

CONFERÈNCIES FME  
Recull del Curs Bernhard Riemann  
2007-2008



# CONFERÈNCIES FME

VOLUM V

Curs Bernhard Riemann  
2007-2008

FACULTAT DE MATEMÀTIQUES I ESTADÍSTICA  
UNIVERSITAT POLITÈCNICA DE CATALUNYA

© 2007-2008

FACULTAT DE MATEMÀTIQUES I ESTADÍSTICA  
Pau Gargallo, 5 — 08028 Barcelona

Fotocomposició, impressió i enquadernació:

BARCELONA DIGITAL  
c/ Rosselló, 77. 08029 Barcelona  
Tel. 93-3638610

ISBN: 978-84-7653-176-1  
Dipòsit legal: B-39425-2008

*Printed in Spain*

Reservats tots els drets. Queda totalment prohibida la reproducció total o parcial d'aquest llibre per qualsevol procediment electrònic o mecànic, àdhuc fotocòpia, gravació magnètica o qualsevol sistema, sense el permís de la FACULTAT DE MATEMÀTIQUES I ESTADÍSTICA.

# ÍNDEX

<b>Presentació del Curs Riemann (2007-2008)</b> (Sebastià Xambó) .....	9
<b>Presentació del Volum Riemann</b> (Comissió Riemann) .....	11
<b>Índex d'autors</b> .....	15
<b>La funció <math>\zeta</math> de Riemann</b> (Jordi Quer) .....	17
<b>Riemann's Influence in Geometry, Analysis and Number Theory</b> (Michael Atiyah) .....	55
<b>Duality in Mathematics and Physics</b> (Michael Atiyah) .....	69
<b>An overview of Riemann's life and work</b> (Rossana Tazzioli) .....	93
<b>Geometria de Riemann</b> (Joan Porti) .....	103
<b>Riemann and Partial Differential Equations. A Road to Geometry and Physics</b> (Juan Luis Vázquez) .....	139
<b>Riemann and Physics</b> (Emili Elizalde) .....	161

<b>Códigos convolucionales y Geometría algebraica</b> (J.M. Muñoz Porras y J.A. Domínguez Pérez) .....	199
<b>Funciones Theta de Riemann</b> (J.M. Muñoz Porras y E. Gómez González) .....	211
<b>Riemann y su influencia en el Análisis</b> (Antonio Córdoba) .....	223
<b>Annex: Presentation of the Riemann's Academic Year (2007-2008)</b> (Sebastià Xambó) .....	247

## Presentació del Curs Riemann (2007-2008)

La FACULTAT DE MATEMÀTIQUES I ESTADÍSTICA (FME) dedicà el curs 2007-2008 a Bernhard Riemann (1826-1866).

La **Comissió Riemann**, constituïda per Víctor GONZÁLEZ, Eduard RECASENS, Oriol SERRA, Joan SOLÀ-MORALES i Sebastià XAMBÓ, fou l'encarregada de donar contingut acadèmic a la decisió, incloent-hi l'elaboració del programa de conferències. És l'ocasió per reiterar l'agraïment de l'FME als seus membres pel temps i esforç que hi van dedicar.

La difusió de les conferències s'ha fet posant els articles al Butlletí Digital FME, al qual s'accedeix a través del vincle del mateix nom posat a la part superior del Web de l'FME.\* Ara, a més dels materials accessibles a través del Web, l'FME posa a disposició de tots els interessats un recull en paper de les conferències més directament relacionades amb la vida i obra de Riemann. S'afegeix així un Volum Riemann als volums Poincaré, Einstein, Gauss i Euler. En nom de l'FME, és un agradable obligació agrair molt especialment als autors l'esforç addicional de transformar els materials de la conferència en els articles d'aquest recull.

Un agraïment molt sentit per al professor Sir Michael Atiyah, que va saber trobar temps en la seva agenda per visitar l'FME el desembre de 2007. En aquella ocasió va impartir una conferència *keynote*, «Riemann's influence in Geometry, Analysis and Number Theory», i és una gran satisfacció per a nosaltres poder-la incloure en aquest volum. També agraïm al professor Atiyah, al Centre de Recerca Matemàtica (CRM) i a l'Institut de Matemàtica de la Universitat de Barcelona (IMUB) la gentilesa de permetre'ns incloure en aquest recull la transcripció de la conferència *Duality in Mathematics and Physics*.

---

\* <http://www-fme.upc.es/>

Els articles estan escrits en la llengua en què l'autor ens els ha lliurat: dos estan en català, tres en castellà i cinc en anglès.

Cal dir a més que hem usat un mateix estil de document L<sup>A</sup>T<sub>E</sub>X per compondre els diversos articles (`amsart.sty`), encara que l'article lliurat no fos en L<sup>A</sup>T<sub>E</sub>X o que ho fos en un format diferent d'`amsart`, però que tanmateix hem intentat mantenir molts aspectes de la composició original dels autors.

Ens donarem per satisfets si contribuïm a fer que figures com Riemann siguin millor conegudes entre els matemàtics, sobretot si aquest coneixement arribés ben viu a les generacions més joves.

Sebastià Xambó Descamps  
Degà de l'FME  
Barcelona, agost de 2008



## Presentació del Volum Riemann

Aquest volum recull els articles rebuts corresponents a les conferències més directament relacionades amb Riemann que es van impartir durant el Curs Riemann (2007-2008), i un article excepcional del professor Atiyah del qual donem detalls més avall.

La relació de les conferències és la següent:

- Conferència inaugural, el 19 de setembre de 2007, a càrrec del professor Jordi QUER: *La funció  $\zeta$  de Riemann*.
- Conferència *keynote* del professor Sir Michael ATIYAH, el 18 de desembre de 2007: *Riemann's influence in Geometry, Analysis and Number Theory*.

Aquesta conferència fou seguida per la que va impartir el mateix dia a la Facultat de Matemàtiques de la Universitat de Barcelona, amb el títol *Duality in Mathematics and Physics*, d'acord amb una invitació del CRM sorgida de la seva disposició a aprofitar l'ocasió de la visita del professor Atiyah per oferir una activitat a tota la comunitat matemàtica.

- **Jornada Riemann** (20 de febrer de 2008), destinada a aprofundir en els aspectes més importants de la persona i l'obra d'un matemàtic cabdal. Com en els cinc cursos precedents, es va demanar als conferencians de la Jornada (Rossana TAZZIOLI, Joan PORTI, Juan L. VÁZQUEZ, Jaume AMORÓS i Emili ELIZALDE) que procurassin posar de relleu la influència de Riemann fins al dia d'avui.

Els títols de les conferències, en el mateix ordre, foren els següents: *An overview of Riemann's life and work*, *Geometria de Riemann*, *Las ecuaciones en derivadas parciales de Riemann: un camino a la geometría y a la física*, *Riemann i les funcions de variable complexa*, *Riemann i la física*.

- Dues conferències del professor José M. MUÑOZ PORRAS, el dia 3 d'abril de 2008: *Riemann, funciones theta y variedades abelianas* i *Códigos convolucionales de Goppa*.
- Conferència de cloenda del Curs Riemann, el dia 14 de maig de 2008, a càrrec del professor Antonio Córdoba: *Riemann y el análisis matemático*.

## Apunt biogràfic

Bernhard Riemann va néixer el 17 de setembre de 1826 a Breselenz, un poble situat al nord d'Alemanya. De petit no va anar a l'escola; el seu pare, que era pastor luterà, li va fer de mestre. El 1840, pels estudis de secundària, va anar primer a l'Institut de Hannover i dos anys després al de Lüneburg. L'any 1846, d'acord amb el desig del seu pare, es va matricular a la Universitat de Göttingen per estudiar Teologia, però l'any següent va obtenir el permís patern per estudiar Matemàtiques; aquest interès cap a les matemàtiques li provenia de l'època de secundària, quan el director de l'Institut de Lüneberg li havia deixat llibres d'Euclides, Arquímedes, Descartes, Newton i Legendre. En particular, es va llegir el tractat sobre teoria de nombres de Legendre en sis dies.

El 1847 es va traslladar a Berlín, on rebé les lliçons de Dirichlet sobre Anàlisi i Teoria de nombres i les de Jacobi sobre Mecànica i Àlgebra. Amb Eisenstein estudià les funcions el·líptiques. Acabats els estudis, va tornar a Göttingen per entrar com ajudant al laboratori de física que dirigia Wilhelm Weber i és en aquest període quan s'inicià el seu interès per la Física. En aquests mateixos anys, llegeix al filòsof Johann Herbart, el qual l'influència notablement, sobretot per tot allò que fa a la nova manera en què Riemann concebrà l'espai.

L'anàlisi complexa fou un dels camps més rellevants de la producció científica de Riemann. El novembre de 1851 defensà la seva tesi doctoral titulada «Fonaments per una teoria general de funcions d'una variable

complexa», en la qual introdueix mètodes topològics i, per a l'anàlisi de les funcions multivaluades, les superfícies que avui porten el seu nom. L'altra fita fonamental en aquest camp és la publicació sis anys més tard al Journal de Crelle de la «Teoria de funcions abelianes».

El desembre de 1853, es presentà a un concurs per una plaça com a *Privatedozent* a Göttingen i en una de les memòries que va presentar, la que porta per títol «Sobre la representació d'una funció per sèries trigonomètriques», introdueix allò que avui anomenem «Integral de Riemann», i en una altra memòria, la que porta per títol «Sobre les hipòtesis en què es fonamenta la geometria», introdueix i analitza el concepte de varietat  $n$ -dimensional i les relacions mètriques de les que és susceptible.

El 1859, obté una càtedra a Göttingen: era la càtedra que quedà vacant a la mort del seu gran mestre Dirichlet i que aquest havia ocupat des de la seva creació l'any de la mort de Gauss (1855). Aquest mateix any 1859, Riemann és nomenat membre de l'Acadèmia de Göttingen i membre corresponent de la de Berlín. És per aquest motiu que viatja a Berlín, on visita a Weierstrass, Kummer i Kronecker i publica l'article «Sobre el nombre de nombres primers inferiors a un nombre donat», el qual conté la seva famosa conjectura sobre la localització dels zeros no trivials de la «funció zeta», conjectura que a dia d'avui encara resta oberta.

Riemann, que fou un matemàtic de brillant imaginació i els treballs del qual obririen molts dels camins que conduïren cap a la matemàtica del segle XX, era una persona amb salut dèbil i de caràcter reservat. El seu gran amic fou Richard Dedekind, amb qui solia mantenir llargues converses.

El 1862, quan tenia 36 anys, es va casar amb Elise Koch, una amiga de les seves germanes. L'infortuni va fer que, just dos mesos després de casar-se, Riemann sofrís una pleuresia que l'afectà greument. A partir de llavors, començà una nova vida consistent en donar menys classes i passar llargues temporades a Itàlia pel seu clima més benigne. El 1863, va anar a Pisa per trobar-se amb Betti, Beltrami, Briochi i Casaroti i parlar de topologia. Aquest mateix any va néixer Ida, la seva única filla.

En un tercer viatge a Itàlia s'establí a Selasca, un bell llogaret al costat del llac Maggiore i fou en aquest indret on Riemann, quan li faltaven dos mesos per complir quaranta anys, afectat d'una tuberculosi, morí. Era el 20 de juliol de 1866.

Gauss, matemàtic poc donat a lloar els altres matemàtics, el 1851 (tenia llavors 74 anys), com a membre del tribunal que havia de jutjar la tesi de Riemann, va escriure en el seu informe que el treball de Riemann mostrava «una ment creativa, activa i decididament matemàtica, acompanyada d'una fèrtil i meravellosa originalitat».

