temprana representando el cambio infinitesimal, y describió una forma temprana del llamado teorema de Rolle.

También en el siglo XII, el matemático persa Sharaf al-Din al-Tusi descubrió la derivada de la función cúbica, un importante acontecimiento que abre el camino hacia el cálculo diferencial.

En el siglo XIV, Madhava de Sangamagrama, en conjunto con otros matemáticos y astrónomos de la escuela Kerala de matemáticas y astronomía, al sur de la India, describieron casos especiales de lo que hoy se conocen como las series de Taylor.

Las rutas comerciales llevaron a Europa la notación indo-arábiga y los algoritmos para realizar operaciones con ella, desplazando rápidamente el incómodo sistema de números romanos que se había usado durante casi toda la edad media. Es importante reconocer la contribución de Leonardo de Pisa (Fibonacci) (1170 - 1240) en este proceso.

Esto generó un fuerte interés en el álgebra de los árabes que culminó con la creación del álgebra simbólica en la que se utilizan letras para denotar cantidades desconocidas o variables genéricas. La creación de la notación algebraica fue impulsada principalmente por François Viète (1540 - 1603) y dio un nuevo impulso al desarrollo de las matemáticas que por fin encontraron un leguaje simbólico que además de simplificar la notación de las ecuaciones y fórmulas, ayudaba a manipularlas con una eficiencia nunca antes experimentada.

1.6. Siglo de oro, el XVII

Pronto las ideas renacentistas se unieron a la antigua Geometría griega con lo que Pierre de Fermat (1601 - 1665) y René Descartes (1596 - 1650) crearon la Geometría Analítica.

Ya con la notación algebraica y la Geometría analítica en el saco de herramientas, los matemáticos del siglo XVII, es decir, de la generación de Isaac Newton (1643 - 1727) y Gottfried Leibniz (1646 - 1716), estuvieron en posición de domar los procesos continuos como el movimiento y el cálculo de áreas y volúmenes de figuras arbitrarias. Nace así el Cálculo diferencial e integral que será aprovechado por las ciencias, en especial por la Física, para modelar sus leyes y procesos y predecir el comportamiento de muchos sistemas, usando casi siempre los conceptos de función, límite, derivada e integral.

1.7. Expansión en el siglo XVIII

Durante el siglo XVIII la actividad matemática se ocupó principalmente de extender sus conceptos y sus aplicaciones, especialmente las del Cálculo. En este último aspecto destacan los nombres de los matemáticos suizos Jacob Bernoulli (1654 - 1705), Johann Bernoulli (1667 - 1748) y Leonhard Euler (1707 - 1783), los primeros dos hermanos, el tercero discípulo del segundo y todos ellos procedentes de Basilea, que dedicaron sus vidas a profundizar en el estudio del cálculo, las funciones, los límites, las sucesiones infinitas y las ecuaciones diferenciales y a aplicar estas ideas a la modelación de muchos problemas científicos.

A partir del siglo XVIII las Matemáticas no han dejado de avanzar, creando nuevos conceptos, nuevas disciplinas, encontrando nuevos retos y nuevas aplicaciones. Casi todas las actividades humanas del mundo moderno están rodeadas de procesos matemáticos. Desde la Física teórica que confía a las matemáticas todos sus modelos del mundo material, pasando por la Biología que las necesita para modelar muchos aspectos del comportamiento de las poblaciones de animales y plantas, hasta los sistemas informáticos que las usan en sus sofisticados algoritmos.

1.8. Consolidación en el siglo XIX

Desde principios del siglo XIX comenzaron a aparecer signos de una preocupación por los fundamentos, por entender las raíces más profundas de las matemáticas. Una de estas manifestaciones fue la preocupación por determinar si todas las ecuaciones polinomiales tenían o no soluciones, se vio que sí las tenían, siempre y cuando se aceptara que estuvieran definidas en los complejos. Esto dio mayor impulso a los números complejos, que también se analizaron detalladamente, se formalizaron e incluso se creo una fructífera teoría de funciones de variable compleja que complementa el álgebra de polinomios complejos y extiende incluso el Cálculo. Sir William Rowan Hamilton (1805 - 1865) dedicó una buena parte de su vida intelectual a resolver el problema de la existencia o no de números más generales que los complejos, llegando a la conclusión de que si se quería extender el concepto de número más allá de los complejos, era necesario prescindir de la conmutabilidad en la multiplicación y así descubrió los cuaternios.

Otra manifestación del interés en los fundamentos es la preocupación por el quinto postulado de Euclides en la Geometría que llevó al descubrimiento de su independencia y de la existencia de geometría no euclidianas, es decir, que no satisfacen el quinto postulado. Esta preocupación llevó hacia finales del siglo a una formulación axiomática rigurosa, por parte de David Hilbert (1862 - 1943), de la Geometría euclidiana que sentó las bases para el concepto moderno de sistema axiomático. Quizás la mayor preocupación de los matemáticos del siglo XIX fueron los fundamentos del Cálculo. A pesar de sus innegables éxitos en las aplicaciones, el Cálculo sufría de un grave problema, usaba infinitesimales, entes mal definidos que promovían discusiones interminables acerca de su existencia misma. Los matemáticos sintieron necesidad de aclarar estos conceptos y la solución vino de la mano de matemáticos como Augustin Louis Cauchy (1789 - 1857), Karl Weierstrass (1815 - 1897) y Richard Dedekind (1831 - 1916) que acabaron desterrando el concepto y uso de los infinitesimales, sustituyendo en cambio el de límite. Así se desarrolló una extensa y detallada teoría sobre las funciones, sus límites y sus derivadas que dejaron perfectamente aclarados los conceptos y los fundamentos del Cálculo, aunque a costa de perder la simplicidad original con la que en el siglo XVIII se enfrentaban los problemas usando los infinitesimales de manera intuitiva.

1.9. Siglos XX y XXI

Las Matemáticas siguieron avanzando en el siglo XX de una manera impresionante. Lo hicieron creando nuevos conceptos, nuevas disciplinas, encontrando nuevos retos y nuevas aplicaciones y analizando los rincones más recónditos de sus fundamentos.

Hoy en día casi todas las actividades humanas del mundo moderno están rodeadas de procesos matemáticos. Desde la Física teórica que confía en las matemáticas todos sus modelos del mundo material, pasando por la Biología que necesita a las matemáticas para modelar muchos aspectos del comportamiento de las poblaciones de animales y plantas, hasta los cálculos que hace un GPS para determinar su posición exacta sobre la superficie de la Tierra o los cálculos que hacen las computadoras para dibujar a los personajes y objetos en las películas hechas por computadora, sin olvidar los sofisticados algoritmos que Google para encontrar los sitios de interés que pueden ser de mayor interés para quien hace una búsqueda o los que utilizan las redes sociales para ofrecer un funcionamiento útil y eficiente a sus usuarios, utilizan matemáticas.

Sin embargo es preocupante comprobar que al mismo tiempo los estudiantes de todo el mundo, en su mayoría temen a las matemáticas y tratan de huir de ellas eligiendo carreras en función del menor uso que hacen de las matemáticas. En opinión de los autores esto se debe a que en general la enseñanza de las matemáticas no aprovecha sus logros históricos para motivar y guiar el aprendizaje de la materia.

1.10. Presente y futuro de las Matemáticas

Durante toda la historia, las matemáticas han sido un espejo de la civilización. Desde aquellas marcas primitivas que se usaron para anotar el inicio de las estaciones, las matemáticas fueron evolucionando convirtiéndose en un lenguaje cada vez más útil y preciso. Quienes hicieron descubrimientos importante en relación con el movimiento de los cuerpos celestes, como Kepler y Newton, la convirtieron en un lenguaje aún más poderoso que se convirtió en la lingua franca de las Ciencias. Este hecho unido a la creciente utilización de las matemáticas en muchos aspectos de la vida moderna, debe ser motivo suficiente para que los jóvenes se interesen en ellas. Uno de los objetivos de este libro es proporcionar material que ayude a motivar el interés en las matemáticas echando mano de los detalles más apasionantes de su desarrollo histórico y mostrar que coinciden precisamente con lo que la sociedad espera que los jóvenes aprendan de matemáticas, a pesar de que los planes y programas de estudio muchas veces no lo reflejan.

Bibliografía

Sir Thomas Heath: A history of greek mathematics, Vol II, Dover 2014.

Keith Devlin: The Mathematics Gene, Editor YEAR.

Wikipedia en español, inglés y alemán.

Capítulo 2

Los números

Una extensión del lenguaje natural

2.1. Los sistemas de numeración

Cuando una cultura se desarrolla forma una organización más compleja con divisones de trabajo. Se intercambian bienes, se registran cantidades, un sentido para la historia des desarrolla y se anotan fechas. Por ello no debe sorprender que hay diferentes sistemas de anotar números: esta idea se tuvo que inventar varias veces. En lo que sigue veremos los sistemas de enumeración de algunas culturas. Nos interesa en particular, si el sistema facilitó el cálculo o más bien lo limitó.

En casi todos los sistemas de numeración se observan repeticiones, la excepción de los que presentamos aquí es él de los griegos. El uso repetitivo de un símbolo disminuye el número de símbolos usados y así facilita el empleo ya que no hay que memorizar tantos de ellos. Al final de esta sección haremos una comparación y un análisis sobre el uso para hacer cálculos.

La Cuneiform

Lo usaron diferentes culturas en asia occidental, entre ellos los sumerios y los babilonios. Con una herramienta sencilla, como cuña, se grababa en arcilla húmeda. Luego se horneaba la tablilla. La cuneiforma se desarrolló a partir de pictogramas. La escritura se desarrolló alrededor de 3300 a.C. como pictogramas y se transfromó en una forma cada vez más abstracta. Hay tablillas con la forma que cuneiforma que presentaremos abajo de alrededor de 2200 a.C.

Para denotar números, se usaban dos símbolos sencillos. Para el Uno: un triángulo con punta hacia abajo, que a veces se alargaba con una raya y para el Diez un ángulo abierto hacia la derecha.

1 1 10

Números hasta el 59 se obtenían por repeticón, por ejemplo

2 6 13 24 37 59

Números más grandes que 59 se escribían en forma posicional en base 60. Por ejemplo, el número $81 = 60 + 21$ se escribe como se muestra en la siguiente figura.

$1 \cdot 60 \quad 21$

El primer 1 está en la segunda posición e indica que hay una vez 60 y luego se suman los 21 del segundo grupo de símbolos. Más representaciones en cuneiforme de números se ven en la siguiente figura.

$1 \cdot 60 \quad 40$

$1 \cdot 60 + 40 = 100$

$33 \cdot 60 \quad 26$

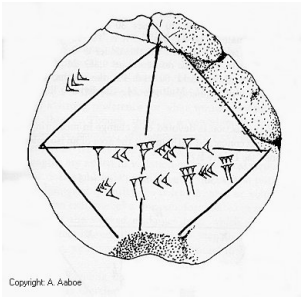
$33 \cdot 60 + 26 = 2016$

$2 \cdot 60^2 \quad 46 \cdot 60 \quad 40$

$2 \cdot 60^2 + 46 \cdot 60 + 40 = 10\,000$

Los babilonios también usaron sesentavos y el número en la figura arriba a la izquierda también podía significar $1 + 40 \cdot 60^{-1} = \frac{5}{3}$. El punto decimal se tenía que concluir del contexto y no se indicaba en la notación.

La siguiente tablilla, que se llama YBC7289 de alrededor de 1800 a.C., muestra un cuadrado con la diagonal.



La anotación sobre la diagonal es interesante. Si se interpreta con un punto decimal después del primer 1:

$1 \quad 24 \cdot 60^{-1} \quad 51 \cdot 60^{-2} \quad 10 \cdot 60^{-3}$

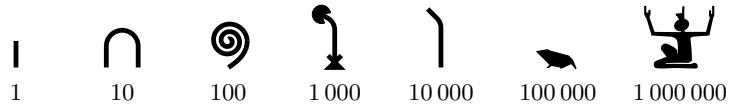
$$1 + 24 \cdot 60^{-1} + 51 \cdot 60^{-2} + 10 \cdot 60^{-3} = 1.41421296\ldots \approx \sqrt{2}$$

se obtiene una aproximación extremadamente buena de raíz de 2.

De los babilonios heredamos la división del tiempo: la hora en 60 minutos, el minuto en 60 segundos. Además también la medición de ángulos en 360 grados.

Los jeroglíficos egipcios

Los egipcios usaron un sistema que se basaba en pocos símbolos y posiblemente muchas repeticiones. Se basaben en la base 10: por cada potencia de diez tenían un símbolo distinto, tal como se muestra en la siguiente figura.

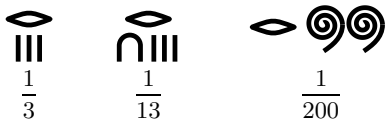


El símbolo del 1 es totalmente abstracto, sobre el significado del 10 hay diversos opiniones, pero el 1000 es representado por una planta de agua como un lirio. Vemos en estos símbolos la transformación que transcurrieron con los jeroglíficos: al principio eran imágenes y por el frecuente uso se abstrajeron cada vez más. El uno, por ser el más usado es una simple raya, mientras el símbolo del millón sigue siendo una imagen.

Para representar números se repetían los símbolos las veces que eran necesarias.

La siguiente ilustración muestra una parte de los anales de Thoutmosis III del año 1437 a. C. que indica que en el momento el rey contaba con 4622 ganados.

El sistema no era posicional. Por ello necesitaron una idea diferente para representar fracciones. Estas las expresaron como suma de fracciones unitarias, es decir con numerador 1. Para indicar que se trataba de una fracción y no de un número escribían el símblo de un ojo encima o a un lado del número. La siguiente figura muestra algunas fracciones unitarias en la denotación egipcia.



$\frac{1}{3}$ $\frac{1}{13}$ $\frac{1}{200}$

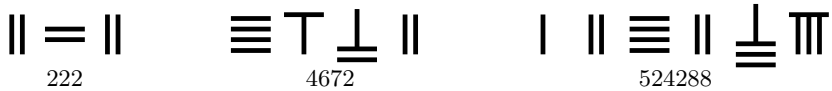
Los egipcios tenían símbolos especiales para las fracciones $\frac{2}{3}$, $\frac{1}{2}$ y en general para fracciones unitarias cuyo denominador es una potencias de 2.

Las varillas de los chinos

En China se desarrolló un sistema posicional en base diez. Se encontró en tumbas del período de los Reinos Combatientes (alrededor desde 480 a.C hasta 220 a.C.). El sistema usa un sólo símbolo: una varilla. Para números hasta 9 una varilla vertical significa 1 y se repetía hasta por máximo 4 veces. Una varilla horizontal significa 5 y se escribía encima de las varillas verticales.

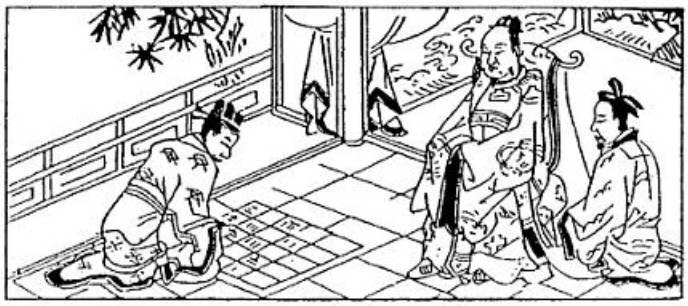
1 2 4 5 8

El sistema era posicional, pero el significado de horizontal y vertical cambiaba de posición en posición. En la posición de las decenas una varilla horizontal significaba 1 decena mientras una vertical representaba 5 decenas. La siguiente figura muestra algunos números.

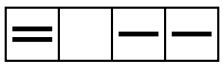


222 4672 524288

No tenían un símbolo para representar el cero, pero como se muestra en la siguiente ilustración, se usaba una rejilla para indicar las posiciones.



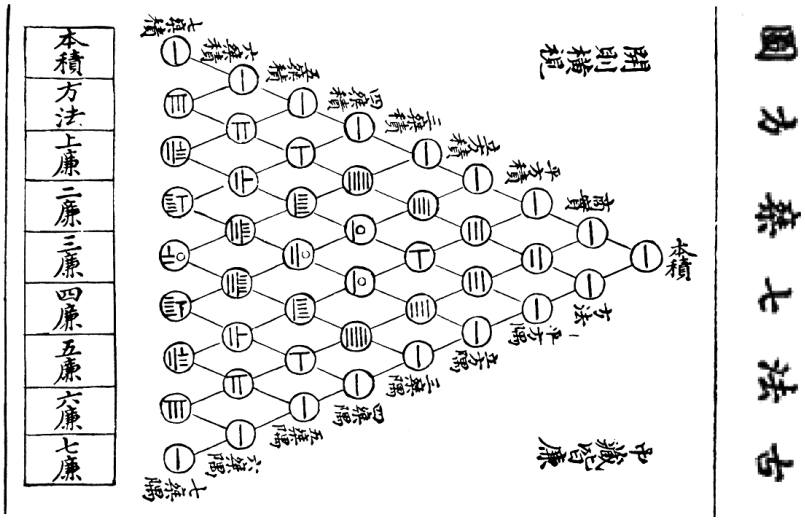
Así, el número 2015 se escribía de la siguiente manera.



2015

La adición con varillas es muy sencillo ya que estas se pueden juntar físicamente. En este sentido, el sistema de las varillas era más práctico que nuestro sistema decimal actual. En la dinastía Song (del año 960 al 1279) se usaban las varillas con eficacia: se tenían algoritmos para llevar a cabo las cuatro operaciones básicas, se conocían los números negativos y las fracciones. La siguiente ilustración, del

año 1303, muestra lo que nosotros conocemos por "triángulo de Pascal" (era conocido siglos antes en China). Los números se representaron con varillas, pero ya se puede ver un pequeño circulito para indicar el cero y que ya no siempre se cambiaba de dirección de las variallas de posición en posición.



Este sistema numérico de varillas se abandonó por el uso del ábaco.

Letras de los griegos

Los griegos es la cultura, que descubrió la argumentación, el sistema axiomático, la observación en medicina, la filosofía, la democracia, la poesía y el teatro tenían el peor sistema de numeración. Adoptaron sus letras de los fenicios, igual que los judíos y con ello también el sistema de numeración. Usaron muchas de sus letras como símbolos para expresar números. Para distinguirlos de las letras las adornaron con un acento, es decir α denota la letra y α' denota el número 1. La siguiente tabla muestra los símbolos que usaron, tres de ellos ya no se usan hoy: digrama, copa y sampi.

α	β	γ	δ	ε	ϝ	ζ	η	ϑ
(alfa)	(beta)	(gama)	(delta)	(épsilon)	(digama)	(dseta)	(eta)	(theta)
1	2	3	4	5	6	7	8	9
ι	κ	λ	μ	ν	ξ	ο	π	ϟ
(iota)	(kappa)	(lambda)	(mi)	(ni)	(xi)	(ómicron)	(pi)	(koppa)
10	20	30	40	50	60	70	80	90
ρ	ς	τ	υ	φ	χ	ψ	ω	ϡ
(rho)	(sigma)	(tau)	(ípsilon)	(fi)	(ji)	(psi)	(omega)	(sampi)
100	200	300	400	500	600	700	800	900

Con ello podían representar números hasta 999. Por ejemplo

ια′	φλβ′	ωπη′
11	532	888

Para números más grandes usaron un símbolo de acento en la linea base. Por ejemplo 2015 se representaba por ,βιε′. No obstante que su sistema de nuemarción era poco útil para hacer operaciones matemáticas, investigaron sobre los números y se sirvieron de la geometría para hacerlo: así descubrieron que cada número podía factorizarse en primos, que había un número infinito de primos y que existían números que no eran fraccionarias que hoy se llaman irracionales.

Más tarde, alrededor de 140 a.C. se sirvieron de un sistema posicional en base de 60, posiblemente adoptado de los babilonios, para denotar fracciones menores que 1. Para ello usaron un símbolo especial para el 0. Esto lo usaron para hacer cálculos astrónomicos.

Letras de los romanos

Los romanos usaron algunas letras para denotar números. La siguiente figura muestra los cinco símbolos que se usaron an la antigüedad.

I	V	X	L	C	⌈ D	M ⌉
1	5	10	50	100	⌊ 500	1000 ⌋

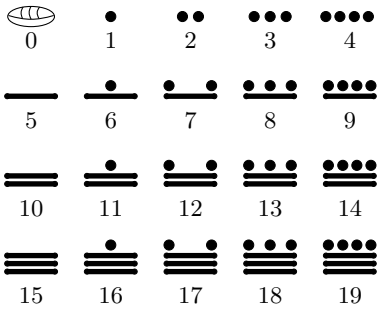
Las letras I, X y C se podían repetir hasta 4 veces. Sólo más tarde se añadieron las letras D para denotar 500 y M para denotar 1 000. También se introdujo la regla que un símbolo de menor valor a la izquierda de uno de mayor valor significa que el primero se tenía que restar del segundo: IX denota el número 9 y CM el número 900. Aquí hay algunos ejemplos.

XXIII	LIX	XCVI	DCCCLXXXVIII	MCMXLV	MMXV
23	59	96	888	1945	2015

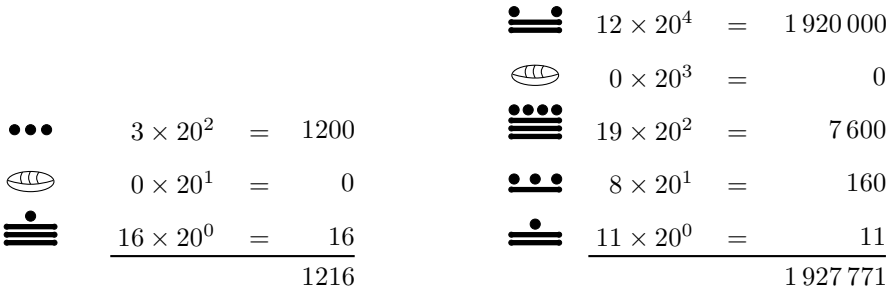
El sistema romano usa pocos símbolos y es parecido al de los egipcios, solo que no se podían denotar números muy grandes. Esto no se hizo hasta en la edad medieval. Los números romanos se encuentran con frecuencia para denotar años y los números en los relojes.

Los jeroglíficos de los mayas

Los mayas usaron sólo dos símbolos para denotar números: un punto y una raya que representaban el número 1 y 5 respectivamente. Estos símbolos lo repitieron para representar números hasta 19. La siguiente figura muestra los primeros números en la escritura maya.



Los mayas usaron un sistema posicional en base de veinte. La siguiente figura mustra algunos números en este sistema numérico.



Los mayas usaron este sistema de numeración alrededor de 300 a. C. hasta la llegada de los conquistadores españoles para anotar fechas. Como anotaron eventos astronómicos, como eclipses solares, fue posible descifrar que su conteo de días iniciaba el 11 de agosto de 3114 a. C. en el calendario Gregoriano proléptico. Contaban día por día en un sistema casi-vigésimal, dado que había una excepción al uso de la base 20: un día se llama *k'in*, 20 *k'in* hacen un *winal*, pero 18 *winal* hacen un *tun*, luego 20 *tun* hacen un *k'atun*, 20 *k'atun* un *b'ak'tun* y hay dos niveles más. La excepción tiene el efecto que un *tun* consiste de 360 días, lo que es cercano a un año.

En un sistema posicional se requiere un cero, que indica que en esta posición no hay ningún número. Los mayas usaron diferentes símbolos. La primera piedra que usa un símbolo data alrededor de 300 d. C. y por ello es el primer testimonio del cero, tres siglos antes de los árabes.

Los símbolos de los hindues y árabes

Durante los siglos I al IV d. C. los hindues desarrollaron el sistema decimal que hoy usamos, sólo que los símbolos que usaron fueran otros. Este sistema lo adoptaron los árabes y contaron con un símbolo para el cero. Los símbolos se transformaron con el uso y por la adopción en otras las culturas y no eran estandardizados.

La numeración moderna

En el siglo X europeos empiezan a estudiar textos árabes y adoptan de la numeración los símbolos del 1 al 9. Más tarde se adopta también el cero y los símbolos toman su forma moderna en el siglo XV, cuando los libros se empiezan a imprimir con letras móviles. El punto decimal se usa por primera vez en 1492 en la obra *Compendio del Abaco* del italiano Francesco Pellos, pero tomó alrededor de 100 años más hasta que se adoptara ampliamente. Hoy día se usa en algunas países también la coma para separar la parte entera del fraccionario.

Comparación

Para denotar números se inventaron diferentes sistemas, muchas veces con rasgos similares. Con excepción de la numeración de los griegos, todos los sistemas usan algún tipo de repetición. Esto permite reducir la cantidad de símbolos. Podemos ver que en diferentes culturas se usaron ideas muy similares. Distinguimos dos clases de sistemas de numeración:

- Los símbolos tienen un valor fijo: egipcios, griegos y romanos.
- Los símbolos cambian su valor según la posición donde se encuentran: esto es el principio del sistema posicional. Lo usaron: los sumerios y babilonios, los chinos, los mayas y finalmente los hindues y árabes.

El sistema posicional tiene claras ventajas: es posible representar números arbitrariamente grandes y las operaciones básicas como la adición o la multiplicación se pueden hacer siguiendo un algoritmo fijo. Los sistemas cuyos símbolos un valorfijo no tienen estas propiedades, pero a cambio son bastante fáciles de leer.

Es interesante ver que los sumerios y babilonios, los chinos y los mayas usaron muy pocos símbolos básicos que repetían para representar "una cifra", es decir, el número en una sola posición.

También es interesante comparar los usos que los testimonios históricos dejan entrever. Los babilonios y egipcios y también los romanos usaron los números sobre todo para fines mercantiles: para contar cantidades de ganado o de bienes en las bóvedas. Los babilonios también los usaron para fines astronómicos: el sistema sexagésimal de los babilonios es el origen de la división del ángulo completa en 360 grados.

Los mayas usaron los números para un calendario preciso. Los chinos, hindues y árabes usaron los números de manera más abstracta: la multiplicación y la división de números es una idea que no se basa en una aplicación directa de la vida cotidiana en aquellos tiempos.

No tiene sentido multiplicar números como 4622 por otro número cualquiera como 7765, si uno denota el número de vacas y otro el número de árboles. Es decir, las operaciones aritméticas en general sólo tienen sentido, si se entiende el número como concepto abstracto, sólo así se pueden motivar las operaciones de multiplicación y división entre todos los números.

Ejercicio. Para comprender mejor las ventajas y desventajas de los distintos sistemas de enumeración sugerimos al lector representar los números 832 y 57 y llevar a cabo la adición y la multiplicación en varios sistemas, sin recurrir al sistema decimal.

Bibliografía

Georges Ifrah: *Historia universal de las cifras*, Editorial Espasa Calpe, 2009.

Johannes Lehmann: *So rechneten Ägypter und Babylonier*, Urania Verlag 1994.

Johannes Lehmann: *So rechneten Griechen und Römer*, Urania Verlag 1994.

Hans Lenz: *Universalgeschichte der Zeit.*, Marix Verlag, 2013.

David Freidel, Linda Schele, Joy Parker: *Maya Cosmos: three thousand years on the shaman's path*, 1993.

Wikipedia en español, inglés y alemán.

2.2. Numerología y Teoría de números.

La teoría de números nace alrededor de 600 A.C. Cuando Pitágoras y sus discípulos comenzaron a estudiar las propiedades de los enteros positivos. Sabían que los tonos musicales estaban relacionadas con razones de números enteros. Por ejemplo si un tono tiene 440 vibraciones por segundo, el que tiene 880 vibraciones por segundo es el mismo tono, solo que una octava más alto. Claro que los pitagóricos no tenían manera de contar el número de vibraciones por segundo de una nota musical, pero sí observaron que si una cuerda producía una nota, al hacer vibrar la misma cuerda a la mitad de su longitud, se producía la misma nota, una octava más alta. Y si se presionaba la cuerda para reducir su tamaño a 3/4 o 2/3, los sonidos que se obtenían eran armónicos, mientras que si se presionaba a distancias intermedias que no correspondían a fracciones simples de la longitud original, podían obtenerse tonos disonantes. Los pitagóricos creían que todo en el universo tenía secretos relacionados con razones de números enteros y que si pudieran descubrir estos secretos entonces todo sería claro y podrían escuchar una gran armonía: la música de las estrellas.

Para buscar esos secretos, los pitagóricos clasificaron a los enteros positivos en varios tipos de números:

los pares 2, 4, 6, 8 10, 12, 14, 16, ...

los impares 1, 3, 5, 7, 9, 11, 13, 15, ...

los primos 2, 3, 5, 7, 11, 13, 17, 19, 23, ...

y los compuestos 4, 6, 8, 9, 10, 12, 14, 15, 16, 18, 20, ...

Los griegos sabían que todo número compuesto es producto de primos. Euclides demostró que hay una infinidad de números primos. Hubieran estado encantados si hubieran sabido que los números primos a la larha iban a utilizarse en criptografía y en la codificación y decodificación de mensajes que se enviarían al espacio exterior.

También hicieron otras clasificaciones:

Números perfectos como el 6=1+2+3 y 28= 1+2+4+7+14, que son iguales a la suma de todos sus factores enteros. ¿Existen números perfectos que sean impares? Acomodando objetos en patrones poligonales, generaron números que llamaron

triangulares, 1, 3, 6, 10, 21, ... cuadrados, 1, 4, 9, 16, 25, 36, ...

pentagonales, 1, 5, 12, 22, 35, ... hexagonales, 1, 6, 15, 28, 45, ...

etcétera.

Siglos más tarde Pierre Fermat descubrió uno de los profundos secretos de la teoría de números: que todo número entero positivo puede ser escrito como la suma de a lo más tres números triangulares, y de a lo más cuatro números cuadrados y de a lo más cinco números pentagonales, etcétera. Por ejemplo:

$30 = 10 + 10 + 10$ suma de tres números triangulares

$30 = 1 + 4 + 9 + 16$ suma de cuatro números cuadrados

$30 = 1 + 1 + 1 + 5 + 22$ suma de cinco números pentagonales

$30 = 1 + 1 + 1 + 6 + 6 + 15$ suma de seis números hexagonales

Las matemáticas y el mundo, una relación mágica o científica. El interés de los pitagóricos en los números y sus propiedades representa quizás la manifestación más antigua de las matemáticas puras. Se trata de entender por deducción las propiedades de unos entes abstractos (los números) que han adquirido una existencia propia, independiente de sus humildes orígenes como utensilios para ayudar en el quehacer cotidiano. Un hecho que es interesante notar, es que esta primera aparición del interés matemático puro surge unido al misticismo, se atribuyen propiedades mágicas a ciertos números. Cabe preguntar qué tanto ayudó esta tendencia la misticismo a desarrollar las matemáticas. Aparte de esta etapa de los pitagóricos, caen en esta misma línea de misticismo la astrología de Ptolomeo y los esfuerzos de Kepler para explicar el sistema solar con los poliedros regulares. Son intentos de entender el mundo usando las matemáticas pero de una manera poco científica, como esperando descubrir algunas decisiones divinas a las que podemos tener acceso solo porque coinciden con relaciones entre números o entre cuerpos geométricos. Es como suponer que Dios es matemático y se comunica con nosotros a través de los objetos matemáticos pero sin que haya relaciones de causa efecto en los caprichos que llevan a organizar "matemáticamente" el mundo. Esta tendencia mística o mágica hacia entender el mundo a través de las matemáticas contrasta notablemente con la postura más científica de Galileo en donde ya las leyes de la naturaleza están ante nosotros como un libro abierto, no hace falta adivinar, basta leerlo, pero antes hay qe conocer las matemáticas y su lenguaje. La gran diferencia es que en esta postura galileana, no hay magia, no hay capricho ni casualidad, es una relación fáctica. El mundo tiene un comportamiento matemático. La idea de Dios deja de ser necesaria, aunque tampoco es necesariamente incompatible.

2.3. La Teoría de Números

Números se ocupaban en todas las culturas avanzadas, ver la sección 2.1 sobre los diferentes sistemas de numeración. Pero fue hasta en los griegos que se empezaron a estudiar los números por si mismo, es decir en propiedades que tienen los números respecto a las operaciones que son la adición y la multiplicación.

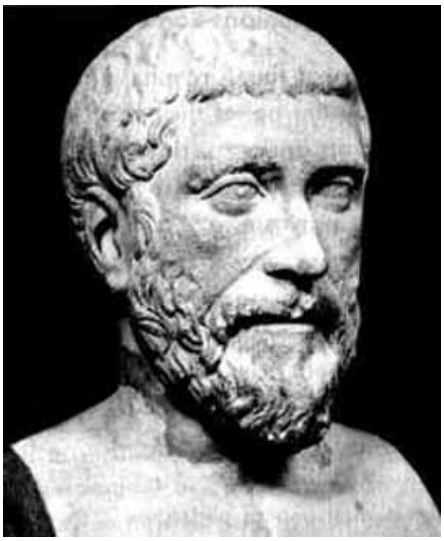
Un número era según Tales una colección de unidades. Por ello no se consideraba al 1 un número. Y Aristóteles dijo que el 1 no debería considerarse número ya que es la unidad y sirve para medir. Y lo que se mide y con lo que se mide no puede ser lo mismo. También había el pensamiento de que el 2, la *díada*, es la característica de ser par, por lo que el 2 a veces tampoco se consideraba un número, sino una característica de números.

De ello se ve como los griegos empezaron a estudiar las propiedades y aunque difieren en las

definiciones, encontraron muchas resultados de la teoría de números.

Los pitagóricos

La teoría de números nace alrededor de 550 a.e.c. cuando Pitágoras de Samos (ca. 569–475 a.e.c) y sus discípulos comenzaron a estudiar las propiedades de los enteros positivos.



Pitágoras de Samos

Sobre la vida de Pitágoras se sabe poco. Nació en la isla de Samos cerca de la costa de la actual Turquía. Su padre era aristócrata o negociante. Se mudó a Croton, en Sicilia donde fundó un tipo de escuela. Sobre la característica de la escuela hay opionen divergentes entre los historiadores. Algunos piensan que era una escuela abierta parecida a la escuela de Platón, otros opinan que era muy cerrada como una secta en las cuales los discípulos tenían que vivir según reglas estrictas. El problema es que las fuentes que hoy se tienen de ella, son romanos y datan alrededor de 500 años posteriores. Hay indicios que aceptaban hombres igual que mujeres entre sus discípulos y que eran vegetarianos.

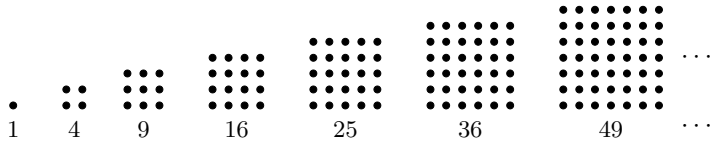
De lo que se cuenta de la escuela de Pitágoras es que se interesaban en las matemáticas, en particular en los números y la geometría y con la relación de estos conceptos con el cosmos. Reducido al máximo, el lema de los pitagóricos era: *todo es números*. Hay indicios que su pensamiento todavía estaba dominado mucho por la mística, es decir en creencias de interrelaciones y conexiones que sólo se asumen. De lo que quedó mucha evidencia es que fue una escuela sumamente influyente, poco tiempo después de fundar la escuela, la hermandad de los pitagóricos dominaban la ciudad y pronto una parte importante de la extensión de los griegos. Hacia el final de su vida, expulsaron a Pitágoras de Croton. Él y sus discípulos se fueron al exilio. Influyeron el desarrollo de las ciencias de manera profunda.

Podemos ver la manera cómo relacionaban la música con los números. Sabían que los tonos musicales estaban relacionadas con razones de números enteros. Por ejemplo si un tono tiene 440 vibraciones por segundo, el que tiene 880 vibraciones por segundo es el mismo tono, solo que una octava más

alto. Claro que los pitagóricos no tenían manera de contar el número de vibraciones por segundo de una nota musical, pero sí observaron que si una cuerda producía una nota, al hacer vibrar la misma cuerda a la mitad de su longitud, se producía la misma nota, una octava más alta. Y si se presionaba la cuerda para reducir su tamaño a $\frac{3}{4}$ o $\frac{2}{3}$, los sonidos que se obtenían eran armónicos, mientras que si se presionaba a distancias intermedias que no correspondían a fracciones simples de la longitud original, podían obtenerse tonos disonantes. Los pitagóricos creían que todo en el universo tenía secretos relacionados con razones de números enteros y que si pudieran descubrir estos secretos entonces todo sería claro y podrían escuchar una gran armonía: la música de las estrellas.

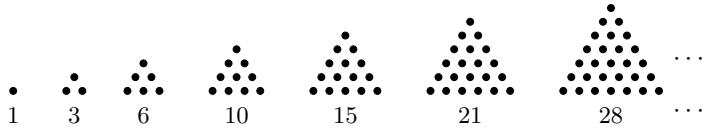
Para buscar esos secretos, los pitagóricos clasificaron a los enteros positivos en varios tipos de números: distingían entre *pares*, esto son los números que se pueden separar en dos partes iguales, e *impares*, en los que esta separación no es posible.

A parte de los números *cuadrados*

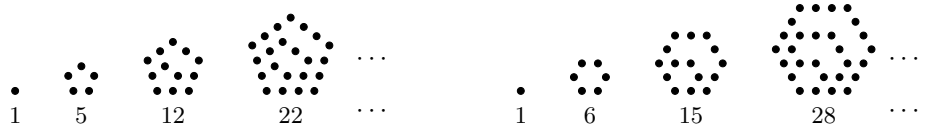


números cuadrados

consideraban también números *triangulares*, *pentagonales* y *hexagonales* etcétera.



números triangulares



números pentagonales y hexagonales

Estos números ocurren cuando se consideran arreglos como si se tomaran canicas o piedras que se colocan para formar figuras. Por ello a estos números también se conocen como números *figurales*.

Los números que no se pueden acomodar en un arreglo rectangular los llamó Phaidos, un discípulo de Pitágoras, números *incompuestos*, Thymarides, otro discípulo, los llamó *rectlineales*. Hoy los llamamos *primos*. A los números que se pueden acomodar en un arreglo rectangular se les llama *compuestos*. El número 2 originalmente no era considerado primo, pero se sabe que Aristóteles sí lo consideró ya que dijo que el 2 era el único número primo que es par. Dos números son *primos relativos* si el único número que divide a ambos es la unidad.

Se consideraban más tipos de números: un número es *perfecto* si es igual a la suma de sus divisores menores que ell número. Un ejemplo de un número perfecto es 6 ya que $6 = 1 + 2 + 3$ y $1, 2, 3, 6$ son todos los divisores, mientras $1, 2, 3$ son quellos que son menores que 6. A esta suma de los divisores

menbores que el número se le llama la suma *alíquota*. Dos números que son uno la suma alíquota del otro se llaman números *amigables*. Es probable que estas definiciones ya se hicieron entre los discípulos de PITÁGORAS, pero la primera fuente en la que aparece algo escrito es en *Los Elementos* de EUCLIDES.

2.4. El estudio de los números naturales en Euclides

Se piensa que el material que trata el estudio de los números naturales en *Los Elementos* de EUCLIDES, esto son los tomos VII, VIII y IX, se deben a ARQUITAS DE TARENTO (430–360). Esto porque se encontró un fragmento de Arquitas exhibiendo una demostración que es esencialemnte idéntica a una demostración en Euclides.

Veamos ahora algunos resultados expuestos en *Los Elementos*. EUCLIDES trató los números como segmentos que son medibles por una unidad dada. Por ello el lenguaje usado en los EUCLIDES es algo difícil de leer. Por esta razón usaremos un lenguaje y una notación moderna para dar las demostraciones, pero trataremos de apegarnos a la idea original lo más posible.

Prop. VII.1: *Dados dos números desiguales y restando sucesivamente el menor del mayor, si el que queda no mide nunca al anterior hasta que quede una unidad, los números iniciales serán primos relativos.*

Demostración. Es por reduccio ad absurdum, es decir, se supone que si hay un número $e > 1$ que mide ambos números dados a y b. Sea $a > b$. Entonces también $a - b$ es divisible entre e. Esto se puede repetir las veces que se quita el menor número del mayor. Cada vez, la diferencia que se obtiene también será divisible entre e. Como la última diferencia que se obtiene es la unidad, e debe dividir la unidad, que es absurdo ya que $e > 1$.

Comentarios. En la demostración original de Euclides, se lleva a cabo el quitar el menor número del mayor tres veces de manera explícita como para dar un ejemplo.

El método de quitar el menor número del mayor – llámemosle *quita-e-intercambia* para poder referirnos a él –, era conocido antes. Así se encontraron por primera vez que hay segmentos que no son *conmensurables*, ver la Sección **??**, como por ejemplo el lado y la diagonal en un pentágono regular o en un cuadrado.

Prop. VII.2: *Dados dos números que no son primos relativos, hallar el máximo número que divide a los dos números (el máximo común divisor).*

Demostración. Sean $a > b$ los dos números dados. Si b divide a a, entonces b es divisor común y también es máximo ya que no puede haber uno mayor que b.

Si b no divide a entonces se quita repetitivamente el menor número del mayor. Sea e el último número que se obtiene. Entonces e tienen que ser mayor que 1, sino a y b serían primos relativos por la Prop. VII.1. Así que $e > 1$ es un divisior comñun de a y b.

Luego demuestra por reducció ad absurdum que e es el máximo de los divisiores comunes, muy similar a la demsotración de la Prop. VII.1.

Comentario. En esencia el argumento aquí es igual que en la Prop. VII.1, pero EUCLIDES tiene que diferenciar los dos casos porque el resultado, hoy lo llamaríamos *máximo común divisor*, en la Prop. VII.1 es la unidad (que no es un número) mientras en la Prop VII.2 es un número.

Este resulatdo se conoce hoy como *Algoritmo de* EUCLIDES. Los algoritmos son procedimientos para calcular algo o para obtener algún resultado. Los algoritmos más simples hoy en día, gracias a la notación algebraica, pueden expresarse mediante fórmulas. Por ejemplo, el algoritmo para calcular el área de un triángulo es $\frac{\text{base}\times\text{altura}}{2}$, una fórmula, y para calcular la superficie de una esfera es $\frac{4}{3}\pi \cdot r^2$, otra fórmula.

Pero no todos los algoritmos pueden expresarse mediante fórmulas cerradas. El cálculo del máximo común divisor de dos números enteros positivos fue quizás el primer algoritmo en la historia que no puede describirse mediante una fórmula.

Veamos ahora cómo funciona el Algoritmo de ECULIDES en un ejemplo: buscamos el máximo común divisior de $a = 263$ y $b = 98$. Denotamos el par como $(273, 98)$. Ahora siempre quitamos el menor del mayor.

(273,98)	$273 - 98 = 175$
(175,98)	$175 - 98 = 77$
(98,77)	$98 - 77 = 21$
(77,21)	$77 - 21 = 56$
(56,21)	$56 - 21 = 35$
(35,21)	$35 - 21 = 14$
(21,14)	$21 - 14 = 7$
(14,7)	$14 - 7 = 7$
(7,7)	$7 - 7 = 0.$

El máximo común divisor de 273 y 98 es 7. Este algoritmo es fácil de implementar y además es muy rápido. Muchas veces se acorta el número de pasos al dividir el mayor entre el menor y tomando el resto en cada paso. En el ejemplo anterior se tendría

(273,98)	$273 = 2 \cdot 98 + 77$
(98,77)	$98 = 1 \cdot 77 + 21$
(77,21)	$77 = 3 \cdot 21 + 14$
(21,14)	$21 = 1 \cdot 14 + 7$
(14,7)	$14 = 2 \cdot 7 + 0.$

La importancia práctica de este algoritmo radica en que es rápido, el resultado puede obtenerse realizando pocas operaciones, aún cuando los números involucrados puedan ser muy grandes.

Prop. VII.31: *Cada número compuesto se divide por algún número primo.*

Demostración. Sea a el número compuesto dado. Como a es compuesto, es divisible por algún número $b > 1$. Si b es primo entonces ya demostramos la afirmación. Si b no es primo, entonces es compuesto y por lo tanto existe algún divisor $c > 1$. Si c es primo, entonces c divide a b y como b divide a a entonces c divide a a y demostramos la afirmación. Si c no es primo entonces se puede seguir de la misma manera. Finalmente se encontrará un número primo que divide a.

Prop. VII.32: *Cada número (mayor que* 1*) o es primo o se divide por algún número primo.*

Demostración. Sea $a > 1$ un número. Entonces a es primo o a es compuesto. En el segundo caso sigue de la Prop. VII.31 que a es divisible entre un número primo.

Comentario. Como consecuencia se obtiene que todo número es producto de primos.

Prop. IX.20: *Hay más números primos que cualquier número dado de números primos.*

Reformulación moderna. Hay una infinidad de números primos. En aquel entonces no se habló nunca de una "infinidad" de cosas.

Demostración. Sean $p_1, \ldots, p_n$ los números primos que están dados. Consideramos el producto $x = p_1 \cdot p_2 \cdot \ldots \cdot p_n$ y su sucesor $y = x + 1$. Por la Prop. VII.32 el número y es o primo o divisible por un primo. Si y es primo, entonces, como $y > p_i$ para cualquier $i = 1, \ldots, n$, hay al menos un número primo más que los que estaban dados. Si y es divisible entre un primo, entonces sea z un primo que divide y.

En el tomo VIII de *Los Elementos* habla de sucesiones geométricos. Esto son sucesiones, en los cuales la proporción entre dos números consecutivos es constante. Por ejemplo $3, 12, 48, 172$ o $32, 8, 2, \frac{1}{2}$ son sucesiónes geométricas de cuatro números. En tiempos de los griegos no había sucesiones infinitas, sino cada sucesión tenía un número finito de miembros.

Prop. IX.36: *Si se forma una suseción geométrica a partrir de la unidad en donde los términos están en la proporción 1:2 y si la suma de todos estos números es un número primo entonces es el producto de esta suma con el último número de la sucesión es un número perfecto.*

Reformulación moderna. Si $p = 1+2+2^2+\ldots+2^{n-1}$ es primo, entonces $x = 2^{n-1}p$ es perfecto.

Demostración. La suma $1 + 2 + 4 + \ldots + 2^{n-1}$ es igual a $p = 2^n - 1$ y se supone que p es primo. El número es entonces $x = (2^n - 1) \cdot 2^{n-1} = 2^{n-1}p$. Sus divisores son

$$1,\ 2,\ \ldots 2^{n-2},\ 2^{n-1},\ p,\ 2p, \ldots,\ 2^{n-2}p,\ 2^{n-1}p = x.$$

La suma alíquota s es

$$\begin{aligned} s &= \underbrace{1 + 2 + \ldots 2^{n-2} + 2^{n-1}}_{=2^n-1=p} + \underbrace{p + 2p + \ldots 2^{n-2}p}_{(2^{n-1}-1)p} \\ &= 2^{n-1}p \\ &= x, \end{aligned}$$

lo que demuestra que x es perfecto.

Comentario. Leonhard Euler demostró que cada número perfecto que es par tiene que ser de esta forma.

Desarrollos en el renacimiento y después

Siglos más tarde Pierre Fermat (1601–1665) descubrió uno de los profundos secretos de la teoría de números: que todo número entero positivo puede ser escrito como la suma de a lo más tres números triangulares, y de a lo más cuatro números cuadrados y de a lo más cinco números pentagonales, etcétera.

Por ejemplo:

$30 = 10 + 10 + 10$	suma de tres números triangulares
$30 = 1 + 4 + 9 + 16$	suma de cuatro números cuadrados
$30 = 1 + 1 + 1 + 5 + 22$	suma de cinco números pentagonales
$30 = 1 + 1 + 1 + 6 + 6 + 15$	suma de seis números hexagonales

Pero este resultado no lo demostró. En 1770 Joseph Louis Lagrange demostró el resultado para cuadrados. En 1796 Carl Friedrich Gauss demostró el caso para números triangulares. La versión general la demostró Augustin Louis Cauchy finalmente en 1813. Esto muestra una característica de las matemáticas: muchas veces sólo por el esfuerzo conjunto de varias personas es posible avanzar en terrenos difíciles.

Fermat también pensaba que había encontrado una fórmula para primos. Pensó que $F_n = 2^{(2^n)} + 1$ es primo para cualquier entero $n \geq 0$. En efecto

$$\begin{aligned} F_0 &= 2^{2^0} + 1 = 2^1 + 1 = 3, \\ F_1 &= 2^{2^1} + 1 = 2^2 + 1 = 5, \\ F_2 &= 2^{2^2} + 1 = 2^4 + 1 = 17, \\ F_3 &= 2^{2^3} + 1 = 2^8 + 1 = 129, \\ F_4 &= 2^{2^4} + 1 = 2^16 + 1 = 65\,537 \end{aligned}$$

todos son primos, pero

$$F_5 = 2^{2^5} + 1 = 2^32 + 1 = 4\,294\,967\,297 = 641 \cdot 6\,700\,417$$

no lo es, como demsotró Leonhard Euler como 100 años después. Hoy se sabe que $F_5, \ldots, F_{32}$ no son primos, pero no se sabe de F_{33}. El número F_{33} es simplemente demasiado grande, para poder decidirilo. No se sabe si hay más *Números de Fermat*, es decir, más números de la forma $F_n = 2^{(2^n)} + 1$ que sean primos.

Hay una serie de enunciados sobre números naturales, que son fáciles de enunciar pero muy difíciles de demostrar.

El último Teorema de Fermat: *La ecuación $a^n + b^n = c^n$ no tiene soluciones en los números naturales si $n > 2$.*
Fue formulado en 1637 por Pierre Fermat y demostrado en 1995 por Andrew Wiles.

La Conjetura de Goldbach: *Cada número par mayor que 2 es suma de dos primos.*
Fue formulado en 1742 por Christian Goldbach y no se ha podido demostrar.

La Conjetura de Catalan: *Solo existen dos potencias de números naturales que están a distancia 1: esto son 3^2 y 2^3.*
Fue formulado en 1844 por Eugène Charles Catalan y demostrado en 2002 por Preda Mihăilescu.

La Conjetura de primos gemelos: *Hay un número infinito de pares de primos que están a distancia 2 como 11 y 13 o 41 y 43.*

En 1849 Alphones de Polignac formuló una conjetura más general: que para cada número natural k hay un número infinito de pares de primos p y p' tales que $p' - p = 2k$. Para $k = 1$ se obtiene la Conjetura de primos gemelos. No se ha podido demostrar.

Primos de Mersenne: *Hay un número infinito de primos de* Mersenne*, esto son primos de la forma* $2^n - 1$. No es difícli ver que si $2^n - 1$ es primo, entonces n tiene que ser primo, ya que

$$2^{ab} - 1 = (2^a - 1)(1 + 2^a + 2^{2a} + 2^{3a} + \ldots + 2^{(b-1)a}).$$

Si se podría demostrar que hay un número infinito de primos de Mersenne, entonces por la Prop. IX.36 habría un número infinito de números perfectos.

Perfectos impares: *¿Existen números perfectos impares?* No se sabe hasta la fecha si esto es cierto o falso.

Hay muchas preguntas más que están abiertas. Algunas involucran también conceptos mucho más recientes como por ejemplo la *Función zeta de* Riemann. Se define como

$$\zeta(s) = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \ldots$$

para cualquier número complejo s. Se sabe que $\zeta(s) = 0$ para cualquier número entero negativo, esto se llaman los "zeros triviales" de la función ζ. La siguiente conjetura conforma uno de los problemas de milenio, cuya resolución estyá dotado de mundial reconocimiento además de un premio de un millión de dólares estadounidenses.

Hypótesis de Riemann: *Los zeros no triviales de* ζ *tienen parte real de* $\frac{1}{2}$. Propuesto en 1859 por Bernhard Riemann. Sigue abierto desde entonces.

Si se podría demostrar la Hipótesis de Riemann, se obtendrá como consecuencia una estimación precisa sobre el número de primos que son menores que un número dado. Sea $\Pi(k)$ el número de primos p con $p \leq k$. Por ejemplo $\Pi(20) = 8$ ya que los primos p con $p \leq 20$ son $2, 3, 5, 7, 11, 13, 17, 19$. Se sabe que para grandes k se tiene $\Pi(k) \approx \frac{k}{\ln(k)}$, más presico, que

$$\lim_{k \to \infty} \frac{\Pi(k)}{\frac{k}{\ln(k)}} = 1.$$

Si la Hipótesis de Riemann sería cierta, entonces se podría estimar la diferencia entre las dos funciones.

La dificultad de encontrar la factorización en primos de un número natural hoy se aprovecha para encriptar y desencriptar mensajes. Se basa en la teoría de números y el hecho que es posible encriptar a partir del produto pq de dos primos p y q pero sólo es posible desencriptar si se conocen los dos factores. De esta manera, teoría de números fue en un inicio el estudio del mundo, se volvió un campo de estudio muy alejado de la realidad y en la era digital se ha vuelto nuevamente un área aplicado e intensamente usado.

2.5. La irracionalidad de la raíz cuadrada de dos

Los pitagóricos creían que todo en el universo estaba relacionado con razones de enteros, es decir, que todo, en particular todas las distancias) se podía medir con fracciones. Sin embargo al intentar

calcular la diagonal de un cuadrado de lado igual a la unidad, descubrieron que esto no era cierto. Era imposible representar la longitud de esa diagonal con una fracción. Hé aquí un par de demostraciones de este hecho.

Los pitagóricos creían que todo podía representarse con números. Por ejemplo, descubrieron que las notas y melodías producidas eran armónicas, y sonaban bien (en su concepción) cuando había relaciones numéricas enteras simples entre las distancias de los huecos de las flautas o las longitudes de las cuerdas. Una cuerda reducida a la mitad de tamaño, produce un sonido que es una octava más alta que el tono original, y las combinaciones sonoras agradables para el gusto occidental guardan relaciones numéricas simples como $\frac{2}{3}$, $\frac{3}{4}$, $\frac{4}{5}$. Hoy en día decimos que se representan con fracciones sencillas. Así, también creían que todos los segmentos que se pudieran construir en la geometría, guardaban relaciones fraccionarias entre ellos.

Sin embargo, un día descubrieron una contradicción en esa creencia. La diagonal de un cuadrado no podía guardar una relación fraccionaria con el lado del cuadrado. Fue un descubrimiento que sacudió sus creencias y que, a la larga, les llevaría a construir una noción más precisa y completa de la longitud, cosa que lograron sin ampliar el concepto de número, que es lo que se hace en la actualidad.

He aquí una demostración de que la razón entre la diagonal y el lado de un cuadrado no puede ser una fracción. Supongamos que la relación entre la diagonal y el lado del cuadrado es $\frac{m}{n}$, donde m y n son enteros positivos. Sin pérdida de generalidad podemos suponer que m y n son primos entre sí, ya que al dividir m y n por todos sus factores comunes, la fracción $\frac{m}{n}$ no se altera. Esta suposición equivale a decir que, usando cierta longitud como unidad, la diagonal del cuadrado mide m y su lado mide n. Del teorema de Pitágoras resulta que $m^2 = n^2 + n^2$, es decir, $m^2 = 2n^2$ y por tanto m^2 es un número par. Pero se sabe que sólo los números pares tienen cuadrados pares, por lo que m debe ser par. Por tanto $m = 2k$ para algún entero positivo k. Entonces $m^2 = 4k^2 = 2n^2$. Dividiendo por 2 obtenemos: $2k^2 = n^2$, por tanto n^2 es par y también lo es n. Hemos concluido entonces que tanto m como n son pares, lo cual es una contradicción, pues habíamos postulado que m y n no tenían múltiplos comunes. Esto demuestra que es contradictorio suponer que la razón entre la diagonal y el lado de un cuadrado es una fracción.

Por el teorema de Pitágoras sabemos que el cuadrado construido sobre la diagonal de un cuadrado de lado 1 es 2 y por ello decimos que esa diagonal mide $\sqrt{2}$. La demostración anterior prueba que $\sqrt{2}$ no es una fracción, en lenguaje moderno, $\sqrt{2}$ no es un número racional. Hoy en día decimos que

$$\sqrt{2} \text{ es un número irracional}$$

Pero para poder decir esto fueron necesarios 20 siglos de maduración intelectual que llevaron a la humanidad a concebir o inventar, poco a poco y culminando el proceso a finales del siglo XIX, los *números reales* que es el sistema numérico que permite incluir a los número irracionales.

Este descubrimiento (de que la razón entre la diagonal y el lado de un cuadrado no es una fracción) destruyó completamente la creencia pitagórica de que todo en el universo se podía expresar con razones de enteros positivos. De hecho, el descubrimiento produjo una gran crisis en el ámbito de los pitagóricos, al grado de que quisieron mantenerlo en secreto e incluso hay rumores de quien lo descubrió, o lo publicó por vez primera, Hipaso de Metaponto (siglo V - siglo IV aec), murió en un naufragio, quizás provocado por sus propios compañeros en venganza por haber destruido la teoría en la que habían basado todo el conocimiento; o que se suicidó para expiar su culpa buscando otra vida en la que resolver el problema que había creado.

Los pitagóricos habían basado toda su filosofía en que todo en el mundo era número. Su descubrimiento de que la música y los números estaban relacionados les entusiasmó de tal manera que creyeron ver en ello una señal mística de que todo en el mundo era así: relaciones entre números. El hecho de que la relación entre dos magnitudes tan simples como el lado y la diagonal de un cuadrado no cumplieran con esta ley, trastornaba toda su filosofía. De hecho la invalidaba. No supieron ver que ese hecho constituía un nuevo reto que hubieran podido enfrentar ampliando el concepto de número. Y es que ellos veían a los números (enteros positivos) como algo dado al hombre junto con su propia existencia, posiblemente por alguna divinidad, en lugar de verlos como una creación humana que, como tal, era perfectamente ampliable. Lamentablemente esa creencia, la de la existencia a priori de los números enteros positivos, sigue vigente hoy en día entre la gente común, lo cual le dificulta mucho apropiarse de las matemáticas modernas. Darse cuenta de que tanto los números enteros positivos como los racionales, los reales, que incluyen a los irracionales, y los complejos son creaciones humanas, abre la puerta de entrada al mundo de las matemáticas modernas.

Quien enfrentó por primera vez y resolvió parcialmente este conflicto, fue el gran matemático Eudoxo de Cnido (390 - 337 aec), creando una aritmética de segmentos que permitía operar con ellos, con absoluta claridad y precisión, sin asignarles valores numéricos como longitudes. Decimos que lo resolvió parcialmente porque no generalizó el concepto de número sino que desarrolló una manera de hablar y manipular de esas magnitudes que podían ser inconmensurables, sin asignarles valores numéricos. Sin embargo el paso que le faltó no era de carácter técnico ino puramente filosófico. Para llegar a la solución moderna sólo faltaba definir esas magnitudes como números, que es, en el fondo, lo que hicieron dos milenios más tarde personajes como Weierstrass, Dedekind y Cauchy para definir los números reales. Pero el hecho de que este paso hubiera tardado dos mileniso en darse nos da una clara indicación de que se trata de un salto conceptual muy importante y nada fácil. Es pues, natural, que su aprendizaje cause tantas dificultades a los estudiantes del mundo moderno.

Los pitagóricos no podían aceptar que raíz cuadrada de 2 es un número ni una razón entre números, pero tampoco podían negar que era la medida de la diagonal del cuadrado de lado uno, lo decía su teorema más sólido. Los problemas lógicos generados por el descubrimiento de la irracionalidad de $\sqrt{2}$ fueron resueltos por Eudoxo de Cnido, uno de los más grandes matemáticos de la antigüedad griega, de una manera muy ingeniosa, creando lo que se conoció como la *Teoría de las proporciones*, que Euclides presenta en el Libro V de *Los elementos*. En esta teoría se construyen *proporciones* que pueden usarse en lugar de los números y que incluyen a la razón entra la diagonal y el lado del cuadrado, o sea, a $\sqrt{2}$. Sin embargo, a cambio, estas proporciones, es decir estos *nuevos números*, no siempre pueden representarse por fracciones. Muchos de ellos no son fracciones, y aún así pueden usarse para realizar operaciones con ellos mediante un *álgebra geométrica* que se presenta en el capítulo de Geometría.

Siglos después, a finales del XIX, una construcción fue utilizada por Dedekind (y otra por Cauchy) para formalizar el concepto de número real, que es el que abarca tanto a los racionales como a los irracionales. También poco mas o menos en la misma época David Hilbert reorganiza los fundamentos de la Geometría usando el Teorema de Pitágoras como fundamento. Así que el Teorema de Pitágoras es una de las piezas más importantes en la historia de las matemáticas. No solo es uno de los primeros teoremas descubiertos por varias civilizaciones, de los primeros en ser demostrados lógica y rigurosamente, sino también fue pieza clave en el descubrimiento de los números irracionales y se convirtió en el fundamento esencial de las matemáticas modernas. Otra pieza clave en esta historia, igualmente importante, tiene que ver no con triángulos rectángulos sino con círculos, se

trata del número π.

Las matemáticas y el mundo: ¿una relación mágica o científica?

El interés de los pitagóricos en los números y sus propiedades representa quizás la manifestación más antigua de las matemáticas puras. Se trata de entender, por deducción lógica, las propiedades de unos entes abstractos (los números) que han adquirido una existencia propia, independiente de sus humildes orígenes como utensilios para ayudar en el quehacer cotidiano. Un hecho que es interesante notar, es que esta primera aparición del interés matemático puro surge unido al misticismo, se atribuyen propiedades mágicas a ciertos números. Cabe preguntar qué tanto ayudó esta tendencia al misticismo a desarrollar las matemáticas. Aparte de esta etapa de los pitagóricos, caen en esta misma línea de misticismo la astrología de Ptolomeo y los esfuerzos de Kepler para explicar el sistema solar con los poliedros regulares. Son intentos de entender el mundo usando las matemáticas pero de una manera poco científica, como esperando descubrir algunas decisiones divinas a las que podemos tener acceso solo porque coinciden con relaciones entre números o entre cuerpos geométricos. Es como suponer que Dios es matemático y se comunica con nosotros a través de los objetos matemáticos pero sin que haya relaciones de causa efecto en los caprichos que llevan a organizar "matemáticamente" el mundo. Esta tendencia mística o mágica hacia entender el mundo a través de las matemáticas contrasta notablemente con la postura más científica de Galileo Galilei en donde ya las leyes de la naturaleza están ante nosotros como un libro abierto, no hace falta adivinar, basta leerlo, pero antes hay que conocer el lenguaje en que están escritas: las matemáticas. La gran diferencia es que en la postura galileana, no hay magia, no hay capricho divino ni casualidad, es una relación fáctica. El mundo tiene un comportamiento matemático. La idea de Dios deja de ser necesaria, aunque no por ello se necesariamente incompatible.

Capítulo 3

Geometría y Medición

Estrategias para cuantificar el espacio

La mayor parte de las aplicaciones de la geometría a la vida productiva se realizan al intentar calcular tamaños, distancias, áreas o volúmenes de objetos del mundo material cuya forma puede describirse en términos de figuras geométricas. En estos casos resulta imprescindible poder asignarles valores numéricos. De esto se ocupa la geometría cuando se combina con la *medición*.

En este capítulo exploramos la relación entre la geometría clásica y la necesidad de medir cosas que se pueden identificar con algunas partes de figuras geométricas. Primero se investigan las relaciones entre los conceptos geométricos de longitud, área y volumen con sus medidas numéricas y más adelante se presentan algunos casos de cálculos específico junto con los métodos que para realizarlos se han desarrollado a través de los siglos.

Cuando las nociones de medida se intentan aplicar a figuras geométricas formadas por curvas que no están acotadas por líneas rectas, o por cuerpos que no están acotados por superficies planas, la medición comienza a necesitar de procesos infinitos. Los esfuerzos de Eudoxo de Cnido (390 - 337 aec) y Arquímedes de Siracusa (287 - 212 aec) en esta dirección representan los primeros avances en lo que dos milenios más tarde, culminaría con la creación del *cálculo*, gracias al trabajo de matemáticos de los siglos XVII y XVIII como Isaac Newton (1643 - 1727), Gottfried Leibniz (1646 - 1716), Jakob Bernoulli (1654 - 1705), Johann Bernoulli (1667 - 1748) y Leonhard Euler (1707 - 1783).

3.1. El uso de los números para cuantificar la realidad

Las necesidades del comercio, el cobro de impuestos, la construcción y muchas otras actividades que realizaban los miembros de las diversas culturas de la antigüedad requirieron no sólo la creación de los números sino también formas de cuantificar, asignando valores numéricos, a ciertos aspectos de la realidad como tamaño, peso, duración, sin lo cual su utilidad era escasa. Este tipo de actividad ha sido probablemente la mayor fuente de conceptos e ideas matemáticas desde la antigüedad hasta

nuestros días. Esta fue, sin duda, una de las primeras actividades humanas que requirió la aplicación de un pensamiento matemático avanzado.

Los registros que tenemos en tabletas y papiros muestran un interés en realizar cálculos que permitieran asignar valores numéricos a ciertas *abstracciones de utilidad práctica*. En particular exhiben el conocimiento de algunos procedimientos y fórmulas para calcular *áreas* y *volúmenes*, a veces de manera exacta y a veces sólo aproximada. Por ejemplo, hay en ellas estimaciones de lo que hoy llamamos *el número* π, que eran útiles en fórmulas o procedimientos que se usaban para calcular *áreas* de terrenos y *volúmenes* de pirámides y contenedores. No hay evidencias de que se hubiera llegado a esas estimaciones, fórmulas y procedimientos mediante razonamientos deductivos, lo cual apuna a un pensamiento mtemático muy primitivo. Esta actividad matemática surge en la antigüedad como una *tecnología* asociada a la cuantificación, es decir, a la asignación de valores numéricos a ciertas abstracciones de interés práctico. Los métodos de cálculo se desarrollan para obtener unos valores a partir de otros aprovechando las propiedades de los números y sus operaciones, las cuales aparecen como algo natural en esos contextos. La notación numérica tan distinta en las diversas culturas atestigua que estamos ante desarrollos independientes que obedecieron a necesidades prácticas similares y dieron lugar a sistemas equivalentes para denotar a los números y realizar operaciones con ellos. Podríamos decir que se trata de una concepción de las matemáticas como una tecnología para cuantificar abstracciones de utilidad práctica. Aún no aparece, al menos de manera explícita, la deducción.

Con el tiempo, esta actividad se fue haciendo cada vez más racional, justificando de manera rigurosa cada fórmula y procedimiento que se inventaba, de manera que acabó integrándose por completo a la actividad matemática deductiva inaugurada por los antiguos griegos con la geometría, aunque sigue dando mayor importancia a la utilidad práctica que al aspecto teórico.

Los métodos de cálculo de los babilonios y los egipcios

Hemos visto que los antiguos babilonios y egipcios ya tenían sorprendentes conocimientos prácticos de matemáticas y que incluso habían desarrollado formas de enseñarlos a través de la escritura en tabletas de barro y papiros, respectivamente. Tales conocimientos se transmitían a través de ejemplos concretos, algunos de los cuales ya se describieron en la Sección **??**. El siguiente ejemplo, que aparece en una de las tabletas babilonias, ilustra la capacidad que tenían para describir algoritmos de cálculo y realizar operaciones con su extraordinaria notación numérica. El ejemplo enseña cómo calcular la *longitud* de uno de los catetos de un triángulo rectángulo cuya hipotenusa y el otro cateto se conocen. Se trata de una viga que primero se apoya verticalmente en un muro y luego su parte inferior es separada de la pared, de manera que la superior desciende una cierta distancia. Para reconocer el significado del contenido de la tableta hay que recordar que los babilonios usaban el sistema sexagésimal y que en su notación $a; b, c$ representa al número $a + b\frac{1}{60} + c\frac{1}{60^2}$. El ejemplo se ve así:

Una viga 0;30	tiene la longitud $30\frac{1}{60} = \frac{1}{2}$
De arriba bajó 0;6	eso equivale a $6\frac{1}{60}$
Abajo, cuánto se alejó?	
0;30 elévalo al cuadrado, 0;15 lo ves	$(30\frac{1}{60})^2 = 900\frac{1}{60^2} = \frac{15}{60}$
0;6 de 0;30 réstalo, 0;24 lo ves	$30\frac{1}{60} - 6\frac{1}{60} = 24\frac{1}{60}$
0;24 elévalo al cuadrado, 0;9,36 lo ves	$(24\frac{1}{60})^2 = \frac{576}{3600} = 9\frac{1}{60} + 36\frac{1}{60^2}$
0;9;36 de 0;15 réstalo, 0;5,24 lo ves	$15\frac{1}{60} - (9\frac{1}{60} + 36\frac{1}{60^2}) = 5\frac{1}{60} + 24\frac{1}{60^2}$
0;5,24 tiene qué como raíz? 0;18 lo es	$(18\frac{1}{60})^2 = 324\frac{1}{60^2} = 5\frac{1}{60} + 24\frac{1}{24^2}$
0;18 se alejó.	

En las fuentes antiguas, como las tablillas de arcilla que se encontraron en Babilonia, no hay ningún indicio de una *justificación* lógica de por qué esta serie de instrucciones funciona *siempre*, es decir aún si se toman otros números de los que se dan en el ejemplo.

Es interesante ver que en las fuentes egipcias y babilonias se encontraron dos métodos diferentes de calcular el volumen de un *tronco*, es decir una pirámide cuadrada a la cual se le quitó la punta con un corte paralelo a la base.

Si se traducen estos dos métodos a nuestro lenguaje algebraico, resultan las siguientes dos fórmulas:

$V = \frac{1}{2}(a^2 + b^2)h$	Babilona (ca. 1900 a.e.c.)
$V = \frac{1}{3}(a^2 + ab + b^2)h$	Egipto (ca. 1850 a.e.c.)

En ambas a y b son los lados de las dos bases y h es la altura. Está claro que las fórmulas no pueden ser ambas correctas.

Ejercicio. Sugerimos al lector determinar rigurosamente si alguna de las dos fórmulas es correcta; en qué casos, si los hay, alguna de ellas puede ser correcta y, si alguna no es correcta, encontrar qué tan buenas o malas aproximaciones produce. Este ejercicio le ayudará a apreciar el nivel matemático de la cultura egipcia.

Como veremos en este capítulo, cuantificar conceptos como longitud, área, volumen, peso y tiempo, es decir aquello que no es *discreto*, sino *continuo*, no fue fácil. Una de las razones por las que se encontraron dificultades para lograrlo es que la humanidad no tenía. en en lenguaje natural, una noción suficientemente amplia de *número*. Carecía de lo que hoy llamamos *números reales*, que son los que se necesitan para cuantificar lo que es continuo. En cierto sentido, este capítulo es una historia del desarrollo paralelo de las ideas del *continuo* y de los *números reales*.

Recordemos que la cultura clásica griega, que es de la que parten estas ideas, encontró muy temprano en su historia que había longitudes inconmensurables, es decir, que los números fraccionarios, que son los que habían llegado a manejar las civilizaciones más antiguas como la babilonia y la egipcia, eran insuficientes para cuantificar lo continuo. La respuesta inicial ante este problema, como veremos en las primeras secciones del capítulo, fue desarrollar los conceptos de manera geométrica, evitando tanto como era posible su cuantificación. De hecho, *Los elementos* de Euclides logran una presentación muy completa de los conceptos de longitud, área y volumen sin asignarles valores numéricos.

3.2. Longitud, área y volumen

Introducción

El volumen de un cuerpo depende de su tamaño y de su forma. Por ejemplo, un cubo tiene mayor volumen que una esfera cuyo diámetro es igual al lado del cubo. Esto lo sabemos porque la esfera cabe dentro del cubo y no lo llena. Tenemos una idea intuitiva de lo que llamamos *volumen* y sabemos que, independientemente de cómo lo midamos, debe tener ciertas propiedades. Concretamente:

- (a) Si un cuerpo A cabe en otro cuerpo B, el volumen de A es menor que el de B.
- (b) Si A y B son cuerpos ajenos, el volumen de la unión de A y B es la suma de los volúmenes de A y B.

Al hacer estas afirmaciones suponemos inconscientemente que el volumen de un cuerpo es un número. (a) dice que uno de esos números es menor que otro y (b) habla de sumar dos de esos números.

Lo mismo que con el volumen ocurre con las longitudes de segmentos y las áreas de figuras planas.

Sin embargo Euclides, en su libro *Los elementos*, hace un tratado muy completo de longitudes, áreas y volúmenes sin asignar valores numéricos a tales conceptos. De hecho ni siquiera introduce palabras equivalentes a longitud, área y volumen.

¿Cómo es eso posible?

Simplemente crea un concepto de igualdad diciendo que dos segmentos son *iguales* si se pueden superponer exactamente, es decir, si uno de ellos se puede mover rígidamente (mediante traslaciones y giros) hasta superponerlo exactamente sobre el otro (en el lenguaje moderno: si son congruentes).

En el caso de las figuras planas dice que dos de ellas son iguales si una se puede cortar en partes que puedan acomodarse, mediante movimientos rígidos (sin deformarlas), hasta cubrir exactamente la otra.

La educación matemática actual tiene una fuerte tendencia a tratar los conceptos de longitud, área y volumen en términos estrictamente numéricos, lo cual propicia un ocultamiento de algunos aspectos geométricos muy profundos de estos conceptos y redunda en una presentación superficial del tema.

Usemos el Teorema de Pitágoras para ejemplificar este hecho. El teorema de Pitágoras dice, según Euclides, que

En un triángulo rectángulo
los cuadrados construidos sobre los catetos
son iguales al cuadrado construido sobre la hipotenusa

En este enunciado los *cuadrados* son figuras geométricas. Lo que quiere decir es que hay una manera de cortar esos dos *cuadrados* cuyos lados son los catetos y acomodar las partes de manera que cubran exactamente el *cuadrado* cuyo lado es la hipotenusa. El enunciado habla de figuras geométricas y de operaciones geométricas consistentes en cortes y movimientos rígidos de los pedazos. En ningún momento habla de números asignados a esos cuadrados ni de sumarlos para compararlos con el

número asignado al otro.

En cambio, el enunciado que se usa en general en la educación matemática actual es:

En un triángulo rectángulo,
la suma de los cuadrados de los catetos
es igual al cuadrado de la hipotenusa

Los *cuadrados* que se mencionan en este enunciado son *números* que se obtienen al elevar al cuadrado otros números (multiplicándolos por sí mismos) que son las longitudes de los catetos. Y la *suma* se entiende como una operación aritmética, suma de números, y no como una operación geométrica.

Sin duda la creación de la Geometría Analítica y en general la aritmetización de la geometría fue un gran logro de la humanidad. Pero al reemplazar la enseñanza de la geometría clásica con su forma aritmetizada se ha perdido profundidad, se han empobrecido los conceptos y ha quedado un estudio que privilegia el cálculo por encima de la comprensión conceptual.

Es perfectamente posible desarrollar los conceptos matemáticos de longitud, área y volumen sin necesidad de asignarles valores numéricos. Para ello es necesario tratarlos como propiedades de los objetos geométricos, que se relacionan unas con otras. Por ejemplo, se pueden comparar y agregar y, cuando uno de ellos está contenido en otro, se puede tomar su diferencia.

En las próximas secciones nos apartamos del tratamiento numérico con la intención de recuperar, al menos en espíritu, el enfoque euclidiano, y exhibir con claridad los aspectos geométricos de la medición. El tratamiento numérico irá apareciendo poco a poco, de manera parecida a como ocurrió históricamente.

3.2.1. Los conceptos de longitud, área y volumen

En el antiguo Egipto, se calculaba el volumen de pirámides y otros cuerpos geométricos, tanto para estimar la cantidad de grano que podían guardar en sus trojes, como la cantidad de material y tiempo que iba a requerir una construcción.

Desarrollaron el concepto de volumen como una generalización del de área, y dejaron claro que el volumen de un prisma es el producto de la superficie de la base por la altura. Esto no se limita a prismas rectos sino que aplica a los oblicuos. Este hecho es una generalización, en tres dimensiones, de la relación ya conocida entre en un paralelogramo y el rectángulo con la misma base y la misma altura.

Los egipcios supieron calcular el volumen de las pirámides, tomando la tercera parte del área de la base por la altura. Se desconoce si tenían argumentos racionales para deducir esta fórmula, pero la usaron incluso para calcular el volumen de pirámides truncadas.

Observaron la analogía entre

$$\text{Area} = \frac{\text{base} \times \text{altura}}{2}$$

para el cálculo del área de un triángulo y

$$\text{Volumen} = \frac{\text{área de la base} \times \text{altura}}{3}$$

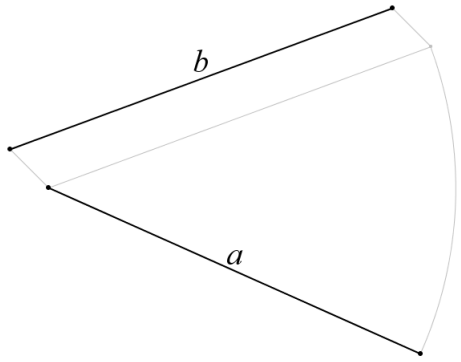
en el cálculo del volumen de una pirámide.

Como veremos más adelante, la primera demostración conocida de esta fórmula la dio Eudoxo de Cnido (390 - 337 aec). Pero para llegar a esto, conviene antes hacer una presentación conceptual de los conceptos de longitud, área y volumen que nos acerque al pensamiento griego de aquella época en relación a estos conceptos.

Longitud de segmentos

Partiremos de las nociones de *segmento recto*, que llamaremos simplemente *segmento* y de *transformación rígida*, entendida como aquella que no cambia ni el tamaño ni la forma de las cosas. Cada segmento queda determinado por dos puntos que se denominan los extremos del segmento. Los segmentos pueden estar sobre una línea recta, en un plano o en el espacio. Un segmento determina una única línea recta, la que determinan los puntos extremos del segmento, según los dos primeros postulados de Euclides.

Decimos que *dos segmentos a y b tienen la misma longitud* si se puede aplicar una transformación rígida a uno de ellos de manera que lo superponga exactamente sobre el otro. Para ello basta que la transformación rígida lleve los dos extremos del primer segmento a los dos extremos del segundo. Cuando dos segmentos a y b tienen la misma longitud, podemos escribir $a \overset{g}{=} b$. [1]



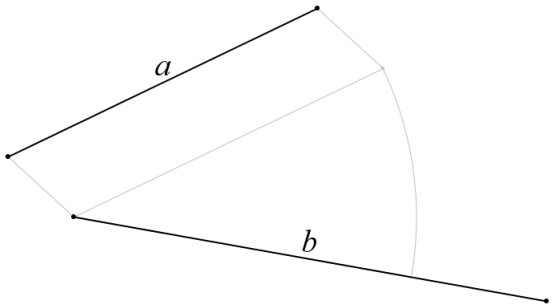
$a \overset{g}{=} b$

Es fácil demostrar que $\overset{g}{=}$ es una *relación de equivalencia*, es decir, que:

- para todo segmento a, $a \overset{g}{=} a$
- si a y b son dos segmentos tales que $a \overset{g}{=} b$ entonces $b \overset{g}{=} b$
- si a, b y c son tres segmentos tales que $a \overset{g}{=} b$ y $b \overset{g}{=} c$, entonces $a \overset{g}{=} c$.

También podemos comparar las longitudes de dos segmentos a y b. Decimos que la longitud de a es menor que la de b si es posible hallar una transformación rígida que lleve a a a estar contenido en b sin que los dos extremos de a coincidan con los de b. Cuando la longitud del segmento a es menor que la del segmento b podemos escribir $a \overset{g}{\leq} b$

[1] La letra g sobre los signos de igualdad y desigualdad indicará en esta sección que no se trata de relaciones numéricas sino geométricas

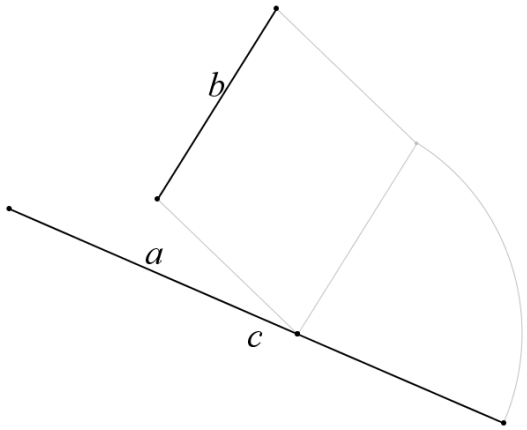


$a \overset{g}{\leq} b$

Es fácil ver que $\overset{g}{\leq}$ es una relación transitiva, es decir, que si $a \overset{g}{\leq} b$ y $b \overset{g}{\leq} c$ entonces $a \overset{g}{\leq} c$.

Bajo ciertas circunstancias dos segmentos se pueden sumar. Concretamente, si dos segmentos a y b se encuentran sobre una misma línea recta y son ajenos excepto por un extremo de a coincide con uno de b, entonces el segmento c determinado por los dos extremos libres (o sea los que no coinciden), se puede llamar *la suma* de a y b, y se puede escribir: $a + b \overset{g}{=} c$.

En general no es posible sumar segmentos. Sin embargo, si a y b son dos segmentos cualesquiera, se puede construir un segmento b' sobre la línea recta determinada por a tal que $b' \overset{g}{=} b$ que tenga la misma longitud que b y quede en posición de poderse sumar con a. Tal construcción se puede hacer con regla y compás como se indica en la siguiente figura.



$c \overset{g}{=} a + b'$

Podemos definir una "función" *Long* que asigne a cada segmento un ente abstracto que podemos llamar *longitud*, entendida no como un número sino como

la clase de equivalencia de todos los segmentos a' tales que $a' \overset{g}{=} a$

Con esta definición queda claro que la longitud $Long(a)$ de un segmento a **no es un número** pero podemos hablar de ella absoluta precisión, como veremos a continuación.

Todos los segmento que constan de un sólo punto, es decir, cuyos extremos coinciden, tiene la misma longitud. En efecto, uno de estos segmentos puede hacerse coincidir exactamente con cualquier otro

de ellos mediante una traslación. A estos segmentos los llamamos *segmentos nulos*. La longitud de los segmentos nulos es única y la consideramos como el *cero aditivo* de las longitudes.

Hemos establecido una manera de hablar de las longitudes de los segmentos sin necesidad de asignarles valores numéricos. Podemos comparar las longitudes de dos segmentos, saber si son iguales o si una e menor que la otra y hasta podemos sumar las longitudes de dos segmentos. Todo ello sin asignarles valores numéricos.

A continuación enunciamos las propiedades más importantes de la longitud de segmentos. Dejamos al lector los detalles de las demostraciones, algunas de las cuales utilizan propiedades de las transformaciones rígidas que aún no hemos especificado y que están íntimamente relacionadas con los axiomas referentes a la homogeneidad del espacio.

Propiedades de la longitud de segmentos

Sean a, b, c, ... segmentos cualesquiera.

Simetría: Si $Long(a) = Long(b)$, entonces $Long(b) = Long(a)$.

Transitividad: Si $Long(a) = Long(b)$ y $Long(b) = Long(c)$,
entonces $Long(a) = Long(c)$.

Conmutabilidad: $Long(a) + Long(b) = Long(b) + Long(a)$.

Cero aditivo: Si o es un segmento nulo (que es vacío o bien consta de un solo punto) entonces $Long(a) + Long(o) = Long(a)$ para cualquier segmento a.