La Comissió Riemann  
Barcelona, agost de 2008

## Índex d'Autors

Sir Michael ATIYAH  
School of Mathematics  
University of Edinburgh  
President of the Royal Society of Edinburgh  
M.Atiyah@ed.ac.uk

Antonio CÓRDOBA  
Departamento de Matemáticas  
Universidad Autónoma de Madrid  
Antonio.Cordoba@uam.es

Emili ELIZALDE  
Consejo Superior de Investigaciones Científicas  
Instituto de Ciencias del Espacio (CSIC)  
Institut d'Estudis Espacials de Catalunya (IEEC/CSIC)  
elizalde@ieec.uab.es, elizalde@math.mit.edu

José M. MUÑOZ PORRAS  
Departamento de Matemáticas  
Facultad de Ciencias  
Universidad de Salamanca  
jmp@usal.es

Joan PORTI  
Departament de Matemàtiques  
Facultat de Ciències  
Universitat Autònoma de Barcelona  
porti@mat.uab.cat

Jordi QUER  
Departament de Matemàtica Aplicada II  
Universitat Politècnica de Catalunya  
Jordi.Quer@upc.edu

Rossana TAZZIOLI  
Dipartimento di Matematica e Informatica  
Facoltà di Scienze Matematiche, Fisiche e Naturali  
Università di Catania  
tazzioli@dmi.unict.it

Juan Luis VÁZQUEZ  
Departamento de Matemáticas  
Facultad de Ciencias  
Universidad Autónoma de Madrid  
JuanLuis.Vazquez@uam.es

# LA FUNCIÓ $\zeta$ DE RIEMANN

JORDI QUER

RESUM. El novembre de 1859 Riemann envià un manuscrit de sis fulls a l'Acadèmia de Berlín titulat *Sobre el nombre de primers menors que una quantitat donada*, el qual seria l'única publicació dedicada a la teoria de nombres de tota la seva producció científica. Aquest treball, sens dubte una de les peces mestres de les matemàtiques de tots els temps, és pioner en l'aplicació de tècniques analítiques per a l'estudi de problemes aritmètics. En ell Riemann introdueix la funció  $\zeta$  i en dóna diverses propietats, de les quals en treu conseqüències sobre l'acumulació dels nombres primers. També hi enuncia la famosa conjectura sobre els seus zeros que ha passat a la història amb el nom d'hipòtesi de Riemann, i que, havent resistit els esforços de molts dels millors matemàtics del segle XX, és considerada avui dia el problema obert més important de les matemàtiques.

L'objectiu d'aquestes notes és explicar el contingut del treball de Riemann i el paper fonamental que ha jugat en l'estudi de la distribució dels nombres primers.

## 1. INTRODUCCIÓ

El matemàtic alemany Bernhard Riemann va néixer l'any 1826 a Breselenz i morí prematurament als 39 anys a Itàlia de tuberculosi. Excepte una estada de dos anys a Berlín, la seva activitat acadèmica va estar sempre vinculada a la Universitat de Göttingen, on va estudiar, i on va acabar ocupant la càtedra que havia estat prèviament de Gauss i de Dirichlet, dos dels seus mestres. Malgrat que l'edició de les seves obres completes requereix un únic volum no gaire gruixut, les contribucions cabdals de Riemann a l'anàlisi i a la geometria el converteixen en la figura central de les matemàtiques de mitjan segle XIX. L'únic article que Riemann publicà sobre aritmètica és la pedra fundacional d'una nova branca de les matemàtiques: la teoria analítica de nombres.

El teorema dels nombres primers

$$(1) \quad \pi(x) := \#\{p : p \text{ primer} < x\} \sim \text{Li}(x) := \int_0^x \frac{dt}{\log t}, \quad x \rightarrow \infty,$$

dóna el creixement asimptòtic de la funció  $\pi$  que compta els nombres primers. Conjecturat per Gauss i Legendre a finals del segle XVIII, va ser abordat per diverses generacions de matemàtics de primera fila fins que finalment l'any 1898 Hadamard i de la Vallée-Poussin van aconseguir demostrar-lo seguint el camí indicat per Riemann en el seu article. La reputació d'aquesta conjectura entre la comunitat matemàtica del segle XIX és comparable amb la que ha gaudit durant més de tres segles l'últim teorema de Fermat o amb l'adquirida per la hipòtesi de Riemann d'ençà que Hilbert la va incloure l'any 1900 en la seva famosa llista de problemes oberts. Fins i tot es va difondre el mite que la seva demostració proporcionaria a l'autor la immortalitat. No fou el cas de Hadamard ni de de la Vallée-Poussin, que en canvi van gaudir d'una longevitat considerable, arribant als 98 i 96 anys respectivament.

El manuscrit que ens ocupa és la primera comunicació que Riemann enviava a l'Acadèmia poc després d'haver-ne estat elegit membre. Volent que el tema del treball i la importància dels resultats estiguessin a l'altura de l'ocasió, decidí enviar un resum de les seves investigacions sobre aquesta famosa conjectura, demostrant així el respecte que l'assumpte li mereixia. En la introducció descriu el treball com «una investigació sobre l'acumulació dels nombres primers; un tòpic que potser no semblarà del tot desmereixedor d'una tal comunicació, donat l'interès que els mateixos Gauss i Dirichlet han mostrat en ell durant un període de temps ben llarg.» El resultat principal del treball és la fórmula explícita

$$(2) \quad \pi(x) = \text{Ri}(x) + \sum_{\zeta(\rho)=0} \text{Ri}(x^\rho), \quad x > 1,$$

que proporciona el valor exacte de la funció  $\pi$  en la forma d'un terme principal dominant  $\text{Ri}(x)$  més un terme d'error que és la suma d'una sèrie infinita de funcions oscil·latòries  $\text{Ri}(x^\rho)$  associades als zeros complexos de la funció  $\zeta$ . L'extensió d'aquesta funció zeta (que Euler i Dirichlet havien considerat prèviament per a valors enters i valors reals de la variable) a tot el domini complex és una de les aportacions fonamentals de Riemann, i des d'aleshores se la coneix pel nom de funció zeta de Riemann. El teorema dels nombres primers és equivalent a



$\pi(x) \sim \text{Ri}(x)$ ; és a dir, al fet que l'error relatiu del terme principal en la fórmula de Riemann tendeixi a zero, i aquesta propietat és la que portarà finalment a la demostració quaranta anys més tard.

La hipòtesi de Riemann assegura que tots els zeros no trivials de la funció  $\zeta$  estan situats damunt la recta crítica  $\Re(s) = 1/2$ . Després de resistir els esforços de molts dels millors matemàtics del darrer segle i mig, ha esdevingut una de les conjectures més famoses i amb més reputació de dificultat de totes les matemàtiques: l'any 1900 Hilbert la va incloure a la seva famosa llista de vint-i-tres problemes presentada al congrés internacional de París; forma part també de la llista dels set «problemes del mil·leni» proposats per la fundació Clay i de la llista de 18 «problemes per al segle XXI» elaborada per Stephen Smale, ambdues publicades l'any 2000 amb motiu de l'Any Mundial de les Matemàtiques. A diferència de moltes altres conjectures, la demostració de les quals no té conseqüències apreciables, llevat potser del reconeixement que rep el matemàtic que ho aconsegueix per primera vegada (és el cas, per exemple, de l'últim teorema de Fermat), la hipòtesi de Riemann i les seves generalitzacions apareixen com a condició en l'enunciat d'innombrables resultats matemàtics teòrics, així com en l'anàlisi de la complexitat d'algorismes importants per a moltes aplicacions. És per aquest motiu que, per referir-s'hi, el costum ha portat a fer servir el terme hipòtesi i no pas a dir-ne conjectura, com és l'habitual. Fins que no sigui demostrada no es podrà considerar que l'article de Riemann ha estat completament entès i que totes les seves potencialitats han estat prou aprofitades.

L'estil d'escriure matemàtiques de Riemann és considerat difícil, ja que acostuma a presentar les seves investigacions de manera molt condensada i de vegades es deixa portar per una certa tendència a la manca de rigor. En algunes afirmacions sense demostració no queda prou clar si es tracta d'un simple exercici que el lector pot resoldre sense dificultat o d'un resultat difícil que requereix un esforç considerable. Moltes vegades es limita a indicar vagament el camí a seguir per poder arribar a una demostració com a única justificació del resultat que presenta. L'article que ens ocupa n'és un bon exemple: tot i la seva curta extensió (la publicació [12] al Butlletí de l'Acadèmia té només 8 pàgines) conté una quantitat ingent de nous conceptes, idees i resultats, i aplica tècniques d'integració i anàlisi complexa sofisticades i innovadores,

limitant-se l'autor en el millor dels casos a donar indicacions o esbossos de demostració. La història demostra que Riemann va sobreestimar en aquesta ocasió les capacitats dels seus coetanis, ja que les demostracions dels resultats més importants de l'article no van ser publicades fins gairebé mig segle més tard per Hadamard i von Mangoldt.

Riemann no va publicar mai més res sobre aquest tema, entre altres raons perquè ja no li quedaven molts anys de vida. Ben entrat el segle XX Siegel decidí revisar el que quedava dels seus papers, dipositats a la biblioteca de la Universitat de Göttingen després d'haver sofert una purga considerable per part de la seva assistenta i de la seva muller, que volien evitar vestigis de tota mena d'informació personal. L'examen d'aquests papers va resultar fructífer tant històricament com científica. D'una banda, va servir per refutar la creença de la comunitat matemàtica de l'època que Riemann, matemàtic eminentment teòric, havia arribat als seus resultats i formulat la seva hipòtesi en base només a una portentosa intuïció i penetració teòrica: en els seus papers s'hi troba el desenvolupament d'una expressió adequada per al tractament numèric de la funció zeta i un munt de càlculs a mà d'aproximacions dels seus zeros. D'una altra, alguns dels resultats trobats per Siegel foren encara, setanta anys després de la seva mort, contribucions originals rellevants de Riemann a una àrea en què alguns dels millors matemàtics de l'època havien estat treballant tot aquest temps.

La funció  $\zeta$  introduïda per Riemann i la manera com es pot fer servir per a l'estudi d'un problema aritmètic és, juntament amb altres funcions i tècniques semblants usades per Dirichlet poc abans per demostrar el teorema de la progressió aritmètica, només la punta de l'iceberg de la teoria de les funcions  $L$ , que des d'aleshores no ha parat de créixer. Aquestes funcions han proliferat i han guanyat importància a partir d'investigacions de molts dels grans matemàtics del segle XX (Hecke, Weil, Selberg, Serre, Langlands, Shimura, Iwaniec, ...), fins al punt que actualment s'han convertit en eines universals, d'ús imprescindible en un ampli ventall de disciplines matemàtiques, en particular i especialment per a la teoria de nombres, la geometria aritmètica o la teoria de representacions, a part del l'interès notable que l'estudi de les funcions  $L$  té per si mateix. Per a la majoria d'elles les propietats més bàsiques demostrades per Riemann en el cas de la funció  $\zeta$  segueixen sent conjectures. Per exemple, la conjectura de Shimura-Taniyama, demostrada

per Wiles i altres, que conduí finalment a l'últim teorema de Fermat, és equivalent al fet que les funcions  $L$  de les corbes el·líptiques sobre  $\mathbb{Q}$  tinguin prolongació analítica i equació funcional.

La hipòtesi sobre la localització dels zeros de la funció  $\zeta$  es generalitza de manera natural al context de les funcions  $L$ , on rep els noms d'hipòtesi de Riemann generalitzada o estesa. Fins ara no s'ha aconseguit demostrar que la compleixi (ni que no ho faci) cap funció de la classe de les funcions  $L$  globals, el prototipus de les quals és la  $\zeta$  de Riemann. En canvi hi ha una altra classe de funcions  $L$  associades a varietats algebraïques sobre cossos finits, anomenades funcions  $L$  locals, per a les quals Deligne va aconseguir demostrar la hipòtesi de Riemann l'any 1973 (la qual, en aquest context, és una de les propietats d'aquestes funcions que, junt amb altres, se segueixen coneixent un cop demostrades amb el nom de conjectures de Weil).

El resultat principal del treball de Riemann, la fórmula explícita, ha estat sovint interpretat en termes de l'anàlisi harmònica de la manera següent. La funció que representa el so d'un instrument musical, potser amb forma de pols rectangular o de dent de serra, que en general presentarà discontinuïtats de salt, pot obtenir-se com a superposició infinita de funcions oscil·latòries (harmònics) amb les tècniques habituals de l'anàlisi de Fourier. De manera anàloga, la funció  $\pi$ , que és una funció en forma d'escala que puja un graó en cada nombre primer a través d'una discontinuïtat de salt, s'obté també per superposició infinita de funcions oscil·latòries, les amplituds i freqüències de les quals estan governades pels zeros de la funció zeta. Per tant (vegi's [8]) «els zeros de la funció zeta contenen la música dels nombres primers.»

## 2. LA DISTRIBUCIÓ DELS NOMBRES PRIMERS

Euclides va demostrar que hi ha infinits nombres primers: donat un nombre primer  $p$  tots els factors primers del nombre  $p! + 1$  són més grans que  $p$ .

El següent avenç significatiu en la comprensió de la seqüència dels nombres primers va ser la demostració d'Euler de que la suma dels seus inversos és divergent, a diferència de la suma dels inversos de molts altres conjunts infinits de nombres, que pot ser finita. Per exemple, la suma

dels inversos de les potències de 2 és finita i igual a 2 o la suma dels inversos dels quadrats és  $\pi^2/6$ . Per tant «hi ha més nombres primers que potències de 2 o que quadrats.»