Tricotomía: Sucede una y sólo una de estas tres posibilidades:
(1) $Long(a) < Long(b)$
(2) $Long(b) < Long(a)$
(3) $Long(a) = Long(b)$.

Transitividad de la desigualdad: Si $Long(a) < Long(b)$ y $Long(b) < Long(c)$, entonces $Long(a) < Long(c)$.

Propiedad arquimediana: Dados dos segmentos a y b existe un entero n tal que $Long(a) < nLong(b)$.

Notas
1)Podemos extender la definición de $nLong(a)$ al caso $n = 0$ definiendo $0Long(a) = Long(o)$ donde o es cualquier segmento nulo.
2) Se puede definir la resta o diferencia $Long(b) - Long(a)$ de longitudes, pero sólo en el caso en que $Long(a) < Long(b)$.
3) En *Los elementos*, Euclides hace esencialmente lo mismo que hemos hecho aquí, pero sin hablar de longitud, sino que dice que *dos segmentos son iguales* si uno se puede superponer sobre el otro. Euclides y sus contemporáneos no distinguieron entre el segmento y su longitud. Hablaban de ambas cosas como si fueran la misma. En nuestra presentación, en cambio, dos segmentos son iguales únicamente si ocupan el mismo sitio, es decir, si tienen los mismos extremos. Si no, pueden tener la misma longitud, pero no ser iguales.

Armados ya con el concepto abstracto, no numérico, de la longitud de un segmento, podremos desarrollar en las siguientes secciones los de área y volumen, y más adelante el de longitud de una curva.

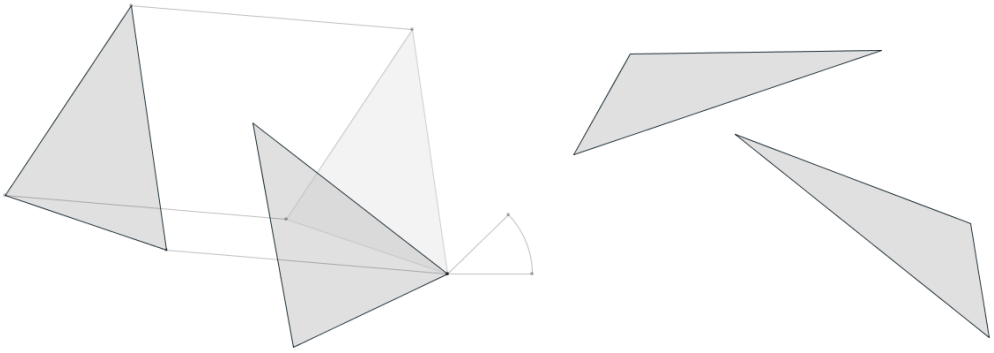
Es interesante observar que el concepto no numérico de longitud, que es esencialmente el que usa la geometría euclidiana y que hemos formalizado con la función $Long$, puede dar lugar al concepto numérico de longitud que se usa en la actualidad con sólo asumir la existencia de los número reales, como veremos en el siguiente párrafo.

Dado un segmento a podemos definir $2Long(a)$ como $Long(a + a')$ para un segmento a' tal que $a' \overset{g}{=} a$ que se encuentre en posición de ser sumado con a, es decir, que sea ajeno con a, excepto por un extremo de cada uno de ellos. Podemos extender esta definición y dar significado a $nLong(a)$ para cualquier entero positivo n. Igualmente se puede extender la definición a $qLong(a)$ para cualquier número racional $q = m/n$ y finalmente, por continuidad, a $rLong(a)$ para cualquier número real r. Si se adopta un segmento cualquiera u como "unidad", se puede demostrar, usando el axioma de Arquímedes, que para cada cada segmento a, existe un único número real r para el cual $Long(a) = rLong(u)$. Así es posible definir una función que asigne a cada segmento a el número real tal que $Long(a) = rLong(u)$. Ésta es la función que se usa en la actualidad para hablar de longitudes de segementos. Lo que hemos intentado mostrar en esta sección es que es posible hablar de longitud sin usar números, y eso es esencialmente lo que hacían los antiguos griegos cuando hablaban de igualdad de segmentos. Lo mismo ocurría cuando hablaban de igualdad de figuras y cuerpos, queriendo significar la igualdad de sus áreas o sus volúmenes, como veremos a continuación.

Área de figuras planas

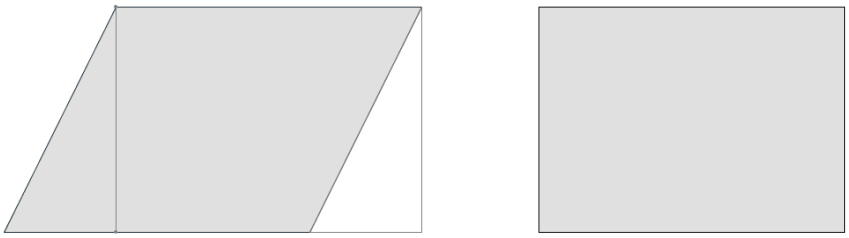
Decimos que dos figuras planas tiene la misma área si:

(a) podemos sobreponer exactamente una sobre la otra mediante traslaciones, rotaciones o reflexiones respecto a una recta.



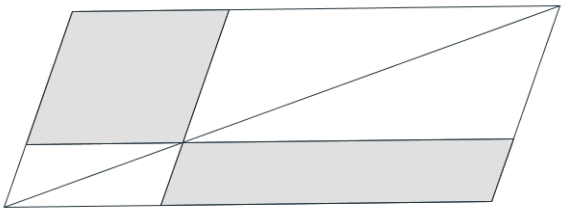
Ejemplos de triángulos con áreas iguales, según (a)

(b) podemos partir una de ellas y mover los pedazos hasta cubrir exactamente la otra



Ejemplo de dos paralelogramos con áreas iguales, según (b)

(c) podemos formarlas a partir de figuras con la misma área quitando de ellas partes que también tienen la misma área.



Ejemplo de dos paralelogramos con áreas iguales, según (c)

Nota: Las reflexiones de una figura plana respecto a un eje pueden considerarse rotaciones en el espacio tridimensional. Sin embargo cuando más adelante hablemos de cuerpos en el espacio, las reflexiones respecto a un plano no podrán considerarse rotaciones, a menos que éstas fueran en un espacio de 4 dimensiones. Por ello resulta más conveniente pensar en las reflexiones respecto a un eje como transformaciones en el plano que conservan el área.

Nota: Las traslaciones, rotaciones y reflexiones se llaman *transformaciones rígidas* o *de congruencia* y forman un grupo no conmutativo respecto a la operación de composición.

Si A y B tiene la misma área, según los criterios (a), (b) y (c), entonces escribimos:

$$A \overset{g}{=} B$$

El criterio (b) nos plantea un problema que es importante resolver de entrada. ¿Qué significa partir una figura? Tomemos el ejemplo de un triángulo que partimos mediante un corte recto que pasa por uno de sus vértices y por el punto medio del segmento opuesto. El segmento que determina el corte ¿en cual de las dos partes queda? ¿En una, en la otra, en ambas o en ninguna? Esta duda se relaciona también con la de si una figura polígona plana incluye o no su orilla. Lo primero que hay que aclarar es que estos problemas surgen porque pensamos en las figuras planas como subconjuntos del plano formados por puntos y, dada una figura, que concebimos como un conjunto de puntos, cada punto del plano debe estar en ella o en su complemento.

La comparación y la suma de áreas puede tratarse de manera análoga a como se hizo con las longitudes.

Decimos que el área de A es menor o igual que la de B si es posible hallar una partición de A tal que sus pedazos puedan meterse (mediante transformaciones de congruencia) dentro de B. Cuando

el área de A es menor o igual que la de B podemos escribir:

$$A \overset{g}{\leq} B$$

De momento nos limitaremos al estudio de figuras poligonales. La orilla o frontera de una figura poligonal es un polígono (segmentos "encadenados") cerrado.

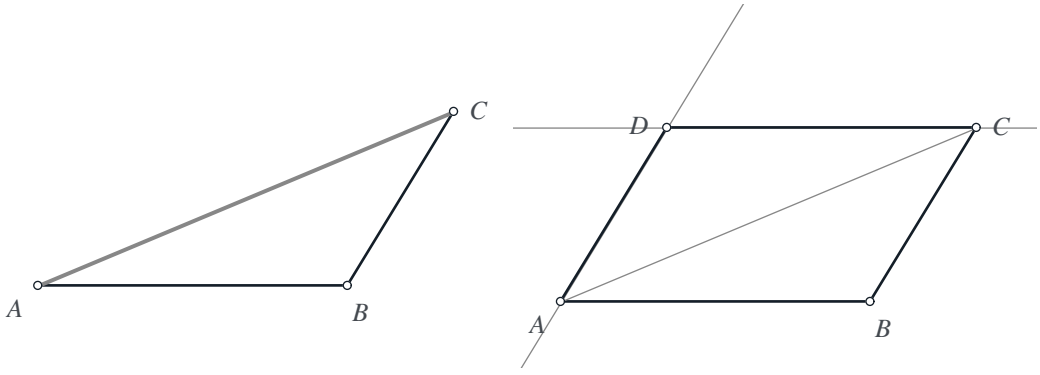
De la misma manera en que se hizo en la sección anterior para la función *Long* aplicada a segmentos, en esta sección podemos definir la función $Area(A)$ para una figura poligonal como *la clase de equivalencia* de todas las figuras poligonales A' tales que $A' \overset{g}{=} A$. Así, si A y B son dos figuras planas ajenas, la figura C que consta de ambas tiene área igual a la suma de las áreas de A y B. Es decir:

$$Area(A) + Area(B) = Area(C)$$

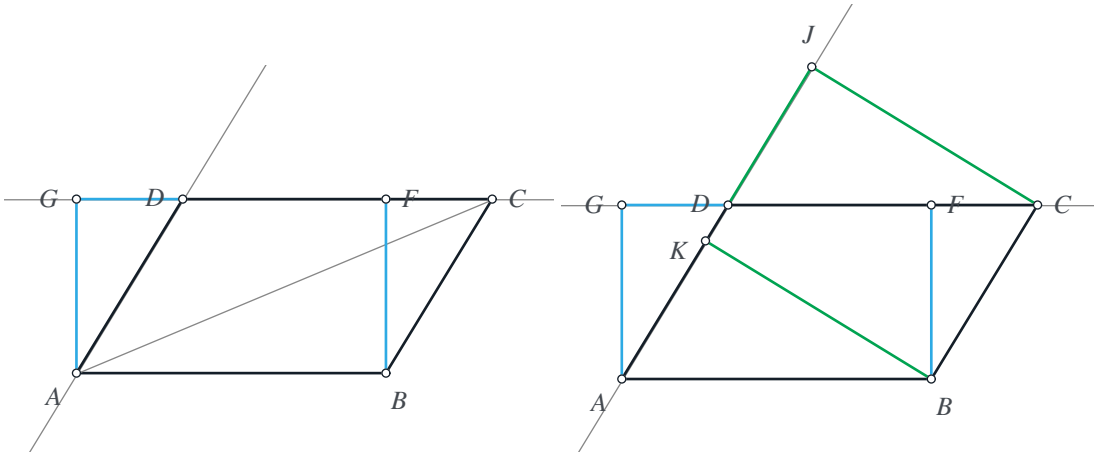
Nota: Igual que con los segmentos, es posible definir para las figuras poligonales una función numérica si se asume la existencia de los números reales y se adopta alguna figura poligonal como "la unidad". En el caso de las áreas es costumbre usar como unidad un cuadrado cuyo lado es congruente con la unidad adoptada para medir longitudes.

Área del triángulo

¿Cuál es el área del triángulo ABC? Podemos decir que es "la mitad" de la del paralelogramo $ABCD$. ¿Pero cuál es el área del paralelogramo $ABCD$?



Si F es la intersección de CD con su perpendicular por B, decimos que el área del paralelogramo $ABCD$ es $AB \cdot BF$. Lo justificamos porque si cortamos el rectángulo BCF y lo trasladamos a la posición ADG, como se indica en la figura, obtenemos un rectángulo de lados AB y BF, al cual asignamos $AB \cdot BF$ como área, por definición.



Pero igualmente podemos construir el rectángulo $BCJK$ y decir que el área del paralelogramo es $BC{\cdot}BK$. ¿Cuál es la buena? ¿Son iguales las áreas de $ABFG$ y $BCJK$? Los triángulos BCF y ABK son congruentes y esto implica que $\frac{BF}{BC} = \frac{BK}{AB}$, que es equivalente a $AB \cdot BF = BC \cdot BK$.[2]

Vemos que para tener una medida de área para triángulos y paralelogramos es necesario usar la ley de las proporciones de los triángulos semejantes. Pero también podríamos demostrar esta ley postulando la existencia de una medida del área de las figuras poligonales, la cual asigna como área a un rectángulo el producto de sus lados y asigna el mismo valor a dos figuras cuando una de ellas puede obtenerse de la otra cortando una parte y pegándola en otro sitio.

Esta situación, en la que podemos formular la teoría de áreas a partir de la semejanza de triángulos o la teoría de los triángulos semejantes a partir de la teoría de áreas, es típica de los modelos matemáticos de la física. Otro ejemplo, muy apreciado por los físicos, es la equivalencia de las formulaciones de la mecánica clásica a partir de las leyes de Newton o a partir de las leyes de conservación o a partir del principio de mínima acción.

Áreas de polígonos

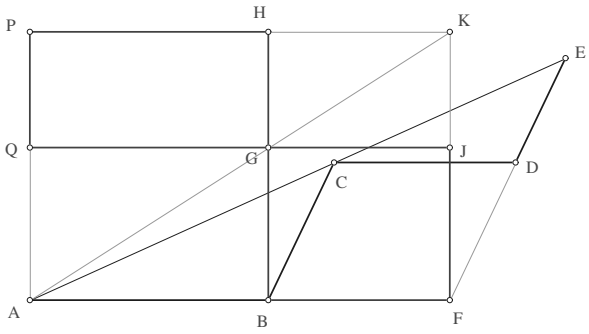
Todo rectángulo puede transformarse en un cuadrado mediante cortes y traslaciones sin superposición. Este hecho, que demostraremos en la siguiente escena, permite demostrar que toda figura

[2] Recordemos la discusión presentada en la sección **??** relacionada con esta figura, en la que se exhibe la equivalencia entre la semejanza de triángulos y el concepto de área para rectángulos, lo cual ayuda a justificar la equivalencia de las igualdades:

$$\frac{AB}{CD} = \frac{BC}{DE}$$

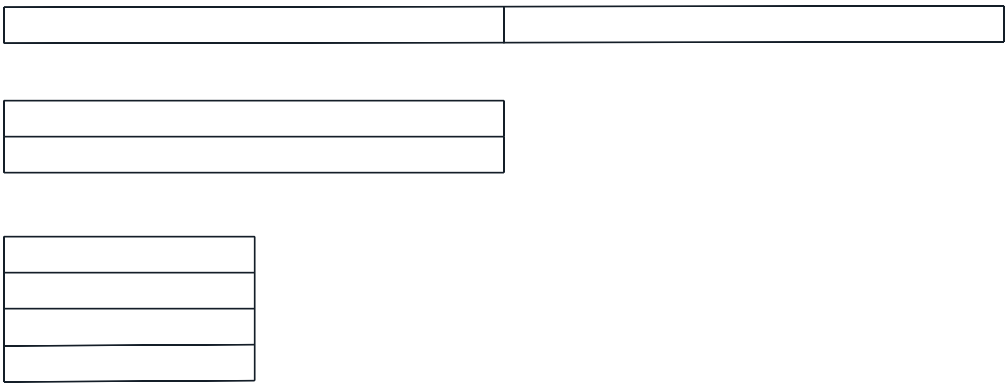
y

$$AB \cdot DE = BC \cdot CD$$

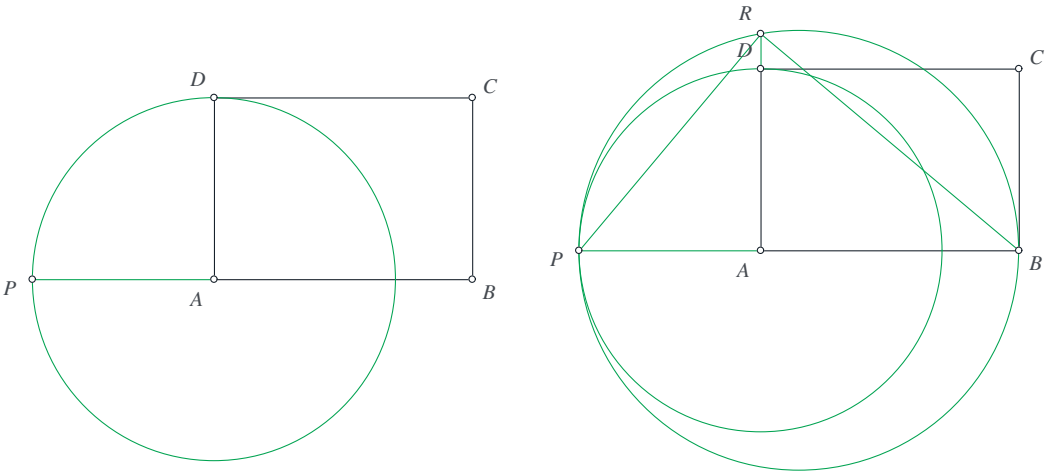


poligonal se puede transformar, mediante cortes y traslaciones sin superposición, en un cuadrado, con lo cual se le pude asignar como medida de su área la del cuadrado. Esto permite completar fácilmente la teoría de áreas de figuras poligonales planas.

La ilustración prueba este lema: Todo rectángulo se puede transformar, mediante cortes y traslaciones sin superposición, en otro cuyo lado más largo no sea mayor que el doble del otro. Si un rectángulo, como el de la figura, tiene un lado más largo que el doble del otro, podemos partirlo a la mitad por su lado más largo, tantas veces como sea necesario hasta llegar a uno en el que el lado más largo no sea mayor que el doble del otro. (En este paso se usa el axioma de Arquímedes o de continuidad.)

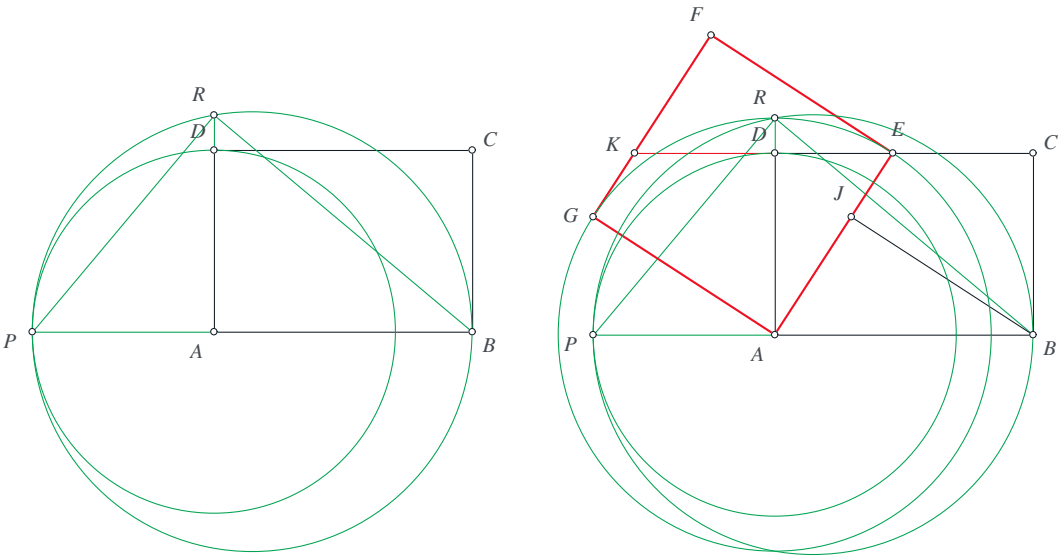


Pero con esta operación no podemos transformar un rectángulo arbitrario en un cuadrado. Para lograr esto hace falta un corte particular, como se muestra en las siguientes figuras.



Probaremos que un rectángulo cuyo lado más largo no es mayor que el doble del otro, se puede transformar en un cuadrado mediante cortes y traslaciones sin superposición. Sea $ABCD$ un rectángulo cuyo lado más largo es AB y éste es menor que el doble de BC. Sea P la intersección de AB con la circunferencia de centro A y radio $BC = AD$. Entonces $PA = BC$.

Sea R la intersección de AD con la circunferencia de diámetro PB. El ángulo PRB es recto (esto se demuestra en la unidad sobre el círculo) y los triángulos PAR y RAB son semejantes. Por lo tanto $\frac{PA}{AR} = \frac{AR}{AB}$, que equivale a $PA \cdot AB = AR^2$. Y como $PA = BC$, se tiene que $AB \cdot BC = AR^2$.



Sea E la intersección de DC con la circunferencia de centro A y radio AR. Entonces $AE = AR$ y el cuadrado de lado AE tiene la misma área que el rectángulo $ABCD$. Pero además $ABCD$ se puede transformar en el cuadrado $AEFG$. Para ello definimos J como la proyección de B sobre AE y K como la intersección de CD con FG. Trasladando el triángulo ABJ a la posición KEF y el cuadrilátero $JBCE$ a la posición $GADK$ se convierte $ABCD$ en $AEFG$.

Dejamos como ejercicio para el lector la demostración de que una figura poligonal se puede transformar en un cuadrado mediante cortes y traslaciones sin superposición. Para ello se sugiere considerar las siguientes proposiciones:

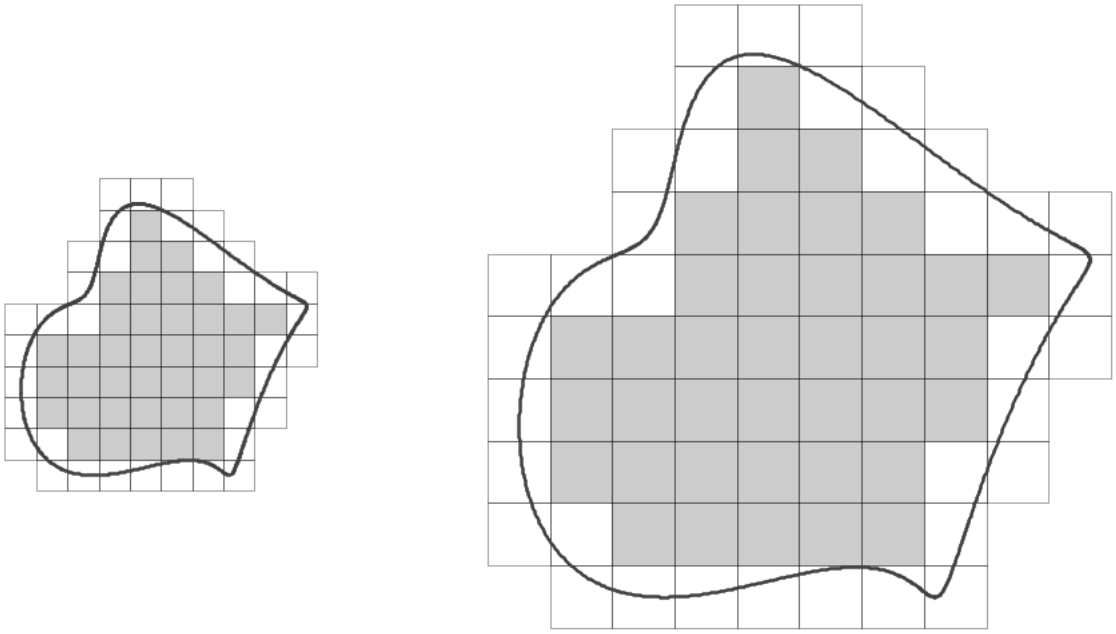
1) Toda figura poligonal se puede partir en triángulos.
2) Todo triángulo se puede transformar en un rectángulo.
3) Un cuadrado se puede transformar en un rectángulo con un lado dado.

3.3. El método de exhaución

Áreas de otras figuras

Para obtener el área de una figura plana que no es poligonal, es decir, que su frontera puede incluir partes curvas, hace falta recurrir al método de exhaución de Eudoxio que, a grandes rasgos, se puede describir así: el área de una figura plana se puede encontrar como el "límite"de las áreas de una sucesión de figuras poligonales contenidas en ella, si cualquier punto del interior de la figura cae en alguna de las figuras de la sucesión.

El resultado más importante con respecto a las áreas las figuras planas que no son poligonales, y que es fácil de demostrar siguiendo la idea sugerida en la siguiente figura,



es que si dos figuras planas F_1 y F_2 son semejantes y la razón de homotecia de (las dimensiones lineales de) F_2 a (las de) F_1 es 2, entonces

$$Area(F_2) = 4Area(F_1)$$

Este resultado puede hallarse en el libro XII de *Los elementos* de Euclides. Ambos cálculos se presentan más adelante en este mismo capítulo.

La demostración es una aplicación simple de la idea descrita en el método de exhausión. Los resultados más importantes que se obtuvieron en la antigua Grecia utilizando el método de exhausión son la *cuadratura de la parábola* y el cálculo del *área del círculo* y ambos deben su realización a Arquímedes.

Volumen

El volumen de un cuerpo se define de manera análoga al área de una figura plana.

Aquí hay que demostrar algunos teoremas sobre volúmenes de prismas. Por ejemplo:

(a) Si las áreas de las bases de dos prismas son iguales y las alturas también son iguales, entonces los dos prismas tienen el mismo volumen.

(b) Si dos prismas tienen alturas iguales y el área de la base de uno de ellos es el doble que la del otro, entonces el volumen del primero es el doble que el del segundo.

(c) Si un prisma cuya base es un paralelogramo se corta en dos prismas triangulares mediante un

plano que pasa por una diagonal de la base y es paralelo a las aristas del prisma que no son de las bases, entonces los prismas resultantes tienen igual volumen y éste, por lo tanto, es la mitad del volumen del prisma original.

El resultado más importante que usaremos, y es fácil de demostrar siguiendo la idea equivalente al caso del área, es que:

(d) Si dos cuerpos C_1 y C_2 son semejantes y la razón de homotecia de (las dimensiones lineales de) C_2 a (las de) C_1 es 2, entonces

$$Vol(C_2) = 8Vol(C_1).$$

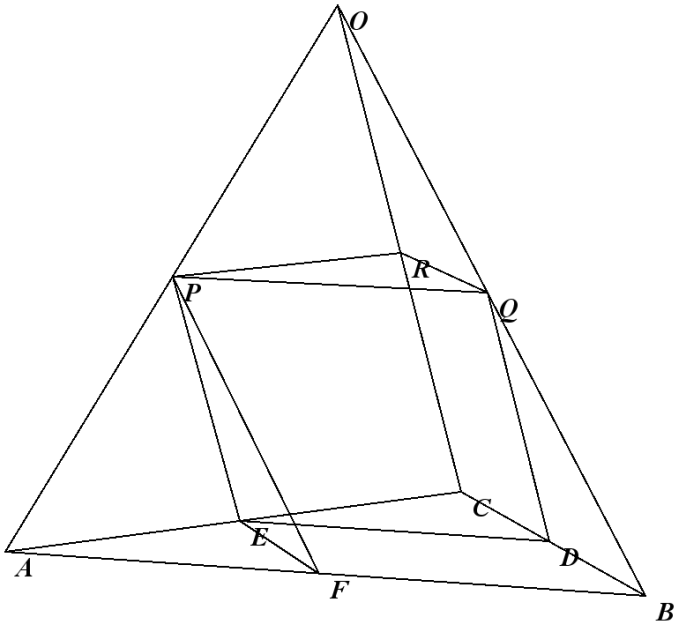
Armados con estos resultados se puede enfrentar uno de los problemas más importantes del cálculo de volúmenes de poliedros: el volumen de una pirámide, que puede reducirse al de un tetraedro. La conocida fórmula

$$\text{Volumen} = \frac{\text{área de la base} \times \text{altura}}{3} \tag{3.1}$$

es análoga a la del área del triángulo: $Area = \frac{base \times altura}{2}$, pero es bastante más difícil de obtener. Sabemos que los egipcios la conocían, pero no hay ninguna prueba de que tuvieran alguna demostración de ella. A continuación presentamos su demostración, que se debe a Eudoxo de Cnido (390 - 337 aec) y aparece en el libro XII de Los elementos de Euclides, donde además se generaliza a las pirámides y al cono.

El volumen de un tetraedro

Sea $ABCO$ un tetraedro.
Sean D, E, F los puntos medios de BC, CA y AB, respectivamente.
sean P, Q, R los puntos medios de OA, OB y OC, respectivamente.



Sea V el volumen del prisma con base triangular ABC y lado OA. Sea T el volumen del tetraedro $ABCO$. Demostraremos que $T = \frac{V}{3}$

El tetraedro puede partirse en cuatro cuerpos: dos tetraedros $AEFP$ y $PQRO$, semejantes al original con razón de homotecia $\frac{1}{2}$; y dos prismas triangulares cuyas bases son DBQ y EDC.

El volumen de cada uno de los tetraedros $AFEP$ y $PQRO$ es $\frac{T}{8}$ y el de cada uno de los prismas triangulares de bases DBQ y EDC es $\frac{V}{8}$. Por tanto:

$$T = 2\frac{T}{8} + 2\frac{V}{8}$$

de donde resulta que

$$T = \frac{V}{3}$$

Es importante reconocer que uno de los hechos más importantes que se utilizan en esta demostración es que el volumen de un cuerpo cuyas dimensiones lineales son el doble de las de otro, tiene un volumen igual a ocho veces el del segundo (que probablemente aprendieron los griegos en relación con el famoso problema de duplicar el cubo) y esto se demuestra usando el método de exhaución.

Nota: En realidad Eudoxio demostró la fórmula $V = \frac{a\,b\,c}{6}$ para un tetraedro rectangular utilizando el método de exhaución y un *proceso recursivo*. Es la demostración que puede hallarse en *Los elementos* de Euclides. La que presentamos aquí no requiere del proceso recursivo pero sí del método de exhaución. Sin embargo, el proceso recursivo que utilizó Eudoxo probablemente inspiró a Arquímedes para resolver el problema de la cuadratura de la parábola, como se ve más adelante.

A partir de este resultado no es difícil demostrar la fórmula 3.1:

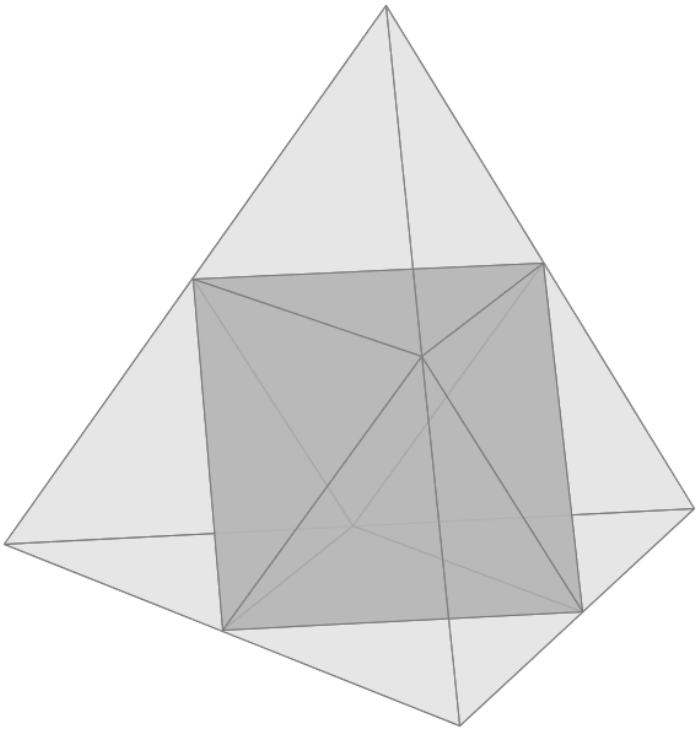
$$\text{Volumen} = \frac{\text{área de la base} \times \text{altura}}{3}$$

para el volumen de cualquier pirámide con base poligonal.

Ejercicio. Demostrar la fórmula 3.1 para el volumen de una pirámide y calcular los volúmenes de un tetraedro, un octaedro y un icosaedro regulares, descomponiendo estos poliedros en pirámides y hallando en cada caso las áreas de sus bases y sus alturas.

Volúmenes de poliedros

Ejemplo. El volumen del octaedro inscrito en un tetraedro, con lados la mitad de los del tetraedro, tiene la mitad de volumen porque resulta de quitarle al tetraedro cuatro tetraedros iguales, cada uno con la octava parte de volumen del original.



Este resultado es válido tanto si el tetraedro inicial es regular como si es irregular. Lamentablemente, este procedimiento para obtener volúmenes de poliedros no es muy general. A diferencia del caso plano, no siempre es posible descomponer un poliedro en partes que puedan recomponerse en un cubo. Este resultado apenas se demostró en el siglo XX. Fue el tercer problema de los que planteó David Hilbert en el congreso de matemáticos del año 1900.

El tercer problema de Hilbert

En el año 1900 David Hilbert presentó a la comunidad matemática 23 problemas que aún no se habían podido resolver para esa fecha. Uno de ellos, el tercero, era la pregunta de si con los poliedros sucedería lo mismo que con los polígonos en el sentido de que toda figura poligonal se puede partir en otras las cuales puedan recomponerse en un cuadrado. Es decir, ¿es cierto que todo cuerpo en el espacio formado por poliedros se puede partir en otros que puedan recomponerse en un cubo?

La formulación original del problema era ésta:

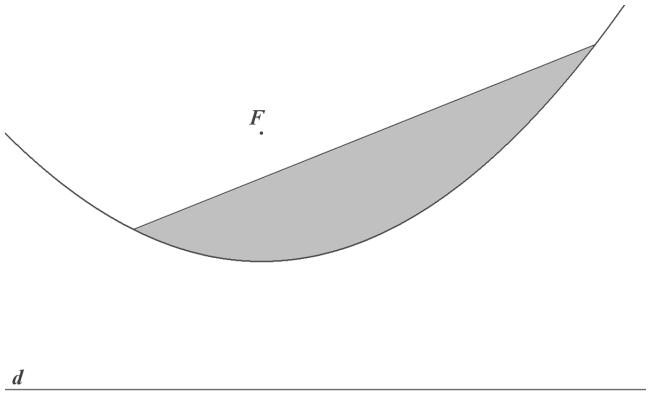
Tercer problema de Hilbert. Dados dos poliedros de igual volumen, ¿es posible cortar el primero en una cantidad finita de piezas poliédricas que puedan ser ensambladas de modo que quede armado el segundo?

Éste fue el primero de los 23 problemas que se resolvió. En 1903 Max Dehn (1878-1952) encontró un invariante ante seccionamientos de poliedros que no es el mismo, por ejemplo, para un tetraedro regular que para un cubo. Por tanto, por mucho que partamos el tetraedro, sus partes tendrán siempre el mismo valor del invariante esas partes no podrán juntarse para formar un cubo.

El invariante de Dehn es lo que se llama el *producto tensorial* de las longitudes de las aristas por el ángulo definido por las caras que se unen en cada arista.

3.4. La cuadratura de la parábola

Sin duda la aplicación estelar del método de exhausión de Eudoxio la realizó Arquímedes al calcular el área de un sector de parábola. Un *sector de parábola* es la región acotada por una parábola y una recta que la corta en dos puntos.



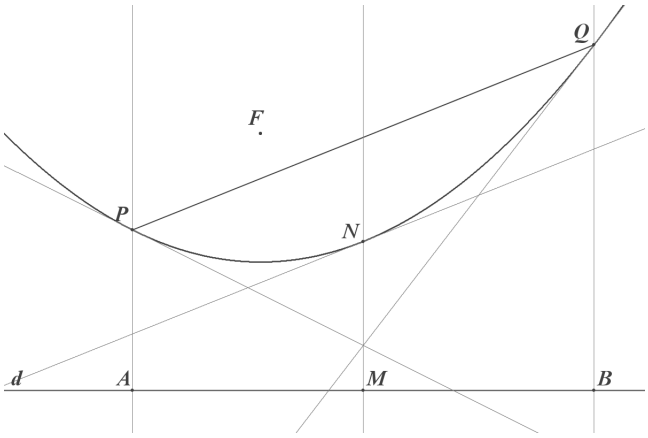
El resultado de Arquímedes puede expresarse de manera muy simple y general:

El área de un sector de parábola es $\frac{2}{3}$ *del mínimo paralelogramo que lo contiene.*

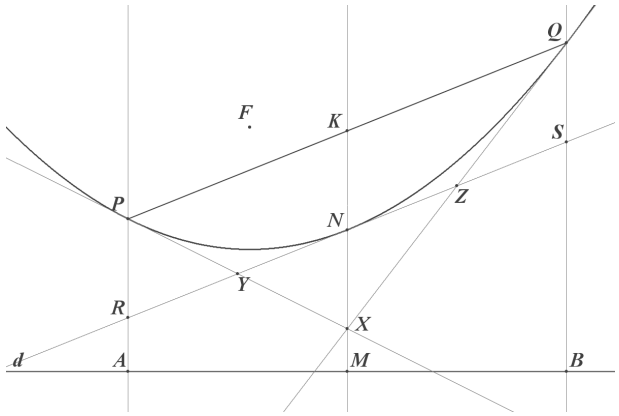
Arquímedes produjo al menos dos demostraciones de este resultado. Aquí nos ocuparemos únicamente de la que se apoya en el método de exhausión. Esta demostración se basa en una propiedad de la parábola que describimos a continuación. Sea F el foco de una parábola y sea d su directriz. Sean P y Q dos puntos sobre la parábola y sea N el punto sobre la parábola cuya proyección

ortogonal M sobre la directriz d es el punto medio de las proyecciones ortogonales A y B sobre d de P y Q, respectivamente. Entonces:

La tangente a la parábola en N es paralela a PQ



Para demostrar esta propiedad definimos X como el punto de intersección de las tangentes a la parábola en P y Q. Sabemos que X está sobre MN porque su distancia a A es la misma que a B, ya que las tangentes a las parábolas en P y Q son las mediatrices de FA y FB, respectivamente. Esta última es un propiedad básica de las tangentes a una parábola que se el lector no, conviene que la demuestra como **ejercicio**.

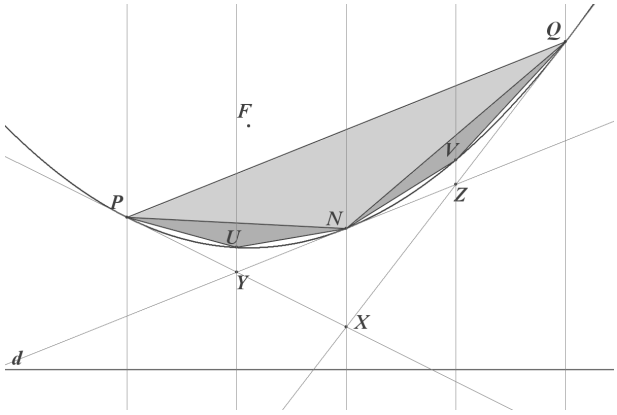


Utilizando el mismo argumento puede verse que el punto Y, definido como la intersección de las tangentes a la parábola en P y N, también se encuentra a la mitad entre las paralelas AP y MN. Por tanto los triángulos PRY y XNY son congruentes y $PR = NX$ – donde R y S son los puntos de intersección de la tangente a la parábola en N con AP y BQ, respectivamente. Análogamente, $NX = QS$. Por lo tanto $PR = QS$. Esto completa la demostración de que la tangente en P es paralela a PQ.

De hecho, como $PR = KN = QS$, donde K es el punto medio de PQ, resulta que N es el punto medio de KX y por lo tanto el área del triángulo PNQ es la mitad de la del triángulo PXQ. También es claro que el área del triángulo PNQ es la mitad de la del paralelogramo $PRSQ$.

Arquímedes demostró que el área del sector de parábola PNQ es igual a $\frac{4}{3}$ del área del triángulo PNQ, o lo que es igual, $\frac{2}{3}$ del área del paralelogramo $PRSQ$.

La demostración de Arquímedes procede por exhausión, agregando primero el triángulo PNQ y luego dos triángulos PUN y NVQ donde U y V son los puntos sobre la parábola que están también sobre las perpendiculares a d en los puntos medios de AM y MB, respectivamente. Y así se continúa dividiendo cada uno de los sectores de parábola en dos, y agregando sus triángulos inscritos.



El área a_0 del triángulo PNQ se puede calcular explícitamente:

$$a_0 = \frac{AB \cdot KN}{2}$$

Las áreas de los dos triángulos PUN y NVQ, sumadas, tienen el valor:

$$a_1 = 2\frac{\frac{AB}{2} \cdot \frac{NK}{4}}{2} = \frac{AB \cdot NK}{8} = \frac{a_0}{4}$$

Procediendo por iteración, los cuatro triángulos siguientes tienen un área conjunta:

$$a_2 = 4\frac{\frac{AB}{4} \cdot \frac{NK}{4^2}}{2} = \frac{AB \cdot NK}{2 \cdot 4^2} = \frac{a_0}{4^2}$$

En general:

$$a_n = 2n\frac{\frac{AB}{2n} \cdot \frac{NK}{4^n}}{2} = \frac{AB \cdot NK}{2 \cdot 4^n} = \frac{a_0}{4^n}$$

Así, agragndo todos estos términos, llegamos a una serie infinita para expresar el área s del sector de parábola PNQ:

$$s = a_0 + a_1 + a_2 + ... = \Big(1 + \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} \cdots\Big)a_0 = t \cdot a_0$$

donde

$$t = 1 + \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} \cdots$$

Para calcular esta suma infinita observamos que

$$t = 1 + \frac{1}{4}\Big(1 + \frac{1}{4} + \frac{1}{4^2} + \cdots\Big) = 1 + \frac{1}{4}t$$

Por tanto, despejando t de esta sencilla ecuación, obtenemos $t = \frac{4}{3}$.