Ja al segle XIV Nicolas d'Oresme va donar una demostració de la divergència de la sèrie harmònica; de fet la seva suma és  $\log \infty$ , en el sentit que la diferència  $\sum_{n=1}^N 1/n - \log N$  tendeix a un nombre  $\gamma = 0.577215665\dots$ , conegut com a constant d'Euler-Mascheroni. La demostració d'Euler es basa en una relació entre la sèrie harmònica i un producte sobre els nombres primers, que es coneix pel nom de producte d'Euler, i que jugarà un paper importantíssim en tots els estudis posteriors sobre la funció zeta i els nombres primers. Tenint en compte la suma d'una progressió geomètrica,

$$1 + \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \dots = \frac{1}{1 - \frac{1}{p}},$$

i gràcies a la descomposició única en producte de primers, es té el famós producte d'Euler

$$\sum_{n=1}^{\infty} \frac{1}{n} = \prod_p \left( 1 + \frac{1}{p} + \frac{1}{p^2} + \dots \right) = \prod_p \left( 1 - \frac{1}{p} \right)^{-1},$$

o, més en general,

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \left( 1 - \frac{1}{p^s} \right)^{-1}$$

per a nombres reals  $s > 1$ , de manera que la identitat prèvia s'interpreta com un pas al límit. Aplicant logaritmes i la sèrie  $-\log(1-x) = \sum x^n/n$  es dedueix que

$$\log \sum_{n=1}^{\infty} \frac{1}{n} = \sum_p -\log \left( 1 - \frac{1}{p} \right) = \sum_p \sum_n \frac{1}{np^n} = \sum_p \frac{1}{p} + \sum_p \sum_{n \geq 2} \frac{1}{np^n}.$$

Ara bé, és immediat comprovar que el segon sumand està fitat. Per tant la divergència de la sèrie harmònica implica la de la suma dels inversos dels primers. L'any 1874 Mertens va precisar aquest resultat demostrant que de fet la seva suma és igual a  $\log \log \infty$ , en el sentit que la diferència  $\sum_{p \leq N} 1/p - \log \log N$  és convergent; el seu límit,  $0.2614972128\dots$ , es coneix pel nom de constant de Mertens.

**La conjectura de Gauss i Legendre.** Cap a finals del segle XVIII Gauss i Legendre s'interessen per la distribució dels nombres primers. Consideren la funció  $\pi(x)$  que compta el nombre de primers fins a  $x$ ,

$$\pi(x) := \#\{p : p \text{ primer } \leq x\} = \sum_{p \leq x} 1.$$

Es tracta d'una funció constant a trossos que en cada primer té una discontinuïtat de salt en què augmenta en una unitat. Per bé que el comportament d'aquesta funció és impredecible quan es consideren intervals curts, d'acord amb la manera erràtica i aparentment aleatòria com van apareixent els nombres primers entre els nombres naturals, l'aspecte de  $\pi(x)$  és en canvi d'una regularitat sorprenent quan s'observa des de prou lluny. A la figura 1 es pot veure la gràfica de la funció  $\pi$  en dos intervals de mides de diferent ordre de magnitud.

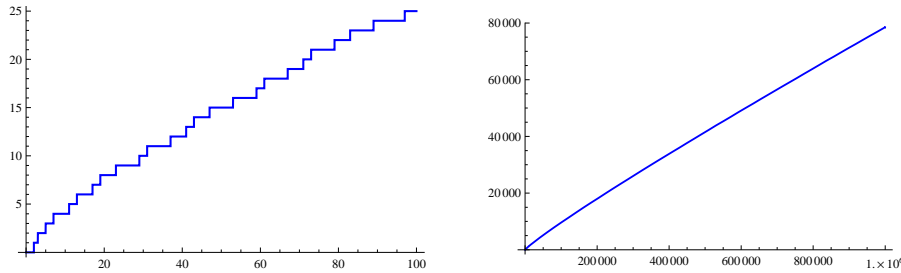


FIGURA 1. La funció  $\pi(x)$  a escales diferents.

Tant Legendre com Gauss van provar d'explicar la regularitat de la funció  $\pi$  buscant una funció simple que en donés el creixement asimptòtic. L'any 1798 Legendre proposà la funció

$$\frac{x}{\log x - B},$$

on  $B = 1.08366$  és una constant que va determinar empíricament a partir de les dades que tenia disponibles; gràcies a la potència de càlcul accessible actualment, és fàcil veure que el valor de  $B$  donat per Legendre proporciona una bona aproximació de  $\pi$  quan  $x$  es mou pels voltants de  $10^6$ , però deixa de fer-ho per valors gaire més grans. El valor que s'assigni a la constant  $B$  és irrellevant de cara a l'equivalència asimptòtica entre aquesta funció i la funció  $\pi$ , però la que va millor no

és la proposada per Legendre: mig segle més tard Chebyshev va demostrar que si el límit de la funció  $\log x - x/\pi(x)$  existeix (ara sabem que sí que existeix, però Chebyshev no ho sabia demostrar), aleshores aquest límit ha de ser igual a 1.

Tot i no haver-ho publicat, Gauss havia arribat a una conclusió semblant sobre el creixement de  $\pi(x)$  abans que Legendre, i sembla que s'arribà a produir una certa discussió entre tots dos sobre a qui se li havia d'atribuir el mèrit. En una carta al seu col·lega Johann Enke escrita l'any 1849, Gauss li explica que, cap a l'any 1792, quan tenia 15 anys, havia rebut com a regal un llibre de logaritmes (l'equivalent durant moltes generacions de les calculadores actuals), que contenia també una taula de nombres primers, i que examinant aquestes llistes de nombres primers notà que la probabilitat que un nombre enter de magnitud  $x$  sigui primer és aproximadament  $1/\log x$ . Basant-se en aquesta observació empírica Gauss proposà que el valor de la funció  $\pi(x)$  es pot aproximar de la manera següent:

$$\pi(x) = \frac{1}{\log 2} + \frac{1}{\log 3} + \frac{1}{\log 4} + \cdots + \frac{1}{\log [x]} = \sum_{n=2}^{[x]} \frac{1}{\log n} \simeq \int_2^x \frac{dt}{\log t}.$$

En la carta a Enke, Gauss també diu que des d'aleshores acostumava a dedicar quarts d'hora en què estava desvagat a comptar primers en intervals de mil enters, arribant així a acumular dades fins a tres milions i, per il·lustrar la coincidència entre la funció  $\pi$  i el valor de la seva predicció, dóna la taula 1, en la qual els valors de la funció  $\pi(x)$  contenen errors menors (entre parèntesi s'ha afegit la modificació que cal per corregir-los). Observi's que amb els valors correctes l'aproximació encara millora per als valors alts de  $x$ .

La funció introduïda per Gauss, però amb la integral calculada des de zero, el qual la modifica només en una constant igual a  $1.04516\dots$ , es coneix amb el nom de logaritme integral, i es denota

$$\text{Li}(x) = \int_0^x \frac{dt}{\log t},$$

on la integral impròpia s'ha d'interpretar com el valor principal de Cauchy per evitar la singularitat en  $t = 1$ .

$x$	Nombre de primers $\leq x$	$\int \frac{dt}{\log t}$	Diferència
500 000	41 556 (-18)	41 606.4	50.4
1 000 000	78 501 (-3)	78 627.5	126.5
1 500 000	114 112 (+43)	114 263.1	151.1
2 000 000	148 883 (+50)	149 054.8	171.8
2 500 000	183 016 (+56)	183 245.0	229.0
3 000 000	216 745 (+71)	216 970.6	225.6

TAULA 1. Taula amb valors de  $\pi$  (corregits) comparats amb  $\int_2^x \frac{dt}{\log t}$  (Gauss, carta a Enke, 1849).

Tant la funció  $\text{Li}(x)$  com la funció  $x/(\log x - B)$  proposada per Legendre són asimptòticament equivalents a la funció  $x/\log x$ , de manera que les investigacions de tots dos condueixen a formular la mateixa conjectura:

$$\pi(x) \sim \frac{x}{\log x} \quad (x \rightarrow \infty), \quad \text{o sigui,} \quad \lim_{x \rightarrow \infty} \frac{\pi(x)}{x/\log x} = 1,$$

que es coneix amb el nom de teorema dels nombres primers. El resultat es pot enunciar amb una o altra funció  $x/\log x$  o  $\text{Li}(x)$  indiferentment. Ara bé, tot i que tots dos enunciats són equivalents, el logaritme integral aproxima millor  $\pi(x)$  i quan hom es refereix a qüestions relatives al terme d'error en el teorema dels nombres primers es refereix a l'error en relació a aquesta segona funció, o sigui, a l'enunciat (1) del teorema dels nombres primers donat a la introducció. En canvi, quan l'únic que importa és el creixement asimptòtic, tot sovint s'enuncia a la literatura amb la funció  $x/\log x$  estalviant-se així d'haver de definir el logaritme integral.

Els únics progressos rellevants en la direcció d'aquesta conjectura abans de Riemann es deuen a Chebyshev, que pels volts de 1850 va aconseguir demostrar amb tècniques elementals que per  $x$  prou gran es tenen desigualtats

$$0.9219 \frac{x}{\log x} < \pi(x) < 1.10555 \frac{x}{\log x}.$$

També va demostrar que si el límit

$$\lim_{x \rightarrow \infty} \frac{\pi(x)}{x/\log x}$$

existeix, aleshores ha de ser necessàriament igual a 1, i que l'aproximació de Gauss sempre serà millor que la de Legendre per  $x$  prou gran, fins i tot si es canvia la constant  $B$  proposada per Legendre per qualsevol altra. A la figura 2 es comparen aquestes dues aproximacions. Observi's que, fins a tres milions, segurament el valor més gran per al qual Gauss havia calculat  $\pi$ , l'estimació de Legendre encara supera la seva, però la situació canvia ben aviat. Després d'un interval en que totes dues s'assemblen força, la funció Li pren la primera posició, i la conservarà per sempre.

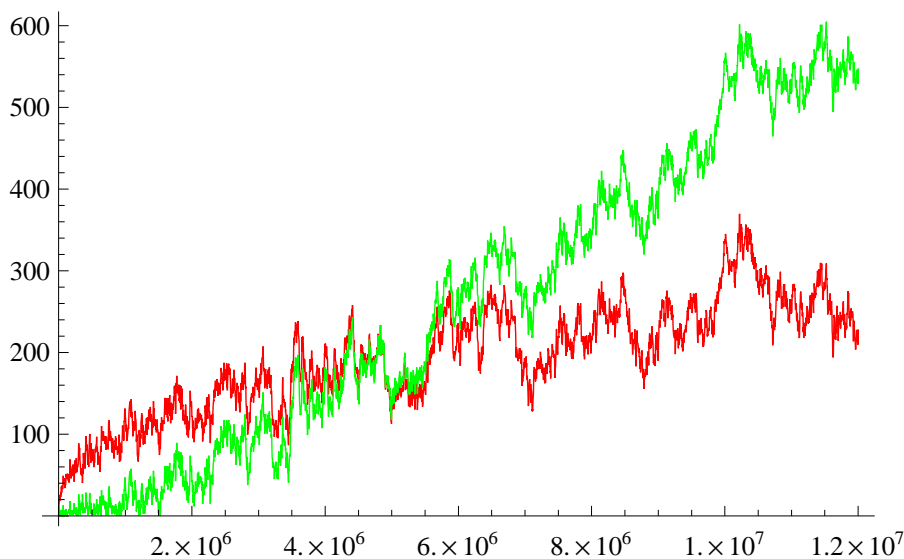


FIGURA 2. Error en les aproximacions de  $\pi$  de Legendre (verd) i Gauss (vermell).

### 3. EL TREBALL DE RIEMANN

El contingut de l'article de Riemann [12] s'organitza en dues parts, cadascuna de les quals ocupa aproximadament la meitat de la seva extensió. La primera està dedicada a definir i estudiar la funció  $\zeta$ . Riemann dóna dos mètodes per obtenir-ne la prolongació analítica i demostrar que satisfà una equació funcional, estudia algunes propietats dels seus zeros, i finalment en troba una expressió com a producte infinit. La segona és l'aplicació del que ha fet prèviament a l'estudi de



la distribució dels nombres primers, comença definint una nova funció  $\pi^*(x)$  per comptar primers que és una variant de  $\pi(x)$ , amb la qual està relacionada per una expressió simple, i dedueix una fórmula explícita que proporciona aquesta funció com la suma d'un terme principal i d'una sèrie infinita de funcions oscil·ladores que depenen dels zeros de la funció zeta. Finalment, la relació entre les funcions  $\pi^*(x)$  i  $\pi(x)$  li permet obtenir una fórmula anàloga per a la funció  $\pi(x)$ , que és el resultat principal del treball. En aquesta secció es donaran alguns detalls i es comentaran breument els aspectes més importants del contingut d'aquest treball de Riemann.

**La funció  $\zeta$ . Prolongació analítica i equació funcional.** Riemann defineix la funció  $\zeta$  a partir del valor comú de la suma de la sèrie harmònica i el producte d'Euler corresponent

$$(3) \quad \zeta(s) := \sum_{n \geq 1} \frac{1}{n^s} = \prod_p \frac{1}{1 - \frac{1}{p^s}}, \quad \Re(s) > 1,$$

totes dues expressions convergents per nombres complexos amb part real  $> 1$ , i que defineixen una funció analítica en aquest semiplà.

La funció zeta i el producte (3) ja havien estat considerats anteriorment per Euler per a valors enters de la variable  $s$ , que havia arribat a una identitat anàloga a l'equació funcional de Riemann per a aquest tipus de valors i havia aconseguit calcular el valor de  $\zeta(n)$  per a tot enter parell  $n \geq 2$ , obtenint la famosa fórmula en que apareixen els nombres de Bernoulli. També havien estat considerats per Dirichlet, un dels mestres de Riemann, que els va fer jugar un paper important en la seva demostració del teorema de la progressió aritmètica, i sembla que Riemann pot haver estat influït per Dirichlet a l'hora d'enfocar-hi la seva atenció.

La veritable innovació de Riemann consisteix a considerar  $\zeta$  com una funció de la variable complexa  $s$  i estendre-la a tot el pla complex, el qual li permetrà descobrir la importància fonamental que tenen els zeros no trivials d'aquesta funció, situats fora de la recta real i a l'esquerra del semiplà de convergència, per a l'estudi de la distribució dels nombres primers.

Riemann obté, per dues vies diferents, una «expressió que dóna el valor de la funció  $\zeta$  per a tot valor de  $s$ .» En terminologia actual es tracta de la prolongació analítica de la funció definida per (3) a una funció meromorfa a tot el pla complex llevat d'un pol (que resulta ser simple, de residu 1) en el punt  $s = 1$ .

La primera expressió la troba de la manera següent: si a la funció gamma d'Euler  $\Gamma(s) = \int_0^\infty e^{-x} x^{s-1} dx$  es fa el canvi de variable de  $x$  per  $nx$  i se suma sobre tots els valors de  $n \geq 1$  s'arriba a

$$\Gamma(s)\zeta(s) = \int_0^\infty \frac{x^{s-1}}{e^x - 1} dx, \quad \Re(s) > 1.$$

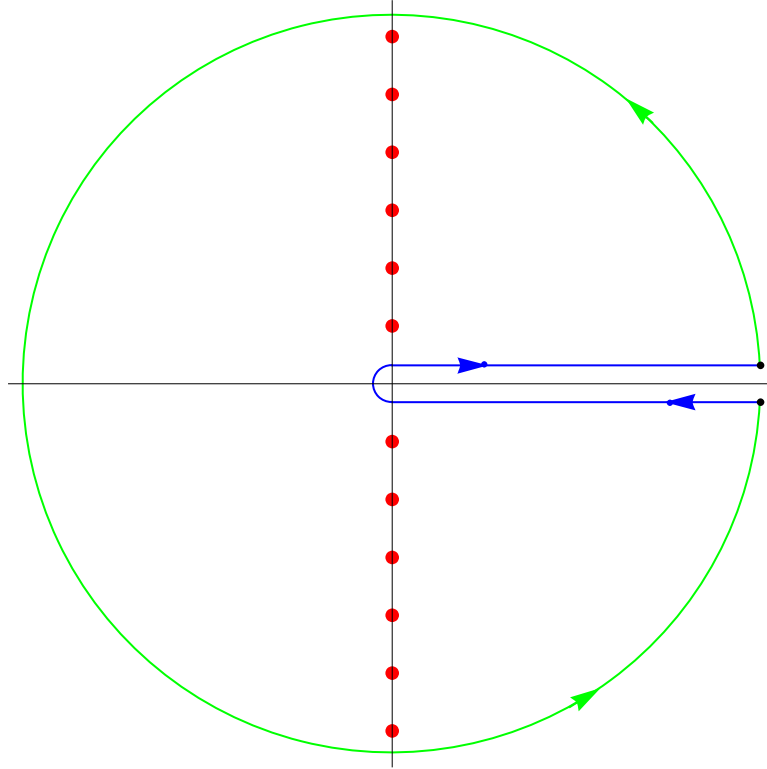
Riemann integra la funció  $\frac{(-x)^{s-1}}{e^x - 1}$  entre  $+\infty$  i  $+\infty$  sobre un camí que envolta el semieix real positiu, dibuixat en blau a la figura 3 (amb els extrems tendint a  $+\infty$ ), i on  $(-x)^{s-1} = e^{(s-1)\log(-x)}$  essent  $\log$  la determinació principal del logaritme fora del semieix real negatiu. Tenint en compte la identitat anterior, obté una expressió

$$(4) \quad \zeta(s) = \frac{\Gamma(1-s)}{2\pi i} \int_{+\infty}^{+\infty} \frac{(-x)^s}{e^x - 1} \frac{dx}{x}$$

vàlida per  $\Re(s) > 1$ . Ara bé, la integral d'aquesta fórmula convergeix per a tot valor complex de la variable  $s$  i, com que la convergència és uniforme sobre compactes, defineix una funció entera. Tenint en compte que la funció gamma és meromorfa a tot  $\mathbb{C}$ , amb pols simples als enters  $\leq 0$ , i que la sèrie harmònica defineix una funció holomorfa per  $\Re(s) > 1$  i és divergent en  $s = 1$ , es dedueix que l'expressió a la dreta de (4) és una funció meromorfa amb un únic pol simple en el punt  $s = 1$ , que és l'extensió de la funció  $\zeta$  a tot el pla complex que es volia. A continuació Riemann integra de nou la mateixa funció d'abans (per valors de la variable  $s$  amb part real negativa) sobre el camí tancat de la figura 3 (amb el radi exterior tendint a infinit). Igualant el resultat amb la suma dels residus a les singularitats, que estan als punts  $2\pi in$  per  $n \neq 0$ , obté l'equació funcional

$$(5) \quad \zeta(s) = \Gamma(1-s)(2\pi)^{s-1} 2 \sin(\pi s/2) \zeta(1-s),$$

que relaciona els valors de la funció  $\zeta$  en els punts  $s$  i  $1-s$ . Tenint en compte identitats ben conegudes de la funció gamma, Riemann observa que l'equació funcional anterior es pot enunciar de manera més elegant

FIGURA 3. Camí d'integració. El radi exterior tendeix a  $\infty$ .

com

la funció  $\Gamma(s/2)\pi^{-s/2}\zeta(s)$  és invariant per  $s \mapsto 1 - s$ .

Aquesta expressió tan simètrica inspira a Riemann la segona demostració de la prolongació analítica i de l'equació funcional: en canviar  $x$  per  $n^2\pi x$  a la integral que defineix la funció gamma i sumar sobre tots els valors  $n \geq 1$  s'obté l'expressió com a transformada integral

$$\Gamma(s/2)\pi^{-s/2}\zeta(s) = \int_0^\infty \psi(x)x^{s/2-1} dx,$$

on  $\psi(x) = \sum_{n=1}^\infty e^{-\pi n^2 x}$  és essencialment la funció theta de Jacobi. Separant la integral anterior en dos intervals pel punt  $x = 1$ , i aplicant l'equació funcional  $2\psi(1/x) + 1 = \sqrt{x}(2\psi(x) + 1)$  deduïda de la que relaciona els valors de la funció theta en  $x$  i en  $1/x$ , al tros de la integral

entre zero i u, s'obté la fórmula

$$\Gamma(s/2)\pi^{-s/2}\zeta(s) = \frac{1}{s(s-1)} + \int_1^\infty \psi(x) (x^{s/2} + x^{(1-s)/2}) \frac{dx}{x},$$

que permet també estendre la funció zeta a tot el pla complex, ja que la integral que hi apareix convergeix en tot nombre complex definit una funció entera. A més, en aquesta expressió l'equació funcional es llegeix directament, ja que la part de la dreta es queda igual en canviar  $s$  per  $1-s$ . Aquest mateix mètode serà utilitzat més tard per Hecke per obtenir la prolongació analítica i l'equació funcional de les funcions  $L$  associades a formes modulars.

**Els zeros de la funció  $\zeta$ .** De la convergència del producte d'Euler (3) per  $\Re(s) > 1$  es dedueix la no anul·lació de la funció  $\zeta$  en aquest semiplà, i de l'equació funcional (5) en resulta que al semiplà  $\Re(s) < 0$  la funció  $\zeta$  té zeros simples exactament en els enters parells negatius  $s = -2, -4, -6, \dots$ , que es coneixen amb el nom de *zeros trivials* de la funció  $\zeta$ . Tots els demés zeros, els *zeros no trivials*, han d'estar per tant en l'anomenada *banda crítica*  $0 \leq \Re(s) \leq 1$ . L'eix de simetria d'aquesta banda, la recta  $\Re(s) = 1/2$  s'anomena la *recta crítica*. Per tal d'estudiar aquests zeros Riemann considera la funció

$$\xi(s) := \frac{1}{2}s(s-1)\Gamma\left(\frac{s}{2}\right)\pi^{-s/2}\zeta(s),$$

que és una funció entera, ja que la multiplicació per  $s-1$  anul·la el pol de  $\zeta$ , la multiplicació per  $s$  anul·la el pol de  $\Gamma(s/2)$  en zero, i els zeros trivials de  $\zeta(s)$  es corresponen exactament amb els demés pols de  $\Gamma(s/2)$ . Per tant, la funció  $\xi$  té per zeros exactament els zeros no trivials de la funció  $\zeta$ . L'equació funcional de la funció  $\zeta$  equival a  $\xi(s) = \xi(1-s)$ . Observant que  $\xi(\bar{s}) = \overline{\xi(s)}$ , es dedueix que la funció  $\xi$  pren valors reals a la recta crítica, i que els zeros no trivials de  $\zeta$  estan situats simètricament tant respecte de la recta real com respecte de la recta crítica.

Integrant la derivada logarítmica de la funció  $\xi$  al voltant del rectangle  $[0, 1] \times [0, T]$ , Riemann diu haver calculat el nombre de zeros  $N(T)$  de la funció  $\zeta$  en aquesta regió i afirma que aquest càlcul proporciona l'estimació següent

$$(6) \quad N(T) = \frac{T}{2\pi} \log \frac{T}{2\pi} - \frac{T}{2\pi} + O(\log T).$$

En fer aquesta afirmació sense donar cap detall sobre la manera d'arribar al resultat, Riemann, que era un mestre consumat en el càlcul d'integrals definides, peca d'optimista sobre les habilitats dels lectors de l'article: la primera demostració rigorosa de (6) la donarà von Mangoldt l'any 1905, i per això aquesta fórmula es coneix de vegades com a fórmula de Riemann-von Mangoldt.

La frase que Riemann escriu a continuació és l'única de l'article que avui dia segueix essent un repte. Referint-se a la funció  $\xi(\frac{1}{2} + it)$  com a funció de la variable  $t$  (complexa) diu: «hom troba efectivament aproximadament aquest nombre d'arrels reals dins de la regió considerada, i és molt probable que totes les arrels siguin reals.» En sentit estricte aquesta frase conté una afirmació: que el nombre de zeros  $N_0(T)$  de part real  $1/2$  en el rectangle  $[0, 1] \times [0, T]$  també admet la mateixa estimació (6), seguida d'una conjectura: que tots els zeros de la funció  $\xi$  tenen part real  $1/2$ , o sigui que per a tot valor de  $T$  es tindrà sempre una igualtat  $N(T) = N_0(T)$ .

Pel que fa a l'afirmació, avui dia encara no ha estat demostrada. Van haver de passar molts anys abans que es trobessin demostracions de resultats molt més dèbils: l'any 1914 Hardy va aconseguir demostrar que  $\xi$  té infinits zeros a la recta crítica, i més concretament que  $N_0(T) > CT$  per una constant positiva  $C$  i  $T$  prou gran; l'any 1943 Selberg va millorar aquesta estimació demostrant que  $N_0(T) > C \frac{T}{2\pi} \log\left(\frac{T}{2\pi}\right)$  per alguna constant positiva menor que 1 i  $T$  prou gran, el qual es pot interpretar com que una fracció positiva  $C$  dels zeros no trivials de la funció zeta cauen sobre la recta crítica. Levinson provà el 1974 que  $C > 1/3$  i Conrey ho millorà el 1989 veient que  $C > 2/5$ . Tot i que fa uns anys alguns investigadors anunciaven haver demostrat una nova fita que implicaria que almenys la meitat dels zeros de  $\zeta$  satisfan la hipòtesi de Riemann, el resultat no ha estat mai publicat.