Sustituyendo en $s = t \cdot a_0$ obtenemos:

$$s = \frac{4}{3}a_0 = \frac{4}{3}\frac{AB \cdot NK}{2} = \frac{2}{3}AB \cdot NK$$

que es lo que queríamos demostrar.

3.5. El número π a través de los siglos

Una de las contribuciones más notables de Arquímedes al pensamiento matemático fue su manera de estimar el valor del número π, o más precisamente, de calcular el perímetro de la circunferencia. Para hacer justicia a este trabajo vamos a presentarlo enmarcado en un relato detallado de las distintas maneras en que se ha enfrentado este problema antes y después de Arquímedes, desde las primeras fórmulas usadas en Babilonia y Egipto hasta los trabajos que probaron que π no es un número racional y los métodos que se usan en la actualidad para estimar con muchísima precisión su valor. En toda esta historia la intervención de Arquímedes es fundamental y marca un hito en la historia del pensamiento matemático porque muestra cómo realizar estimaciones tan precisas como uno necesite, de cantidades que no pueden calcularse con total exactitud. En el fondo el trabajo de Arquímedes respecto al número π marca el nacimiento de lo que hoy llamamos Análisis Matemático.

El número π en la antigüedad

Las culturas antiguas sabían que el perímetro de una circunferencia es un poco mayor que el triple de su diámetro y que el área de un círculo es algo mayor que tres veces la del cuadrado de lado igual al radio. Está más o menos claro que sabían que esas razones eran constantes, pero no que supieran que son iguales. De hecho las estimaciones que usaban para una y otra no siempre eran las mismas.

Los babilonios usaban unas veces 3 y otras $3\frac{1}{8} = 3.125$. No está claro si uno de estos valores era para la razón de los perímetros y el otro para la de las áreas o si usaban 3 en los ejercicios escolares y $3\frac{1}{8}$ cuando querían mayor precisión. No se sabe cómo obtuvieron esas aproximaciones. Nada hay que indique un intento para justificarlas. Es probable que las obtuvieran experimentalmente (por ejemplo rodeando una columna cilíndrica con una cuerda o midiendo las marcas que deja una protuberancia en el perímetro de una rueda).

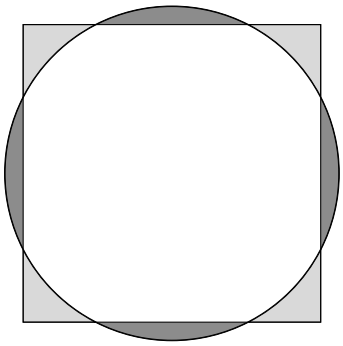
El número π entendido como *la razón entre el perímetro y el diámetro de una circunferencia* o como *la razón entre el área de un círculo y el cuadrado de su radio*, no aparece en las primeras civilizaciones, pero en cambio casi todas ellas conocían fórmulas para calcular, de manera aproximada, los perímetros y las áreas de los círculos. Un ejemplo es la fórmula

$$A = (8/9d)^2$$

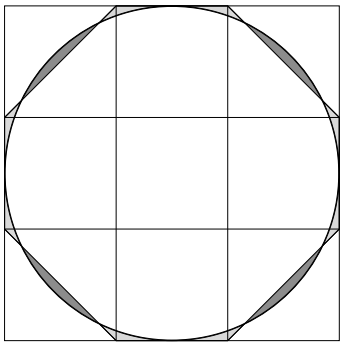
que usaban los egipcios para el área de un círculo de diámetro d. Como puede apreciarse por su aspecto, esta fórmula es más bien una "cuadratura del círculo"(solo aproximada, por supuesto), es decir, exhibe un cuadrado (de lado 8/9 d) cuya área es más o menos igual a la del círculo. Esta

fórmula es una muestra clara de lo que interesaba a los egipcios: tener un método simple de calcular, de manera aproximada, el área de un círculo. Muchos otros ejemplos de los ítems encontrados en los papiros egipcios que datan de más de 1500 años A.C. Muestran que su interés era tener procedimientos o fórmulas para calcular algo. Se desconoce si el procedimiento era producto de un razonamiento o fue encontrado de manera empírica y su justificación residía únicamente en que funcionaba bien en la práctica. Justificar procedimientos de cálculo por su utilidad y no por su racionalidad, ha subsistido hasta nuestros días y constituye una de las maneras no ortodoxa de hacer matemáticas y que muchos justifican por los beneficios prácticos que aporta.

En su afán por justificar racionalmente las fórmulas de la antigüedad, los historiadores de las matemáticas han inventado a veces argumentos factibles que hubieran podido haber usado quienes las inventaron. Por ejemplo, la figura muestra un círculo con diámetro d y un cuadrado con lado $\frac{8}{9}d$. A simple vista puede apreciarse que la diferencia entre sus áreas es pequeña. El exceso en las esquinas del cuadrado parece compensarse con los sectores del círculo que le sobresalen. De acuerdo con esto, el área del círculo (πr^2 donde $r = \frac{d}{2}$) es "casi" igual a $(\frac{8}{9}d)^2$. Tomando $d = 2$ (o sea $r = 1$) obtenemos una estimación para π de $(\frac{8}{9}2)^2 = 256/81$, que en decimales nos da $\pi \simeq 3.16$.



El razonamiento es éste: si se parte el cuadrado circunscrito al círculo de diámetro d en nueve cuadrados, cada uno de lado $\frac{1}{3}d$, como muestra la figura.



Salta a la vista que el octágono irregular en la figura debe tener un área muy parecida a la del círculo. Los 8 excesos alrededor de sus vértices parecen compensarse con los 4 de la circunferencia sobre los lados diagonales del hexágono. El área del octágono es igual a 7 veces la de los cuadrados

de lado $\frac{d}{3}$, o sea

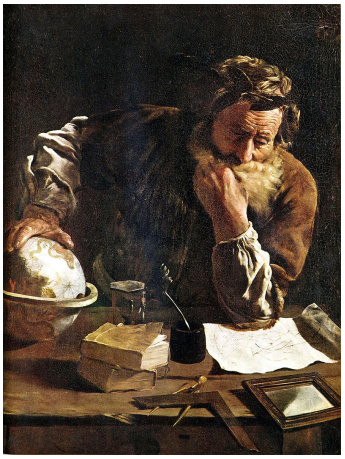
$$7\left(\frac{d}{3}\right)^2 = \frac{7}{9}d^2 = \frac{63}{81}d^2$$

que es un valor muy cercano a $\frac{64}{81}d^2$. Pero $\frac{64}{81}$ tiene la ventaja de ser el cuadrado exacto de $\frac{8}{9}$ y proporciona una *cuadratura* aproximada del círculo que ayuda a simplificar los cálculos.

Que la razón entre la circunferencia y el diámetro es la misma para todos los círculos parecía ser conocimiento común en Babilonia y Egipto. La primera demostración documentada de que la razón del área del círculo al cuadrado de su radio es constante corresponde a Euclides en el libro XII de sus Elementos. Sin embargo Euclides no proporciona el valor de dicha razón. Y es que esta constante fundamental de la naturaleza que denotamos por π es difícil de calcular, además de que la notación numérica de los griegos no fomentaba el análisis numérico [3].

El tratamiento de Arquímedes del número π

La primera estimación razonada de π, con cotas bien definidas, la realizó Arquímedes (Siracusa, Sicilia 287 – 212 a.e.c.). Arquímedes fue a la vez físico, ingeniero, inventor, astrónomo y, sobre todo, matemático. Se le considera uno de los científicos más importante de la antigüedad clásica. Pero por lo que más se le admira es por sus trabajos matemáticos, que son los únicos que él mismo consideró suficientemente importantes como para escribirlos y legarlos a la posteridad.



«Domenico-Fetti Archimedes 1620» de Domenico Fetti.

Arquímedes murió durante el sitio de Siracusa (214–212 a.e.c.), ejecutado por un soldado romano. Se dice que el soldado le ordenó levantarse y seguirlo, pero Arquímedes se mantuvo absorto en sus pensamientos dibujando en la arena, lo cual irritó al soldado que, enfadado, le asestó un golpe de espada, a pesar de que tenía órdenes de no hacerle daño.

[3] Los primeros en usar la letra π para denotar esta constante fueron los matemáticos ingleses Oughtred (1647), Isaac Barrow (1664) y David Gregory (1697).

«Cara de la medalla Fields, Archimedes» de Stefan Zachow.

Sus ideas, sus métodos y el rigor de sus demostraciones siguen siendo hoy motivo de asombro y admiración, la cual ha quedado patente en el hecho de que la medalla Fields, que es el premio más importante de las matemáticas, lleva su efigie.

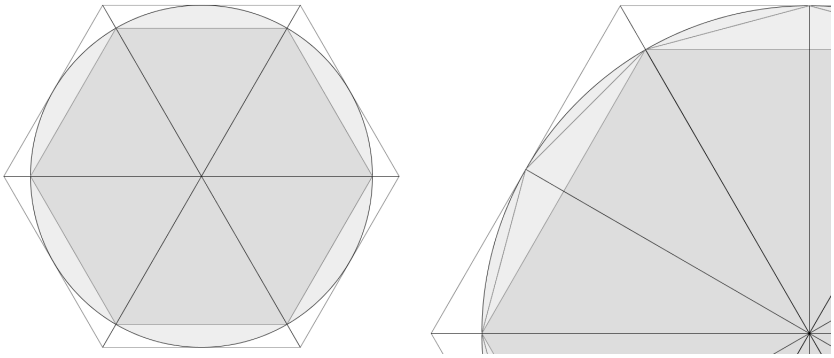
Si p_n y P_n representan los perímetros de los polígonos regulares inscrito y circunscrito de $6 \cdot 2^n$ lados, entonces, según Arquímedes

$$\frac{p_n}{2} < \pi < \frac{P_n}{2}$$

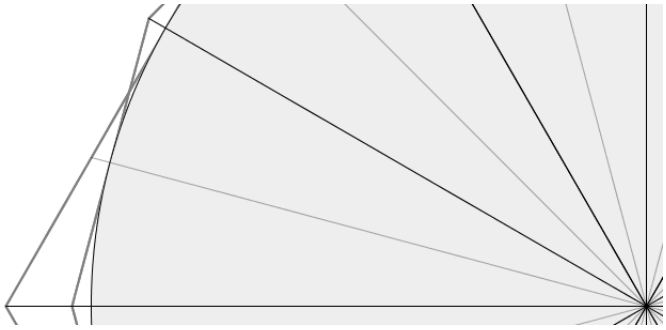
Esto nos dice que el número π se encuentra en todos los intervalos

$$\left[\frac{p_n}{2}, \frac{P_n}{2}\right]$$

Además Arquímedes da argumentos geométricos para demostrar que $\{p_n\}$ es creciente y $\{P_n\}$ es decreciente. Lo primero es bastante evidente porque dos segmentos consecutivos del polígono inscrito de $6 \cdot 2^{n+1}$ lados une los extremos de un solo lado del de $6 \cdot 2^n$ lados.



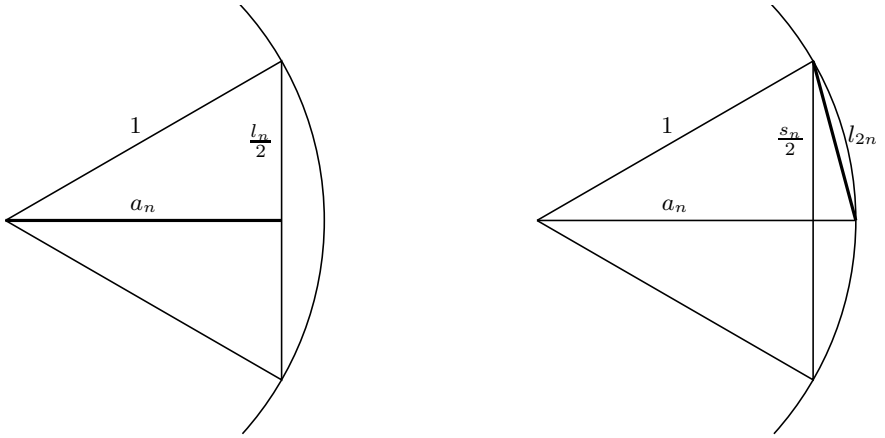
Para lo segundo basta observar que en la siguiente figura



el lado del polígono anterior puede partirse en dos partes, una de ellas igual al lado del polígono siguiente y la otra claramente mayor por poder convertirse en el lado opuesto a un ángulo obtuso en un triángulo en el que uno de los lados adyacentes al ángulo obtuso es precisamente el lado del polígono anterior.

Si probáramos que la longitud de esos intervalos tiende a cero cuando n tiende a infinito, obtendríamos una manera de estimar el valor de π con tanta precisión como quisiéramos. Para ellos bastaría tomar polígonos con un número de lados n suficientemente grande. Para lograr esto Arquímedes calculó explícitamente esos perímetros para los polígonos regulares inscritos y circunscritos de $6 \cdot 2^n$ lados y demostró que su diferencia tiende a cero cuando n tiende a infinito.

Sea l_n el lado del polígono inscrito de $6 \cdot 2^n$ lados y a_n el apotema correspondiente. Entonces, usando el Teorema de Pitágoras, obtenemos



$$a_n = \sqrt{1 - \left(\frac{l_n}{2}\right)^2} \tag{3.2}$$

$$l_{n+1} = \sqrt{\left(\frac{l_n}{2}\right)^2 + (1 - a_n)^2} \tag{3.3}$$

El lado L_n del polígono circunscrito de $6 \cdot 2^n$ lados puede hallarse por triángulos semejantes a partir de l_n y a_n:

$$L_n = \frac{l_n}{a_n} \tag{3.4}$$

Sumando la mitad de los lados de los polígonos, Arquímedes obtiene estas cotas para π que permiten estimar su valor con la precisión que uno quiera

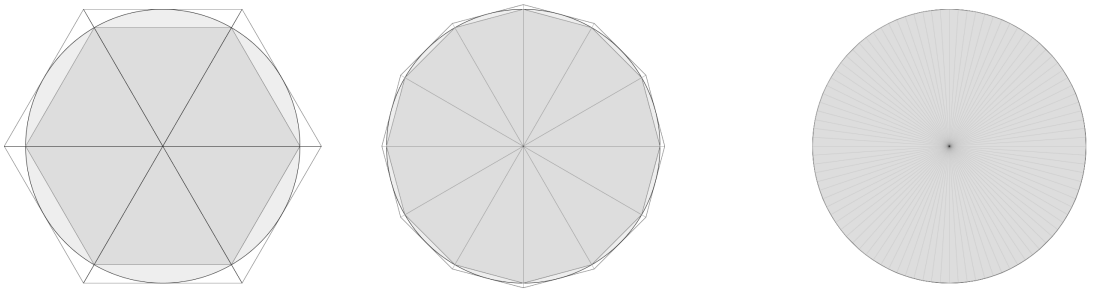
$$\frac{p_n}{2} = 3 \cdot 2^n \cdot l_n < \pi < 3 \cdot 2^n \cdot L_n = \frac{P_n}{2}$$

ya que la diferencia entre estos valores es

$$3 \cdot 2^n (L_n - l_n) = 3 \cdot 2^n \cdot L_n (1 - a_n)$$

que tiende a cero cuando n tiende a infinito porque $3 \cdot 2^n \cdot L_n$ está acotado (por ejemplo por $3 \cdot 2^1 \cdot L_1$ que es el semi-perímetro del hexágono circunscrito) y obviamente a_n tiende a 1 cuando n tiende a infinito.

Las fórmulas recursivas (3.2) y (3.3) involucran raíces cuadradas, que son difíciles de calcular manualmente, y aún con los excelentes algoritmos de nuestro sistema de decimal, hoy no haríamos estos cálculos sin contar con una calculadora. Las raíces cuadradas de números racionales suelen ser irracionales y sólo podemos expresarlas de manera aproximada en nuestro sistema decimal.



6 lados 12 lados $\cdots$ 96 lados

A pesar de estas dificultades, Arquímedes, con algoritmos aritméticos menos eficientes que los nuestros, logró obtener buenas cotas racionales para el número π. Usando los polígonos inscrito y circunscrito de 96 lados obtuvo

$$3 + \frac{10}{71} = \frac{223}{71} < \pi < \frac{220}{70} = 3 + \frac{10}{70}$$

Estas cotas dan 2 cifras decimales de precisión

$$3.1408 < \pi < 3.1429 \tag{3.5}$$

Con una calculadora hoy podemos obtener

$$3.1410 < \pi < 3.1428 \tag{3.6}$$

Observemos que el promedio pesado con $\frac{2}{3}$ de la cota inferior y $\frac{1}{3}$ de la superior es 3.1416, que es una excelente aproximación para π con cuatro cifras decimales de precisión. Esto no es una casualidad, como veremos más adelante.

El método de Arquímedes fue aplicado por Ptolomeo alrededor del año 100, hasta alcanzar un polígono de $1536 = 6 \cdot 2^8$ lados, obteniendo la estimación $\pi \simeq 317/120 \simeq 3.1416$ que es el valor que usamos todavía en las escuelas.

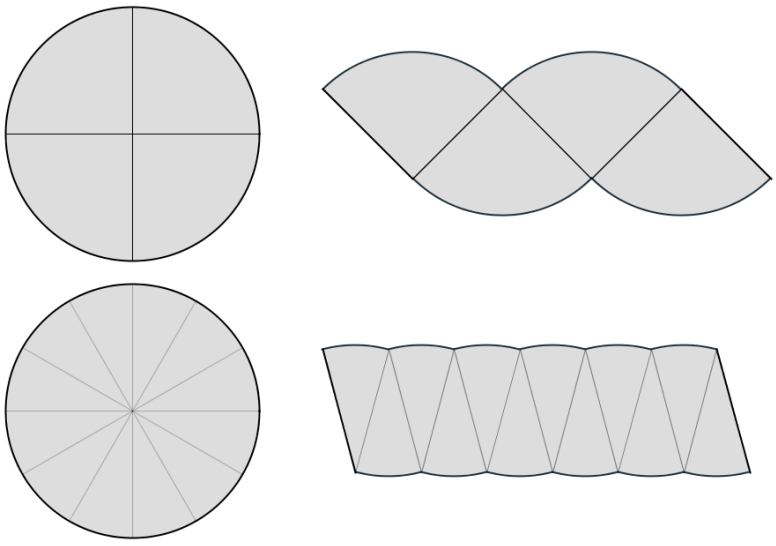
Algunos siglos más tarde y de manera independiente, los chinos usaron el mismo método para obtener una aproximación aún más precisa $\pi \simeq 355/113$, que en decimales corresponde a $\pi \simeq 3.141592$ y es exacta en seis cifras decimales. Este resultado constituye una hazaña extraordinaria de cálculo, dadas las limitaciones aritméticas de aquella época, y se puede considerar que dio a los chinos el récord mundial durante mil años. No fue sino hasta finales de la edad media europea que el uso de los números arábigos permitió mejorar los métodos aritméticos y obtener mejores estimaciones de π. Como veremos más adelante, Arquímedes pudo haber obtenido la misma aproximación que los chinos y con casi las mismas operaciones aritméticas con que logró la suya, que fue mucho más modesta.

El área del círculo

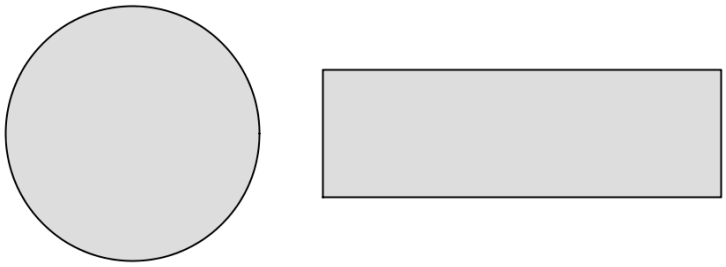
Euclides demostró que la razón entre el perímetro de la circunferencia y su diámetro es constante y también que la razón entre el área del círculo y el cuadrado de su radio es constante. Sin embargo no demostró que ambas constantes son iguales. Quien logra demostrarlo es, para no variar, Arquímedes. Lo hace unos 50 años después de que Euclides escribiera sus *Elementos*. Este resultado es el que nos permite escribir las fórmulas $p = 2\pi r$ para el perímetro de la circunferencia y $A = \pi r^2$ para el área del círculo de radio r.

Las siguientes figuras ilustran el método que usó Arquímedes para obtener el área de un círculo en términos del número π, que de hecho consiste precisamente en encontrar la relación entre el área, el perímetro y el radio:

$$A = \frac{p \cdot r}{2}.$$



Lo que hizo fue dividir el círculo en un número par de sectores y acomodar la mitad de ellos alineados y paralelos, tocándose por un punto, y la otra mitad de la misma manera, pero invertidos, encajados exactamente en los primeros. Luego fue aumentando el número de sectores y observó que la figura que se obtiene, y que por construcción siempre tiene la misma área que el círculo, se va convirtiendo en un rectángulo cuyo lado es la mitad del perímetro y cuya altura es el radio.



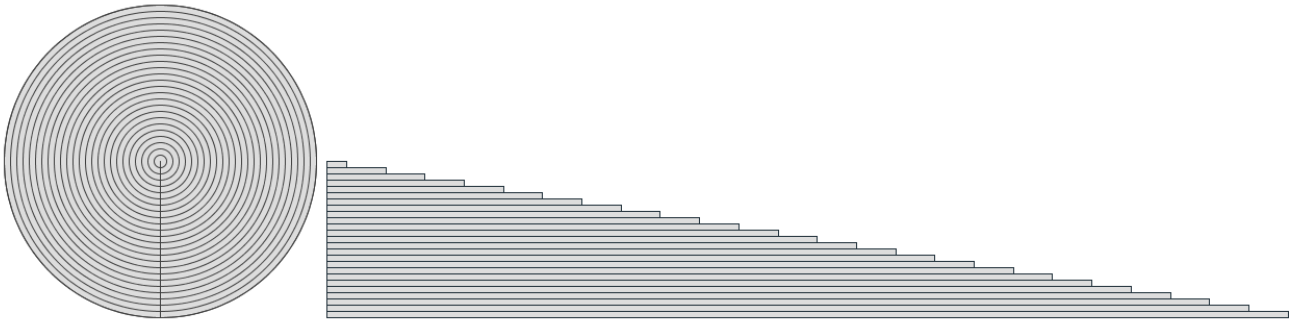
Así, resulta claro que el área del círculo es la mitad del perímetro por el radio. Dado que el perímetro, por definición del número π, es $2\pi r$, esto demuestra que el área del círculo es πr^2. Cabe mencionar que Johannes Kepler (1571 - 1630) utilizó una idea similar para demostrar que el volumen V y la superficie S de una esfera de radio r guardan la relación

$$V = \frac{S \cdot r}{3}.$$

Sugerimos al lector demostrar esta fórmula mediante la idea de partir la superficie de la esfera en muchos pequeños triángulos e interpretarlos como las bases de pirámides con un vértice en el centro de la esfera y cuyas alturas se aproximan al radio de la esfera a medida que los lados de los lados de los triángulos tienden a cero.

Otra manera de demostrar la fórmula para el área del círculo de radio r es partiéndolo en muchos anillos concéntricos del mismo espesor y desenrollarlos, sin estirarlos ni encogerlos, hasta dejarlos rectos y horizontales. Se forma así una figura muy cercana al triángulo de base $2\pi r$ y altura r, cuya área es, por supuesto, πr^2. Cuantos más anillos se usen, mejor es la aproximación.

Así se consigue ontiene demostración de que el área del círculo de radio r es πr^2.



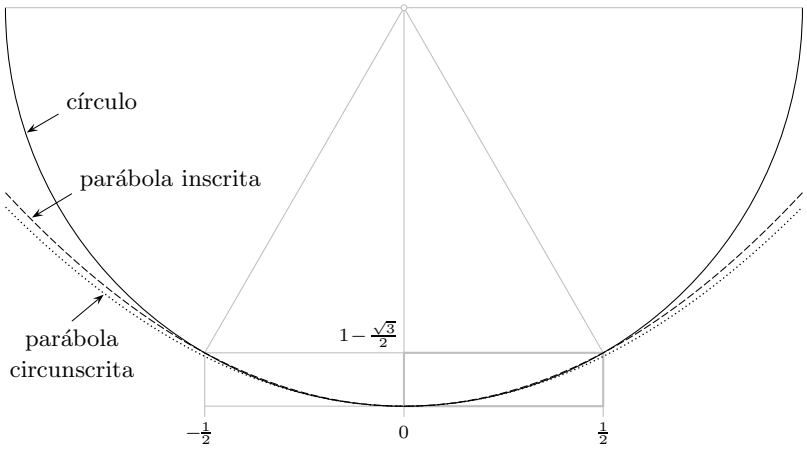
Este argumento aparece en un texto medieval del matemático y astrónomo judío-catalán Abraham Bar Hiyya (o Abraham Iudaeus Savasorda) (Barcelona, siglo XII, 1065-1070 – 1136 ó 1145), precursor de la escuela de Toledo.

«Sefer tzuret, abraham bar»
Photo of Exhibit at the Diaspora Museum, Tel Aviv - en:Beit Hatefutsot.

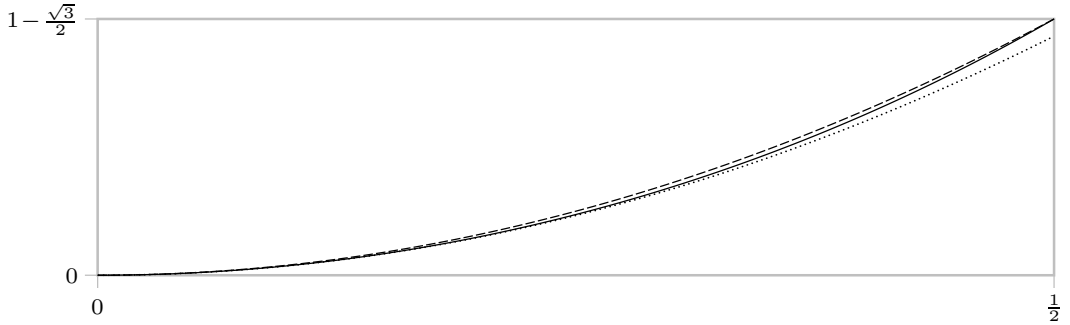
Ya que que π puede obtenerse como la razón del área del círculo al cuadrado de su radio, resulta que es igual al área del círculo unitario (el de radio 1). Así calcularemos π de aquí en adelante.

El número π como el área del círculo unitario

Usando la misma idea de Arquímedes de los polígonos de $6 \cdot 2^n$ lados, pero ahora buscando cotas para sus áreas en lugar de sus perímetros, se llega a estimaciones similares a las que se obtienen con perímetros. Pero usar áreas tiene ventajas. Nos permite aprovechar otro de los importantes resultados de Arquímedes: *la cuadratura de la parábola* (ver 3.4), es decir, que el área de un sector de parábola es igual a $\frac{2}{3}$ de la del paralelogramo o rectángulo más pequeño que la contiene. El promedio pesado con $\frac{2}{3}$ de la cota inferior más $\frac{1}{3}$ de la exterior, que como antes vimos da una estimación muy buena de π, nos sugiere usar parábolas para obtener mejores resultados.



Consideremos dos parábolas con sus ejes sobre la mediatriz del triángulo formado por el centro del círculo y los extremos de uno de los hexágonos inscritos, ambas con vértice en la intersección de la mediatriz con la circunferencia. Sea la primera de ellas la que corta a la circunferencia en los extremos del lado elegido del hexágono y sea la segunda la que tiene el foco en el punto medio del radio que coincide con la mediatriz. Se puede demostrar que la segunda es exterior al círculo y la primera es interior, entre los extremos del lado del hexágono. Las tres curvas coinciden en el vértice de las parábolas.



Se puede obtener una cota inferior para el área del círculo unitario como 6 veces el área del triángulo de base 1 y apotema $a = \sqrt{1-(\frac{1}{2})^2} = \frac{\sqrt{3}}{2}$ más $\frac{2}{3}$ del rectángulo de base 1 y altura $1-a$. O sea que

$$\pi > 6\left(\frac{a}{2} + \frac{2}{3}(1-a)\right) = 6\left(\frac{\sqrt{3}}{4} + \frac{2}{3}\left(1-\frac{\sqrt{3}}{2}\right)\right) > 3.141104 \tag{3.7}$$

Si colocamos el origen en el vértice de las parábolas y hacemos coincidir el eje y con los ejes de las parábolas de manera que el centro del círculo quede en el punto $(0,1)$, la ecuación de nuestra parábola exterior resulta ser $y = \frac{x^2}{2}$. Sumando las áreas del triángulo y el rectángulo y restando la que está bajo la parábola, obtenemos

$$\pi < 6\left(\frac{a}{2} + (1-a) - \frac{1}{3}\left(\frac{0.5^2}{2}\right)\right) = 6\left(\frac{\sqrt{3}}{4} + \left(1-\frac{\sqrt{3}}{2}\right) - \frac{1}{3}\left(\frac{0.5^2}{2}\right)\right) < 3.142603 \tag{3.8}$$

Reuniendo (3.7) y (3.8) llegamos a

$$3.141104 < \pi < 3.142603$$

Esta estimación es ligeramente mejor que la de Arquímedes correspondiente al polígono de 96 lados, pero ésta se obtuvo con muchas menos operaciones aritméticas y, entre ellas, sólo una raíz cuadrada, la de 3.

La misma idea se puede aplicar a los polígonos de $6 \cdot 2^n$ lados aprovechando las fórmulas recursivas (3.2) y (3.3) para a_n y l_n que se usaron en el método de Arquímedes. Con esos valores las cotas para π son

$$6 \cdot 2^n \left(\frac{a_n}{2} + \frac{2}{3} \cdot \frac{l_n^2}{2} \right) < \pi < 6 \cdot 2^n \left(\frac{a_n}{2} + (1 - a_n) - \frac{1}{3} \cdot \frac{l_n^2}{2} \right)$$

En el caso $n = 4$, que corresponde al polígono de 96 lados, esto da 7 dígitos de precisión

$$3.14159264 < \pi < 3.14159267$$

que son 5 más que los 2 que dio el método de Arquímedes.

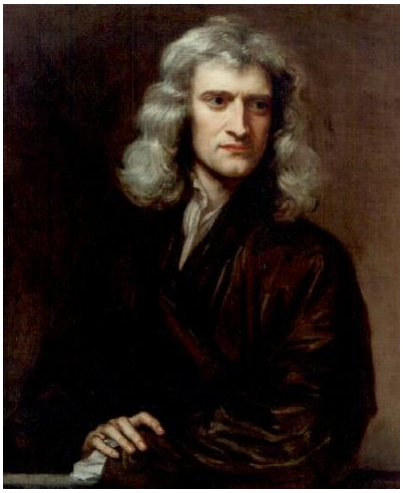
Resulta bastante extraño que Arquímedes no haya usado esta idea, siendo que él mismo obtuvo la cuadratura de la parábola. Tal vez su trabajo sobre la cuadratura de la parábola haya sido posterior al del cálculo de π.

Aunque este camino nos permite calcular muchos decimales de π, tiene el gran inconveniente de las raíces cuadradas, que nos obligan a realizar operaciones complejas con resultados inexactos.

Si se hacen promedios pesados con $\frac{3}{5}$ de las cotas inferiores y $\frac{2}{5}$ de las superiores, se obtienen aproximaciones con un 50 % más de cifras decimales correctas. Esto sugiere que podría haber métodos aún mejores. Uno de ellos lo descubrió Isaac Newton. En la siguiente sección lo redescubriremos nosotros mismos usando argumentos geométricos y del cálculo diferencial e integral.

Cálculo de π según Newton

El gran científico Sir Isaac Newton (1643 - 1727) encontró un método para calcular π que es mucho más eficiente que cualquiera de los que le precedieron.



«Sir Isaac Newton (1643-1727)» de Sir Godfrey Kneller
- http://www.phys.uu.nl/ vgent/astrology/images/newton1689.jpg].
Disponible bajo la licencia Dominio público vía Wikimedia Commons -
https://commons.wikimedia.org/wiki/File:Sir_Isaac_Newton_(1643-1727).jpg
#/media/File:Sir_Isaac_Newton_(1643-1727).jpg

Su resultado puede escribirse así

$$\pi = 3\sum_{n=0}^{\infty} \frac{(2n)!}{16^n \cdot (n!)^2 \cdot (2n+1)} \tag{3.9}$$

Esta fórmula expresa π en términos de una serie rápidamente convergente de números racionales en la que cada término tiene una expresión explícita. De ella podemos obtener un algoritmo recursivo muy sencillo para cada término de la serie a partir del que le precede. Efectivamente, definiendo

$$c_n = \frac{(2n)!}{16^n \cdot (n!)^2}$$

la fórmula (3.9) se puede escribir así

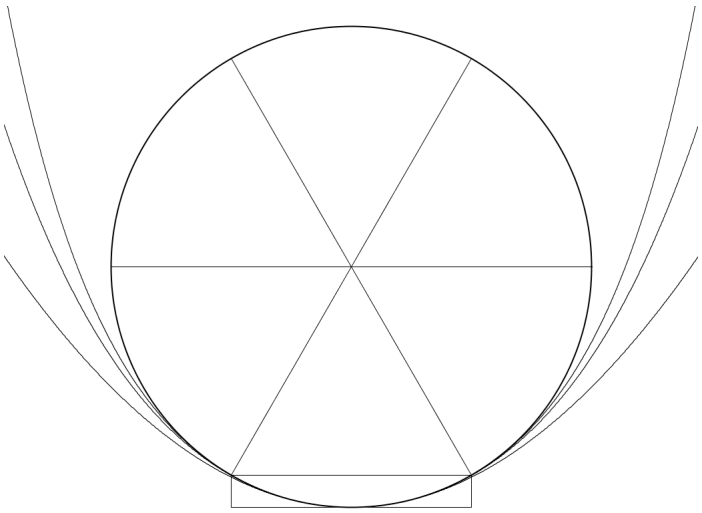
$$\pi = 3\sum_{n=0}^{\infty} \frac{c_n}{2n+1} \tag{3.10}$$

donde $c_0 = 1$ y, para $n = 1, 2, \ldots$,

$$c_n = c_{n-1}\frac{(2n-1)(2n)}{16n^2} = c_{n-1}\frac{(2n-1)}{8n} \tag{3.11}$$

La serie (3.10) brinda una manera sencilla y muy rápida de obtener buenas aproximaciones de π. De hecho, con este método se pueden obtener 16 decimales de π con sólo 22 términos. Manualmente este cálculo puede hacerse en una hora, en minutos usando una calculadora y en fracciones de segundo con una computadora.

Para obtener este resultado calcularemos π como el área del círculo unitario y usaremos el hexágono inscrito y la aproximación a una función mediante una serie de potencias, concepto que surgió apenas en el siglo XVII gracias, entre otros, al propio Newton.



Aproximando $1-\sqrt{1-x^2}$ por $\sum_{n=0}^{N} a_n x^n$, $N = 1, 2, 3, \ldots$

El área puede calcularse así

$$\pi = 6(T + R - 2C) \tag{3.12}$$

donde:
T es el área del triángulo equilátero de lado 1: $T = \frac{\sqrt{1-(\frac{1}{2})^2}}{2}$,
R es el área del rectángulo de base 1 y altura $1 - \sqrt{1-(\frac{1}{2})^2}$ y
C es el área bajo la gráfica[4] de la función $1-\sqrt{1-x^2}$ entre 0 y $\frac{1}{2}$.
Entonces

$$T = \tfrac{1}{2} - \tfrac{1-\sqrt{1-(\frac{1}{2})^2}}{2} \quad R = 1 - \sqrt{1-\left(\tfrac{1}{2}\right)^2} \quad C = \int_0^{\frac{1}{2}} 1 - \sqrt{1-x^2}dx$$

Definiendo $f(x) = 1 - \sqrt{1-x^2}$, estas fórmulas pueden escribirse así

$$T = \tfrac{1}{2} - \tfrac{f(\frac{1}{2})}{2} \quad R = f(\tfrac{1}{2}) \quad C = \int_0^{\frac{1}{2}} f(x)dx$$

Sustituyendo en (3.12) obtenemos

$$\pi = 6\Big(\frac{1}{2} - \frac{f(\frac{1}{2})}{2} + f\Big(\frac{1}{2}\Big) - 2\int_0^{\frac{1}{2}} f(x)dx\Big)$$

que se simplifica a

$$\pi = 3\Big(1 + f\Big(\frac{1}{2}\Big) - 4\int_0^{\frac{1}{2}} f(x)dx\Big) \tag{3.13}$$

La función $f(x)$ se puede escribir como

$$f(x) = \sum_{n=1}^{\infty} a_n x^{2n} \tag{3.14}$$

[4]El origen se colocó una unidad abajo del centro para que la gráfica coincida con la circunferencia.

donde los coeficientes a_n, que más adelante obtendremos usando la serie binomial, son positivos. Sustituyendo (3.14) en (3.13)

$$\pi = 3\Big(1 + \sum_{n=1}^{\infty} a_n \Big(\frac{1}{2}\Big)^{2n} - 4\int_0^{\frac{1}{2}} \sum_{n=1}^{\infty} a_n x^{2n} dx\Big)$$

Calculando la integral término a término

$$\pi = 3\Big(1 + \sum_{n=1}^{\infty} a_n \Big(\frac{1}{2}\Big)^{2n} - 4\sum_{n=1}^{\infty} a_n \frac{(\frac{1}{2})^{2n+1}}{2n+1}\Big)$$

Reuniendo las dos sumas y sacando como factores comunes a $a_n(\frac{1}{2})^{2n}$

$$\pi = 3\Big(1 + \sum_{n=1}^{\infty} a_n \Big(\frac{1}{2}\Big)^{2n} \Big(1 - 4\frac{\frac{1}{2}}{2n+1}\Big)\Big)$$

Lo cual podemos reescribir así

$$\pi = 3\Big(1 + \sum_{n=1}^{\infty} a_n \frac{1}{2^{2n}} \Big(1 - \frac{2}{2n+1}\Big)\Big)$$

Usando $2n+1$ como denominador común, podemos simplificar y escribir

$$\pi = 3\Big(1 + \sum_{n=1}^{\infty} a_n \frac{2n-1}{4^n \cdot (2n+1)}\Big) \tag{3.15}$$

Para obtener los coeficientes a_n desarrollamos $(1-x^2)^{\frac{1}{2}}$ en serie binomial

$$f(x) = 1 - (1-x^2)^{\frac{1}{2}} = 1 - 1 + \frac{1}{2}x^2 + \frac{\frac{1}{2}\cdot\frac{1}{2}}{2!}x^4 + \frac{\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{3}{2}}{3!}x^6 + \ldots$$

Simplificando

$$f(x) = \frac{1}{2}x^2 + \frac{\frac{1}{2}\cdot\frac{1}{2}}{2!}x^4 + \frac{\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{3}{2}}{3!}x^6 + \ldots$$

De lo cual se puede ver que $a_1 = \frac{1}{2}$ y, para $n > 1$

$$a_n = \frac{1 \cdot 3 \cdot \cdots \cdot (2n-3)}{2^n n!} = \frac{1 \cdot 3 \cdot \cdots \cdot (2n-1)}{2^n \cdot n! \cdot (2n-1)}$$

Multiplicando numerador y denominador por $2 \cdot 4 \cdot \cdots \cdot 2n$ obtenemos

$$a_n = \frac{(2n)!}{(2 \cdot 4 \cdot \cdots \cdot 2n) \cdot 2^n \cdot n! \cdot (2n-1)} = \frac{(2n)!}{2^n \cdot n! \cdot 2^n \cdot n! \cdot (2n-1)}$$

y finalmente

$$a_n = \frac{(2n)!}{4^n \cdot (n!)^2 \cdot (2n-1)} \tag{3.16}$$

para $n = 1, 2, \ldots$ Sustituyendo (3.16) en (3.15)

$$\pi = 3\Big(1 + \sum_{n=1}^{\infty} \frac{(2n)!}{4^n \cdot (n!)^2 \cdot (2n-1)} \cdot \frac{(2n-1)}{4^n \cdot (2n+1)}\Big)$$

Lo cual puede escribirse finalmente como

$$\pi = 3\Big(1 + \sum_{n=1}^{\infty} \frac{(2n)!}{16^n \cdot (n!)^2} \cdot \frac{1}{(2n+1)}\Big) \tag{3.17}$$

Como la fórmula para los coeficientes en (3.17) es válida para $n = 0$ y da 1, esto nos lleva a (3.9), que es lo que queríamos demostrar.

Cálculos de π con muchos decimales

Utilizando (3.10) y (3.11) en una computadora con aritmética de alta precisión se puede obtener la siguiente aproximación de π con 600 cifras decimales de precisión.

$$\pi = 3.$$
141592 653589 793238 462643 383279 502884 197169 399375 105820 974944
592307 816406 286208 998628 034825 342117 067982 148086 513282 306647
093844 609550 582231 725359 408128 481117 450284 102701 938521 105559
644622 948954 930381 964428 810975 665933 446128 475648 233786 783165
271201 909145 648566 923460 348610 454326 648213 393607 260249 141273
724587 006606 315588 174881 520920 962829 254091 715364 367892 590360
011330 530548 820466 521384 146951 941511 609433 057270 365759 591953
092186 117381 932611 793105 118548 074462 379962 749567 351885 752724
891227 938183 011949 129833 673362 440656 643086 021394 946395 224737
190702 179860 943702 770539 217176 293176 752384 674818 467669 405132

El cálculo tarda unos 25 segundos.

Para obtener 1,200 decimales se requieren unos 3 minutos, para 7,200 decimales, 4 horas y para 12,000 decimales, 22 horas.