Quant a la conjectura que tots els zeros de  $\xi$  estan situats sobre la recta crítica  $\Re(s) = 1/2$ , Riemann escriu: «Naturalment, hom desitjaria una demostració estricta d'aquest fet; he deixat de banda la recerca d'una tal prova després d'alguns intents no reeixits, ja que no cal per a l'objectiu immediat de la meua investigació.» Aquesta conjectura ha passat a formar part de la llista dels problemes oberts més difícils de les matemàtiques amb el nom d'*hipòtesi de Riemann*.

**Expressió en producte infinit.** En l'obra de Riemann apareix sovint el problema de la caracterització de funcions analítiques en termes de les seves singularitats. Naturalment, la funció  $\xi$  també és objecte d'una anàlisi d'aquesta mena. Argumentant que la funció  $\log \xi$  té les mateixes singularitats logarítmiques que la suma  $\sum_{\xi(\rho)=0} \log(1 - s/\rho)$ , i que totes dues funcions tenen un comportament semblant a l'infinit, Riemann justifica que han de diferir en una constant i en exponenciar obté l'expressió en producte infinit

$$(7) \quad \xi(s) = \xi(0) \prod_{\substack{\rho \in \mathbb{C} \\ \xi(\rho)=0}} \left(1 - \frac{s}{\rho}\right),$$

on se sobreentén que cada factor apareix en el producte tantes vegades com la multiplicitat de  $\rho$  com a zero de la funció, i que el producte s'ha de fer en ordre creixent del mòdul de les arrels per assegurar-ne la convergència. Aquesta factorització és anàloga a la que es donaria si  $\xi$  fos un polinomi.

Tot i que les raons que Riemann esgrimeix són certament les que acaben portant a la fórmula (7), la justificació que en fa al seu article està molt lluny de poder-se considerar una demostració. En particular la convergència de la suma  $\sum \log\left(1 - \frac{s}{\rho}\right)$  requereix d'una estimació de les magnituds de les arrels  $\rho$ , i la que Riemann addueix és conseqüència de (6), fórmula que no es demostraria fins gairebé mig segle més tard. D'altra banda els arguments que dóna per justificar el comportament de  $\xi(s)$  per  $s \rightarrow \infty$  són massa vagues per poder-los acceptar com a una demostració.

La representació com a producte infinit (7) va ser demostrada per primera vegada per Hadamard l'any 1893 en un treball en que investiga i resol el problema general de trobar expressions en producte infinit per a funcions enteres.

**La fórmula explícita.** Immediatament després de trobar la fórmula (7), cap al final de la pàgina 3 del seu manuscrit, Riemann dóna per acabat el seu estudi de la funció  $\zeta$  i comença a ocupar-se de la distribució dels nombres primers. Diu: «Amb aquests fets preparatoris, ara es pot determinar el nombre de primers menors que  $x$ .»

Tot i que l'objectiu és estudiar la funció  $\pi$ , Riemann introdueix una altra funció anàloga, que compta no només els nombres primers sinó també les potències de primer, tot i que cada potència  $p^n$  compta només amb un pes igual a  $1/n$ ,

$$\pi^*(x) = \sum_{p^n \leq x} \frac{1}{n}.$$

De fet, per tal que les fórmules que apareixeran més endavant per aquesta funció siguin correctes, tan a les funcions  $\pi$  i  $\pi^*$  com en totes les funcions amb discontinuïtats de salt, el valor en els punts on hi ha la discontinuïtat s'han de redefinir com la semisuma dels límits laterals, com és habitual en l'anàlisi de Fourier.

La relació entre aquesta funció i la funció  $\zeta$  s'obté de la manera següent: substituint cada sumand  $p^{-ns}$  per l'expressió  $s \int_{p^n}^{\infty} x^{-s-1} ds$  en el logaritme del producte d'Euler,

$$\log \zeta(s) = - \sum_p \log(1 - p^{-s}) = \sum_p \sum_{n \geq 1} \frac{1}{n} p^{-ns}$$

Riemann arriba a l'expressió

$$\log \zeta(s) = s \int_1^{\infty} \pi^*(x) x^{-s} \frac{dx}{x}$$

de  $\log \zeta$  com a transformada integral de la funció  $\pi^*$ . Aleshores, per inversió de Fourier, obté la fórmula

$$(8) \quad \pi^*(x) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \log \zeta(s) x^s \frac{ds}{s}, \quad (a > 1),$$

que dona la funció  $\pi^*$  en termes de la funció  $\log \zeta$ .

De la relació entre les funcions  $\zeta$  i  $\xi$  i de la descomposició en producte d'aquesta segona funció resulta que  $\log \zeta(s)$  és la suma

$$-\log(s-1) + \sum_{\xi(\rho)=0} \log\left(1 - \frac{s}{\rho}\right) - \log \Gamma\left(\frac{s}{2} + 1\right) + \frac{s}{2} \log \pi + \log \xi(0),$$

i Riemann obté la fórmula explícita per a la funció  $\pi^*$  com el resultat de substituir  $\log \zeta(s)$  per cadascun dels sumands d'aquesta expressió en la fórmula (8), i calculant les integrals corresponents. De fet el

procediment es complica una mica ja que fer-ho així donaria lloc a integrals divergents i en realitat la substitució es fa en la fórmula

$$\pi^*(x) = \frac{-1}{2\pi i} \frac{1}{\log x} \int_{a-i\infty}^{a+i\infty} \frac{d}{ds} \left[ \frac{\log \zeta(s)}{s} \right] x^s ds$$

que es dedueix de (8) integrant per parts.

Ara, introduint en aquesta expressió cadascun dels sumands en què s'ha expressat  $\log \zeta(s)$ , es van obtenint els termes de la fórmula explícita: el primer és

$$\frac{-1}{2\pi i} \frac{1}{\log x} \int_{a-i\infty}^{a+i\infty} \frac{d}{ds} \left[ \frac{-\log \zeta(s)}{s} \right] x^s ds = \int_0^x \frac{1}{\log t} dt = \text{Li}(x).$$

El terme corresponent a la suma sobre les arrels de  $\xi$  dona

$$\frac{-1}{2\pi i} \frac{1}{\log x} \int_{a-i\infty}^{a+i\infty} \frac{d}{ds} \left[ \frac{\sum \log \left( 1 - \frac{s}{\rho} \right)}{s} \right] x^s ds = \sum_{\xi(\rho)=0} \text{Li}(x^\rho),$$

on  $\text{Li}(x^\rho)$  s'ha d'interpretar com el valor obtingut, fixant  $x$ , per prolongació analítica de la funció  $\text{Li}(x^\beta)$  dels  $\beta$  reals positius a tot el semiplà complex  $\beta > 0$ . Aquí cal dir que Riemann no justifica prou els passos que segueix per arribar a aquest resultat. En particular, reconeix que el fet que es pugui avaluar terme a terme per a cada sumand requeria una discussió més acurada. D'altra banda els termes  $\text{Li}(x^\rho)$  només tenen sentit si  $\Re(\rho) > 0$  i Riemann no havia demostrat que els zeros no poguessin estar a l'eix imaginari, fet que va ser més endavant clau en la demostració del teorema dels nombres primers.

El terme corresponent al sumand de la funció  $\Gamma$  dona

$$\frac{-1}{2\pi i} \frac{1}{\log x} \int_{a-i\infty}^{a+i\infty} \frac{d}{ds} \left[ \frac{-\log \Gamma \left( \frac{s}{2} + 1 \right)}{s} \right] x^s ds = \int_x^\infty \frac{dt}{t(t^2 - 1) \log t}.$$

Quant al terme  $\frac{s}{2} \log \pi$ , la divisió per  $s$  el converteix en constant i al derivar s'anul·la. Finalment, la integral sobre la constant  $\xi(0)$  dona com a valor aquesta mateixa constant, que és igual a  $-\log(2)$ .



Sumant tots els resultats anteriors s'arriba a la fórmula explícita per a la funció  $\pi^*$

$$(9) \quad \pi^*(x) = \text{Li}(x) - \sum_{\substack{\rho \in \mathbb{C} \\ \xi(\rho)=0}} \text{Li}(x^\rho) + \int_x^\infty \frac{dx}{x(x^2-1)\log x} - \log 2.$$

Aquesta fórmula és el resultat principal de l'article de Riemann i aconseguix donar el valor exacte de la funció esglaonada  $\pi^*(x)$  com la suma d'un terme principal  $\text{Li}(x)$ , dos termes residuals fitats, i una suma infinita de termes oscil·latòris corresponents a les arrels no trivials de  $\zeta$ . És fàcil comprovar que el terme fitat que s'escriu com una integral és igual a la suma  $\sum_{n \geq 1} \text{Li}(x^{-2n})$ , que té la mateixa forma que el sumatori de (9) però per les arrels trivials de  $\zeta$  en els enters negatius parells, de manera que la fórmula de Riemann es pot escriure, de manera més simple,

$$\pi^*(x) = \text{Li}(x) - \log(2) - \sum_{\substack{\rho \in \mathbb{C} \\ \zeta(\rho)=0}} \text{Li}(x^\rho).$$

Les gràfiques de les figures 4, 5 i 6 serveixen com a il·lustració de la fórmula explícita, mostrant l'aspecte que tenen els termes oscil·latòris corresponents a algunes arrels (en agrupar-les per parelles simètriques es cancel·len les parts imaginàries) i dibuixant les aproximacions de  $\pi^*$  que s'obtenen a partir de sumes parcials en dos intervals diferents.

**La funció Ri de Riemann.** En els darrers paràgrafs de l'article Riemann considera la funció  $\pi$  i enuncia les conseqüències per aquesta funció que s'obtenen dels seus resultats. Només a partir de les definicions la relació entre  $\pi^*$  i  $\pi$  s'obté immediatament:

$$\pi^*(x) = \pi(x) + \frac{1}{2}\pi(x^{1/2}) + \frac{1}{3}\pi(x^{1/3}) + \dots = \sum_{n=1}^{\infty} \frac{1}{n}\pi(x^{1/n}),$$

i amb la fórmula d'inversió de Möbius s'obté la relació recíproca

$$\begin{aligned} \pi(x) &= \pi^*(x) - \frac{1}{2}\pi^*(x^{1/2}) - \frac{1}{3}\pi^*(x^{1/3}) - \frac{1}{5}\pi^*(x^{1/5}) + \frac{1}{6}\pi^*(x^{1/6}) + \dots \\ &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n}\pi^*(x^{1/n}), \end{aligned}$$

on  $\mu$  és la funció de Möbius que val zero en els enters divisibles pel quadrat d'algun primer i val  $(-1)^m$  en els enters divisibles exactament

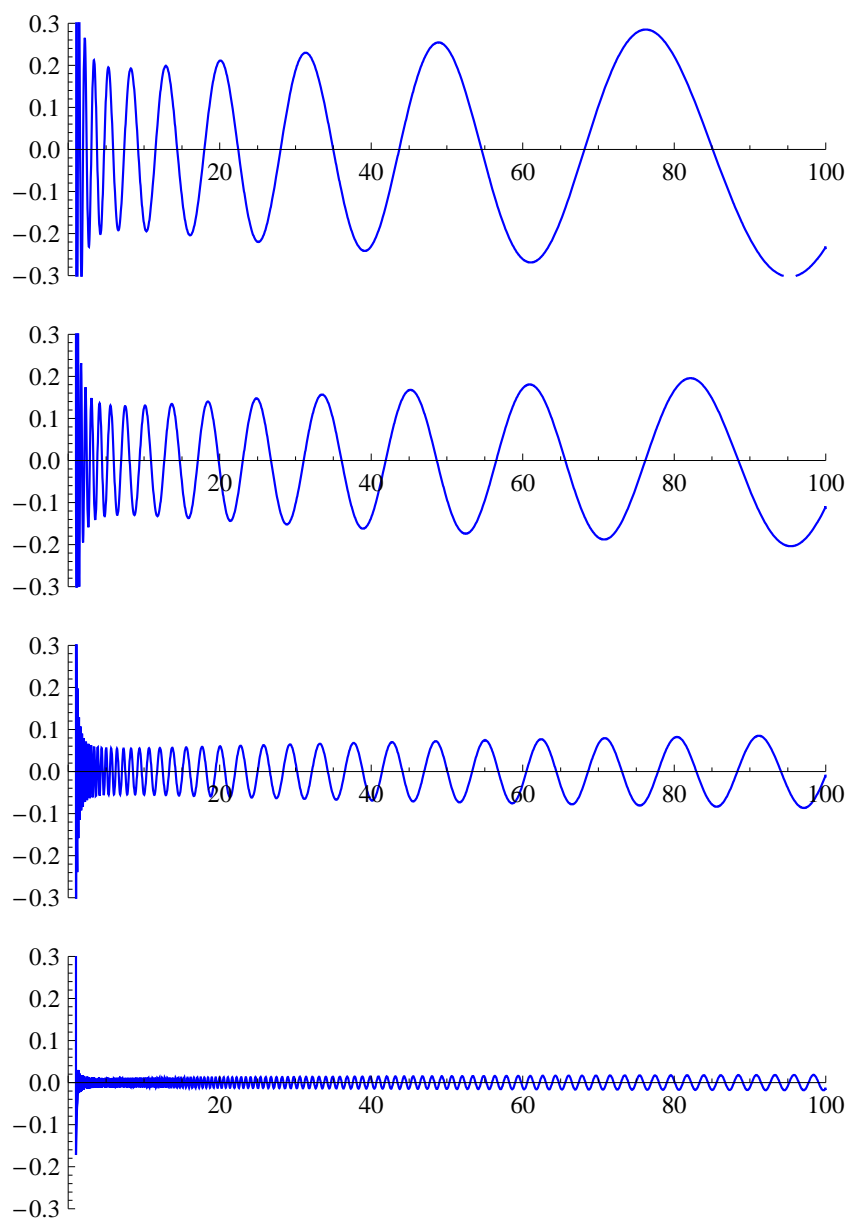


FIGURA 4. Termes oscil·latòris  $-\text{Li}(x^\rho) - \text{Li}(x^{1-\rho})$  corresponents als parells d'arrels números 1,2,10 i 100.

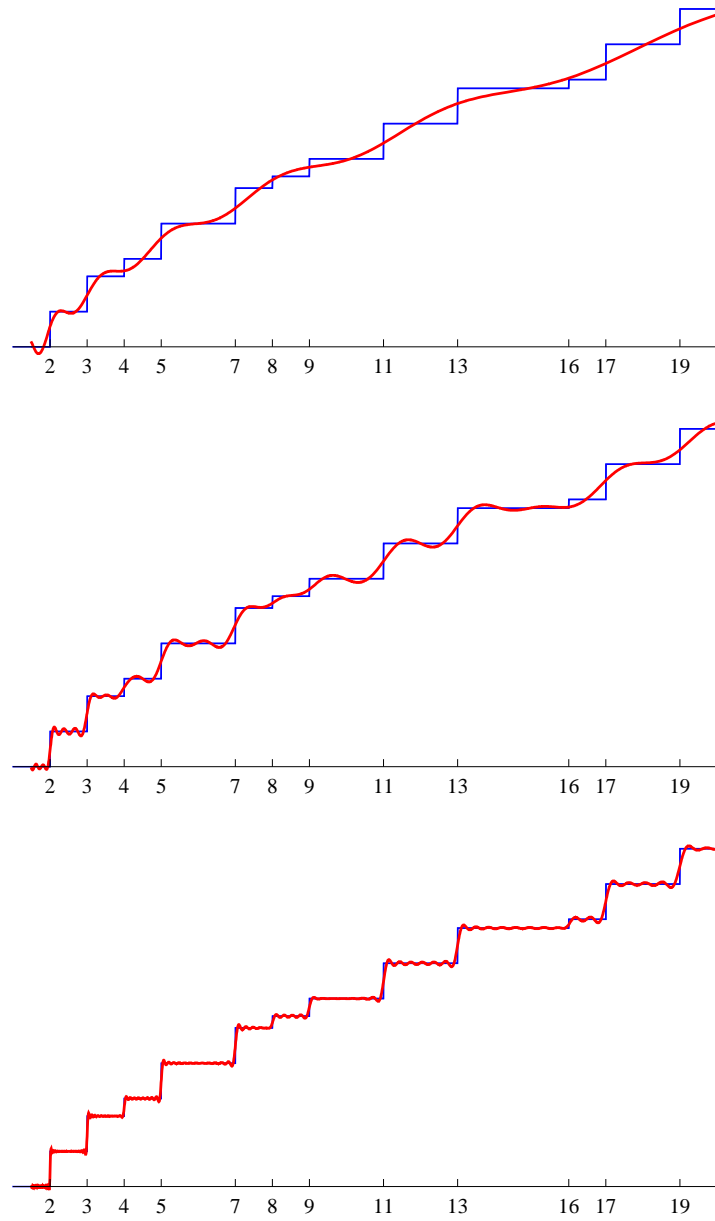


FIGURA 5. Aproximació de  $\pi^*$  a l'interval  $(1, 20]$  amb la fórmula explícita de Riemann usant 1, 10 i 100 zeros de  $\zeta$ , respectivament.

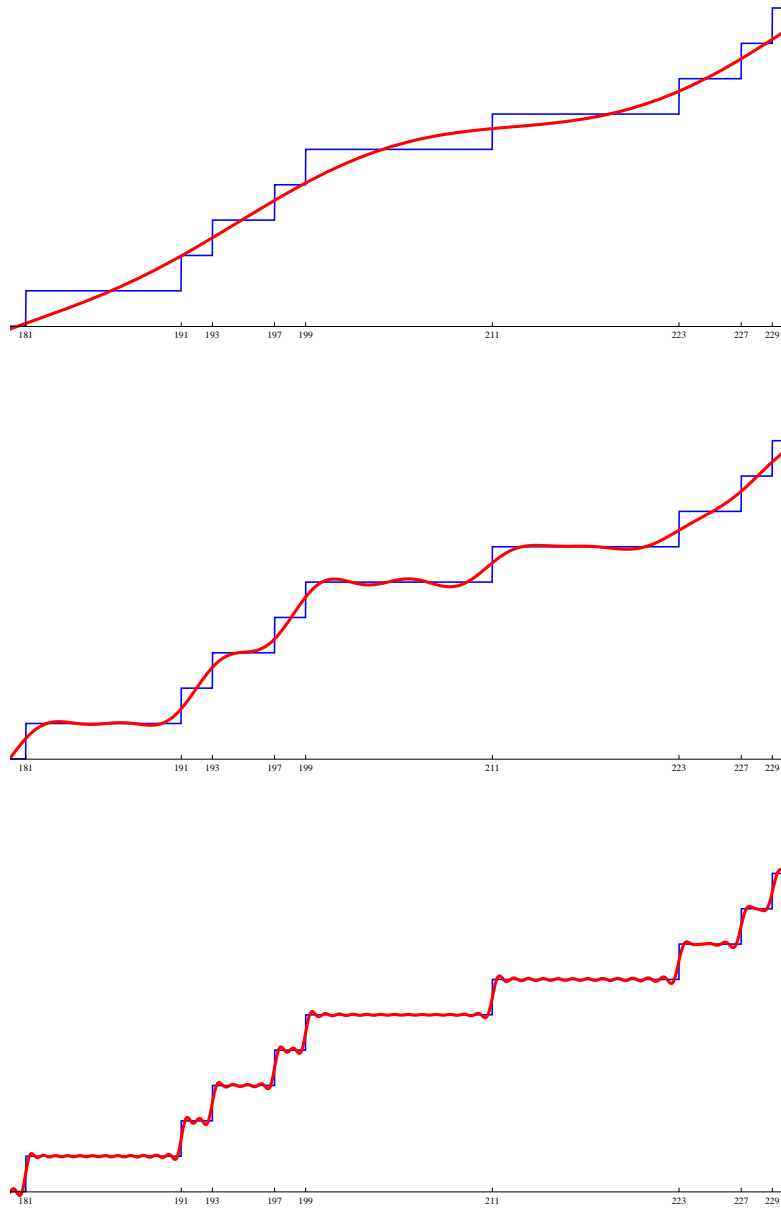


FIGURA 6. Aproximació de  $\pi^*$  a l'interval  $[180, 230]$  amb la fórmula explícita de Riemann usant 10, 100 i 1000 zeros de  $\zeta$ , respectivament.

per  $m$  primers diferents. Atès que totes dues funcions  $\pi$  i  $\pi^*$  valen zero per  $x < 2$  les sumes anteriors són en realitat finites per a cada valor de  $x$ , ja que eventualment serà  $x^{1/n} < 2$ .

Ara, substituint  $\pi^*$  per la sèrie de la fórmula explícita (9) en el sumatori anterior i arrançant la suma terme a terme s'arriba a una fórmula explícita semblant per a la funció  $\pi$ . El terme principal d'aquesta fórmula és la funció

$$\begin{aligned} \text{Ri}(x) &= \text{Li}(x) - \frac{1}{2} \text{Li}(x^{1/2}) - \frac{1}{3} \text{Li}(x^{1/3}) - \frac{1}{5} \text{Li}(x^{1/5}) + \frac{1}{6} \text{Li}(x^{1/6}) + \dots \\ &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \text{Li}(x^{1/n}), \end{aligned}$$

que es coneix com a *funció de Riemann*. Per a valors reals  $x > 1$  aquesta funció admet la sèrie ràpidament convergent següent, anomenada *sèrie de Gram*:

$$\text{Ri}(x) = 1 + \sum_{n=1}^{\infty} \frac{(\log x)^n}{n!n\zeta(n+1)}.$$

Els termes corresponents al sumatori sobre les arrels de  $\zeta$  de les funcions  $\text{Li}(x^p)$  donen lloc a un sumatori de funcions  $\text{Ri}(x^p)$  i el terme constant  $\log(2)$  desapareix gràcies a la identitat  $\sum_{n \geq 1} \mu(n)/n = 0$  (que, poca broma!, és equivalent al teorema dels nombres primers). D'aquesta manera s'obté l'elegant fórmula explícita (2) per a la funció  $\pi$  enunciada a la introducció.

Per aquestes consideracions Riemann ha aconseguit arribar a una nova funció  $\text{Ri}(x)$  que s'afegeix a les funcions ja conegudes  $x/\log x$  i  $\text{Li}(x)$  com a candidata a aproximar la funció  $\pi$ . En l'últim paràgraf Riemann fa referència a les dades calculades per Gauss. Dels valors de  $\pi(x)$  fins a tres milions, observa que la funció  $\text{Li}$  sobreestima la quantitat de nombres primers cometent un error de magnitud comparable a l'arrel quadrada, i que el segon terme  $-\frac{1}{2} \text{Li}(\sqrt{x})$  en la suma que defineix  $\text{Ri}$ , que és el terme dominant de la diferència entre totes dues, sembla corregir aquesta sobreestimació, tal i com es pot constatar a les gràfiques de les figures 7 i 8. Des d'aleshores les taules de nombres primers i de valors de la funció  $\pi$  s'han anat ampliant, sobretot des de la proliferació de les calculadores electròniques. A més, i pel que fa al càlcul de  $\pi$ , l'any 1871 l'astrònom alemany Meissel va proposar un

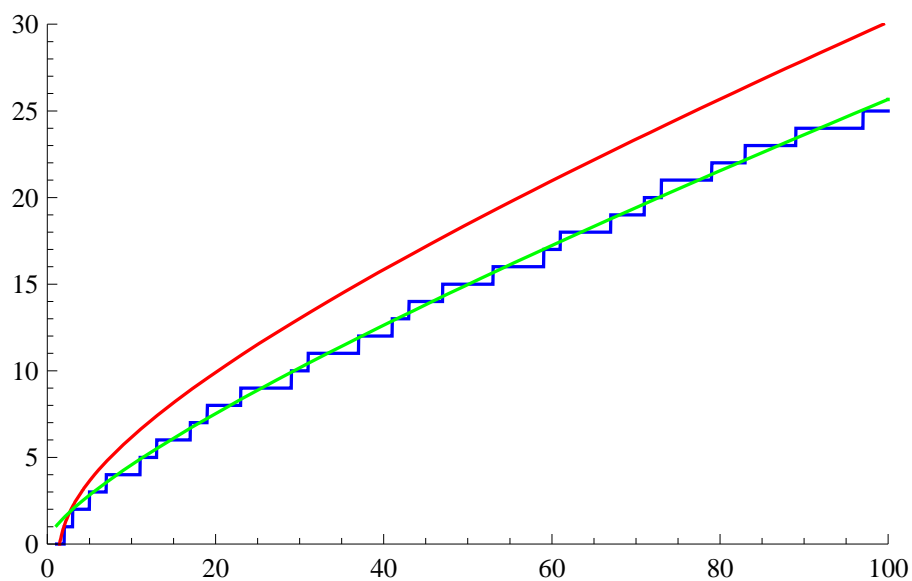


FIGURA 7. Les funcions  $\text{Li}$  (vermell) i  $\text{Ri}$  (verd) comparades amb  $\pi$  per  $x \leq 100$

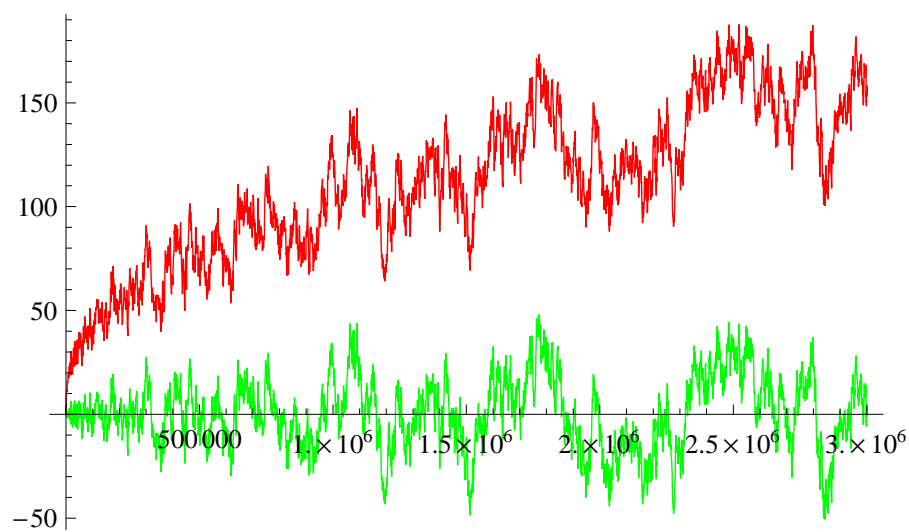


FIGURA 8. Diferències  $\text{Li}(x) - \pi(x)$  (vermell) i  $\text{Ri}(x) - \pi(x)$  (verd) a l'interval  $x \leq 3\,000\,000$

mètode combinatori que millora sensiblement el càlcul basat en el garbell d'Eratòstenes, i que ha estat perfeccionat i simplificat per Lehmer i altres.