Actualmente se conocen métodos aún más eficientes con los que se han llegado a calcular miles de millones de cifras decimales de π. Esos cálculos tardan meses en realizarse. Hace mucho tiempo había interés en esto porque no se sabía si π era racional y se intentaba descubrir un patrón que se repitiera, en la esperanza de descubrir la fracción que lo representaba exactamente. Pero desde que se descubrió que π no puede ser un número racional (siguiente sección), este tipo de esfuerzos perdió relevancia porque para la mayoría de los cálculos que hay que realizar, basta conocer unos 16 decimales de π, que es lo que ofrece la aritmética de una computadora personal. Para cálculos más delicados, un centenar de decimales serían más que suficientes. Así que en realidad el cálculo de π con miles o millones de cifras decimales no tienen ninguna utilidad práctica, se trata de un mero desafío. Algunos investigadores tienen interés en estudiar la distribución estadística de los dígitos en una expansión larga de π, por ejemplo, en saber si todos aparecen con la misma probabilidad. Con la riqueza de las matemáticas es imposible saber si ese tipo de investigaciones dará alguna vez un resultado interesante.

La irracionalidad de π

Durante siglos los matemáticos se preguntaron si π sería o no racional. Si lo fuera habría una periodicidad en sus decimales. Esto fomentó el interés de obtener muchos decimales de π. A medida que se calculaban más y más se iban buscando patrones que se repitieran. Pero nunca se encontró uno. En el siglo XVIII, el matemático suizo JOHANN HEINRICH LAMBERT dio la primera demostración de que nunca se encontraría tal patrón porque π es un número irracional. A continuación presentamos una demostración de la irracionalidad de π publicada por IVAN NIVEN de Purdue University en 1946, la cual requiere únicamente de conceptos elementales de cálculo.

La demostración procede por contradicción. Supongamos que $\pi = \frac{a}{b}$, donde a y b son enteros positivos. Definamos

$$f(x) = \frac{x^n(a-bx)^n}{n!}$$

y

$$F(x) = f(x) - f^{(2)}(x) + f^{(4)}(x) - ... + (-1)^n f^{(2n)}(x)$$

donde n es un entero positivo cualquiera. Entonces $n!f(x)$ es un polinomio con coeficientes enteros de grado no mayores que n, así que $f(x)$ y sus derivadas $f^{(i)}(x)$ toman valores enteros en $x = 0$ y en $x = \frac{a}{b} = \pi$. Por cálculo elemental sabemos que

$$\frac{d}{dx}\big[F'(x)\operatorname{sen}(x) - F(x)\cos(x)\big] = \big[F''(x) + F(x)\big]\operatorname{sen}(x) = f(x)\operatorname{sen}(x)$$

y por tanto

$$\int_0^\pi f(x)\operatorname{sen}(x)dx = \big[F'(x)\operatorname{sen}(x) - F(x)\cos(x)\big]_0^\pi = F(\pi) + F(0) \tag{3.18}$$

$F(\pi)+F(0)$ es un entero porque $f^{(i)}(\pi)$ y $f^{(i)}(0)$ son enteros para todo i. Pero para $0 < x < \pi$

$$0 < f(x)\operatorname{sen}(x) < \frac{\pi^n a^n}{n!}$$

así que el valor de la integral (3.18) es un entero positivo y sin embargo es menor que $\pi\frac{(\pi a)^n}{n!}$ para todo entero positivo n. Pero $\frac{c^n}{n!}$ tiende a cero cuando n tiende a infinito, lo cual nos lleva a una contradicción. Por lo tanto es imposible que π sea racional.

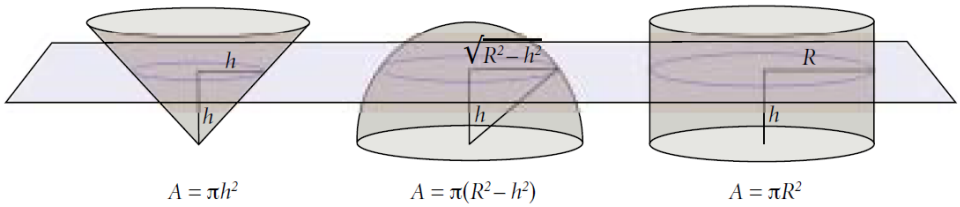
Ahora se sabe que el número π no solo no es racional sino que tampoco es algebraico, es decir, que no es solución de alguna ecuación algebraica con coeficientes racionales. Los números reales que no son algebraicos se llaman trascendentes. La primera demostración de que π es un número trascendente la dio CARL LOUIS FERDINAND VON LINDEMANN en 1882.

El enorme interés que se puede apreciar a través de lo expuesto en esta sección por comprender al número π y los grandes esfuerzos realizados por estimar con gran precisión su valor, se deben a la enorme importancia que ha tenido para el desarrollo de las matemáticas. No sólo el perímetro y el área del círculo están íntimamente relacionados con el número π sino también el volumen y la superficie de la esfera guardan una íntima relación con este número, como demostró el mismo Arquímedes.

3.6. El volumen y la superficie de la esfera

Arquímedes obtuvo el volumen y la superficie de una esfera y demostró que el volumen de la esfera es igual a 2/3 del volumen del mínimo cilindro que la contiene. También demostró que la superficie de la esfera es igual a 2/3 de la superficie de ese mismo cilindro. Estaba tan orgulloso de este resultado que pidió se inscribiera en su tumba. Gracias a que aparente se cumplió su deseo, fue posible localizar su tumba, que estuvo perdida durante varios siglos.

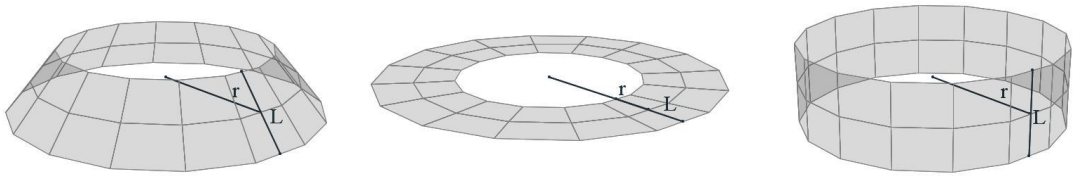
El cálculo del volumen de la esfera utiliza el Teorema de Pitágoras de una manera muy ingeniosa. Consideremos tres cuerpos: media esfera, un cono circular recto y un cilindro circular. Las bases del cono y el cilindro tiene el mismo radio que la esfera ambos tiene altura también igual al radio. Lo que Arquímedes demuestra es que el volumen de la semiesfera más el del cono es igual al del cilindro. Como ya sabía que el volumen del cono es $\frac{1}{3}$ del del cilindro, resulta que el de la semiesfera debe ser $\frac{2}{3}$ del del cilindro.



Sea R el radio común de la semiesfera y de las bases del cono y el cilindro. Coloquemos los tres cuerpos apoyados en un plano horizontal, con los ejes del cono y el cilindro perpendiculares al plano el vértice del cono a la altura de la base de la semiesfera.

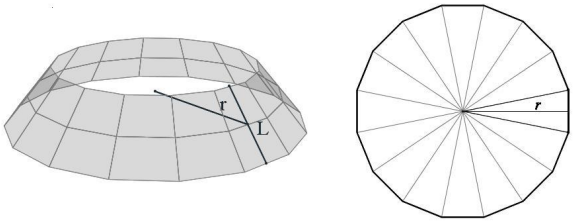
Ahora cortamos los tres cuerpos por un plano horizontal a una altura h de la base de la semiesfera. Los cortes del plano en el cono y en la semiesfera son círculos de radios h y $\sqrt{R^2-h^2}$ respectivamente. Por tanto la suma de sus áreas es $\pi h^2 + \pi(R^2 - h^2) = \pi R^2$, que es el área del corte en el cilindro. Como esto ocurre para todos los planos horizontales, lo que queríamos probar resulta evidente. Esta idea fue utilizada sistemáticamente por Bonaventura Cavalieri (1598 - 1647) calcular los volúmenes de otros cuerpos a partir de las áreas de sus secciones paralelas a un plano, por lo que en la actualidad se reconoce como el *Principio de Cavalieri*.

El cálculo de la superficie de la esfera depende de la siguiente observación. Consideraremos sectores de pirámides cuyas bases son polígonos regulares. Entonces sus caras son trapecios iguales. Todos los sectores de piramidales que tienen el mismo perímetro medio y la misma longitud lateral L, tienen la misma superficie.

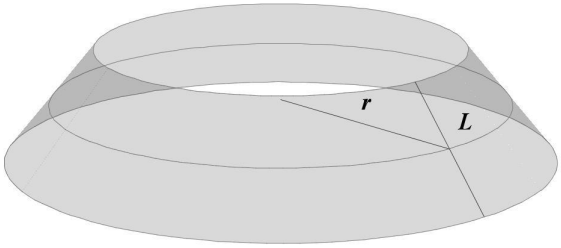


Esto se debe a que el área de los trapecios que los forman depende sólo de L y de la anchura media. Por tanto el área del sector piramidal depende sólo de la longitud del polígono cuyos vértices son

los puntos medios de las aristas laterales. De hecho el área del sector de pirámide es PL, donde P es el perímetro del polígono.



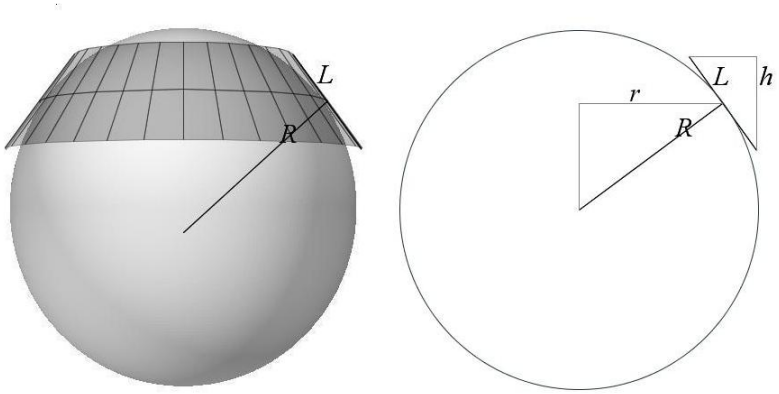
Como esto sucede independientemente del número de lados que tengan las orillas del sector piramidal, lo mismo va a ocurrir con los sectores cónicos.



Los que tengan el mismo radio medio r y el mismo tamaño de sus elementos laterales L, tendrán la misma superficie: $2\pi rL$.

Para calcular la superficie de la esfera Arquímedes la cubre con N sectores de cono como el de la figura. La superficie de la cubierta será igual a

$$\sum_{n=1}^{N} 2\pi r_n L_n$$



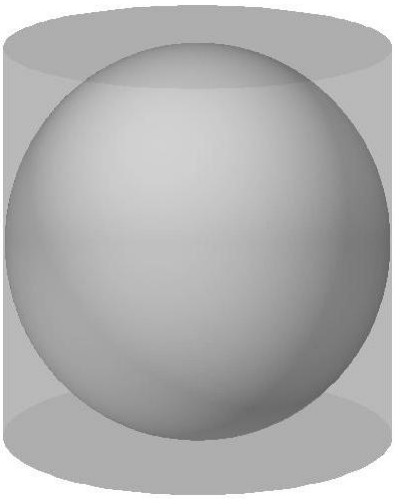
Luego observa que, por semejanza de triángulos:

$$\sum_{n=1}^{N} 2\pi r_n L_n = \sum_{n=1}^{N} 2\pi R h_n = 2\pi R \sum_{n=1}^{N} h_n = 2\pi RH$$

donde H es la altura total de la cubierta. A medida que la cubierta se hace más fina, H tiende a $2R$ y por lo tanto, la superficie de la esfera debe ser

$$4\pi R^2$$

En realidad el argumento de Arquímedes es más riguroso, acota la superficie de la esfera por una sucesión que converge a $4\pi R^2$ desde arriba y por otra que converge al mismo valor desde abajo. Su método satisface los criterios de rigor más estrictos.



La superficie del cilindro circunscrito a la esfera es la del rectángulo cuyos lados son el perímetro $2\pi R$ y la altura $2R$, más las de las dos tapas, que son círculos de radio R:

$$2\pi R \cdot 2R + 2\pi R^2 = 6\pi R^2$$

Por lo tanto la razón entre la superficie de la esfera y la del cilindro circunscrito a ella es $\frac{2}{3}$, la misma que entre el volumen de la esfera y el del cilindro circunscrito. Éste es el resultado del que Arquímedes estaba más orgulloso, por lo que pidió se inscribiese en su tumba.

3.7. Cálculos de distancias celestes

Ideas sobre el cosmos

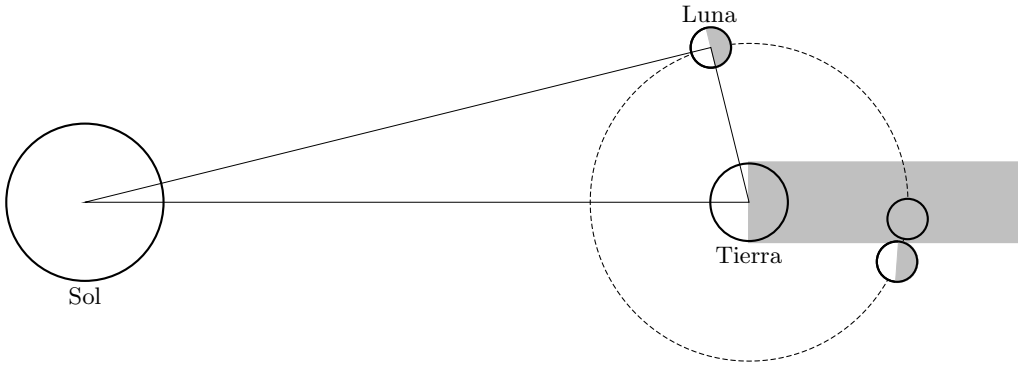
Para poder medir distancias entre objetos celestes es necesario tener una concepción de estos como cuerpos físicos. Mientras se piensa que el Sol es una rueda de fuega que es jalado por una yunta

uno no puede pensar en medir la distancia hacia el Sol. El requisito es entonces que el pensamiento se haya liberado en cierta medida de las explicaciones mísiticas. Esto inció con Tales de Mileto (c. 625–547 a.C.), quién – por ejemplo – trató de explicar las inundaciones anuales del Nilo por vientos fuertos. Una explicación que ciertamente es equivocada, pero trata de explicar un fenómeno físico, y en esto se diferencía mucho de los demás explicaciones en aquel entonces, por razones de la misma naturaleza. Anaximandro (c. 610–547 a.C.), considerado como "alumno" de Tales, dió una explicación diferente para estas inundaciones que era igualmente equivocada. Este pequeño ejemplo muestra como empezó una era diferente: se busca explicar los fenómenos de la naturaleza por la misma naturaleza. Con ello se puede especular y discutir sobre las razones de los fenómenos y se empieza dudar de las explicaciones dogmáticas basadas en creencias.

Para retornar al tema de esta sección: para poder medir distancias entre objetos celestes, es necesario tener primero una concepción natural del cosmos. La palabra "cosmos" origina de la palabra griega para "adornar" ya que se pensaba que la bóveda celeste estaba adornada con el Sol, la Luna, las estrellas y los planetas.

Tales pensaba que la Tierra flota sobre agua, sin decirnos sobre qué se apoyaba el agua. Los hindues a cambio opinaban que la Tierra reposaba sobre una tortuga, esta sobre un elefante, y este último sobre una hoja de Lotus. Nos podemos preguntar en qué se aboyaba la hoja. Anaximandro fue el primero que resolvió el "problema de soporte", es decir la cuestión, qué soportaba la Tierra. Anaximandro pensó que la Tierra flotaba en el aire, aunque la pensaba todavía en forma de un cilindro. Fue Pitágoras (ca. 569–475 a.e.c.) el primero que pensó la Tierra como una esfera. Anaxágoras, él que afirmó que la Luna y el Sol eran rocas esféricos y que la Luna reflejaba la luz del Sol. Aristóteles dió varios argumentos en favor de la forma esférica de la Tierra misma. Dos de ellos son: que más en el Norte o en el Sur se podían ver nuevos astros, que en eclipses lunares, cuando la Luna entra a la sombra que proyecta la Tierra, la sombra siempre se delimita por una arco de circumferencia. La sombra de la Tierra es redonda, por lo que se concluye que la Tierra misma lo es también.

El primero que desarrolló un sistema celeste, que colocaba el Sol en el centro fue Aristarco de Samos (c. 310–230 a.C.). En este sistema se pensó a la Tierra y a los cinco planetas en aquel entonces conocidos: Mercurio, Venus, Marte, Jupiter y Saturno como esferas que giran alrededor del Sol. Veremos en esta sección primero las ingeniosas ideas de Aristarco: cómo estimó los tamaños relativos entre la Tierra, la Luna y el Sol y las distancias entre ellos. Sus estimaciones daban que el Sol era varias veces más grande que la Tierra y por ello colocó a él en el centro del universo y no la Tierra. Eso fue son casi dos mil años antes de Nicolás Copérnico (1473–1543), que acredita de manera justa el sistema heliocéntrico a Aristarco.



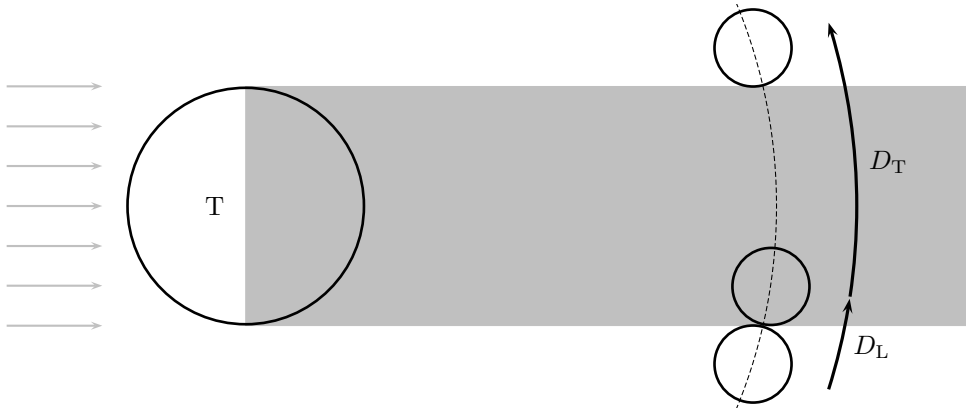
Después veremos la idea de Eratóstenes de Cyrene (c. 275.–194 a.C.), de medir el tamaño de la Tierra y finalmente una idea de Galileo Galilei (1564–1642) para medir la altura de las montañas en la Luna.

Todas estas ideas tienen varios aspectos en conjunto: primero, se basan en mediciones que se hicieron en una sola posición de la Tierra, segundo, se mide una distancia inalcanzable de manera indirecta y tercero, las ideas son muy sencillos pero no por eso menos ingeniosos.

Tamaño relativo entre la Tierra y la Luna

Aristarco comparó el diámetro de la Tierra D_{T} con el diámetro de la Luna D_{L} observando un eclipse lunar. En un eclipse lunar la Tierra se encuentra en medio del Sol y la Luna y esta última pasa en su órbita la sombre que proyecta la Tierra. En esta posición se observa una Luna llena y durante la ecplise la aparación de un disco lleno empieza reducirse poco a poco hasta que la Luna quede casi totalmente oscurecida.

Aristarco comparó dos tiempos: por un lado, el tiempo t_{L} que tardaba la Luna desde el primer momento que empieza oscurecerse hasta que que quede totalmente oscurecido. Por otro lado observó el tiempo t_{T}, que tardó en volverse como luna llena completo despuúes de haberse oscurecido por completo. La situación se muestra esquemáticamente en la siguiente figura.



Como la Luna se mueve más o menos con la misma velocidad todo el tiempo, se tiene que

$$\frac{D_{\mathrm{T}}}{D_{\mathrm{L}}} \approx \frac{t_{\mathrm{T}}}{t_{\mathrm{L}}}.$$

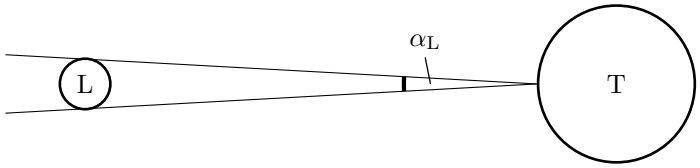
Aristarco estimó o midió que $\frac{t_{\mathrm{T}}}{t_{\mathrm{L}}} \approx 0.36$. En aquel entonces no había herramientas de medición muy precisas, lo que tal vez explica la diferencia con el valor que hoy se tiene: $\frac{t_{\mathrm{T}}}{t_{\mathrm{L}}} \approx 0.27$.

Aristarco obtuvo entonces la estimación

$$D_{\mathrm{L}} \approx 0.36 D_{\mathrm{T}}.$$

Tamaño relativo entre el diámetro de la Luna y la distancia hacia ella

El Sol y la Luna se ven desde la Tierra bajo un cierto *diámetro angular*. Ambos cuerpos celestes, el Sol y la Luna, tienen aproximadamente el mismo diámetro angular, pero sólo aproximadamente, ya que ambos varian ligeramente durante el año, ya que la distancia a la Tierra varía.



Aristarco midió un diámetro angular de ambos cuerpos celestes de 2°. En realidad, el diámetro angular es aproximadamente 0.53°. La medición se puede hacer con una moneda chica: se tapa el Sol, o lo que para el ojo es más agradable, la Luna con la moneda con el brazo extendido y se le pide a otra persona que mida la distancia a de la moneda al ojo. Luego se mide el diámetro de la moneda d. Por semejanza se tiene

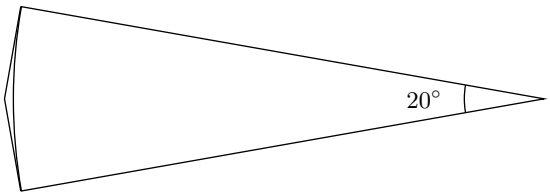
$$d : a = D_{\mathrm{L}} : A_{\mathrm{TL}},$$

donde A_{TL} denota la distancia de la Tierra a la Luna.

Si se toma la medición de Aristarco, entonces se tiene

$$\frac{D_{\mathrm{L}}}{A_{\mathrm{TL}}} \approx \frac{2^\circ}{360^\circ} 2\pi,$$

en buena aproximación, ya que la longitud del arco es casi igual al diámetro para ángulos chicos, como se muestra en la siguiente ilustración con un ángulo mucho más grande: 20°.



Dado que el diámetro angular del Sol y la Luna son casi iguales se tiene

$$\frac{A_{\mathrm{TL}}}{A_{\mathrm{TS}}} \approx \frac{D_{\mathrm{L}}}{D_{\mathrm{S}}},$$

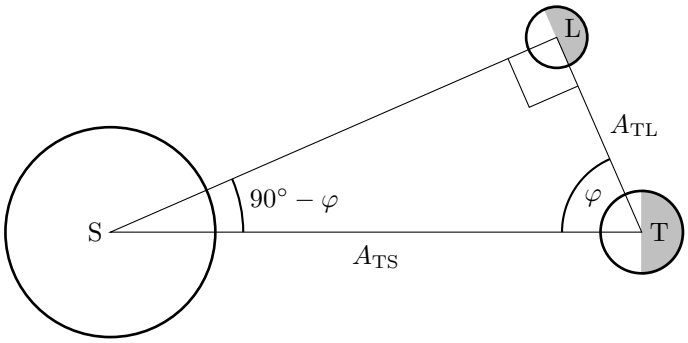
donde D_{S} denota el diámetro del Sol y A_{TS} la distancia de la Tierra al Sol.

Si se toma a $\pi \approx 3$, entonces se obtiene con las mediciones de Aristarco las siguientes estimaciones:

$$A_{\mathrm{TL}} \approx \frac{360^\circ}{2^\circ \cdot 2\pi} D_{\mathrm{L}} \approx 30 \cdot D_{\mathrm{L}} \approx 10.8 \cdot D_{\mathrm{T}},$$
$$A_{\mathrm{TS}} \approx \frac{360^\circ}{2^\circ \cdot 2\pi} D_{\mathrm{S}} \approx 30 \cdot D_{\mathrm{S}}.$$

Tamaño relativo entre la distancia hacia el Sol y la Luna

Para poder relacionar todas las distancias, Aristarco requirió una idea más. Para ello se imaginó la constelación de Tierra, Luna y Sol cuando es media Luna. En esta situación se ve la Luna dividida en dos mitades: una iluminada y otra oscura, luz y sombra, dividida por una línea circular. El plano que contiene esta línea circular pasa por el ojo del observador en la Tierra. El triángulo LTS formada por los tres cuerpos, tiene en la Luna L un ángulo recto. En este momento, Aristarco midió el ángulo



La idea de Aristarco es muy buena, pero a la vez también muy difícil de llevar a cabo con precisión. Medir un ángulo chico es bastante fácil, pero medir un ángulo muy cercano a 90° es bastante difícil si no se cuenta con herramientas de medición de precisión. Además es difícil determinar el momento preciso cuando sucede que el ángulo SLM es exactamente un ángulo recto. Aristarco midió un ángulo de $\varphi \approx 87^\circ$, cuando en realidad es $\varphi = 89.95^\circ$. Eso condujo a que el Sol y la distancia hacia ella se estimó muy por debajo del valor real.

Con la aproximación de ángulos chicos se tiene ahora

$$\frac{A_{\mathrm{TL}}}{A_{\mathrm{TS}}} \approx \frac{90^\circ - \varphi}{360^\circ} 2\pi.$$

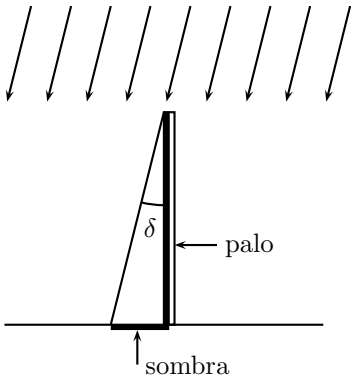
Con la medición de Aristarco se obtiene entonces

$$A_{\mathrm{TS}} \approx 20 \cdot A_{\mathrm{TL}} \approx 216 \cdot D_{\mathrm{T}}$$
$$D_{\mathrm{S}} \approx \frac{1}{30} \cdot A_{\mathrm{TS}} \approx 7.2 \cdot D_{\mathrm{T}}$$

Se ve que las relaciones están bien observados: el Sol es más grande que la Tierra, la Luna más chica, y el Sol más lejos que la Luna. Eso explica por qué ARISTARCO colocó en su sitsema el Sol en el centro del universo. Si ARISTARCO hubiera tenido unas herramientas e medición, entonces habría obtenido estimaciones muy buenas con las sencillas, pero igeniosas ideas de relacionar los tamaños. Con ello concluimos ahora lo que se refiere al trabajo de ARISTARCO.

El tamaño de la Tierra

La primera medición del tamaño de la Tierra se atribuye a ERATÓSTENES (276–194 a. C.). Escuchó que el día del solsticio de verano un palo vertical en Siena (una ciudad que hoy se llama Asuan, en Egipto) no deja sombra, es decir el Sol se encuentra justo en el Cenit. Siena se encuentra casi exactamente en el Sur de Alejandría, lugar en el cual se encontraba ERATÓSTENES. Él midió entonces el ángulo δ entre un palo vertical y la línea recta entre la punta del palo y su sombra, es decir el ángulo de los rayos de Sol con la vertical en Alejandría.



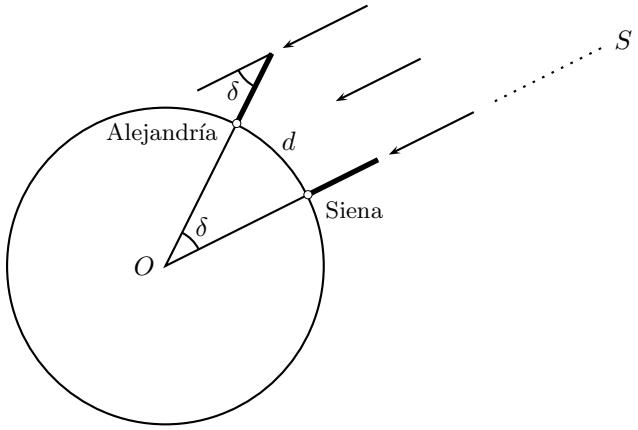
ERATÓSTENES midió que el ángulo δ era $\frac{1}{50}$ de un ángulo completo, es decir

$$\delta = \frac{1}{50} 360^\circ = 7.2^\circ.$$

Si se conoce la distancia d entre Alejandría y Siena, entonces se puede calcular la circunferencia U_T de la Tierra por

$$\frac{\delta}{360^\circ} = \frac{d}{U_\mathrm{T}}.$$

Esto lo explica la siguiente figura.



Para estimar la distancia entre Alejandría y Siena, Eratóstenes se basó en los relatos de los comerciantes que contaban que recorrían la distancia en caravanas en 50 días y estimaban que se trasladaban como 100 "estadios" cada día. El "estadio" es una medida antigua y hoy no se sabe bien cuánto valía, pero se piensa que debería ser entre 157 m y 211 m.

La distancia entre Alejandría y Siena se estimaba entonces en 5000 estadios y por ello la circunferencia terrestre por

$$U_{\mathrm{T}} = 50 \cdot 5000\,\text{estadios} = 250\,000\,\text{estadios},$$

es decir entre 39 250 km y 52 750 km, una estimación muy buena si uno piensa en la sencillez de la medición.

Más tarde se perdió el conocimiento de Eratóstenes y Cristobal Colón emprendió su viaje hacia la India pensando que la Tierra era mucho más chica. Si hubiera conocido la medición de Eratóstenes, tal vez nunca hubiera zarpado. Se volvió la imaginación nuevamente más antropocentrista y con ello geocentrista.

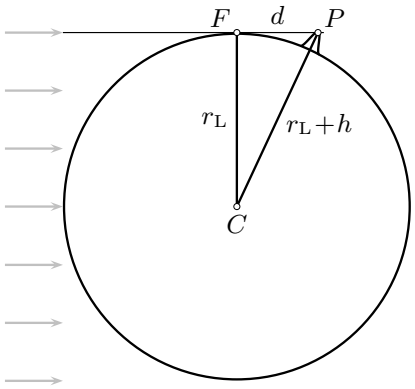
La altura de las montañas en la Luna

Muy influyente para la Edad Media eran los libros de Aristóteles. En ciertas asuntos, él se basaba en Platón. Así, al menos en parte, en su cosmología. Según una especulación de Platón, la matería está compuesta por triángulos que forman sólidos regulares. Así se explicó la existencia de los cuatro elementos y el quinto sólido (el dodecaedro) lo atribuyeron al cosmos, a la composición de las estrellas, la Luna, el Sol y los planetas.

Por ello, cuenado Galileo Galilei usa por primera vez, en enero de 1610, de observar el cielo nocturno con un telescopio sus descubrimientos chocan con toda esta teoría, ya que observó que Jupiter tenía cuatro lunas, que la vía láctea estaba compuesta por miles de estrellas, que el Sol tenía manchas y que la superficie de la Luna no estaba lisa sino presentaba irregularidades, ver las siguientes dos ilustraciones. La izquierda es de Galilei misma, la derecha es moderna.

Galilei interpretó estas irregularidades como las sombras que proyectan montañas. Midió la altura de estas montañas al observar la media luna de la siguiente manera: Los rayos de Sol iluminan los picos más altas aunque las valles ya se oscurcieron por completo. Por ello se puede observar los picos más altos más alla que el diámetro. En la ilustración arriba se señaló el diámetro y los picos más altos iluminados del lado oscuro. Galilei estimó que la distancia de los picos que todavía se iluminan distan un ventésimo del diámetro de la Luna.

La siguiente lisutración muestra un esquema de la situación y explica el cálculo correspondiente.



Con la estimación de la distancia $d \approx \frac{1}{10} r_L$ se puede calcular la hipotenusa del triángulo rectángulo CPF:

$$r_L + h = \sqrt{r^2 + d^2} \approx \sqrt{r^2 + \frac{1}{100} r^2} = \sqrt{1.01} r \approx 1.005 r.$$

Por ello se obtiene

$$h = 0.005 r,$$

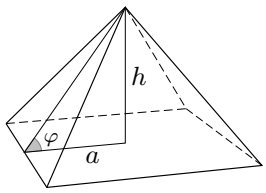
lo que equivale a una altura de $h = 8268\,\mathrm{m}$. Galilei afirma que en la Tierra no hay montaña de semejante altura, pero esto hay que atribuir al hecho que las mediciones de la Tierra apenas se empiezan a hacer unos 100 años después y el Himalaya se midió a mediados del siglo XIX. Actualmente se sabe que las montañas más altas en la Luna se erigen unos $5500m$ sobre la superficie. El cálculo realizado por Galileo dio un valor correcto en cuanto al orden de magnitud.

3.8. La trigonometría

Antecedentes

Con *trigonometría* nos referimos hoy día al cálculo de longitudes y ángulos en triángulos. La etimología es clara: *gon* significa en griego "ángulo", *tri* significa "tres" y *metría* significa "medición", así que el *trigon* es el triángulo, y la trigonometría es la medición de los triángulos. Pero la palabra es una creación de la edad media: aparece por primera vez en un libro de 1590.

Los inicios de lo que hoy llamamos trigonometría son muy antiguos, provienen de los egipcios: ellos calculaban el *sequed* de pirámides, un número ínimamente relacionado con nuestra noción de *pendiente*: el sequed es el número de palmas (anchos de mano) horizontales para subir 7 palmas verticales.



En la figura anterior, el sequed de la pirámide es $s = \dfrac{7a}{h}$.

Hoy podemos ver el sequed –si ignoramos el factor 7 en la definición– como la cotangente del ángulo φ entre los planos que forman un costado de la pirámide y su base. Pero esto es una interpretación moderna y da una falsa impresión. En realidad el *sequed* sólo utiliza la semejanza de triángulos, no relaciona los catetos con el ángulo, como lo hacen las modernas funciones trigonométricas.

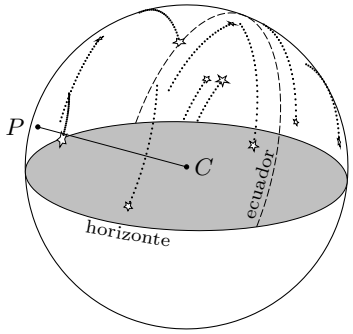
La trigonometría, como muchas áreas de las matemáticas se desarrolló para poder llevar a cabo ciertos cálculos con precisión y eficacia. En este caso, el promotor principal del desarrollo fue la astronomía. Se pensaba que los objetos celestes se movían en círculos alrededor de la Tierra y los astrónomos querían poder calcular sus posiciones y hacer predicciones de los futuros eclipses. Para ello era necesario poder relacionar los ángulos con las proporciones entre los lados correspondientes de los triángulos.

Trigonometría esférica

La trigonometría que conocemos hoy es la trigonometría plana. Es interesante saber que su origen está en la trigonometría esférica, que se desarrolló mucho antes. Los griegos como muchas culturas observaban el cielo y crearon diferentes sistema para explicar los movimientos de los cuerpos celestes.

Al observar el cielo nocturno desde un llano, uno se percata de que las estrellas se mueven respecto a los árboles o las montañas. Nuestra vista nos muestra una bóveda celeste delmitida por el horizonte.

No es difícil observar una regularidad en este movimiento: las estrellas se mueven como se estuvieran pegadas sobre una esfera gigante, el *orbe celeste*, que gira lentamente por un eje que pasa por el punto C en donde está el observador.

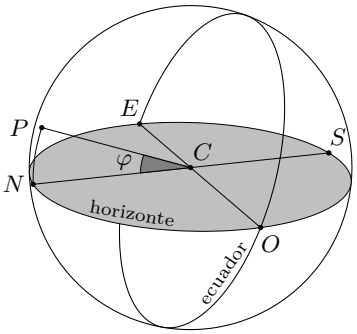


El movimiento es de 15° a la hora, por lo que el orbe celeste da una vuelta completa al día.

Hay una dirección que queda fija: el *polo celeste* P, que es el punto del orbe celeste por el cual este gira. El plano perpendicular al eje de rotación intersecta el orbe en el *ecuador celeste*.

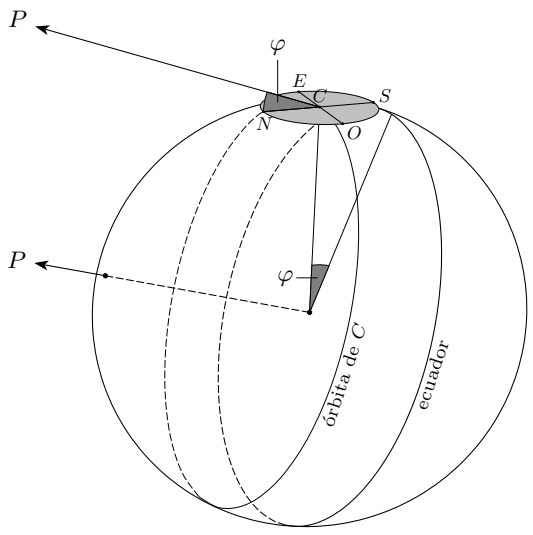
Una estrella cerca del polo celeste casi no se mueve. En el hemisferio norte hay una estrella muy cercana a este punto ideal fijo: la *estrella polar*. En el hemisferio sur no hay una estrella muy luminosa cerca, pero una constelación, llamada *Crux* (o Cruz del Sur) que permite la determinación del polo celeste visible con facilidad.

Dado que el polo magnético de la Tierra se encuentra muy cercano al polo norte de la rotación de la Tierra, la estrella polar en el hemisferio norte sirve para ubicarse tanto como una brújula.



El orbe celeste gira de tal manera que las nuevas estrellas aparecen en el este y otras desaparacen en el oeste.

Si cambiamos en nuestro pensamiento el punto de vista y pensamos que el orbe celeste es fijo y la Tierra se rota por su eje, entonces hay una explicación fácil para el ángulo λ del polo celeste P sobre el horizonte: $\varphi = \sphericalangle PCN$.



El ángulo φ es la *latitud* del punto C donde se encuentra el observador, es decir el ángulo entre la recta CZ, donde Z es el centro de la Tierra y el plano que contiene el ecuador terrestre.

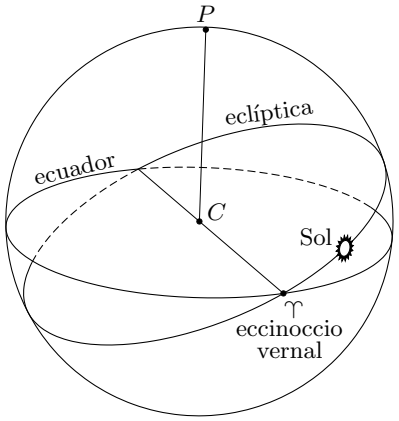
La figura lateral muestra la situación. La esfera es ahora la Tierra y el disco gris es el horizonte con los cuatro puntos cardinales. Dado que las estrellas están muy lejos, el eje de rotación de la Tierra es en buena medida paralela a la recta que une el espectador con el polo celeste o la estrella polar.

En el primer paso de simplificar la situación, se puede uno imaginar de estar en el polo norte. Ahí las estrellas siempre están a la misma altura. El horizonte coincide con el ecuador celeste. Dicho de

otro modo: podemos así ignorar la rotación de la Tierra sobre su propio eje.

Ahora veremos el movimiento aparente del Sol. Si el Sol fuera opaco, podríamos ver las estrellas a un lado del Sol, al igual que lo podemos cuando está la Luna. Dado que el Sol es mucho más luminoso que las estrellas, no podemos verlas de día. Pero podemos calcular la posición del Sol en el orbe celeste.

Resulta que el Sol se mueve muy despacio por el orbe celeste. Un poco menos que 1° cada día. Tarda un año en completar una vuelta. Claro, el origen de este movimiento aparente es que la Tierra se mueve alrededor del Sol y por ello cambian las estrellas, que están muchísmo más lejos, más allá del Sol.

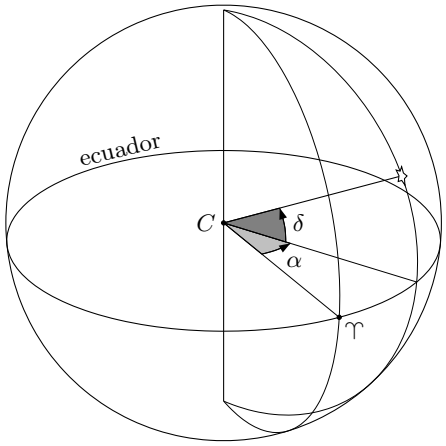


El Sol sigue una círculo máximo que se llama *eclíptica* y que está inclinado por 23.44° respecto al ecuador. Recordemos: el ecuador resulta por el movimiento de la Tierra por su propio eje, la eclíptica por el movimiento de la Tierra alrededor del Sol.

El eje de rotación de la Tierra está inclinado por 23.44° respecto al plano en donde se mueve la Tierra alrededor del Sol.

Los dos planos, el que contiene el ecuador y el que contiene la eclíptica se intersectan en una recta. Cuando el Sol está en uno de los dos puntos de esta recta, en la Tierra ocurren los *equinoccios* una vez en primavera, el Sol está entonces en el punto ♈, y otra vez en otoño. En cada uno de los dos equinoccios, el día y la noche son igual de largos en toda la tierra. Dado que el eje de rotación está inclinada respecto a la eclíptica, la Tierra expone medio año más el hemisferio norte y la otra mitad más el hemisferios sur. Esta desigualdad produce las estaciones en el año.

La ubicación de las estrellas en el orbe celeste se suele dar con respecto al ecuador. Se dan dos coordenadas: la declinación δ y la ascención recta α.



La *declinación* δ es el ángulo que la estrella se eleva sobre el ecuador celeste, visto desde el centro C. Para estrellas en el hemisferio norte y sur la latitud es positiva resp. negativa.