Actualment se sap el valor exacte de  $\pi(x)$  per a molts valors de  $x$  de l'ordre de les 20 xifres decimals, en particular per a les potències de 10 fins a la 23-ena. La taula 2 conté aquests valors i els compara amb els proporcionats per les tres aproximacions de Legendre, Gauss i Riemann, respectivament. En aquesta taula s'observa com les funcions

$n$	$\pi(x)$	$\pi(x) - x/\log x$	$\text{Li}(x) - \pi(x)$	$\pi(x) - \text{Ri}(x)$
3	168	23	9	0
4	1229	143	16	-2
5	9592	906	37	-5
6	78498	6116	129	29
7	664579	44158	338	88
8	5761455	332774	753	97
9	50847534	2592592	1700	-79
10	455052511	20758029	3103	-1828
11	4118054813	169923159	11587	-2318
12	37607912018	1416705193	38262	-1476
13	346065536839	11992858452	108970	-5773
14	3204941750802	102838308636	314889	-19200
15	29844570422669	891604962452	1052617	73218
16	279238341033925	7804289844393	3214632	327052
17	2623557157654233	68883734693928	7956590	-598255
18	24739954287740860	612483070893536	21949508	-3501366
19	234057667276344607	5481624169369952	99878336	23884333
20	2220819602560918840	49347193044659712	222745856	-4891825
21	21127269486018731928	446579871578169344	597340160	-86432204
22	201467286689315906290	4060704006019645440	1932787712	-127132665
23	1925320391606818006727	37083513766592905216	7236222976	1019262049

TAULA 2. Valors de  $\pi$  en  $x = 10^n$  fins al rècord actual i comparació amb les aproximacions de Legendre, Gauss i Riemann.

$x/\log x$  i  $\text{Li}(x)$  donen aproximacions respectivament per defecte i per excés en tots els casos, i que l'error comès per la segona és clarament inferior. El nombre de xifres de  $x$  és notòriament el doble que el nombre de xifres de l'error  $\text{Li}(x) - \pi(x)$ , el qual ratifica que la magnitud de l'error és, com ja havia observat Riemann, de l'ordre de l'arrel quadrada.

Per a tots els valors de la taula la funció Ri aproxima  $\pi$  millor que Li i la diferència  $\text{Ri}(x) - \pi(x)$  va canviant de signe. Aquesta aproximació,

tot i ser millor, no ho és significativament en magnitud, ja que el nombre de xifres de tots dos errors sembla comportar-se de manera molt similar. Tal i com es veurà a la secció següent en parlar de l'error en el teorema dels nombres primers, tot i que la experimentació numèrica i la observació de pautes i correlacions en les dades és una pràctica molt útil en matemàtiques, s'ha d'anar amb molt de compte en deixar-se convèncer d'un fet només basant-se en una propietat observada numèricament, fins i tot si es disposa d'una quantitat ingent de dades. Vegeu al respecte els «philosophical comments» de Peter Sarnak sobre la verificació numèrica de la hipòtesi de Riemann al final de la seva presentació del problema del mil·leni [14].

#### 4. DESPRÉS DE RIEMANN

Els efectes del treball de Riemann van trigar a arribar: després de la publicació es va fer el silenci durant gairebé 30 anys. Després, a l'última dècada del segle XIX floreix una edat d'or de la teoria analítica de nombres que continuarà amb menys intensitat durant el primer quart del segle XX. Es publiquen demostracions de tots els resultats enunciats per Riemann (excepte de la seva estimació del nombre de zeros sobre la recta crítica, que encara no ha estat provada), es demostra el teorema dels nombres primers seguint la via que ell havia obert, i s'obtenen nombrosos resultats sobre el terme d'error i la seva relació amb la hipòtesi de Riemann i amb altres propietats sobre la posició dels zeros no trivials. Des d'aleshores, tot i el rang mític que la hipòtesi de Riemann va assolir des de 1900 gràcies a Hilbert, de les diverses connexions proposades amb diferents branques de les matemàtiques i la física, i de la ingent informació numèrica obtinguda gràcies a la revolució informàtica, es fa difícil destacar progressos espectaculars.

Aquesta secció està dedicada a repassar alguns resultats i idees posteriors directament relacionats amb el treball de Riemann.

**Demostració de la fórmula explícita.** El resultat principal del treball de Riemann és la fórmula explícita (9), i va ser demostrat per von Mangoldt l'any 1895. En realitat el que fa von Mangoldt és demostrar una fórmula anàloga, que en molts aspectes és més simple que la de Riemann, tot i que essencialment es tracta de dues versions diferents d'un



mateix fenomen. Un cop demostrada la seva fórmula, von Mangoldt aconseguí sense massa dificultat passar a la de Riemann.

En les seves investigacions sobre el teorema dels nombres primers, cap al 1850, Chebyshev havia introduït dues noves funcions per comptar primers, anàlogues a les funcions  $\pi$  i  $\pi^*$ . Les funcions

$$\vartheta(x) = \sum_{p \leq x} \log p \quad \text{i} \quad \psi(x) = \sum_{p^n \leq x} \log p,$$

que es coneixen com a *funcions de Chebyshev*. Es tracta també de funcions en forma d'escala, que tenen els graons en els nombres primers i en les seves potències respectivament, exactament igual que  $\pi$  i  $\pi^*$ , amb l'única diferència que l'altura del graó en cada primer  $p$  o en qualsevol potència seva és  $\log p$ . Hi ha fórmules que expressen l'una en funció de l'altra completament anàlogues a les que donen la relació entre  $\pi$  i  $\pi^*$ . Chebyshev havia observat que la funció  $\psi(x)$  creix com  $x$  i havia demostrat sense massa dificultat que tant l'equivalència asimptòtica  $\vartheta(x) \sim x$  així com també  $\psi(x) \sim x$  són totes dues equivalents al teorema dels nombres primers  $\pi(x) \sim \text{Li}(x)$ .

Von Mangoldt arribà a la seva fórmula seguint pas a pas la construcció de Riemann, però partint de la derivada logarítmica  $-\zeta'(s)/\zeta(s)$  en comptes de la funció  $\log \zeta(s)$  i de la funció de Chebyshev  $\psi(x)$  en comptes de la funció  $\pi(x)$ . En expressar la primera com a transformada integral de la segona i per inversió de Fourier es tenen les expressions

$$-\frac{\zeta'(s)}{\zeta(s)} = s \int_0^\infty \psi(x) x^{-s} \frac{dx}{x},$$

$$\psi(x) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \left( -\frac{\zeta'(s)}{\zeta(s)} \right) x^s \frac{ds}{s} \quad (a > 1).$$

Aplicant la derivada logarítmica a la descomposició en producte de  $\zeta$  (que afortunadament havia estat demostrada per Hadamard tot just dos anys abans) von Mangoldt obté

$$-\frac{\zeta'(s)}{\zeta(s)} = \frac{s}{s-1} - \sum_{\zeta(\rho)=0} \frac{s}{\rho(s-\rho)} - \frac{\zeta'(0)}{\zeta(0)}.$$

Finalment, substituint cadascun dels sumands d'aquesta expressió en la fórmula integral per a la funció  $\psi$  i prèvia justificació de la convergència

en cadascun dels passos, arriba a la fórmula

$$(10) \quad \psi(x) = x - \sum_{\xi(\rho)=0} \frac{x^\rho}{\rho} + \frac{1}{2} \log \left( \frac{x^2}{x^2 - 1} \right) - \log 2\pi, \quad x > 1.$$

En comparar-la amb la fórmula explícita de Riemann (9) s'observa que l'analogia entre totes dues és total. La funció  $x$  és el terme principal. La funció  $\frac{1}{2} \log \left( \frac{x^2}{x^2 - 1} \right)$  és monòtona decreixent i tendeix a zero, de manera que la seva contribució és irrellevant; en escriure-la com a la suma de la sèrie  $-\sum_{n \geq 1} \frac{x^{-2n}}{2n}$  es veu que la seva presència a la fórmula correspon als zeros trivials de la funció  $\zeta$ . Cadascun dels zeros no trivials dona lloc a un terme oscil·latòri  $x^\rho/\rho$ , en el benentès que la suma s'ha de fer aparellant cada zero  $\rho$  amb  $1 - \rho$  o bé simplement com a límit de les sumes parcials fent tendir  $|\rho|$  a infinit.

**El teorema dels nombres primers.** El Teorema dels Nombres Primers, conjeat per Gauss i Legendre, va ser finalment demostrat l'any 1898 simultàniament i de manera independent per Hadamard i de la Vallée-Poussin, de manera molt semblant, i tots dos seguint el camí indicat per Riemann.

Segons la versió equivalent del teorema dels nombres primers donada per Chebyshev, el que s'ha de demostrar és que  $\lim_{x \rightarrow \infty} \psi(x)/x = 1$ . Tenint en compte la fórmula explícita de von Mangoldt (10) es dedueix que el teorema dels nombres primers equival al límit

$$\lim_{x \rightarrow \infty} \frac{\sum_{\rho} x^\rho/\rho}{x} = \lim_{x \rightarrow \infty} \sum_{\rho} \frac{x^{\rho-1}}{\rho} = 0.$$

Ara bé, si es demostra que el límit es pot calcular terme a terme, aleshores n'hi ha prou a comprovar que  $x^{\rho-1} \rightarrow 0$  per a tot zero  $\rho$  a la banda crítica, el qual és equivalent a la condició  $0 < \Re(\rho) < 1$ . Atesa la posició simètrica dels zeros respecte la recta crítica, resulta que el teorema dels nombres primers és conseqüència de que la funció zeta de Riemann no tingui zeros sobre la recta  $\Re(s) = 1$ .

Tant Hadamard com de la Vallée-Poussin van evitar haver de justificar l'intercanvi del límit amb el sumatori partint de fórmules que són lleugeres variacions de la fórmula explícita de von Mangoldt i que són expressions no pas per a  $\psi$  sinó per a les funcions  $\int_0^x t^{-k} \psi(t) dt$  amb  $k = 1$  (Hadamard) o  $k = 2$  (de la Vallée-Poussin). Les demostracions

segueixen els mateixos passos però són força més fàcils; de fet sembla que Hadamard volia que la demostració del teorema dels nombres primers no depengués del treball de von Mangoldt publicat tot just feia poc, amb aspectes que encara no havien aconseguit convèncer alguns experts.

Un cop arribats aquí l'únic que cal és demostrar que  $\zeta(s) \neq 0$  per a tot complex  $s$  amb  $\Re(s) = 1$ . Tots dos ho aconsegueixen en veure, combinant la funció zeta, el seu logaritme i la seva derivada logarítmica, i fent servir l'equació funcional, que si hi hagués un zero al punt  $s = 1 + it$  aleshores seria  $|\zeta(1 + 2it)| = \infty$ , que no pot ser ja que  $\zeta$  és holomorfa fora de  $s = 1$ .