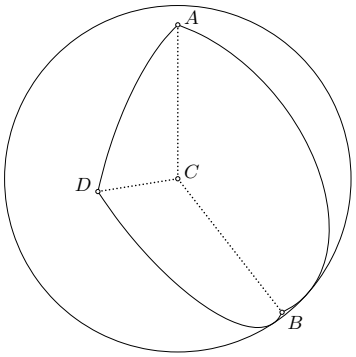
La *ascención recta* α es el ángulo que se mide a partir del equinoccio de primavera ♈ en sentido matemático, es decir en contra de las manecillas del reloj si vemos el orbe desde arriba.

Las dos coordenadas, declinación y ascención recta, conforman las coordenadas ecuatoriales.

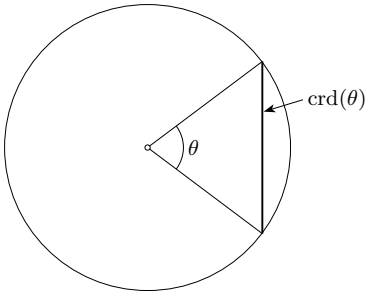
En nuestro sistema solar, todos los planetas se mueven en planos que tienen poca inclinación respecto al plano en el que se mueve la Tierra. Por ello, los planetas se ven desde la Tierra cercanos a la eclíptica, en una cinta que va desde 5° abajo a 5° arriba. Esta cinta alrededor de la eclíptica se llama *Zodiaco*.

Para dar la posición de un planeta se dan por ello coordenadas con respecto a la eclíptica. Estas coordenadas se llaman *latitud* y *longitud*. Calcular de unas coordenadas a las otras constituye uno de los principales problemas de la trigonometría esférica. Veamos un problema sencillo: la longitud del Sol indica el día del año en el cual estamos. Se debe calcular la declinación y la ascención recta del Sol a partir de su longitud λ.

El primero que escribió un libro para resolver estos problemas es Menelao de Alejandría (ca. 70–130). De su vida se sabe poco. Ptolomeo registró observaciones que hizo Menelao en Roma en el año 98. Escribió el libro *Sphaerica* en el que introduce el concepto de triángulo esférico.



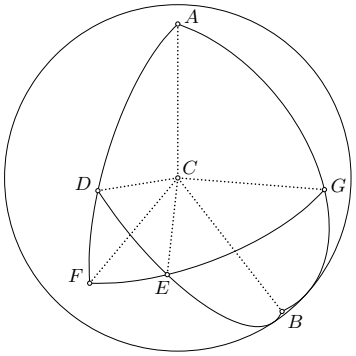
Los lados de un triángulo esférico son arcos de círculos máximos. La figura muestra un triángulo esférico ABD. El punto C es el centro de la esfera.



Para entender lo que ahora se llama el Teorema de Menelao, hay que saber que los griegos calculaban las cuerdas en función del ángulo que las subtiende.

$\mathrm{crd}(\theta)$ denota la cuerda que se ve desde el centro del círculo bajo un ángulo θ.

En el tercer tomo de su libro demuestra una proposición, que en la actualidad se conoce como el Teorema de Menelao.



Se intersecta el triángulo con un arco de círculo máximo y obtiene así los puntos E sobre BD, F sobre AD y G sobre AB.

El Teorema de Menelao dice lo siguiente:

$$\frac{\mathrm{crd}(2\sphericalangle ACG)}{\mathrm{crd}(2\sphericalangle GCB)} = \frac{\mathrm{crd}(2\sphericalangle ACF)}{\mathrm{crd}(2\sphericalangle FCD)} \cdot \frac{\mathrm{crd}(2\sphericalangle DCE)}{\mathrm{crd}(2\sphericalangle ECB)}.$$

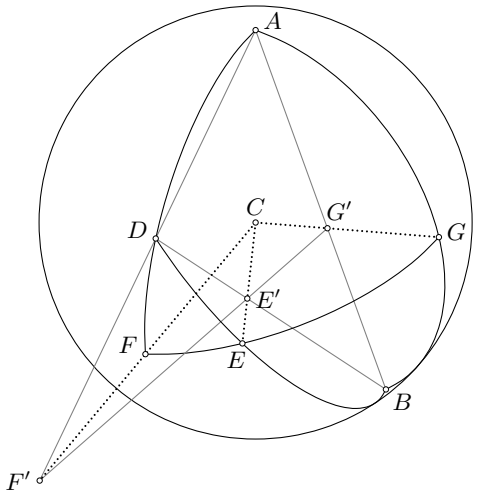
Hoy en día usaríamos la función seno, en vez de las cuerdas y el Teorema se ve así:

$$\frac{\mathrm{sen}(\sphericalangle ACF)}{\mathrm{sen}(\sphericalangle FCD)} \cdot \frac{\mathrm{sen}(\sphericalangle DCE)}{\mathrm{sen}(\sphericalangle ECB)} \cdot \frac{\mathrm{sen}(\sphericalangle BCG)}{\mathrm{sen}(\sphericalangle GCA)} = 1.$$

Similarmente demuestra

$$\frac{\mathrm{crd}(2\sphericalangle ACB)}{\mathrm{crd}(2\sphericalangle ACG)} = \frac{\mathrm{crd}(2\sphericalangle DCB)}{\mathrm{crd}(2\sphericalangle DCE)} \cdot \frac{\mathrm{crd}(2\sphericalangle FCE)}{\mathrm{crd}(2\sphericalangle FCG)}. \tag{3.19}$$

Para demostrarlo, considera el plano por los puntos del triángulo original ABD y proyecta los otros tres puntos desde C a este plano. Se obtienen así los puntos E', F' y G'.



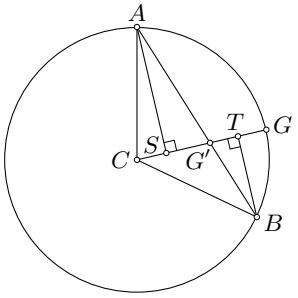
Menelao usa sin demostrar la afirmación análoga en el plano:

$$\frac{AG'}{G'B} = \frac{AF'}{F'D} \cdot \frac{DE'}{E'B}.$$

El hecho de que la usó sin demostrar indica que ya se conocía antes. Pero como es la primera fuente donde aparece, se considera hoy también como Teorema de Menelao. La demostración es fácil y elemental al usar homotecias.

Luego usa que $\dfrac{AG'}{G'B} = \dfrac{\mathrm{crd}(2\sphericalangle ACG)}{\mathrm{crd}(2\sphericalangle GCB)}$. Esto se ve al considerar el círculo que pasa por A, G' y B: Se erigen perpendiculares a CG por A y por B, entonces se obtienen triángulos semejantes ASG' y BTG'. Por ello se tiene

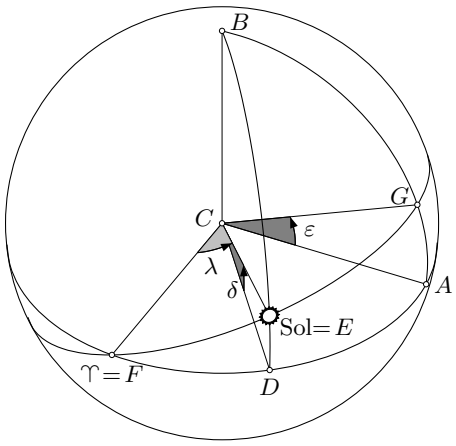
$$\frac{AG'}{G'B} = \frac{AS}{BT} = \frac{2AS}{2BT} = \frac{\mathrm{crd}(2\sphericalangle ACG)}{\mathrm{crd}(2\sphericalangle GCB)}.$$



Similarmente se obtiene $\dfrac{AF'}{F'D} = \dfrac{\mathrm{crd}(2\sphericalangle ACF)}{\mathrm{crd}(2\sphericalangle FCD)}$ y también $\dfrac{DE'}{E'B} = \dfrac{\mathrm{crd}(2\sphericalangle BCG)}{\mathrm{crd}(2\sphericalangle GCA)}$. Con ello queda demostrado el Teorema de Menelao.

Véamos ahora cómo se aplica este último resultado para determinar la declinación δ del Sol a partir de su latitud λ.

Para ello se usa el polo norte celeste B, el Sol será el punto E y $F = \Upsilon$ el punto vernal de primavera. El círculo máximo por F y D es el ecuador y aquel por F y E la eclíptica. Además se usa el círculo máximo que es perpendicular al ecuador y a la eclíptica y que intersecta a estos en A y G respectivamente.



Entonces se tiene la situación del Teorema de Menelao. Nótese que $\sphericalangle ACG = \varepsilon = 23.44°$ y que $\sphericalangle ACB = \sphericalangle FCG = \sphericalangle DCB = 90°$. Al sustituirlo en (3.19) se obtiene

$$\frac{\mathrm{crd}(180°)}{\mathrm{crd}(2\varepsilon)} = \frac{\mathrm{crd}(180°)}{\mathrm{crd}(2\delta)} \cdot \frac{\mathrm{crd}(2\lambda)}{\mathrm{crd}(180°)},$$

de donde se obtiene:

$$\operatorname{crd}(2\delta) = \frac{\operatorname{crd}(2\varepsilon) \cdot \operatorname{crd}(2\lambda)}{2r}.$$

En la notación moderna esto es

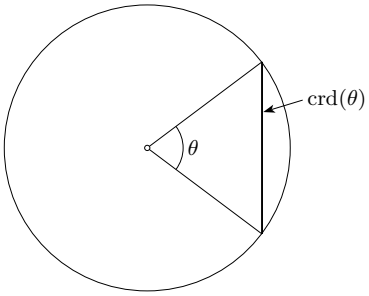
$$\operatorname{sen}(\delta) = \operatorname{sen}(\varepsilon)\operatorname{sen}(\lambda).$$

Las primeras tablas trigonométricas

Del primero que se sabe que construyó algo parecido a lo que hoy llamamos *tablas trigonométricas* fue Hiparco de Nicea (190–120 a.e.c). A Hiparco se le considera uno de los más grandes astrónomos de la antigüedad. Desarrolló parte de la trigonometría esférica, la cual consideraremos más adelante en este capítulo, y se le conoce por haber descubierto la precesión de los equinoccios, es decir que el eje de rotación de la Tierra no está fijo sino que gira lentamente. Sin duda Hiparco tenía una extrema destreza matemática.

Hiparco de Nicea

Se sabe que calculó una tabla de cuerdas, es decir, para diferentes ángulos θ, la longitud de la cuerda que abarca el ángulo θ en un círculo de cierto radio.



$\operatorname{crd}(\theta)$ denota la cuerda que se ve desde
el centro del círculo bajo un ángulo θ.

Lamentablemente casi todos los trabajos de Hiparco se perdieron. Se piensa que la siguiente tabla podría ser una buena reconstrucción de su tabla de cuerdas.

θ	crd(θ)
0°	0
7°30′	450
15°	897
22°30′	1341
30°	1780

θ	crd(θ)
37°30′	2210
45°	2631
52°30′	3041
60°	3438
67°30′	3820

θ	crd(θ)
75°	4186
82°30′	4533
90°	4862
97°30′	5169
105°	5455

θ	crd(θ)
112°30′	5717
120°	5954
127°30′	6166
135°	6352
142°30′	6511

θ	crd(θ)
150°	6641
157°30′	6743
165°	6817
172°30′	6861
180°	6876

Hay varios aspectos que deben observarse. Los cálculos se hacían sobre un círculo con un radio muy grande. Podemos obtener el radio de ese círculo usando $r = \text{crd}(60°)$, que arroja el valor $r = 3438$. Los valores de crd(θ) aparecen redondeados a números enteros. La tabla no resulta muy útil para fines prácticos ya que la diferencia entre dos ángulos consecutivos en ella es de 7°30′, que es un intervalo demasiado grande para realizar cálculos con buena precisión. Probablemente se usaba combinada con interpolaciones lineales.

No sabemos cómo llegó Hiparco a estos valores. Sobre el número 3438 hay una especulación: Si se subdivide el ángulo completo de 360° en minutos se obtiene 21 600 minutos y

$$\frac{21\,600}{2\pi} \approx 3\,438.$$

Esto permite comparar la longitud de arcos medidos en minutos directamente con la longitud de las cuerdas. Por ejemplo, la cuerda correspondiente al arco de $60° = 3600'$ tiene una longitud de 3438, mientras el arco tiene una longitud de 3600.

Del primero que tenemos indicios claros de cómo calculó los valores en una tabla de cuerdas es de Claudio Ptolomeo, (98–162).

Claudius Ptolomeo

Se sabe poco de la vida de Ptolomeo. Trabajó como astrónomo en Alejandría y elaboró un atlas de las estrellas. Para ello retomó datos de Hiparco. Se sabe que hizo observaciones del año 127 al 141. El nombre Claudio, en latin Claudius, indica que su familia, de ascendencia griega, era respetada por considerárseles ciudadanos romanos. Escribió un libro cuyo título original era "La Compilación Matemática", pero pronto se le llamó "La Gran Compilación". Al traducirlo al árabe se le nombró "al-majisti" y hoy lo conocemos como el "Almagest". En 13 tomos describe el movimiento del Sol, la Luna y los planetas conocidos en aquel entonces, todo ello basado en una teoría geocéntrica, que

perduró hasta que Copernico redescubriera la teoria heliocéntrico de Aristarco.

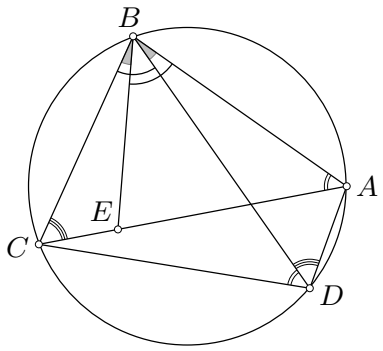
También escribió un libro llamado "Geografía" en donde describe, en 8 tomos, cómo hacer un mapa y da la ubicación de los lugares más importantes del mundo conocido entonces, con sus longitudes y latitudes. Por mucho tiempo, la suya fue la única manera conocida de hacer un mapamundi.

Para elaborar una mejor tabla de cuerdas, Ptolomeo usa un resultado que ahora lleva su nombre: el *Teorema de* Ptolomeo, aunque es muy posible que éste fuera conocido desde antes: En un cuadrilátero $ABCD$ inscrito en un círculo, el producto de las diagonales $AC \cdot BD$ es igual a la suma de los productos de los lados opuestos $AB \cdot CD + BC \cdot DA$, es decir

$$AC \cdot BD = AB \cdot CD + BC \cdot DA.$$

El Teorema de Pitágoras resulta un caso particular de este teorema, al tomar el cuadrilátero como rectángulo. El Teorema de Ptolomeo se puede demostrar de manera elemental, como veremos a continuación.

Sea E sobre la diagonal AC aquel punto que satisface $\sphericalangle CBE = \sphericalangle DBA$. De ahí se sigue que $\sphericalangle CBD = \sphericalangle EBA$, que son los ángulos indicados con un arco sencillo en la figura.



Además se tiene que $\sphericalangle BAC = \sphericalangle BDC$, dado que son ángulos sobre la cuerda BC. Estos ángulos se muestran con arcos dobles en la figura. Similarmente se tiene $\sphericalangle ACB = \sphericalangle ADB$, dado que son ángulos sobre la cuerda AB.

De ahí se sigue que los triángulos BAE y BDC son semejantes y también lo son BCE y BDA, ya que coinciden en los tres ángulos.

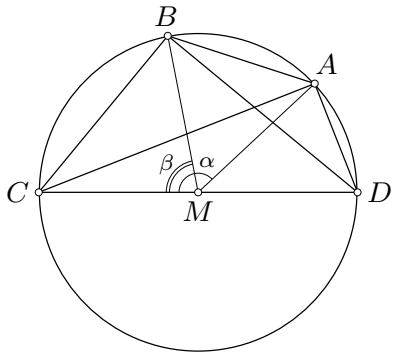
Por tanto

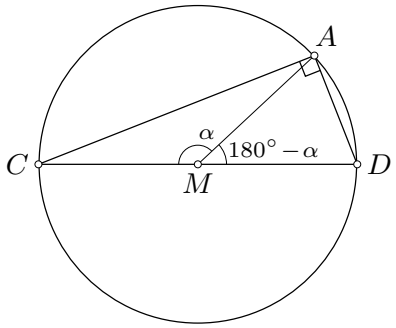
$$\frac{AB}{AE} = \frac{DB}{DC} \qquad \text{y} \qquad \frac{BC}{CE} = \frac{BD}{DA}$$

de donde

$$\begin{aligned} AB \cdot CD &= BD \cdot AE \\ BC \cdot DA &= BD \cdot CE \\ AB \cdot CD + BC \cdot DA &= BD \cdot (AE + CE) = BD \cdot AC. \end{aligned}$$

Con este Teorema, Ptolomeo podía calcular la cuerda correspondiente a la diferencia de dos ángulos cuyas cuerdas ya conocía, es decir, si ya había calculado $a = \text{crd}(\alpha)$ y $b = \text{crd}(\beta)$ entonces podía obtener $\text{crd}(\alpha - \beta)$. Para ver esto consideramos un diámetro CD de un círculo con centro M y colocamos A y B sobre la circunferencia de tal manera que $\sphericalangle CMA = \alpha$ y $\sphericalangle CMB = \beta$, como se puede ver en la figura del lado izquierdo.





Se tiene ahora que

$$AB = \operatorname{crd}(\alpha-\beta), \qquad AC = \operatorname{crd}(\alpha), \qquad BC = \operatorname{crd}(\beta),$$
$$CD = 2R, \qquad DA = \operatorname{crd}(180^\circ-\alpha), \qquad BD = \operatorname{crd}(180^\circ-\beta).$$

Nótese que $\operatorname{crd}(180^\circ-\alpha)$ y $\operatorname{crd}(180^\circ-\beta)$ se pueden calcular con el Teorema de Pitágoras – véase la figura anterior a la derecha – como sigue:

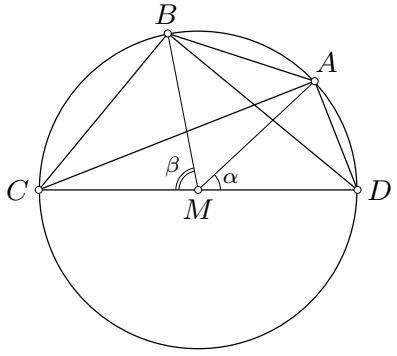
$$\operatorname{crd}(180^\circ-\alpha) = \sqrt{4r^2-\operatorname{crd}(\alpha)^2}.$$

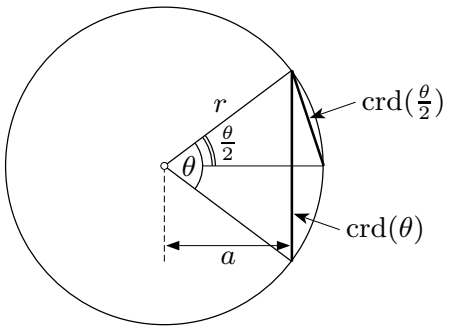
De la relación de Ptolomeo se sigue que

$$\operatorname{crd}(\alpha-\beta) = AB = \frac{AC\cdot BD - BC\cdot DA}{CD} = \frac{\operatorname{crd}(\alpha)\operatorname{crd}(180^\circ-\beta) - \operatorname{crd}(\beta)\operatorname{crd}(180^\circ-\alpha)}{2r},$$

donde el lado derecho se puede calcular cuando se conocen $\operatorname{crd}(\alpha)$ y $\operatorname{crd}(\beta)$.

Un argumento completamente similar muestra que se puede calcular $\operatorname{crd}(180^\circ-(\alpha+\beta))$, y por lo tanto $\operatorname{crd}(\alpha+\beta)$, si se conocen $\operatorname{crd}(\alpha)$ y $\operatorname{crd}(\beta)$. Para ello, considerando la siguiente situación





se obtiene la fórmula

$$\operatorname{crd}(180^\circ-(\alpha+\beta)) = AB = \frac{\operatorname{crd}(180^\circ-\alpha)\operatorname{crd}(180^\circ-\beta) - \operatorname{crd}(\alpha)\operatorname{crd}(\beta)}{2r}.$$

Además, desde la época de Arquímedes, se sabía cómo calcular el apotema de una cuerda y de ahí la cuerda de la mitad del ángulo dado – ver la figura anterior derecha. Este cálculo queda así:

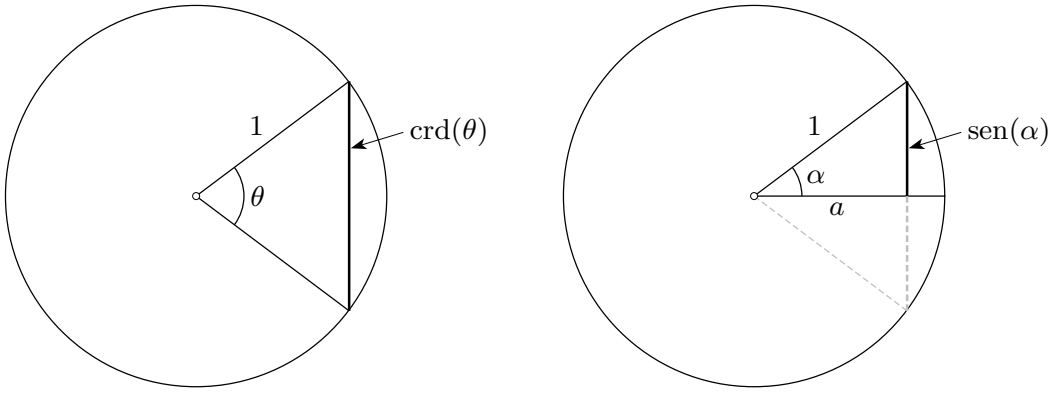
$$a = \sqrt{r^2 - \left(\frac{\operatorname{crd}(\theta)}{2}\right)^2},$$

$$\operatorname{crd}(\tfrac{\theta}{2}) = \sqrt{\left(\frac{\operatorname{crd}(\theta)}{2}\right)^2 + (1-a)^2}.$$

Los griegos conocían $\operatorname{crd}(72°)$ y $\operatorname{crd}(60°)$. A partir de esto podían obtener $\operatorname{crd}(18°)$ y $\operatorname{crd}(15°)$. Con la fórmula para la subtracción de ángulos podían calcular $\operatorname{crd}(3°)$ y con la fórmula para la adición podían obtener $\operatorname{crd}(k \cdot 3°)$ para cualquier número natural $k = 0, 1, \ldots, 60$.

Finalmente Ptolomeo estimó $\operatorname{crd}(1°)$ y elaboró después una tabla de cuerdas para los múltiplos de medio grado, es decir para los ángulos $0°, 30', 1°, \ldots, 179°30', 180°$. Su tabla daba las cuerdas en el sistema sexagesimal, no en enteros como la de Hiparco.

En la India se empezaron a calcular tablas del seno a partir del siglo V. El primero en hacerlo fue Aryabhata (ca. 476–550). Se elaboraron sofisticados algoritmos para calcular muchos valores a partir de unos pocos, similarmente a lo que vimos para la tabla de cuerdas de Ptolomeo. La relación entre las dos tablas es bastante directa.



Se tiene que

$$\operatorname{crd}(\theta) = 2\operatorname{sen}\left(\frac{\theta}{2}\right) \qquad \text{y} \qquad \operatorname{sen}(\alpha) = \tfrac{1}{2}\operatorname{crd}(2\alpha).$$

Por ello realmente no importa cuál de las dos tablas se usa. Hoy día ya sólo se conoce el seno, las cuerdas han pasado al olvido. Además se definieron nuevas funciones como el coseno

$$\cos(\alpha) = a, \qquad \text{y} \qquad \operatorname{vers}(\alpha) = 1 - a$$

donde a es el apotema en la figura de arriba a la derecha. Si reemplazamos las cuerdas por los senos en la fórmula

$$\operatorname{crd}(\alpha - \beta) = \frac{\operatorname{crd}(\alpha)\operatorname{crd}(180° - \beta) - \operatorname{crd}(\beta)\operatorname{crd}(180° - \alpha)}{2r}$$

y usamos $\varphi = \frac{\alpha}{2}$, $\psi = \frac{\beta}{2}$ y $r = 1$, obtenemos:

$$2\operatorname{sen}(\varphi - \psi) = \frac{2\operatorname{sen}(\varphi)2\cos(\psi) - 2\operatorname{sen}(\beta)2\cos(\alpha)}{2}$$

porque $\operatorname{crd}(180^\circ - \beta) = 2\operatorname{sen}(\frac{180^\circ - \beta}{2}) = 2\operatorname{sen}(90^\circ - \psi) = 2\cos(\psi)$. De ahí se sigue que

$$\operatorname{sen}(\varphi - \psi) = \operatorname{sen}(\varphi)\cos(\psi) - \operatorname{sen}(\beta)\cos(\alpha),$$

que es la fórmula conocida del seno de la diferencia de dos ángulos. De la fórmula para $\operatorname{crd}(\alpha + \beta)$ se obtiene análogamente

$$\cos(\varphi + \psi) = \cos(\varphi)\cos(\psi) - \sin(\varphi)\sin(\psi).$$

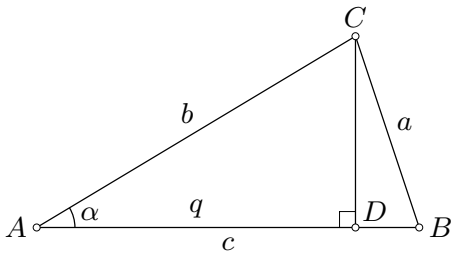
Ley de los senos y Ley del coseno

Estas dos leyes forman parte esencial de la trigonometría moderna. Pero son de procedencia relativamente reciente.

La ley del coseno aparece por primera vez en *Los Elementos* de Euclides como Proposición II.13 en la siguiente forma:

> *En los triángulos acutángulos, el cuadrado del lado que subtiende al ángulo agudo es menor que* [la suma de] *los cuadrados de los lados que comprenden el ángulo agudo en dos veces el rectángulo comprendido por uno de los lados del ángulo agudo sobre el que cae la perpendicular y la* [recta] *interior cortada por la perpendicular hasta el ángulo agudo.*

La redacción se aclara con un dibujo.



Sea α un ángulo agudo del triángulo ABC. En el texto, el cuadrado del lado que subtiende el ángulo agudo, a^2, es igual a la suma de los cuadrados de los lados adyacentes, esto es $b^2 + c^2$, disminuido por el doble de un rectángulo. Un lado de ese rectángulo es uno de los lados adyacentes, tomamos c, y el otro es lo que la altura corta hacia el ángulo agudo, es decir $q = AD$ en el dibujo anterior. Escrito como una fórmula esto queda así

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

Este resultado no se considera parte de la trigonometría, es sólo una comparación entre áreas. Sin embargo, si se observa que $q = b\cos(\alpha)$, da lo que hoy llamamos Ley del coseno

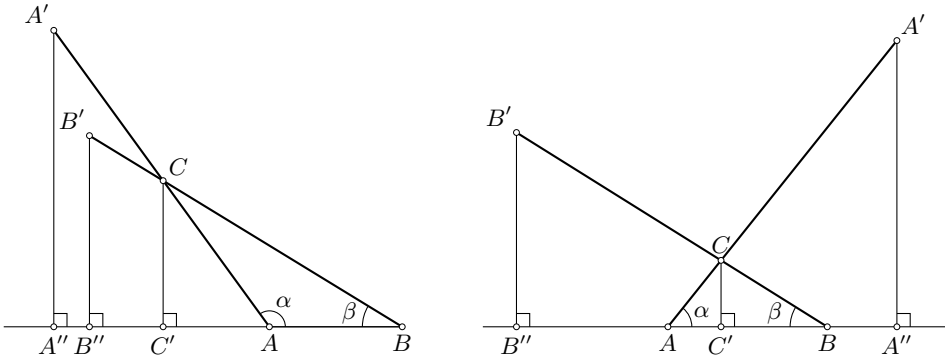
$$a^2 = b^2 + c^2 - 2bc\cos(\alpha).$$

La ley de los senos aparece explícitamente en la literatura en el libro *Tratamiento sobre el cuadrilátero* de Nasir al-Din al-Tusi (1201–1274). Al-Tusi nació en Tus, una ciudad en el norte de lo que hoy es Iran. Su tío le enseñó lógica y física a una edad temprana. Se mudó a Nishapur y allí estudió filosofía, medicina y matemáticas. Vivió en un tiempo de guerra: los mongoles invadían

grandes espacios, sembraban la destrucción y cometían terribles crueldades. AL-TUSI aceptó una invitación a unirse a los "asesinos", una secta islámica con fuerza militar que defendía las fortalezas en esa región. En esa época al-Tusi escribió varios textos sobre filosofía, lógica, matemáticas y astronomía. Cuando los mongoles bajo el mando de Hulegu destruyeron el castillo Alamut, incorporaron a al-Tusi en sus tropas como asesor científico. Así, AL-TUSI acompañó a Hulegu en la conquista de Bagdad y le propuso construir un observatorio. En 1262 se completó el observatorio y AL-TUSI empezó a elaborar tablas muy precisas sobre el movimiento de los planetas. Tradujo varios textos del griego al árabe.

NASIR AL-DIN AL-TUSI

La demostración de la Ley de los senos de AL-TUSI es como sigue. Primero se prolonga AC sobre C a A' de tal manera que AA' mida r. Luego se erige una perpendicular $A'A''$ a AB por A' con pie A''. Por definición se tiene $A'A'' = \text{sen}(\alpha)$. Lo mismo se hace con B y se obtiene $B'B'' = \text{sen}(\beta)$.



Esto es válido aún si uno de los ángulos es obtuso – como se ve en la ilustración anterior izquierda – dado que $\text{sen}(\varphi) = \text{sen}(180° - \varphi)$. Sea además CC' la perpendicular a AB por C. Entonces los triángulos $AC'C$ y $AA''A'$ son semejantes y por ello se tiene $A'A'' : A'A = CC' : CA$, es decir $\text{sen}(\alpha) \cdot AB = r \cdot CC'$. De manera similar $BC'C$ y $BB''B'$ son semejantes, por lo que $\text{sen}(\beta) \cdot BC = r \cdot CC'$. Al igualar los lados derechos se obtiene $\text{sen}(\alpha) \cdot AB = \text{sen}(\beta) \cdot BC$, es decir

$$\frac{\text{sen}(\alpha)}{\text{sen}(\beta)} = \frac{BC}{AC}.$$

La demostración de AL-TUSI es similar a la moderna, ambas se basan en calcular la altura CC' de dos maneras, usando los triángulos rectángulos $AC'C$ y $BC'C$. Pero AL-TUSI primero amplía estos triángulos para que su hipotenusa mida r. Esto deja entrever que $\operatorname{sen}(\varphi)$ no era para él una proporción sino una longitud en un círculo de radio r.

Una vez que la comprensión de las funciones seno y coseno queda claramente establecida – sea esto en el sentido de hoy en un círculo de radio 1, o como se hizo antes en un círculo con un radio grande como $r = 3438$ – resulta inmediato interpretar el enunciado de la Proposición II.13 de *Los Elementos* de Euclides como lo que en la actualidad conocemos como la Ley del coseno.

Se reconoce a Hiparco como padre de la trigonometría porque fue el primero en concebir la importancia de conocer las longitudes de las cuerdas en función del ángulo. Con eso la astronomía obtuvo una herramienta muy poderosa. Pero fue AL-TUSI, quien escribió el primer libro sobre trigonometría y con ello la elevó a un asunto de interés matemático, independiente de la astronomía.

Mejoramiento de las tablas

En la Edad Media se reanudaron los esfuerzos para calcular tablas cada vez más precisas del seno. Las herramientas eran esencialmente las mismas que las que ya vimos en funcionamiento en manos de PTOLOMEO. Para mantener las fórmulas lo más sencillas posible y cercanas a lo que se conoce hoy dia, adoptamos esas herramientas aquí con $r = 1$, aunque en la Edad Media esto no era común:

(a) Calcular $\cos(\varphi)$ a partir de $\operatorname{sen}(\varphi)$ según la fórmula

$$\cos(\varphi) = \sqrt{1 - \operatorname{sen}(\varphi)^2}.$$

(b) Calcular $\operatorname{sen}(\frac{\varphi}{2})$ a partir de $\operatorname{sen}(\varphi)$, según (a) y la fórmula

$$\operatorname{sen}(\tfrac{\varphi}{2}) = \sqrt{\frac{1 - \cos(\varphi)}{2}}.$$

(c) Calcular $\operatorname{sen}(\varphi + \psi)$ a partir de $\operatorname{sen}(\varphi)$ y $\operatorname{sen}(\psi)$, según la fórmula

$$\operatorname{sen}(\varphi + \psi) = \operatorname{sen}(\varphi)\cos(\psi) + \cos(\varphi)\operatorname{sen}(\psi).$$

(d) Calcular $\operatorname{sen}(\varphi - \psi)$ a partir de $\operatorname{sen}(\varphi)$ y $\operatorname{sen}(\psi)$, según la fórmula

$$\operatorname{sen}(\varphi - \psi) = \operatorname{sen}(\varphi)\cos(\psi) - \cos(\varphi)\operatorname{sen}(\psi).$$

Veamos cómo RICHARD OF WALLINGFORD (1292–1336) describió cómo calcular $\operatorname{sen}(1^\circ)$ con la precisión que uno desee. Se parte de que uno conoce $\operatorname{sen}(3^\circ)$. Para ello se usa (d).

1. A partir de $\operatorname{sen}(3^\circ)$ se calcula $\operatorname{sen}(\frac{1}{16}3^\circ)$ usando cuatro veces (b).
2. Se calcula $\operatorname{sen}(5 \cdot \frac{3}{16}^\circ) = \operatorname{sen}(\frac{15}{16}^\circ)$ usando varias veces (c).
3. Se calcula $\operatorname{sen}(\frac{1}{4} \cdot \frac{3}{16}^\circ) = \operatorname{sen}(\frac{3}{64}^\circ)$ usando (b) dos veces.
4. Se calcula $\operatorname{sen}(\frac{3}{64}^\circ + \frac{15}{16}^\circ) = \operatorname{sen}(\frac{63}{64}^\circ)$ usando (c).

5. Se calcula $\operatorname{sen}(\frac{1}{4}\cdot\frac{3}{64}^\circ) = \operatorname{sen}(\frac{3}{256}^\circ)$ usando (b) dos veces.
6. Se calcula $\operatorname{sen}(\frac{3}{256}^\circ + \frac{63}{64}^\circ) = \operatorname{sen}(\frac{255}{256}^\circ)$.

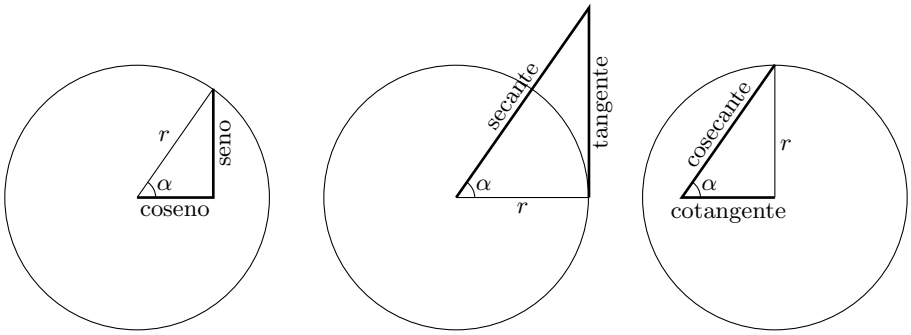
Este proceso se puede continuar tantas veces uno quiera. En este proceso se usa que

$$\frac{3}{4^{n+1}} + \frac{4^n-1}{4^n} = \frac{3}{4^{n+1}} + \frac{4\cdot(4^n-1)}{4^{n+1}} = \frac{3+4^{n+1}-4}{4^{n+1}} = \frac{4^{n+1}-1}{4^{n+1}}.$$

Conforme n aumenta, se obtienen cada vez mejores aproximaciones de 1.

En 1596 se publicó el libro *Opus Palatinum de Triangulis* (Obra palatina sobre triángulos) de Georg Joachim de Porris (1514–1574), también conocido como Rheticus, completado 22 años después de la muerte de Rheticus por su discípulo Lucius Valenton Otho. Rheticus nació en Feldkirch, en el extremo Oeste de Austria, estudió en Zürich y después con Copérnico y es considerado su único discípulo.

Elaboró tablas de las seis funciones trigonométricas que indicamos en el siguiente esquema, aunque no usó estos nombres.



Él y su discípulo usaron un radio $r = 10\,000\,000\,000$ y calcularon los valores de las seis funciones en intervalos de $10'' = \frac{1}{360}^\circ$. La obra consta de 1400 páginas tamaño folio.

Una implementación en las computadoras: el método CORDIC

La función seno es una función *trascendente*, es decir, es imposible calcular $\operatorname{sen}(x)$ con exactitud en un número finito de pasos, salvo para valores especiales como por ejemplo $x = 0$.

Con la invención de las computadoras en los años 50 del siglo pasado, se produjo una revolución en cómo se calcula. Las tablas de logaritmos, seno y tangente desaparecieron y se hicieron común el uso de calculadoras de bolsillo. Pero ¿cómo calcula una calculadora el valor de $\operatorname{sen}(20^\circ)$?

Una idea sería expresar $\operatorname{sen}(x)$ como una seria de potencias:

$$\operatorname{sen}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \frac{x^{11}}{11!} + \dots,$$

donde $n!$ es la *factorial* de n, que está definida como el producto de los primeros n números naturales, es decir, $n! = 1\cdot 2\cdot \ldots (n-1)\cdot n$. Es importante que aquí se use *radianes*, es decir la longitud del arco de un sector con ángulo x y radio 1.

Como $\operatorname{sen}(\pi - x) = \operatorname{sen}(x)$ y $\operatorname{sen}(x + 2\pi) = \operatorname{sen}(x)$ es suficiente calcular $\operatorname{sen}(x)$ para $-\frac{\pi}{2} \leq x \leq \frac{\pi}{2}$. Por ello basta que se tome los primeros 6 sumandos en la serie arriba para obtener una exactitud de 8 dígitos.

Pero para las calculadores, llevar a cabo una multiplicación es "costoso", es decir cuesta mucho tiempo. Con el *esquema de Horner* de William George Horner (1789–1837) se reduce el número de multiplicaciones considerablemente:

$$\begin{aligned}\operatorname{sen}(x) &\approx x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \frac{x^{11}}{11!} \\ &= x \cdot \Big(1 - x^2 \cdot \Big(\frac{1}{3!} - x^2 \cdot \Big(\frac{1}{5!} - x^2 \cdot \Big(\frac{1}{7!} - x^2 \cdot \Big(\frac{1}{9!} - \frac{1}{11!}x^2\Big)\Big)\Big)\Big)\Big)\end{aligned}$$

Hay que calcular x^2 una vez y luego hay que hacer 5 multiplicaciones. Los valores de $\frac{1}{3!}, \ldots \frac{1}{11!}$ se pueden calcular de antemano y guardar como constantes.

Pero en 1959, aún eso era demasiado lento y se buscaba una manera de hacerlo sin multiplicación alguna, excepto multiplicación o división por 2. En las computadoras y calculadoras de bolsillo, los números representan en el *sistema binario*. En este sistema el número decimal 10 se expresa como 1010_2 ya que

$$\mathbf{10} = \mathbf{1} \cdot 2^3 + \mathbf{0} \cdot 2^2 + \mathbf{1} \cdot 2^1 + \mathbf{0} \cdot 2^0.$$

El número 0.2 se expresa como $0.00110011\overline{0011}_2$. El subíndice 2 indica que se trabaja en base 2 y no en la base 10.

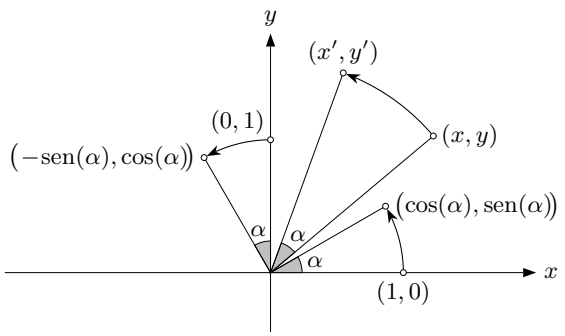
En el sistema binario la multiplicación o la división con 2 no cuesta trabajo, ya que es simplemente es movimiento a la izquierda resp. a la derecha de todos los dígitos.

En 1959 el ingeniero Jack E. Volder desarrolló un algoritmo que lograra calcular $\operatorname{sen}(x)$ sin hacer una sola multiplicación excepto con 2. El método fue bautizado CORDIC que es un acrónimo de coordinate rotation digital computer.

El nombre expresa parte de la idea: se calcula la rotatión en coordenadas.

Se tiene que al rotar por el ángulo α en contra de las manecillas del reloj el punto $(1, 0)$ da al punto $\big(\cos(\alpha), \operatorname{sen}(\alpha)\big)$ y el punto $(0, 1)$ se rota al punto $\big(-\operatorname{sen}(\alpha), \cos(\alpha)\big)$. Un punto en general (x, y) se rota en (x', y') donde

$$\begin{aligned} x' &= \cos(\alpha)x - \operatorname{sen}(\alpha)y, \\ y' &= \operatorname{sen}(\alpha)x + \cos(\alpha)y. \end{aligned}$$



Esto se puede expresar de manera compacta al usar el lenguaje de matrices:

$$\underbrace{\begin{bmatrix} x' \\ y' \end{bmatrix}}_{p'} = \underbrace{\begin{bmatrix} \cos(\alpha) & -\operatorname{sen}(\alpha) \\ \operatorname{sen}(\alpha) & \cos(\alpha) \end{bmatrix}}_{M_\alpha} \cdot \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{p}.$$

Algunas propiedades que serán importantes:

- Las rotaciones se pueden acumular: $M_\beta M_\alpha = M_{\alpha+\beta}$.
- Las rotaciones son lineales, en particular se tiene: $M_\alpha(\lambda p) = \lambda(M_\alpha p)$, para cada punto p y cada número λ.

Volder conjuntó varias ideas para lograr calcular el valor de $\operatorname{sen}(\alpha)$ sin usar multiplicaciones excepto con potencias de 2:

- En vez de M_α se usa

$$R_\alpha = \frac{1}{\cos(\alpha)} M_\alpha = \frac{1}{\cos(\alpha)} \begin{bmatrix} 1 & -\tan(\alpha) \\ \tan(\alpha) & 1 \end{bmatrix};$$

- Se usan ángulos especiales $\varphi_0, \varphi_1, \ldots$ para los cuales se tiene $\tan(\varphi_k) = \frac{1}{2^k}$. Estos ángulos se llaman *ángulos de CORDIC* y se pueden calcular de antemano y guardar en constantes. Para un ángulo de CORDIC φ_k se tiene $\cos(\varphi_k) = \cos(\tan^{-1}(\frac{1}{2^k})) = \frac{1}{\sqrt{1+(\frac{1}{2^k})^2}} = \kappa_k$

$$R_k = \frac{1}{\kappa_k} \begin{bmatrix} 1 & -\frac{1}{2^k} \\ \frac{1}{2^k} & 1 \end{bmatrix}.$$

  Por las propiedades mencionadas se tiene

$$R_k R_\ell = \frac{1}{\kappa_k \kappa_\ell} \begin{bmatrix} 1 & -\frac{1}{2^k} \\ \frac{1}{2^k} & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2^\ell} \\ \frac{1}{2^\ell} & 1 \end{bmatrix}.$$

- El ángulo α del cual se debe calcular el valor $\operatorname{sen}(\alpha)$, se aproxima como

$$\alpha \approx \varepsilon_0\varphi_0 + \varepsilon_1\varphi_1 + \varepsilon_2\varphi_2 + \ldots + \varepsilon_n\varphi_n,$$

  donde $\varepsilon_k = \pm 1$ para $k = 0, 1, \ldots n$. Además, el número n se puede calcular de antemano, ya que de él depende la precisión del valor que se calculará.

Si se juntan estas ideas se obtiene

$$\begin{bmatrix} \cos(\alpha) \\ \operatorname{sen}(\alpha) \end{bmatrix} \approx \frac{1}{\kappa_0\kappa_1\kappa_2\ldots\kappa_n} \begin{bmatrix} 1 & -\varepsilon_n\frac{1}{2^n} \\ \varepsilon_n\frac{1}{2^n} & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & -\varepsilon_1\frac{1}{2^1} \\ \varepsilon_1\frac{1}{2^1} & 1 \end{bmatrix} \begin{bmatrix} 1 & -\varepsilon_0\frac{1}{2^0} \\ \varepsilon_0\frac{1}{2^0} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Se calcula $K_n = 1/\kappa_0\kappa_1\cdots\kappa_n$ de antemano y entonces se tiene

$$\begin{bmatrix} \cos(\alpha) \\ \operatorname{sen}(\alpha) \end{bmatrix} \approx \begin{bmatrix} 1 & -\varepsilon_n\frac{1}{2^n} \\ \varepsilon_n\frac{1}{2^n} & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & -\varepsilon_1\frac{1}{2^1} \\ \varepsilon_1\frac{1}{2^1} & 1 \end{bmatrix} \begin{bmatrix} 1 & -\varepsilon_0\frac{1}{2^0} \\ \varepsilon_0\frac{1}{2^0} & 1 \end{bmatrix} \begin{bmatrix} K_n \\ 0 \end{bmatrix}.$$

Falta explicar cómo se calculan los signos $\varepsilon_0, \varepsilon_1, \ldots \varepsilon_n$. Se pone

$$\varepsilon_0 = \begin{cases} 1, & \text{si } \alpha \geq 0, \\ -1, & \text{si } \alpha < 0, \end{cases}$$

y luego de manera iterativo

$$\varepsilon_{k+1} = \begin{cases} 1, & \text{si } \alpha \geq \varepsilon_0\varphi_0 + \ldots \varepsilon_k\varphi_k, \\ -1, & \text{si } \alpha < \varepsilon_0\varphi_0 + \ldots \varepsilon_k\varphi_k. \end{cases}$$

El siguiente pseudo-códigio muestra el algoritmo

```
cordic(alpha):
   a = 0
   x = K_n
   y = 0
   for k = 0 to n
      if alpha < a then
         e = -1
      else
         e = 1
      endif
      a = a + e·φ_k
      x = x - e·2^-k·y
      y = y + e·2^-k·x
      k = k + 1
   endfor
   cos(alpha) = x
   sen(alpha) = y
```

La idea de Volder luego se amplió a que puede calcular también exponenciales, logaritmos y otras funciones trascedentales. Fue un éxito de tal manera que hoy día ya no nos damos cuenta de cómo los aparatos que nos rodean llevan a cabo los cáclculos de manera eficiente y confiable. De alguna manera, las matemáticas se alejaron de nosotros al ser empacados en el interior íntimo de los aparatos.

Bibliografía para la sección de trigonometría esférica

Glen van Brummelen: The mathematics of the heavens and earth, the early history of trigonometry. Princeton University Press, 2009.

J. L. Berggren: Episodes in the Mathematics of medieval Islam, Springer, 1986.

Scriba, Schreiber: 5000 Jahre Geometrie, Springer 2000.

Lambacher, Schweizer: Ebene Trigonometrie, Klett 1968

MacTutor

Wikipedia inglés, alemán y español

David Maletinsky: Wie rechnen Taschenrechner?, Mentorierte Arbeit, ETH Zürich, 2009

Capítulo 4

Cálculo, funciones y límites

Herramientas para modelar el cambio continuo

4.1. El Cálculo Diferencial e Integral

¿Qué es el cálculo?

El cálculo infinitesimal, cálculo de infinitesimales o simplemente "Cálculo" constituye una parte muy importante de las matemáticas modernas. Incluye el estudio de las funciones, los límites, las derivadas, las integrales, las series infinitas y las ecuaciones diferenciales. El cálculo se puede definir como el estudio del cambio en el tiempo, de la misma manera que la geometría es el estudio del espacio.

El cálculo infinitesimal tiene amplias aplicaciones en la ciencia y la ingeniería y se usa para resolver problemas para los cuales el álgebra por sí sola es insuficiente. Este cálculo se construye con base en el álgebra, la trigonometría y la geometría analítica y como ya se mencionó, incluye dos campos principales, cálculo diferencial y cálculo integral, que están relacionados por el teorema fundamental del cálculo. En matemáticas más avanzada, el cálculo se suele llamar análisis matemático y puede definirse como el estudio de las funciones.

La historia del cálculo

Isaac Newton y Gottfried Leibniz son los nombres más famosos vinculados al desarrollo del cálculo. La aplicación más importante en la historia del cálculo es la formulación y solución de las leyes del movimiento y de la gravitación universal, y se deben a Newton.

Durante toda la historia, desde la antigüedad remota hasta el siglo XVII los matemáticos trataron de descubrir maneras de calcular el área de las figuras planas delimitadas por curvas, En el siglo IV A.C. Eudoxo inventó el método de exhaución que consistía en ir llenando la figura cuya área se desea calcular, con polígonos simples como rectángulos y triángulos, cuyas áreas son fáciles de calcular, acomodando figuras cada vez más pequeñas para ir cubriendo los huecos que quedaban. La teoría

es que si el procedimiento se continúa indefinidamente, se obtiene el área deseada. Pero los procesos infinitos no son fáciles de manejar. Arquímedes logró utilizarlos de una manera increíblemente eficaz y rigurosa para obtener el área acotada por una parábola y una recta, el área de un círculo y los volúmenes y superficies de cilindros, conos y esferas. Sobre todo su trabajo para calcular la superficie de una esfera es una joya en elegancia y rigor matemático. Pero el método de exhaución no dio resultados más allá de los obtenidos por Arquímedes, hasta que el desarrollo de la Geometría Analítica renovó el interés en él. La posibilidad de describir a las curvas con ecuaciones permitía aplicar el método de exhaución de una forma simplificada, analizando los problemas de cálculo de áreas y volúmenes utilizando ecuaciones algebraicas. Este método se desarrolló poco a poco hasta formar una nueva disciplina llamada Cálculo Integral.

Al mismo tiempo, las nociones de velocidad y aceleración que gracias a las leyes de Newton permitían analizar matemáticamente el movimiento de los cuerpos, llevaron a la creación del concepto de derivada, ligado al de velocidad instantánea o razón de cambio instantáneo, lo que da lugar a otra nueva rama de las matemáticas: el Cálculo Diferencial, que permite obtener la velocidad y la aceleración de un cuerpo conociendo su posición en cada instante.

El cálculo antes del siglo XVII

En la antigüedad griega se introdujeron algunas de las ideas del cálculo integral, pero no se desarrollaron de una manera sistemática, aunque sí de manera bastante rigurosa. En el cálculo de áreas y volúmenes, la función básica del cálculo integral puede ser rastreada en el tiempo hasta los papiros matemáticos de Moscú que datan del año 1890 a. C, en los que un egipcio calculó satisfactoriamente el volumen del tronco de una pirámide, sin embargo esto se hizo sin dar argumentos racionales para justificar el método de cálculo.

De la escuela de los matemáticos griegos, Eudoxo usó el método exhaustivo, el cual prefiguraba el concepto de límite, para calcular áreas y volúmenes, mientras que Arquímedes (287 - 212 aec) desarrolló más allá su idea inventando un método heurístico, denominado exahución, que se asemeja al cálculo ya bastante al cálculo integral, con lo cual logró obtener el volumen y la superficie de la esfera. El método de exhación fue usado en el siglo III aec y de manera independiente en China por Liu Hui para encontrar el área de un círculo, obteniendo una mejor aproximación que la de Arquímedes, gracias a la superior aritmética de los chinos.

En China, el siglo V, Zu Chongzhi usó lo que más tarde sería llamado la teoría de los indivisibles por el matemático italiano Bonaventura Cavalieri para encontrar el volumen de una esfera.

Cerca del año 1000, el matemático islámico Alhacén fue el primero en derivar la fórmula para la suma de la cuarta potencia de una progresión aritmética, usando un método a partir del cual es fácil encontrar la fórmula para la suma de cualquier potencia integral de mayor orden. Esto sentó las bases para obtener más adelante el área bajo las gráficas de las funciones polinomiales.

En el siglo XI, el polímata chino Shen Kuo desarrolló ecuaciones que se encargaban de integrar. En el siglo XII, el matemático indio, Bhaskara II, desarrolló una derivada temprana representando el cambio infinitesimal, y describió una forma temprana del llamado teorema de Rolle.

También en el siglo XII, el matemático persa Sharaf al-Din al-Tusi descubrió la derivada de la función cúbica, un importante acontecimiento que abre el camino hacia el cálculo diferencial.

En el siglo XIV, Madhava de Sangamagrama, en conjunto con otros matemáticos y astrónomos

de la escuela Kerala de matemáticas y astronomía, al sur de la India, describieron casos especiales de lo que hoy se conocen como las series de Taylor.

Las rutas comerciales llevaron a Europa la notación indo-arábiga y los algoritmos para realizar operaciones con ella, desplazando rápidamente el incómodo sistema de números romanos que se había usado durante casi toda la edad media. Es importante reconocer la contribución de Leonardo de Pisa (Fibonacci) (1170 - 1240) en este proceso.

En la época moderna, descubrimientos independientes relacionados con el cálculo se estaban llevando a cabo por la matemática japonesa del siglo XVII, gracias al aporte de matemáticos como Seki Kōwa, quien expandió el método exhaustivo.

En Europa, el trabajo fundacional fue un tratado del matemático italiano Bonaventura Cavalieri, quien argumentó que los volúmenes y áreas deberían ser calculados como las sumas de los volúmenes y áreas de delgadas secciones infinitesimales. Estas ideas eran similares a las expuestas en el trabajo El método de los teoremas mecánicos de Arquímedes, el cual estuvo perdido hasta principios del siglo XX. El trabajo de Cavalieri no fue bien recibido por la comunidad científica ya que se observó que a veces sus métodos pueden llevar a resultados erróneos, y porque las cantidades infinitesimales que introdujo no parecían estar definidas con suficiente claridad.

El cálculo en el siglo XVII

El estudio formal del cálculo combinó los infinitesimales de Cavalieri con el cálculo de diferencias finitas desarrollado en Europa más o menos al mismo tiempo. La combinación fue lograda por John Wallis, Isaac Barrow y James Gregory, probando estos últimos el teorema fundamental del cálculo integral cerca del año 1675.

La regla del producto y la regla de la cadena, la noción de derivada de mayor orden, las series de Taylor, y las funciones analíticas fueron introducidas por Isaac Newton en una notación idiosincrásica que usó para resolver problemas de física, principalmente de mecánica. En sus publicaciones, Newton reformuló sus ideas para acomodar el idioma matemático de la época, reemplazando cálculo con infinitesimales por argumentos geométricos equivalentes, los cuales estaban más allá de reproches. Usó los métodos del cálculo para resolver el problema del movimiento planetario, la forma de la superficie de un fluido rotante, y se refirió a lo achatada que es la tierra por los polos, así como a muchos otros problemas, los cuales discutió en su Principia mathematica. En otro trabajo, desarrolló una serie de expansiones para las funciones, incluyendo las potencias fraccionarias e irracionales. Estaba claro que Newton entendía perfectamente los principios de las series de Taylor. Sin embargo no publicó muchos de sus descubrimientos. En su tiempo los sistemas infinitesimales tenían mala reputación.

Las ideas del Cálculo fueron descubiertas independientemente por Gottfried Wilhelm Leibniz, quien además las sistematizó de manera magistral, arrebatando a Newton gran parte del mérito por el descubrimiento del cálculo, gracias a su magnífica organización del tema a su excelente y sugestiva notación. Leibniz fue acusado de plagio por Newton, pero se le reconoce sin ninguna duda como inventor independiente del cálculo y como gran contribuyente a su formalización y organización. Su principal contribución fue el proveer un conjunto de reglas claras para la manipulación de cantidades infinitesimales, permitiendo el cómputo de derivadas de segundo orden y de orden superior, y estableciendo la regla del producto y regla de la cadena en su forma diferencial e integral. A diferencia de Newton, Leibniz le puso mucha atención al formalismo y a menudo le dedicaba varios

días a determinar los símbolos apropiados para cada concepto.

La controversia entre Newton y Leibniz sobre la invención del cálculo

Usualmente se acredita tanto a Leibniz como a Newton con la invención del cálculo. Newton fue el primero en aplicar el cálculo a la física general y Leibniz desarrolló la notación usada en cálculo hasta nuestros días. Las ideas principales que ambos Newton y Leibniz estipularon fueron las leyes de diferenciación e integración, las segundas derivadas, las derivadas de orden superior, y la noción de una aproximación de series de potencias, conocida actualmente como las series de Taylor.

Cuando Newton y Leibniz primero publicaron sus resultados, hubo gran controversia sobre qué matemático (y por ende qué país) merecía el crédito por la invención de esta disciplina. Newton llegó primero a sus resultados, pero Leibniz publicó primero. Newton acusó a Leibniz de robar sus ideas de sus notas inéditas, las cuales Newton había compartido con unos cuantos miembros de la Royal Society. Esta controversia dividió a los matemáticos de habla inglesa de los matemáticos continentales por varios años, causando un retraso de las matemáticas inglesas. Un cuidadoso examen de los papeles de ambos matemáticos demuestra que ellos llegaron a sus resultados independientemente, con Leibniz empezando primero con la integración y Newton con la diferenciación. Hoy, se les da crédito a ambos matemáticos por desarrollar el cálculo independientemente. Fue Leibniz, sin embargo, quien le dio el nuevo nombre a su disciplina. Newton llamó su cálculo el *método de las fluxiones*. La simbología usada por Newton, tal como $\dot{x}$ (derivada primera), $\ddot{x}$ (derivada segunda), no fue muy afortunada y retrasó el avance del cálculo en Gran Bretaña. En cambio, la simbología de Leibniz fue manejable y apuntaló el progreso del cálculo en Europa.

Desde los tiempos de Leibniz y Newton, muchos matemáticos han contribuido al desarrollo del cálculo. En el siglo XIX, el cálculo comenzó a ser planteado más rigurosamente por matemáticos como Cauchy, Riemann y Weierstrass. También fue en este período que las ideas del cálculo fueron generalizadas al espacio euclidiano y al plano complejo. Lebesgue generalizó la noción de la integral de tal manera que virtualmente cualquier función tenga una integral, mientras que Laurent Schwartz extendió la diferenciación casi de la misma manera.

La importancia del cálculo

En la actualidad el cálculo es un tema omnipresente en la mayoría de los programas de educación superior y en las universidades. Los matemáticos alrededor del mundo continúan contribuyendo al desarrollo de esta disciplina, la cual ha sido considerada como uno de los logros más grandes del intelecto humano. El desarrollo de las ecuaciones diferenciales, inconcebibles sin el cálculo, ha jugado un gran papel de cambio cualitativo en la ciencia y la tecnología, comparable con el control del fuego en la época primitiva, las ecuaciones diferenciales son un salto enorme para la ciencia.

Mientras que algunas ideas del cálculo fueron desarrolladas tempranamente en las matemáticas griegas, chinas, indias, islámicas y japonesas, el uso moderno del cálculo comenzó en Europa, durante el siglo XVII, cuando Isaac Newton y Gottfried Leibniz construyeron con base al trabajo de antiguos matemáticos los principios básicos de esta disciplina. El desarrollo del cálculo fue constituido con base en los conceptos de movimiento instantáneo y el área bajo las curvas.

Las aplicaciones del cálculo diferencial incluyen cómputos que involucran velocidad, aceleración, la pendiente de una recta tangente a una curva y optimización. Las aplicaciones del cálculo integral

están en cómputos que incluyen elementos de área, volumen, centro de masa, longitud de arco, trabajo y presión. Aplicaciones más avanzadas incluyen series de potencias y series de Fourier. El cálculo puede ser usado para computar la trayectoria de una nave acoplándose a una estación espacial o la cantidad de nieve en una calzada para coches.

El cálculo es también usado para obtener un entendimiento más preciso de la naturaleza del espacio, el tiempo y del movimiento. Por siglos, matemáticos y filósofos lucharon con paradojas que involucraban la división por cero o sumas de series infinitas de números. Estas preguntas surgen en el estudio del movimiento y área. El antiguo filósofo griego Zenón dio varios ejemplos famosos de tales paradojas. El cálculo provee herramientas que pueden resolver tales paradojas, especialmente los límites y las series infinitas.

Límites vs infinitesimales. Los fundamentos del cálculo.

En matemáticas, los fundamentos se refieren al desarrollo riguroso de un tema a partir de axiomas y definiciones precisas. El obtener un fundamento riguroso para el cálculo ocupó a los matemáticos durante la mayor parte del siglo que siguió a Leibniz y Newton y es aún un área activa en el siglo XXI. Los fundamentos del cálculo fueron objeto de diversas especulaciones filosóficas e interpretaciones informales, la falta de rigor y laxitud con que fueron afrontados ciertos problemas de fundamentación contribuyeron a la crisis de los fundamentos de las matemáticas.

Sin embargo, ya durante el siglo XIX se empezó a trabajar en una aproximación rigurosa para los fundamentos del cálculo. El más usual hoy en día es el concepto de límite definido en la continuidad de los números reales (el concepto de límite es esencialmente un concepto topológico). Una alternativa es el análisis no estándar, en el cual el sistema de números realeses aumentado con infinitesimales y números infinitos, como en la concepción original de Newton y Leibnitz. Los fundamentos del cálculo son incluidos en el campo del análisis real, el cual contiene las definiciones completas y pruebas matemáticas de los teoremas del cálculo, así como también generalizaciones tales como la teoría de la medida y la teoría de distribuciones.

El cálculo se desarrolló mediante la manipulación de "cantidades pequeñas". Históricamente, el primer método para lograr eso se basaba en los infinitesimales. Estos son objetos que pueden ser tratados como números pero que son, en algún sentido, "infinitamente pequeños". Tratándose de números, éstos serían puntos que no son cero, pero que tienen una distancia cero del número 'cero'. Desde este punto de vista, el cálculo es una colección de técnicas para manipular infinitesimales. Este punto de vista perdió terreno en el siglo XIX porque era difícil lograr una noción precisa del infinitesimal. El concepto cobró fuerza nuevamente en el siglo XX con la introducción del análisis no estándar y del "análisis infinitesimal suave" (del inglés smooth infinitesimal analysis) , los que proporcionaron fundamentos sólidos para la manipulación de infinitesimales.

En el siglo XIX, los infinitesimales fueron reemplazados por los límites. Los límites describen el valor de una función en un cierto valor de entrada en términos de sus valores en un punto cercano. Capturan el comportamiento a pequeña escala, como los infinitesimales, pero usan el sistema ordinario de los números reales. En este contexto, el cálculo es una colección de técnicas usadas para la manipulación de ciertos límites. Los infinitesimales son reemplazados por números muy pequeños y el comportamiento infinitamente pequeño de la función es encontrado mediante el comportamiento límite para números cada vez más pequeños. Los límites son fácil de poner en fundamentos, y por esta razón son usualmente considerados como el acercamiento estándar al cálculo.

Ambos cálculos, el diferencial y el integral fueron el resultado de muchas ideas acumuladas durante siglos, de pensadores como Eudoxo, Arquímedes, Nicole Oresme, Galilei, Kepler, Descartes, Fermat, Cavallieri, Torricelli, e Isaac Barrow, maestro de Newton. Apoyados en las ideas acumuladas durante tantos siglos, Isaac Newton y Wilhelm Leibniz, casi simultáneamente y de manera independiente uno del otro, dieron origen al Cálculo Diferencial e Integral. A Newton se deben sobre todo las aplicaciones de esta rama de las matemáticas a la física, sobre todo al movimiento de los cuerpos. A Leibniz se deben sobre todo la notación que se usa en el cálculo para denotar derivadas e integrales, que no es un logro menor. Pero a ambos se debe la unificación de ambas ramas, el Cálculo Diferencial e Integral con la aplicación sistemática del llamado Teorema Fundamental del Cálculo para generar una tecnología que permite resolver problemas de integración utilizando el proceso inverso a la derivación.

4.2. El cálculo

El cálculo diferencial

Línea tangente en $(x, f(x))$. La derivada $f'(x)$ de una curva en un punto es la pendiente de la línea tangente a la curva en ese punto. El cálculo diferencial es el estudio de la definición, propiedades, y aplicaciones de la derivada de una función, o lo que es lo mismo, la pendiente de la tangente a lo largo de su gráfica. El proceso de encontrar la derivada se llama derivación o diferenciación. Dada una función y un punto en su dominio, la derivada en ese punto es una forma de codificar el comportamiento a pequeña escala de la función cerca del punto. Encontrando la derivada de una función para cada punto en su dominio, es posible producir una nueva función, llamada la *función derivada* o simplemente *la derivada* de la función original. En lenguaje técnico, la derivada es un operador lineal, el cual toma una función y devuelve una segunda función, de manera que para cada punto de la primera, la segunda proporciona la pendiente de la recta tangente a la gráfica en ese punto.

Para entender la derivada es importante familiarizarse con su notación. Un símbolo común para la derivada de una función es una marca parecida a un acento o apostrofo llamada símbolo primo. Así la derivada de f es f' (pronunciado "efe prima"). En lo siguiente la segunda función es la derivada de la primera:

$$\text{si } f(x) = mx + b, \text{ entonces } f'(x) = m$$
$$\text{si } f(x) = x^2, \text{ entonces } f'(x) = 2x$$
$$\text{si } f(x) = sen(x), \text{ entonces } f'(x) = cos(x)$$
$$\text{si } f(x) = log(x), \text{ entonces } f'(x) = \frac{1}{x}$$
$$\text{si } f(x) = e^x, \text{ entonces } f'(x) = e^x$$

Si la variable de la función representa tiempo, entonces la derivada representa el cambio de la función con respecto del tiempo. Por ejemplo, si f es una función que tiene al tiempo como variable y sus valores dan la posición de una pelota en cada momento, entonces la derivada de f determina cuánto está cambiando en el tiempo la posición de la pelota, esto es, la velocidad de la pelota.

El cálculo integral y su teorema fundamental

El cálculo integral es el estudio de las definiciones, propiedades, y aplicaciones de dos conceptos relacionados, la integral indefinida y la integral definida. El proceso de encontrar el valor de una integral es llamado integración. En lenguaje técnico, el cálculo integral estudia dos operadores lineales relacionados.

La integral indefinida es la antiderivada, es decir, la operación inversa de la derivada. La función F es una integral indefinida de la función f cuando f es una derivada de F. (El uso de mayúsculas y minúsculas para distinguir entre la función y su integral indefinida es común en el cálculo).

La integral definida transforma una función definida en un intervalo $[a, b]$ en un número que se denota por

$$\int_a^b f(x)dx$$

y que es el valor del área entre una curva definida por $y = f(x)$ y el eje-x, en el intervalo $[a, b]$. La definición técnica de la integral se realiza como el límite de sumas de áreas de rectángulos, que se llaman sumas de Riemann:

$$\int_a^b f(x)dx = \lim_{||P||\to 0} \sum_{n=1}^{N} f(x_n)\Delta x_n$$

donde P representa una partición del intervalo $[a, b]$ en subintervalos de longitudes $\{\Delta x_n\}_{n=1}^{n=N}$, cada x_n es un punto del intervalo de longitud Δx_n y $||P||$ es la *norma de la partición* P que se define como el máximo de los valores Δx_n.

El teorema fundamental del cálculo establece que la diferenciación y la integración son operaciones inversas. Más precisamente, relaciona los valores de las antiderivadas para definir las integrales. Ya que es normalmente más fácil computar una antiderivada que aplicar la definición de una integral definida, el teorema fundamental del cálculo provee una forma práctica de computar integrales definidas. También puede ser interpretado como una declaración precisa del hecho de que la diferenciación es la inversa de la integración.

El teorema fundamental del cálculo establece: Si una función f es continua en el intervalo $[a, b]$ y si F es una función cuya derivada en el intervalo (a, b) es f, entonces

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

Así entonces, para cada x en el intervalo (a, b), es cierto que:

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x)$$

Este hecho, descubierto tanto por Newton como Leibniz, quienes basaron sus resultados en el trabajo previo de Isaac Barrow, fue clave para la masiva proliferación de resultados analíticos luego que su trabajo fuese conocido. El teorema fundamental provee un método algebraico para calcular

muchas integrales definidas – sin necesidad de calcular límites – mediante el encuentro de fórmulas apropiadas para las antiderivadas.

Otra versió de Teorema nos dice que si a partir de una función $f(x)$ creamos otra $F(x)$ definiéndola como el área bajo la gráfica de $f(x)$ entre un punto fijo a y uno variable x, entonces $F(x)$ se puede derivar, es decir, se pueden calcular las pendientes de su gráfica y están dadas por la función original $f(x)$. En la notación de Leibniz,

$$\text{si } F(x) = \int_a^x f(x)dx, \text{ entonces } \frac{dF}{dx} = f(x)$$

Como ya dijimos, este resultado permite obtener las áreas bajo muchas curvas mediante el procedimiento de representarlas como funciones que son las derivadas de otras, de manera que para calcular el área bajo la gráfica de ellas basta tomar la diferencia de las segunas entre los dos extremos del intervalo de integración: Si $f(x) = \frac{dF}{dx}$, entonces la integral de a a b de $f(x)$ es igual a $F(b) - F(a)$. Y esto crea un método muy general para obtener las áreas bajo curvas, para lo cual basta tener una larga lista de funciones y sus derivadas, es decir una tabla de derivadas o, vista en orden inverso, una tabla de integrales. Cabe aclarar que el proceso de obtener las derivadas de muchas funciones es relativamente simple, gracias a unos cuantos resultados básicos que se pueden aplicar recursivamente para cubrir muchos casos de interés.

El aprendizaje del cálculo

Conocer y dominar el cálculo consiste en lo siguiente:

1) Descubrir cómo obtener el área bajo la gráfica de una función escalonada (es decir que es constante en intervalos consecutivos) sumando términos de la forma $f(\xi_i)(x_{i+1} - x_i)$ donde $f(\xi_i)$ es el valor constante de f entre x_i y x_{i+1}.

2) Utilizando una función $f(x)$ que es la derivada de otra función $F(x)$ descubrir el Teorema Fundamental del Cálculo que le permite obtener el área bajo la gráfica de $f(x)$ entre $x = a$ y $x = b$ como

$$\int_a^b f(x)dx = F(b) - F(a)$$

. Para ello aproximar la función $f(x)$ en el intervalo $[a, b]$ por una función escalonada con valores

$$f(\xi_i) = \frac{F(x_{i+1}) - F(x_i)}{x_{i+1} - x_i}$$

en los intervalos de la partición $a = x_0 < x_1 < x_2 < ... < x_n < x_{n+1} = b$. El área bajo la gráfica de esta función es

$$\sum_{i=0}^{n} f(\xi_i)(x_{i+1} - x_i) = \sum_{i=0}^{n} F(x_{i+1}) - F(x_i) = F(b) - F(a)$$

Tomando particiones cada vez más finas se puede concluir que

$$\int_a^b f(x)dx = F(b) - F(a)$$

Esto demuestra el Teorema Fundamental del Cálculo que permite calcular fácilmente el área bajo la gráfica de una función $f(x)$ si se conoce su antiderivada $F(x)$, es decir una función tal que

$$F'(x) = \frac{dF}{dx} = f(x)$$

3) Conocer el concepto de integral indefinida o antiderivada y saber que dos antiderivadas de una función sólo difieren en una constante, por lo cual en general éstas se denotan como:

$$\int f(x)dx = F(x) + c$$

donde c es una constante arbitraria.

4) Entender el significado histórico de este descubrimiento. El Cálculo se convirtió, a partir del siglo XVIII, en una herramienta matemática esencial que permitía resolver problemas como el cálculo de áreas de figuras curvas, volúmenes de todo tipo de cuerpos, longitudes de todo tipo de curvas, trayectorias de partículas sujetas a fuerzas conocidas y otras cantidades de interés científico como el trabajo realizado por una fuerza, o la media y la desviación estándar de una distribución de probabilidad, que anteriormente eran intratables o ni siquiera podían definirse con precisión.

Gracias al *Teorema Fundamental del Cálculo* que relaciona la integral con la derivada, el Cálculo Diferencial e Integral se convirtió en una poderosa herramienta que ha sido por siglos el lenguaje común de las Ciencias y además abrió una nueva era de las matemáticas participando íntimamente en el desarrollo de varias nuevas áreas como la teoría de Ecuaciones Diferenciales, el Análisis Real y Complejo, el Análisis Funcional y la Geometría Diferencial. Los detalles de todos los desarrollos descritos brevemente en este capítulo, así como algunos de los desarrollos posteriores al Cálculo podrán encontarse en el resto de los capítulos del libro.

Las ecuaciones diferenciales y las aplicaciones del cálculo

Las ecuaciones diferenciales relacionan a una función con sus derivadas, y son omnipresentes en las ciencias pues se utilizan para plantear modelos cuantitativos del comportamiento de diversos sistemas.

La Espiral logarítmica de la concha del Nautilus es una clásica imagen usada para representar el crecimiento y cambio relacionados con el cálculo. El cálculo es usado en cada rama de las ciencias físicas y de informática, estadística, ingeniería, economía, negocios, medicina, demografía y en otras áreas donde un problema pueda ser modelado matemáticamente y una solución óptima sea deseada.

La física hace un particular uso del cálculo; todos los conceptos en la mecánica clásica están relacionados a través del cálculo. La masa de un objeto, el momento de inercia de los cuerpos, la energía total de un objeto dentro de un campo conservativo, todo esto se define usando el cálculo. Un ejemplo históricamente más impactante del uso del cálculo en la física son las leyes del movimiento de Newton, donde se usa expresamente la tasa de cambio de la posición para definir la velocidad y la tasa de cambio de la velocidad para definir la aceleración y finalmente para enunciar la segunda ley del movimiento de Newton: la tasa de cambio de momentum (el momentum es el producto e la masa por la velocidad) de un cuerpo es igual a la fuerza que actúa sobre el cuerpo. Incluso la expresión común de la segunda ley de Newton como fuerza = masa $\times$ aceleración involucra el cálculo

diferencial porque la aceleración debe entenderse como la razón de cambio, o la derivada respecto al tiempo, de la velocidad.

En la electricidad y el magnetismo, el cálculo se usa para describir los campos electromagnéticos. Gracias a tal descripción Maxwell descubrió las ondas electromagnéticas ya que las ecuaciones del campo electromagnético pueden manipularse para obtener una ecuación de onda. También la Teoría de la relatividad general de Einstein y la mecánica cuántica se expresan en el lenguaje del cálculo y las ecuaciones diferenciales.

El cálculo también puede ser usado en conjunto con otras disciplinas matemáticas. Por ejemplo, puede ser usado con el álgebra lineal para encontrar la mejor aproximación lineal para un conjunto de puntos en un dominio. También puede ser usado en la teoría de la probabilidad para determinar la probabilidad de una variable continúa al azar desde una densidad de función asumida.

El teorema de Green, el cual establece la relación entre una integral lineal alrededor una simple curva cerrada C y una doble integral sobre el plano de región D delimitada por C, es aplicado en un instrumento conocido como planímetro, el cual es usado para calcular el área de una superficie plana en un dibujo. Por ejemplo, puede ser usado para calcular la cantidad de área que toma una piscina cuando se bosqueja el diseño de un pedazo de propiedad.

En geometría analítica, el estudio de los gráficos de funciones, el cálculo es usado para encontrar puntos máximos y mínimos, la tangente, así también como para determinar la concavidad y los puntos de inflexión.

La química también usa el cálculo para determinar los ritmos de las reacciones y el decaimiento radioactivo.

En la medicina, el cálculo puede ser usado para encontrar el ángulo de ramificación óptimo de vaso sanguíneo para maximizar el flujo.

En economía, el cálculo permite determinar el beneficio máximo por medio del costo marginal y del ingreso marginal.

El cálculo también puede ser usado para encontrar soluciones aproximadas para ecuaciones, usando métodos como por ejemplo el método de Newton, la iteración de punto fijo y la aproximación lineal. Por ejemplo, las naves espaciales usan una variación del método de Euler para aproximar trayectorias curvas dentro de entornos de gravedad cero.

4.3. Las funciones en el cálculo

Todo el lenguaje del cálculo se basa en el concepto de función. Concretamente, se basa en las llamadas funciones reales de variable real, que son aquellas cuyas gráficas pueden representarse en un plano cartesiano como lugares geométricos consistentes en aquellos puntos (x, y) del plano para los cuales $y = f(x)$.

La geometría analítica de Fermat y Descartes había preparado el terreno para la creación del concepto de función como se describió en el párrafo anterior. Además la propia geometría analítica proporcionó muchos ejemplos concretos de funciones que pudieron aplicarse directamente al cálculo. Para empezar, los de función lineal $y = ax + b$ y función cuadrática $y = ax^2 + bx + c$ sirven al cálculo como puntos de partida para comprobar la validez de sus métodos en unos casos en que las derivadas y las integrales pueden encontrarse fácilmente.

Por ejemplo, si $y = mx$ entonces la razón de cambio $\frac{\Delta y}{\Delta x}$ es igual a m en cualquier parte de la gráfica. Y es que la recta tangente a la gráfica de la función $f(x) = mx$ coincide exactamente con la recta $y = mx$. En otras palabras, si $f(x) = mx$, entonces $f'(x) = m$. Por otro lado, en este caso la región entre la curva y el eje-x es siempre un trapecio y su área puede calcularse directamente:

$$\int_a^b f(x)dx = \frac{f(a) + f(b)}{2}(b - a) = \frac{(mb + ma)(b - a)}{2} = \frac{mb^2}{2} - \frac{ma^2}{2}$$

De manera que si se define la función $F(x) = \frac{mx^2}{2}$, se tiene que

$$\int_a^b f(x)dx = F(b) - F(a)$$

Finalmente, es fácil comprobar que $F'(x) = mx$. La gráfica de $y = \frac{mx^2}{2}$ es una parábola con foco en el punto $(0, \frac{1}{2m})$ y directriz $y = -\frac{1}{2m}$. Sea $(a, \frac{ma^2}{2})$ un punto cualquiera de esta parábola. La recta tangente a la parábola en ese punto es la mediatriz de los puntos $(a, -\frac{1}{2m})$ y$(0, \frac{1}{2m}$ (el foco) y por tanto es ortogonal a la recta que los une, la cual tiene pendiente $-\frac{1}{ma}$). Por tanto la pendiente de la recta tangente a la gráfica de $y = \frac{mx^2}{2}$ en el punto $(a, \frac{ma^2}{2})$ es ma. Esto demuestra que la derivada de $f(x) = \frac{mx^2}{2}$ es $f'(x) = mx$.

Otra manera de obtener la recta tangente en el punto $(a, \frac{ma^2}{2})$ es considerar directamente la ecuación de la mediatriz de los puntos $(0, \frac{1}{2m})$ y $(a, -\frac{1}{2m})$. Su ecuación es: $x^2+(y-\frac{1}{2m})^2 = (x-a)^2+(y+\frac{1}{2m})^2$, que al desarrollarse y tras simplificar los cuadrados de x, y y $\frac{1}{2m}$ (los tres desaparecen) queda

$$-\frac{2y}{2m} = -2ax + a^2 + \frac{2y}{2m}$$

que puede escribirse como

$$y = max - \frac{ma^2}{2}$$

que es la ecuación de una recta con pendiente ma, lo que constrituye otra prueba de que la derivada de $f(x) = \frac{mx^2}{2}$ es $f'(x) = mx$.

Pero hay otra prueba de este hecho que muestra un método más general para obtener la derivada de ua función y es la que corresponde a lo que propiamente se considera el cálculo diferencial. La idea es ésta: la pendiente de la recta tangente a la gráfica $y = f(x)$ de una función cualquiera $f(x)$ en el punto $(x, f(x))$ es el límite de las pendiente de la recta (secante) que pasa por los dos puntos cercanos $(x, f(x))$ y $(x+h, f(x+h))$, cuando h tiende a cero. En símbolos esto lo escribimos así

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}$$

A pesar de que aparentemente esta definición de la derivada es más difícil de entender que alguna de las versiones alternativas más intuitivas, resulta ser la única de carácter completamente general. Y

además, su aplicación para obtener derivadas es más sencilla que los métodos que nos proporciona la geometría analítica y que solo aplican a funciones muy particulares, como es el caso anterior de $f(x) = \frac{mx^2}{2}$.

Efectivamente, si aplicamos la definición a esta función vemos que debemos obtener el límite cuando $h \to 0$ de

$$\frac{\frac{m(x+h)^2}{2} - \frac{mx^2}{2}}{h}$$

Desarrollando el binomio en el numerador, simplificando los términos en x^2 que se anulan y dividiendo término a término por h, se tiene

$$\frac{\frac{m(x+h)^2}{2} - \frac{mx^2}{2}}{h} = \frac{\frac{mx^2}{2} + \frac{2mxh}{2} + \frac{mh^2}{2} - \frac{mx^2}{2}}{h} = \frac{mxh + \frac{mh^2}{2}}{h} = mx + \frac{mh}{2}$$

Así resulta muy fácil ver que

$$\lim_{h\to 0} \frac{\frac{m(x+h)^2}{2} - \frac{mx^2}{2}}{h} = \lim_{h\to 0}(mx + \frac{mh}{2}) = mx$$

y por tanto, que la derivada de $f(x) = \frac{mx^2}{2}$ es $f'(x) = mx$.

Las dos grandes ventajas del método del cálculo diferencial para obtener derivadas son que se puede aplica a muchas funciones y que sus resultados son muy generales, de manera que solo hay que aplicarlos una vez a ciertas funciones y, utilizando algunas reglas fáciles de aprender y de justificar, resulta muy fácil obtener las derivadas de un gran número de funciones.

La tecnología del cálculo diferencial tiene dos aspectos:

(1) Encontrar las derivadas de las principales funciones de uso común como las potencias de x, las funciones trigonométricas, la función exponencial y la función logaritmo.

(2) Obtener las reglas para derivar funciones que se forman por medio de sumas, multiplicaciones y divisiones de otras funciones, por la composición de funciones y por la definición de una función inversa.

En las dos secciones siguientes se presentan, en orden inverso, y de manera resumida, estos dos aspectos del cálculo diferencial. Todo esto se hizo lo hicieron los creadores del cálculo diferencial sin tener una definición precisa del concepto de límite. Y es que como se verá, para los propósitos planteados, los límites que se calculan son bastante sencillos y pueden obtenerse sin necesidad de profundizar mucho en el concepto de límite. Sin embargo, el de límite es un concepto tan fundamental para el desarrollo de as matemáticas contemporáneas, que su comprensión es imprescindible, por lo que se estudiará su desarrollo detenidamente en las secciones posteriores.

Reglas para derivar las funciones formadas por operaciones con otras

A continuación presentamos las reglas de derivación para las funciones definidas mediante operaciones de otras funciones. Demostraremos sólo dos de estas reglas, las que resultan un poco más

complicadas de probar. Casi todas ellas las enunció y demostró Gottfried Leibniz. Al enunciar la reglas usaremos las dos notaciones, la de f' y la de $\frac{df}{dx}$ para denotar las derivadas.

Regla de la suma

$$(f+g)'(x) = f'(x) + g'(x) \text{ equivalentemente: } \frac{d(f+g)}{dx} = \frac{df}{dx} + \frac{dg}{dx}$$

Regla del producto

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x) \text{ equivalentemente: } \frac{d(fg)}{dx} = \frac{df}{dx}g + f\frac{dg}{dx}$$

Regla de la cadena para composición de funciones:

$$(f(g))'(x) = f'(g(x))g'(x) \text{ equivalentemente: } \frac{d(f(g))}{dx} = \frac{df}{dx}(g)g'$$

La demostración de la regla de la cadena se hace dividiendo y multiplicando expresión cuyo límite hay que obtener por $g(x+h) - g(x)$:

$$\frac{f(g(x+h)) - f(g(x))}{h} = \frac{f(g(x+h)) - f(g(x))}{g(x+h) - g(x)} \frac{g(x+h) - g(x)}{h}$$

Si la función $g(x)$ es constante en una vecindad de x, la división no es válida, y en ese caso está claro que $f(g(x+h)) - f(g(x))$ es igual a 0 para h suficientemente pequeña y por tanto en ese caso tanto la derivada de $g(x)$ como la de $f(g(x))$ son cero y por tanto la regla de la cadena se umple de manera trivial en ese caso. Si $g(x)$ no es constante en una vecindad de x, entonces $g(x+h)$ es un número cercano a $g(x)$ (aquí, claro, se está suponiendo que g es una función bien portada, es decir, continua y derivable) y por tanto el primer factor del segundo término de la igualdad anterior tiende a $f'(g(x))$ cuando $h \to 0$. El segundo factor, por definición, tiene a $g'(x)$ cuando $h \to 0$. Con esto queda demostrada la regla de la cadena.

Regla de la función inversa

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}$$

Ésta última regla se obtiene derivando ambos lados de la expresión $f(f^{-1}(x)) = x$, que es la que define a la función inversa. Al aplicar la regla de la cadena al miembro izquierdo y derivar diretamente el derecho, se obtiene:

$$f'(f^{-1}(x))f^{-1}(x) = 1$$

La regla se obtiene despejando $f^{-1}(x)$ de esta igualdad.

La **regla de la división de funciones**

$$(\frac{f}{g})'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2} \text{ equivalentemente: } \frac{d}{dx}(\frac{f}{g}) = \frac{\frac{df}{dx}g - f\frac{dg}{dx}}{g^2}$$

es muy útil también y puede obtenerse aplicando sucesivamente la regla del producto a $f\frac{1}{g}$, luego la regla para derivar potencias, que aparece en la siguiente subsección, y la regla de la cadena. Éste es un magnífico ejemplo de cómo funciona la tecnología del cálculo diferencial. Dejamos los detalles como ejercicio para el lector.

Derivadas de las funciones comunes

Derivadas de las potencias x^p

La fórmula

$$\frac{dx^p}{dx} = px^{p-1}$$

para derivar potencias x, es decir, funciones de la forma $f(x) = x^p$ es muy general. En realidad, aplica a cualquier número real p, pero sólo vamos a demostrarla para valores enteros, tanto positivos como negativos de p y para valores fraccionarios de $p = \frac{m}{n}$ con m y n números enteros arbitrarios. La demostración se basa en la fórmula binomial:

$$(a+b)^n = a^n + na^{n-1}b + \frac{n(n-1)}{2!}a^{n-2}b^2 + + \frac{n(n-1)(n-2)}{3!}a^{n-3}b^3 + ... + b^n$$

Si aplicamos la fórmula binomial al primer sumando del numerador de la expresión cuyo límite hay que calcular para obtener la derivada de x^n, tendremos:

$$\frac{(x+h)^n - x^n}{h} = \frac{x^n + nx^{n-1}h + \frac{n(n-1)}{2!}x^{n-2}h^2 + \frac{n(n-1)(n-2)}{3!}x^{n-3}h^3 + ... + h^n - x^n}{h}$$

Cancelando x^n con $-x^n$ y dividiendo por h término a término, obtenemos

$$\frac{(x+h)^n - x^n}{h} = nx^{n-1} + \frac{n(n-1)}{2!}x^{n-2}h + \frac{n(n-1)(n-2)}{3!}x^{n-3}h^2 + ... + h^{n-1}$$

Todos los términos de esta suma, excepto el primero, tienen como factor alguna potencia positiva de h, por lo tanto todos ellos, salvo el primero, tienden a 0 cuando $h \to 0$. Por lo tanto

$$\lim_{h \to o} \frac{(x+h)^n - x^n}{h} = nx^{n-1}$$

lo cual demuestra la fórmula $\frac{dx^p}{dx} = px^{p-1}$ para el caso en que p es un entero positivo.

Para el caso en que p es un entero negativo $p = -n$ aplicamos la regla del producto a la función $f(x) = x^n x^{-n} = 1$ y obtenemos

$$nx^{n-1}x^{-n} + x^n \frac{dx^{-n}}{dx} = 0$$

igualdad de la cual podemos despejar $\frac{dx^{-n}}{dx}$, con lo que obtenemos:

$$\frac{dx^{-n}}{dx} = -\frac{nx^{n-1}x^{-n}}{x^n} = -nx^{-n-1}$$

lo cual demuestra nuestra fórmula para el caso $p = -n$.

Para demostrar el caso en que $p = \frac{1}{n}$ aplicamos la regla de la cadena a la función

$$f(x) = (x^{\frac{1}{n}})^n = x$$

obteniendo

$$n(x^{\frac{1}{n}})^{n-1}\frac{dx^{\frac{1}{n}}}{dx} = 1$$

Despejando $\frac{dx^{\frac{1}{n}}}{dx}$ obtenemos

$$\frac{dx^{\frac{1}{n}}}{dx} = \frac{1}{nx^{1-\frac{1}{n}}} = \frac{1}{n}x^{\frac{1}{n}-1}$$

que demuestra la fórmula para el caso $p = \frac{1}{n}$.

Finalmente, para el caso $p = \frac{m}{n}$ basta aplicar la regla de la cadena a la función $f(x) = (x^{\frac{1}{n}})^m = x^{\frac{m}{n}}$. Dejamos los detalles de este último paso como ejercicio para el lector.

Es importante observar que la demostración de cada caso casi siempre se realiza utilizando un caso anterior y la aplicación de una varias de las reglas de derivación para funciones formadas con operaciones de otras funciones. La clave está, claro en la inventar estas funciones de manera adecuada para que nos permitan obtener la derivada buscada, en general, despejándola de una ecuación a la que se llega aplicando las reglas mencionadas.

Derivadas de las funciones trigonométricas

Las fórmula de derivación para el seno y el coseno

$$\frac{dsen(x)}{dx} = cos(x) \quad \text{y} \quad \frac{dcos(x)}{dx} = -sen(x)$$

se obtiene aplicando las fórmulas para el seno y el coseno de una suma

$$sen(\alpha+\beta) = sen(\alpha)cos(\beta) + cos(\alpha)sen(\beta)$$
$$cos(\alpha+\beta) = cos(\alpha)cos(\beta) - sen(\alpha)sen(\beta)$$

a las expresiones cuyos límites hay que calcular para obtener las derivadas correspondientes

$$\frac{sen(x+h)-sen(x)}{h} = \frac{sen(x)cos(h)+cos(x)sen(h)-sen(x)}{h} = sen(x)\frac{cos(h)-1}{h} + cos(x)\frac{sen(h)}{h}$$

y

$$\frac{cos(x+h)-cos(x)}{h} = \frac{cos(x)cos(h)-sen(x)sen(h)-cos(x)}{h} = cos(x)\frac{cos(h)-1}{h} - sen(x)\frac{sen(h)}{h}$$

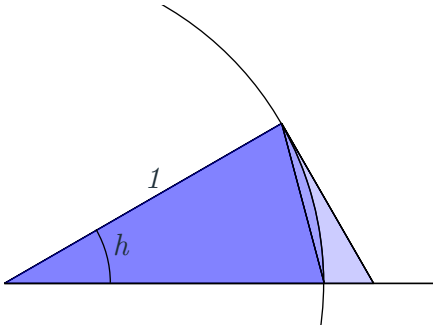
Por tanto, para obtener las derivadas del seno y el coseno basta demostrar que

$$\lim_{h\to 0} \frac{1-cos(h)}{h} = 0 \tag{4.1}$$

y

$$\lim_{h\to 0} \frac{sen(h)}{h} = 1 \tag{4.2}$$

Estos resultados se pueden demostrar compando las tres áreas sombreadas en este esquema



$$\frac{sen(h)}{2} < \frac{h}{2} < \frac{tan(h)}{2}$$

Si multiplicamos por 2 y dividimos por $sen(h)$

$$1 < \frac{h}{sen(h)} < \frac{1}{cos(h)}$$

Invertimos la desigualdad y obtenemos

$$cos(h) < \frac{sen(h)}{h} < 1$$

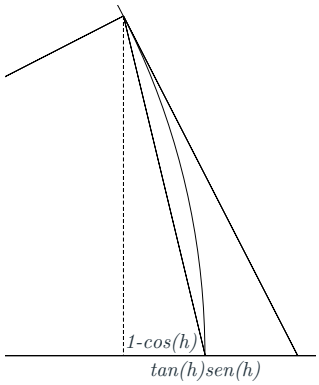
Esto demuestra (4.2). Por otro lado, comparando segmentos sobre el eje x, se ve que

$$1-cos(h) < tan(h)sen(h)$$

Dividiendo por h:

$$\frac{1-cos(h)}{h} < tan(h)\frac{sen(h)}{h} < tan(h)$$

lo cual demuestra (4.1).



Derivadas de las funciones logaritmo y exponencial

Hay varias maneras de definir las funciones exponencial y logaritmo. Una de ella es definir la exponencial aritméticamente, que aparentemente es la más natural. Para esto se parte de

$$a^m = a \cdot a \cdot ... \cdot a \quad (m \text{ factores})$$

para m entero positivo. A continuación se define $b = a^{\frac{m}{n}}$ como el número para el cual $b^n = a^m$. Luego se define para una fracción p, $b = a^{-p}$ como el número para el cual $b \cdot a^p = 1$. Finalmente, se comprueba que $a^{p+q} = a^p a^q$ para cualquier par de fracciones p y q.

Pero aquí aparece una dificultad. ¿Cómo definir, por ejemplo, $a^{\sqrt{(2)}}$ o a^π? Definir a^b para números irracionales b requiere de realizar una extensión de la definición de exponenciación. Se expresar b como el límite de una sucesión de números racionales $\{p_n\}$ tal que $b = \lim_{n\to\infty} p_n$, y *definir*

$$a^b = \lim_{n\to\infty} a^{p_n}$$

.

Para justificar esta definición hace falta demostrar que la exponenciación en suna función *continua* en los racionales. Dejamos los detalles al lector. Finalmente, se puede ver que

$$\frac{da^x}{dx} = \lim_{h\to 0} \frac{a^{x+h} - a^x}{h} = a^x \lim_{h\to 0} \frac{a^h - 1}{h}$$

Para completar la construcción de la función exponencial por este camino, se define el número e como aquel para el cual

$$\lim_{h\to 0} \frac{e^h - 1}{h} = 1$$

En particular,

$$\lim_{n\to\infty} \frac{e^{\frac{1}{n}} - 1}{\frac{1}{n}} = 1$$

Si definimos $e_n = (1 + \frac{1}{n})^n$, vemos que para todo n,

$$\frac{e_n - 1}{\frac{1}{n}} = 1$$

lo cual sugiere que podríamos definir

$$e = \lim_{n\to\infty} e_n = \lim_{n\to\infty} (1 + \frac{1}{n})^n$$

Finalmente se define la función exponencial como

$$exp(x) = e^x$$

Todos estos pasos pueden justificarse rigurosamente, y se hace en libros de análisis y en algunos libros de cálculo.

Otro camino que se puede seguir para definir la función exponencial es usar el límite

$$exp(x) = \lim_{n\to\infty}(1+\frac{x}{n})^n$$

este camino presenta una dificultad de inicio ya que aunque a partir de ella resulta natural definir el número e como

$$e = exp(1) = \lim_{n\to\infty}(1+\frac{1}{n})^n$$

nada nos garantiza que

$$exp(x) = e^x$$

Otra definición equivalente de la función exponencial es

$$exp(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + ...$$

que presenta el mismo problema de entrada que la anterior.

En cualquiera de las opciones que hemos mencionado hasta ahora, la definición de la función logaritmo se hace como la función inversa de la exponencial. Es decir,

$$log(x) = y \text{ significa que } x = exp(y)$$

Finalmente, hay una manera de definir primero la función logaritmo como

$$log(x) = \int_1^x \frac{1}{z}dz$$

y entonces definir la función exponencial como la función inversa del logaritmo, es decir,

$$exp(x) = y \text{ significa que } x = log(y)$$

Explorar todos estos caminos enriquece nuestra comprensión de estas dos importantes funciones.

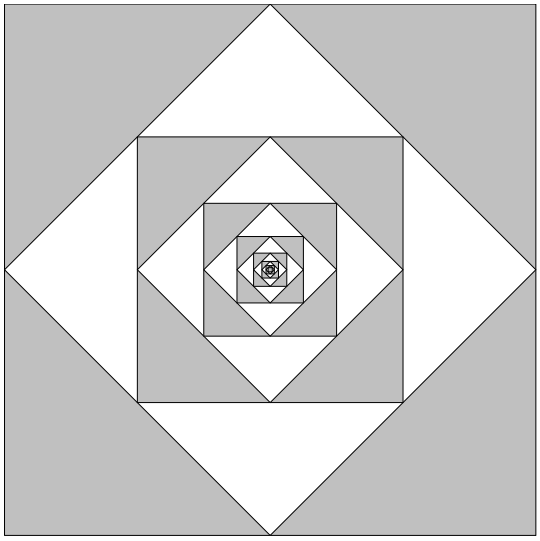
4.4. Desarrollo del concepto de límite

En lo que sigue el concepto de *límite* jugará un papel importante. En el Capítulo 3 se vieron algunos ejemplos donde se usó la idea de la aproximación: en la Sección 3.5 se vió cómo se puede aproximar la longitud de una circunferencia por polígonos inscritos y circunscritos, en la Sección 3.6 se obtuvo la superficie de la esfera por aproximación y en la Sección 3.4 se obtuvo el área de un segmento de parábola, también por aproximación.

Aquí se quiere retomar esta idea, analizarla y generalizarla a todo tipo de sucesiones. Veamos ahora algunos ejemplos de sucesiones.

Ejemplos de sucesiones convergentes y divergentes

En la siguiente figura se muestra un cuadrado, que incluye algunos triángulos sombreados. Deseamos calcular el contenido A del área sombreada suponiendo que el lado del cuadrado grande es 1.



Calculemos esta área paso a paso. Vamos a juntar las áreas de los triángulos de igual tamaño para luego sumarlas todas. El área de los primeros cuatro es la mitad del cuadrado mayor, es decir $F_1 = \frac{1}{2}$. La de los siguiente cuatro triángulos es a su vez la mitad del cuadrado de lado $\frac{1}{2}$, es decir es $\frac{1}{2}(\frac{1}{4})$. Con esto obtenemos la siguiente aproximación para A que es $F_2 = \frac{1}{2} + \frac{1}{2}(\frac{1}{4})$. Luego juntamos los cuatro triángulos del siguiente tamaño y obtenemos una nueva aproximación $F_3 = \frac{1}{2} + \frac{1}{2}\left(\frac{1}{4}\right) + \frac{1}{2}\left(\frac{1}{4}\right)^2$. De esta manera obtenemos cada vez mejores aproximaciones para A:

$$F_n = \frac{1}{2}\left[1 + \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \ldots + \left(\frac{1}{4}\right)^{n-1}\right]. \tag{4.3}$$

La pregunta inicial acerca del área A se reformula ahora como: a qué valor tiende F_n cuando n se hace más y más grande? El paréntesis del lado derecho tiene la forma

$$t_n = 1 + x + x^2 + x^3 + \ldots + x^{n-1} \qquad \text{con } x = \tfrac{1}{4}.$$

Si se multiplica t_n por $1 - x$ se obtiene

$$\begin{aligned} t_n \cdot (1-x) &= (1 + x + x^2 + x^3 + \ldots + x^{n-1}) \cdot (1-x) \\ &= 1 + x + x^2 + x^3 + \ldots + x^{n-1} + \\ &\quad - x - x^2 - x^3 - \ldots - x^{n-1} - x^n \\ &= 1 - x^n. \end{aligned}$$

De ahí que

$$t_n = \frac{1-x^n}{1-x}.$$

En el caso del área sombreado de arriba se tiene $x = \frac{1}{4}$ por lo que

$$F_n = \frac{1}{2} \cdot \frac{1 - \left(\frac{1}{4}\right)^n}{1 - \frac{1}{4}} = \frac{1}{2} \cdot \frac{1 - \left(\frac{1}{4}\right)^n}{\frac{3}{4}}. \tag{4.4}$$

Como $\left(\frac{1}{4}\right)^n$ tiende a cero cuando n se crece, el límite al que tiende F_n es:

$$A = \frac{1}{2} \cdot \frac{1 - 0}{\frac{3}{4}} = \frac{1}{2} \cdot \frac{4}{3},$$

es decir

$$A = \frac{2}{3}.$$

Veamos de nuevo la diferencia de las dos representaciones de F_n como se dan en las ecuaciones (4.3) y (4.4). En la ecuación (4.3) se aproxima A por una suma y de antemano no es evidente qué valor se obtendrá. En cambio, en (4.4), se transformó la expresión algebraicamente de manera que se puede ver bien a qué tiende cuando n se hace grande. Se podría decir que la expresión (4.4) está "madura para obtener el límite", pero (4.3) no.

El ejemplo anterior es algo especial en el sentido de que de antemano se sabe que la sucesión F_n tiene que converger: por un lado crece cada vez que n aumenta porque se añade más área, por otro la suma no puede crecer demasiado dado que el área sombreado debe ser menor que 1. En este tipo de ejemplos la noción de *convergencia* es intuitivamente clara. De hecho, durante mucho tiempo a lo largo de la historia de las matemáticas, el desarrollo de esta noción fue guiado por la intuición.

Hay muchos casos en los que la intuición también nos permite determinar si la sucesión converge o no. Por ejemplo la sucesión a_n:

$$a_1 = 1, \quad a_2 = 2, \quad a_3 = 3, \quad a_4 = 4, \quad \ldots, \quad a_n = n, \quad \ldots$$

crece indefinidamente, es decir rebasa cualquier cota, y por lo tanto no converge. Por otro lado, es evidente que la sucesión b_n:

$$b_1 = 1, \quad b_2 = \tfrac{1}{2}, \quad b_3 = \tfrac{1}{3}, \quad b_4 = \tfrac{1}{4}, \quad \ldots, | \quad b_n = \tfrac{1}{n}, \quad \ldots$$

converge a 0. Finalmente la sucesión c_n:

$$c_1 = 1, \quad c_2 = 3, \quad c_3 = 1, \quad c_4 = 3, \quad \ldots, \quad c_n = 1 + (-1)^n, \quad \ldots.$$

no converge porque oscila indefinidamente entre dos valores fijos.

Pero no siempre es tan fácil determinar si una sucesión converge o no. Veamos el siguiente ejemplo:

$$d_1 = 1, \quad d_2 = 1 + \tfrac{1}{2}, \quad d_3 = 1 + \tfrac{1}{2} + \tfrac{1}{3}, \quad \ldots, \quad d_n = 1 + \tfrac{1}{2} + \tfrac{1}{3} + \ldots + \tfrac{1}{n}, \quad \ldots$$

Utilizando una computadora se obtiene

$$d_{10\,000} \approx 9.79, \qquad d_{20\,000} \approx 10.48, \qquad d_{30\,000} \approx 10.89.$$

Aparentemente la sucesión no va a crecer mucho más. Pero ¿nos basta esto para concluir que la sucesión converge?

No olvidemos que la sucesión d_n se construye sumando cada vez más términos. Así que se requiere de un argumento más sólido para determinar si la sucesión converge o no. Como la sucesión crece, es claro que sólo hay dos situaciones posibles: o la sucesión es acotada y entonces converge a algún valor, o no es acotada y tiende a infinito.

El siguiente argumento demuestra que la sucesión d_n diverge. Construimos una segunda sucesión s_n sustituyendo algunos sumandos de d_n por otros más pequeños, de manera que $s_n \leq d_n$ para todo n. Sustitumos el sumando $\frac{1}{k}$ por $\frac{1}{2^p}$ cuando k satisface $2^{p-1} < k < 2^p$. Por ejemplo

$$d_{10} = 1 + \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{4} + \tfrac{1}{5} + \tfrac{1}{6} + \tfrac{1}{7} + \tfrac{1}{8} + \tfrac{1}{9} + \tfrac{1}{10},$$
$$s_{10} = 1 + \tfrac{1}{2} + \underbrace{\tfrac{1}{4} + \tfrac{1}{4}}_{=\frac{1}{2}} + \underbrace{\tfrac{1}{8} + \tfrac{1}{8} + \tfrac{1}{8} + \tfrac{1}{8}}_{=\frac{1}{2}} + \tfrac{1}{16} + \tfrac{1}{16}.$$

Por ello se tiene para $n = 2^k$

$$d_{2^k} \geq s_{2^k} = 1 + k \cdot \tfrac{1}{2},$$

lo que muestra que la sucesión d_n eventualmente rebasa cualquier cota dada y por lo tanto no converge.

El estilo informal del siglo XVIII

Los ejemplos anteriores muestran dos cosas: por un lado en algunos casos se pueden usar argumentos intuitivos para determinar si una sucesión converge y por otro lado que debemos ser muy cuidadosos porque a veces la intuición puede malaconsejarnos. En el siglo XVIII se pudieron determinar los límites de muchas sucesiones utilizando la intuición y el ingenio. Por ejemplo Johannes Bernoulli infirió que

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \ldots = 1. \tag{4.5}$$

¿Como lo hizo? Utilizó la iguadad

$$\frac{1}{k} - \frac{1}{k+1} = \frac{1}{k(k+1)},$$

con la que puede escribir

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \ldots = \left(1 + \frac{1}{2} + \frac{1}{3} + \ldots\right) - \left(\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots\right)$$

que es la diferencia de dos infinitos, pero observó que todos los sumandos se cancelan mutuamente excepto el 1, lo que prueba (4.5). ¿Es válido este argumento? Este argumento no pasaría la exigencia de rigor de hoy día. Sin embargo produjo el resultado correcto.

Ejercicio. Para profundizar en este tema sugerimos al lector producir un ejemplo en el que un argumento similar llevaría a un error. Esto muestra que hace falta un argumento más riguroso para demostrar (4.5). Sugerimos al lector producir tal argumento.

Un resultado sorprendente de esta época, que se debe a Leonhard Euler, es el siguiente:

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}. \tag{4.6}$$

La manera como llegó a ello muestra a la vez su genio y su atrevimiento para descubrir resultados.

Dado que

$$\frac{1}{(k+1)^2} < \frac{1}{k(k+1)}$$

concluyó que

$$\frac{1}{1} + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \ldots < 1 + \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \ldots = 2$$

por lo que la serie de los recíprocos de los cuadrados está acotada y por lo tanto converge. Para obtener el límite observó que si una ecuación polinomial tiene término constante igual a 1 entonces el coeficiente del término lineal es la suma de los recíprocos de sus raíces, con el signo invertido. Por ejemplo para el grado dos, la ecuación polinomial se puede escribir así

$$(x-a)(x-b) = 0,$$

donde las raíces son a y b. Si expandimos el producto y dividimos entre ab obtenemos

$$\tfrac{1}{ab}x^2 - (\tfrac{1}{a} + \tfrac{1}{b})x + 1 = 0,$$

que es una ecuación polinomial con las mismas soluciones y con término constante 1. El coeficiente del término lineal es en efecto la suma de los recíprocos de las raíces, con el signo invertido. Esto se puede generalizar a polinomios de cualquier grado.

Ahora Euler parte de la ecuación $\operatorname{sen}(x) = 0$ y desarrolla $\operatorname{sen}(x)$ en serie de potencias:

$$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots = 0.$$

Al dividir entre x obtiene

$$1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \ldots = 0$$

y al sustituir $x^2 = y$

$$1 - \frac{y}{3!} + \frac{y^2}{5!} - \frac{y^3}{7!} + \ldots = 0. \tag{4.7}$$

Como las raíces de $\sin(x) = 0$ son $x = \pm k\pi$, donde k recorre los números naturales, al dividir entre x se pierde la solución $x = 0$ y al sustituir $x^2 = y$ obtiene la ecuación (4.7) cuyas raíces son $y = k^2\pi^2$, con k natural.

Ahora manipula la ecuación (4.7) como si fuera un polinomio. Dado que el término constante es 1, el coeficiente del término lineal, es decir $-\frac{1}{3!} = \frac{1}{6}$, debería ser igual a la suma de los recíprocos de las raíces, con el signo invertido:

$$-\frac{1}{6} = -\left(\frac{1}{\pi^2} + \frac{1}{4\pi^2} + \frac{1}{9\pi^2} + \frac{1}{16\pi^2} + \ldots\right).$$

Al multiplicar por $-\pi^2$ obtiene (4.6).

Vemos dos aspectos importantes en el desarrollo anterior. Por un lado, Johannes Bernoulli y Leonhard Euler y sus contemporáneos descubrieron resultados muy importntes al manipular

"sumas infinitas" de manera ingeniosa. Por otro lado extiraban, sin justificación alguna, la validez de las propiedades de las sumas finitas a las infinitas. Hoy día sabemos que hacer esto es algo muy delicado y que puede conducir a errores. Curiosamente nunca se ha descubierto que EULER haya cometido un error manipulando sumas infinitas. Tenía una intuición muy profunda que le permitía actuar con gran atrevimiento sin caer nunca en el error. Lamentablemente no todos somos como Euler. Por ello necesitamos navegar por las aguas más seguras que nos brinda el rigor desarrollado en siglos posteriores.

Podemos resumir esta lección de historia diciendo que con la intuición y el atrevimiento se descubre y con el rigor se demuestra. Para hacer matemáticas conviene aprovecharlo todo.

El camino hacia una definición rigurosa de límite

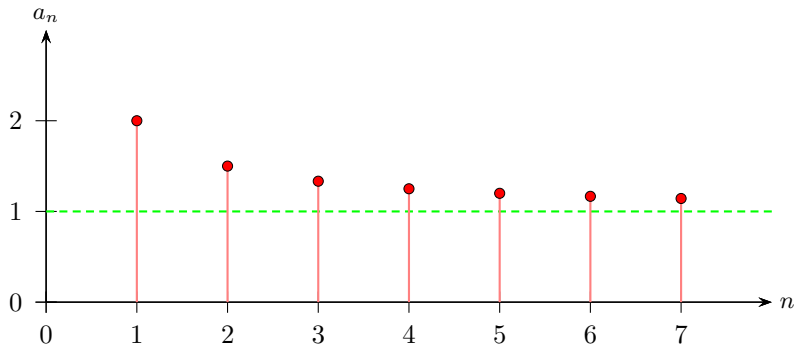
Aunque a lo largo de los dos siglos en los que se trabajó con una noción intuitiva de límite se obtuvieron muchos resultados importantes, durante ese tiempo también se encontraron ejemplos que desafiaban la intuición. Por eso algunos matemáticos del siglo XIX llegaron a la conclusión de que era conveniente aclarar el concepto de límite y definirlo de manera rigurosa. Esto no fue un camino fácil.

A continuación vamos a recrear el proceso que llevó a una definición rigurosa de límite basándonos en algunas formulaciones históricas e inventando otras que ayudan a dar continuidad al desarrollo. En el proceso usaremos algunas sucesiones creadas artificialmente para señalar las debilidades de los sucesivos intentos de definición. La definición que buscamos debe poder aplicarse a cualquier caso, aún al más artificial, es decir, debe poder determinar, para cualquier sucesión que se nos ocurra, si ésta converge o diverge.

Representaremos las sucesiones gráficamente. Por ejemplo, a la sucesión

$$\frac{2}{1}, \frac{3}{2}, \frac{4}{3}, \frac{5}{4}, \ldots \tag{4.8}$$

la representamos con este dibujo:



Intuitivamente es claro que esta sucesión *tiende a* 1, o que *converge a* 1. Observemos que ambas palabras, *tiende* y *converge*, reflejan un pensamiento *dinámico*. En lo que sigue veremos que esta descripción dinámica, que es tan natural y sugerente, es muy engañosa y será necesario reemplazarla por una más *estática*.

A continuación daremos una lista de formulaciones tentativas de lo que quiere decir que una sucesión *tiende* a 1, junto con una crítica que muestra que la formulación aún no capta el sentido correctamente.

1. **La sucesión se acerca cada vez más a 1.** Esto significa que la distancia de a_n a 1 *es cada vez más pequeña*, es decir que disminuye en todos los pasos.

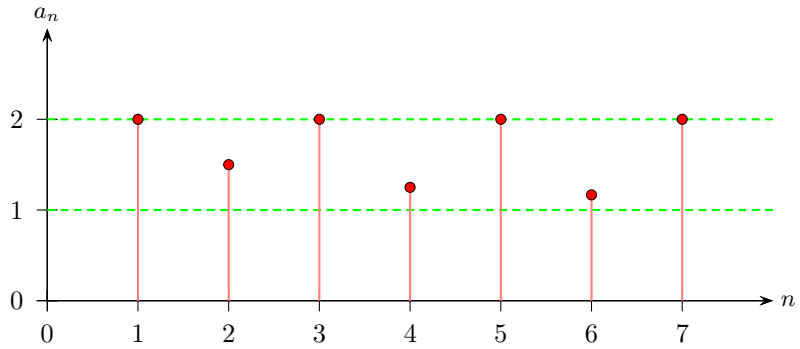
   Crítica: La distancia de a_n a 0.5 también disminuye en cada paso, y sin embargo es claro que la sucesión no converge a 0.5.

2. **La distancia de a_n a 1 se vuelve arbitrariamente pequeña.** Esta formulación ya es más acertada que la anterior, sin embargo aún es insuficiente.

   Crítica: Consideremos ahora la sucesión a_n definda por:

$$a_n = \begin{cases} 1 + \frac{1}{n}, & \text{si } n \text{ es par,} \\ 2, & \text{si } n \text{ es impar.} \end{cases}$$

   Su representación gráfica es:



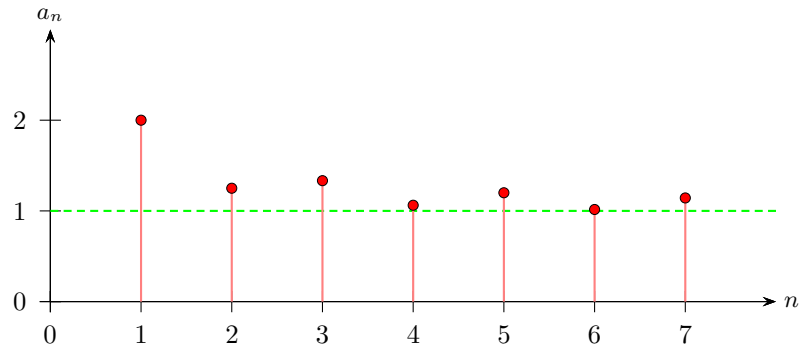
   Se puede decir, y no sin razón, que en esta sucesión la distancia de a_n a 1 se hace arbitrariamente pequeña para muchos miembros de la sucesión, pero no para todos. Para los índices impares la distancia es siempre igual a 1. De hecho esta sucesión no tiene *un* límite, en todo caso tiene *dos*. Aún no hemos captado bien una parte importante del concepto de límite.

3. **La distancia de a_n a 1 disminuye cada vez más y se hace arbitrariamente pequeña.** Aquí se conjuntan las ideas de los intentos 1 y 2 y con ello evitamos la crítica a la definición anterior.

   Crítica: Según esta definición la siguiente sucesión no converge:

$$a_n = \begin{cases} 1 + \frac{1}{n^2}, & \text{si } n \text{ es par,} \\ 1 + \frac{1}{n}, & \text{si } n \text{ es impar.} \end{cases}$$

   Mostramos aquí la representación gráfica de esta sucesión:



Se trata de una sucesión que converge a 1, pero que no cumple la propiedad expresada en 3, ya que la distancia de a_n a 1 no disminuye conforme el índice n aumenta. En una sucesión que converge es posible que la distancia al valor límite aumente temporalmente y más tarde disminuya.

4. **Para índices N suficientemente grandes, la distancia entre a_N y 1 se hace arbitrariamente pequeña y después no vuelve a rebasar esta distancia.** Lo último quiere decir que la distancia de a_n a 1 para $n > N$ nunca es mayor que la de a_N a 1.

   Crítica: En esta formulación hay un problema: para cada índice N el miembro a_N tiene una distancia fija a 1. Por ello no tiene sentido decir que si escogemos N suficientemente grande, a_N estará "arbitrariamente" cerca de 1, ya que es un valor fijo. Un valor que está "arbitrariamente" cerca de otro es idéntico a éste.

   Ahora llegamos al punto clave: en todos los intentos anteriores se ha tratado de definir el límite por cómo se comporta la distancia de a_n a 1 cuando n crece. En el último tratamos de describir qué pasa con la *distancia* a partir de un *índice* N "suficientemente grande". En la siguiente formulación invertimos el orden entre *índice* y *distancia*: tratamos de describir cómo deben ser los índices para que la distancia sea suficientemente pequeña.

5. **Para cada distancia a 1 que consideremos, hay un índice N tal que los miembros a_n con $n > N$ distan de 1 menos que ella.** Esta es la definición correcta. Para llegar a ella tuvimos que fijar primero una distancia y después dar una cota del índice tal que todos los elementos de la sucesión a partir de esa cota distarán de 1 menos que la distancia dada.

   Es una definición compleja y formularla solamente con palabras resulta torpe y difícil de leer. Es un caso típico en donde la introducción de una variable, aquí ε para "la distancia dada", ayuda a presentar la idea en forma más concisa y legible. Así es como se llega a la definición que se encuentra en los libros de texto.

Definición. Una sucesión a_n converge a 1 si

$$\boxed{\textbf{para cada } \varepsilon > 0 \textbf{ existe } N \textbf{ tal que } |a_n - 1| < \varepsilon \textbf{ para todo } n \geq N.}$$

Esta formulación la dió Bernard Bolzano en 1817 en palabras y después Karl Weierstrass de manera formal con la ε, tal como la dimos arriba.

Observemos que esta formulación es meramente estática, en ella ha desaparecido por completo el aspecto dinámico. La expresión "la sucesión tiende a 1" se sigue usando porque refleja la noción intuitiva que tenemos del concepto de límite. Pero su uso tiene ahora el significado, más riguroso,

de la formulación de Bolzano y Weierstrass.

Puntos de acumulación, sucesiones de Cauchy y números reales

Se dice que a es un *punto de acumulación* de una sucesión s_n si en cualquier intervalo abierto que contenga a hay un número infinito de miembros de la sucesión. La definición formal de *punto de acumulación* la dio primero Georg Cantor en 1872, pero una década antes Karl Weierstrass ya había usado el concepto, de manera implícita, en sus estudios sobre la convergencia.

Hay dos maneras de formular que una sucesión converge a a basadas en el concepto de *punto de acumulación* :

a es el único punto de acumulación de la sucesión.

Cada intervalo abierto que contiene a a, contiene a todos los puntos de la sucesión, excepto a un número finito de ellos.

La siguiente formulación, que se debe a Augustin Louis Cauchy, expresa cuando una sucesión converge sin mencionar en ningún momento el valor al que converge.

Si para cada distancia $\varepsilon > 0$ existe una cota N tal que para todos los $m, n \geq N$ se tiene $|s_m - s_n| \leq \varepsilon$, entonces la sucesión converge.

Las sucesiones que tienen esta propiedad, es decir para las cuales "dada una distancia $\varepsilon > 0$, existe una cota N tal que para todos los $m, n \geq N$ sucede que $|s_m - s_n| \leq \varepsilon$" se llaman *sucesiones de Cauchy*, de manera que la formulación anterior se puede abreviar diciendo que "las sucesiones de Cauchy convergen".

A veces los miembros de una sucesión pertenecen a un conjunto M, como por ejemplo los números racionales $M = \mathbb{Q}$, en donde es posible hablar de la distancia entre sus elementos y de las sucesiones de Cauchy formadas con sus elementos, pero para los cuales los límites de estas sucesiones de Cauchy no siempre existen como miembros del conjunto. Éste es el caso de las sucesiones de números racionales que convergen a valores irracionales, como por ejemplo la sucesión $s_n = \left(1 + \frac{1}{n}\right)^n$, cuyos miembros son números racionales, pero cuyo límite es el número de Euler e, que es irracional y no pertenece a M. Las sucesiones de Cauchy permiten determinar números que no se encuentran en M y así ampliar el conjunto M. La cerradura $\overline{M}$ de un conjunto M se define como el conjunto de todos sus puntos de acumulación. Es fácil ver que para cada punto de acumulación de M se puede construir una sucesión de Cauchy que converja a él. En otras palabras, la cerradura de un conjunto M consta de todos los límites de las sucesiones de Cauchy formadas con elementos de M. Esta idea es la que se utiliza para definir el conjunto de los números reales como $\mathbb{R} = \overline{\mathbb{Q}}$, es decir, como la cerradura del conjunto de los números racionales. No desarrollaremos formalmente la construcción de los números reales, que involucra detalles sutiles como el concepto de "clases de equivalencia de sucesiones de Cauchy", pero en realidad la idea central de tal construcción ya está expuesta.

Conclusiones

Las matemáticas no siempre se desarrollan con el mismo rigor. Sobre todo al principio del desarrollo de una nueva área, cuando se aprovecha el uso de conceptos que apenas se están analizando, el rigor es de segunda importancia. Mucho más atractivos parecen los logros que se pueden obtener con esos nuevos conceptos y sólo más tarde empieza una fase de reflexión sobre sus fundamentos. Esto es muy

marcado en el desarrollo de toda el área que hoy llamamos "análisis" y se ejemplifica perfectamente con la historia del concepto de límite.

4.5. Leyes del movimiento Planetario

Copérnico, Kepler, Newton

4.6. Expansión del Cálculo

El Cálculo diferencial e integral como tecnología.

4.7. Ecuaciones diferenciales

como herramienta de modelación de procesos y fenómenos. Ejemplo significativo: el oscilador armónico forzado y amortiguado.

4.8. Series de Potencias

Una herramienta útil en sitios insospechados. Euler.

4.9. Formalización del Cálculo

Descubrimiento de funciones monstruosas: necesidad de formalizar el cálculo. El análisis real, complejo y funcional.

Cauchy y Weierstrass

Cantor y los números infinitos

Una función inyectiva del cuadrado $(0,1)\times(0,1)$ al intervalo $(0,1)$

Representaremos a los números reales del intervalo abierto $I=(0,1)$ con un desarrollo binario de manera que la representación de cada uno sea única. Lo haremos evitando las representaciones que tiene sólo ceros a partir de algún índice. Si $x\in I$, entonces $x\in(0,\frac{1}{2}]$ o $x\in(1/2,1)$. En el primer caso tomamos $a_1=0$ y en el segundo $a_1=1$. Si en el primer caso además $x=\frac{1}{2}$ entonces definimos $a_i=1$ para todo $i>1$. Para los siguientes pasos procedemos por inducción. Supongamos que ya tenemos definidos los elementos a_i para $i<n$ pero aún no han quedado definidos los elementos a_i para $i\geq n$. Entonces $x\in(\frac{2k}{2^n},\frac{2k+1}{2^n}]$ o $x\in(\frac{2k+1}{2^n},\frac{2k+2}{2^n})$ para algún entero no negativo $k<2^n-1$. En el primer caso definimos a_n=0 y en el segundo dos a_n=1. Y si en el primer caso además $x=\frac{2k+1}{2^n}$,

entonces definimos $a_i = 1$ para todo $i > n$. Esto define una sucesión a_n de ceros y unos para cada $x \in I$ tal que

$$x = \sum_{n=1}^{\infty} \frac{a_n}{2^n}$$

La representación es única porque si dos representaciones de este tipo a_n y b_n no son idénticas, deben diferir en algún índice, es decir, $a_k \neq b_k$ para algún índice k. Entonces alguno de los dos elementos a_k y b_k es igual a 0 el otro es igual a 1. Podemos suponer, sin pérdida de generalidad, que $a_k = 1$ y $b_k = 0$. La única manera de que ambas sumas infinitas pudieran ser iguales, es decir, de que

$$\sum_{n=1}^{\infty} \frac{a_n}{2^n} = \sum_{n=1}^{\infty} \frac{b_n}{2^n}$$

es que $b_j = 1$ y $a_j = 0$ para todo $j > k$. Pero esto último es imposible porque ninguna de las representaciones que definimos tiene ceros en la cola.

Ahora definimos la siguiente función $f : I \times I \to I$. Para cada $(x, y) \in I \times I$, tomamos las representaciones binarias únicas de x y y. Supongamos que

$$x = \sum_{n=1}^{\infty} \frac{a_n}{2^n}; \qquad y = \sum_{n=1}^{\infty} \frac{b_n}{2^n}$$

Con ellas definimos $c_{2n-1} = a_n$ y $c_{2n} = b_n$ para todo entero positivo n y definimos

$$f(x, y) = \sum_{n=1}^{\infty} \frac{c_n}{2^n}$$

La función así definida es inyectiva porque si $(x', y') \in I \times I$ no es igual a (x, y), entonces $x \neq x'$ o $y \neq y'$. En cualquiera de los dos casos resultaría que la representación de $f(x', y')$ tendría algún índice diferente de la correspondiente representación de $f(x, y)$ y por tanto $f(x, y) \neq f(x', y')$.

La existencia de esta función nos dice que, en cierto sentido, el conjunto $(0, 1) \times (0, 1)$ tiene al menos tantos elementos como $(0, 1)$, y esto contradice nuestra primera intuición.

4.9.1. Sumas de Riemann y teoría de la medida

4.10. Métodos numéricos en el Cálculo