Les investigacions de de la Vallée-Poussin sobre els zeros de la funció  $\zeta$  li permeteren arribar a un resultat més fort, en què aconsegueix trobar una regió lliure de zeros dins de la banda crítica. Més endavant aquest tipus de resultats sobre regions lliures de zeros s'han anat millorant successivament, però tots ells són del tipus

$$\zeta(\sigma + it) \neq 0 \quad \text{per} \quad 1 - \frac{C}{\lambda(t)} \leq \sigma \leq 1,$$

per  $t$  suficientment gran, on  $C$  és una constant positiva i  $\lambda(t)$  una funció que tendeix a infinit ( $\log t$  en de la Vallée-Poussin, millorada fins a  $(\log t)^{2/3}(\log \log t)^{1/3}$  per Richert el 1963). A pesar d'això avui dia encara no se sap demostrar que  $\zeta$  no tingui zeros a la banda vertical  $1 - \epsilon < \Re(s) < 1$  per cap nombre positiu  $\epsilon$ .

Provocant un gran enrenou en la comunitat matemàtica, Erdős i Selberg van obtenir l'any 1949 una demostració del teorema dels nombres primers basada només en enginyoses fites de funcions aritmètiques, que no requereix tècniques analítiques ni té cap relació amb resultats de Riemann. Se la coneix com a demostració elemental del teorema dels nombres primers, tot i que la paraula elemental només es pot fer servir aquí en el sentit de no necessitar cap eina analítica, doncs la demostració és intricada i requereix de varies desigualtats fines que no s'obtenen pas sense grans dificultats i demandant una considerable virtuositat tècnica. A pesar d'una certa expectació sobre les conseqüències que podria tenir per a l'estudi del terme d'error per altres vies, i d'una agra disputa per la precedència de l'autoria, la realitat és que les tècniques emprades en aquestes i altres «demostracions elementals» posteriors

no han donat fins ara el fruit obtingut amb la potència de les eines analítiques.

**El terme d'error en el teorema dels nombres primers.** Gauss va observar que la funció  $\text{Li}(x)$  aproxima el valor de  $\pi(x)$  sempre per sobre en tot l'interval on havia fet els seus càlculs, i això s'ha seguit complint per a tots els valors de  $x$  dels quals s'ha pogut calcular el valor de la funció  $\pi$  fins ara (vegi's la taula 2). Per més d'un segle tots els matemàtics (Gauss i Riemann inclosos) pensaren que la diferència  $\text{Li}(x) - \pi(x)$  seguiria sent positiva indefinidament. L'any 1916 Littlewood va demostrar un resultat sorprenent: la diferència  $\text{Li}(x) - \pi(x)$  alterna el signe infinites vegades quan  $x$  va creixent! Més precisament, el que Littlewood demostra és que

$$\text{Li}(x) - \pi(x) = \Omega_{\pm} \left( \sqrt{x} \frac{\log \log \log x}{\log x} \right),$$

on la notació indica que el terme de l'esquerra supera en magnitud el de la dreta, per  $x$  tant gran com es vulgui, i prenent tant valors positius com negatius. Observi's que aquest resultat de Littlewood té una altra conseqüència que contradiu el que les dades numèriques porten a creure: que l'aproximació de Riemann  $\text{Ri}(x)$  del valor de  $\pi(x)$  sigui sempre millor que l'aproximació de Gauss  $\text{Li}(x)$  (vegeu la figura 2 i la taula 2). En efecte, al voltant dels infinits punts  $x$  on la diferència  $\text{Li}(x) - \pi(x)$  travessa l'eix real l'aproximació de Gauss és imbatible!

Fins ara no s'ha pogut donar cap valor numèric concret de  $x$  en què  $\pi(x)$  superi  $\text{Li}(x)$ , i pel que en sabem de la magnitud dels nombres amb aquesta propietat és gairebé segur que ni per més capacitat de càlcul de què es disposi en el futur, ni per molt que es millorin els algorismes, mai serà possible calcular el nombre  $\pi(x)$  exactament en cap d'aquests valors. En efecte, l'any 1933 Skewes demostrà, suposant la hipòtesi de Riemann, que existeixen nombres  $x < 10^{10^{34}}$  amb  $\pi(x) > \text{Li}(x)$ . Aquesta fita es cita sovint com el nombre més descomunal que mai hagi jugat algun paper seriós en un article matemàtic. Mes endavant el mateix autor va aconseguir rebaixar la fita canviant el 34 per un 3 i que resultat fos independent de la hipòtesi de Riemann. Amb el temps la fita s'ha anat rebaixant gràcies a disposar de més zeros de  $\zeta$  calculats cada vegada amb una precisió millor; en un article recent [2] es donen arguments (basats en un gran nombre de zeros) a favor de que

el primer canvi de signe en l'error del teorema dels nombres primers té lloc al voltant de  $1.398 \times 10^{316}$ . L'any 1994 Rubinstein i Sarnak [13] demostraren (mòdul hipòtesis plausibles sobre els zeros de la funció zeta que inclouen la hipòtesi de Riemann) un resultat que d'alguna manera explica perquè cal anar tant lluny i es tan difícil trobar valors negatius a pesar dels infinits canvis de signe que Littlewood ens garanteix que hi ha d'haver: la diferència  $\text{Li}(x) - \pi(x)$  és positiva per al 99.999973... per cent dels nombres!

Es recomana al lector encuriós per aquests resultats que llegeixi l'excel·lent article *Prime number races* d'Andrew Granville i Greg Martin [7] on es descriu un fenomen anàleg que es dona en comptar nombres primers en diferents progressions aritmètiques de la mateixa raó, i que s'explica gràcies als resultats de Littlewood i de Rubinstein i Sarnak aplicats a funcions  $L$  de Dirichlet.

Un altre tòpic que guarda molta relació amb els resultats sobre el terme d'error és el de l'estudi estocàstic de la distribució dels nombres primers, que aprofundeix en la observació de Gauss sobre la variació de la densitat dels nombres primers. A finals dels anys 30 Cramér va proposar fer servir el model estocàstic consistent en una successió de variables aleatòries  $\{X_n\}_{n \geq 2}$  que prenen valors 0, 1 amb probabilitat  $P(X_n = 1) = 1/\log n$ , la qual és un model de la successió dels nombres primers. La idea és que qualsevol propietat d'aquesta successió que es compleixi gairebé segur (amb probabilitat 1) la complirà la successió dels nombres primers. El fet és que molts resultats coneguts (teorema dels nombres primers, oscil·lacions de Littlewood) o conjecturats (hipòtesi de Riemann, conjectura dels primers bessons) s'adiuen plenament amb les prediccions d'aquest model, el qual ha estat sovint utilitzat com a argument en favor de propietats conjecturades de la successió dels nombres primers i també com a motor per a predir-ne de noves.

**Mètodes numèrics i zeros de la funció zeta.** El primer càlcul de zeros el va publicar Gram el 1903, tot i que Siegel trobaria més tard càlculs d'aquesta mena entre els papers de Riemann. Emprant la fórmula sumació d'Euler-Maclaurin per calcular els valors de la funció zeta i de la funció gamma, Gram va aconseguir donar els 10 primers zeros de la funció  $\zeta$  amb 6 decimals de precisió i demostrar la hipòtesi de Riemann fins a  $T = 50$ . Vegeu la figura 9, en què, a part del valor

numèric dels zeros, es mostra la seva posició sobre la recta crítica. També s'hi representa gràficament el recorregut de  $\zeta(\frac{1}{2} + it)$  quan el paràmetre  $t$  recorre l'interval  $0 \leq t \leq 50$ , on es veu com passa deu vegades pel punt zero.

La determinació numèrica dels zeros a la recta crítica  $s = \frac{1}{2} + it$  es redueix a l'estudi de la funció  $\xi(\frac{1}{2} + it)$ , que és una funció real de la variable real  $t$ . Utilitzant les tècniques numèriques habituals per a aquest tipus de càlculs es poden aproximar els zeros amb la precisió desitjada, sempre que es sàpiga calcular el valor d'aquesta funció amb prou precisió. En comptes de treballar directament amb  $\xi(\frac{1}{2} + it)$  és habitual considerar la descomposició

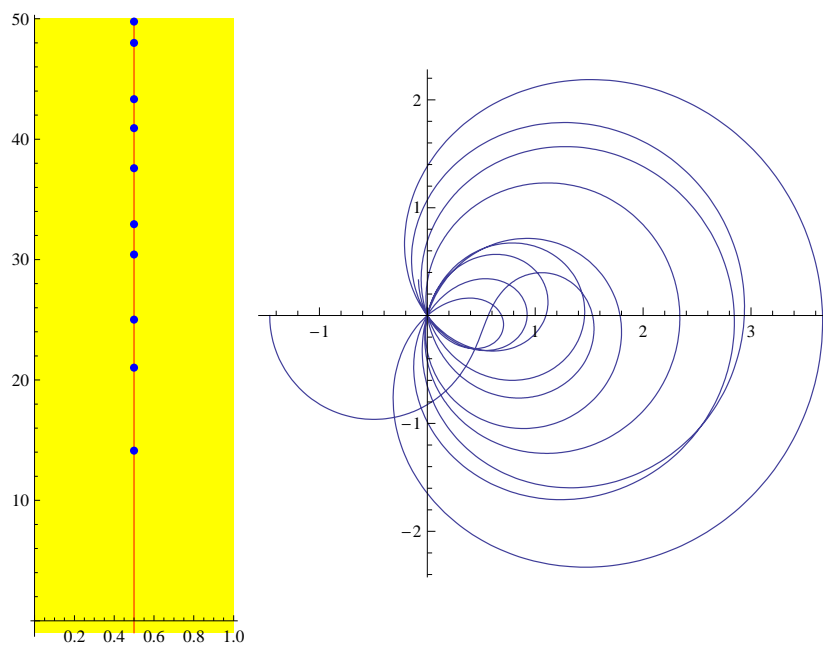
$$\xi\left(\frac{1}{2} + it\right) = \left[-e^{\Re \log(\Gamma(s/2))} \pi^{-1/4} \left(\frac{t^2}{2} + \frac{1}{8}\right)\right] \times \left[e^{\Im \log(\Gamma(s/2))} \pi^{-it/2} \zeta(s)\right],$$

on el primer factor pren sempre valors negatius i la funció del segon factor es coneix com a *funció de Riemann-Siegel* i es denota

$$Z(t) = e^{i\vartheta(t)} \zeta\left(\frac{1}{2} + it\right), \quad \vartheta = \Im \log\left(\Gamma\left(\frac{1}{4} + i\frac{t}{2}\right)\right) - \frac{t}{2} \log \pi.$$

Així, tant els zeros com els canvis de signe de la funció de Riemann-Siegel coincideixen amb els de la funció  $\xi(\frac{1}{2} + it)$ . La figura 10 presenta la gràfica de la funció  $Z$  de Riemann-Siegel en intervals tots de mida 50 contenint nombres de diferents ordres de magnitud. S'hi observa com, en augmentar  $x$ , els zeros es van atapeint com a conseqüència de la fórmula de Riemann-von Mangoldt (6), apareixen alternats amb els extrems locals (conjectura equivalent al fet d'ésser tots simples), i les oscil·lacions augmenten en amplitud.

A més de calcular numèricament valors aproximats dels zeros de la funció  $\zeta$ , també es pot demostrar amb tècniques numèriques que tots els zeros de  $\zeta$  en una regió fitada  $[0, 1] \times [0, T]$  de la banda crítica cauen exactament sobre la recta crítica, de la manera següent. En primer lloc s'apliquen tècniques numèriques per integrar la derivada logarítmica de  $\xi$  sobre la frontera d'aquesta regió; el resultat serà el nombre de zeros  $N(T)$  dins de la regió i, tractant-se d'un enter, pot calcular-se exactament a partir només d'un valor aproximat de la integral. A continuació es busca una seqüència de valors  $0 \leq t_0, t_1, \dots, t_n \leq T$  del paràmetre  $t$  tals que el signe de la funció  $Z(t)$  canviï de l'un al següent,



$$\begin{aligned} \rho_1 &= 0.5 + 14.1347251417347i \\ \rho_2 &= 0.5 + 21.0220396387716i \\ \rho_3 &= 0.5 + 25.0108575801457i \\ \rho_4 &= 0.5 + 30.4248761258595i \\ \rho_5 &= 0.5 + 32.9350615877392i \\ \rho_6 &= 0.5 + 37.5861781588257i \\ \rho_7 &= 0.5 + 40.9187190121475i \\ \rho_8 &= 0.5 + 43.3270732809150i \\ \rho_9 &= 0.5 + 48.0051508811672i \\ \rho_{10} &= 0.5 + 49.7738324776723i \end{aligned}$$

FIGURA 9. Els 10 primers zeros no trivials calculats per Gram l'any 1905, amb altura  $\leq 50$ .

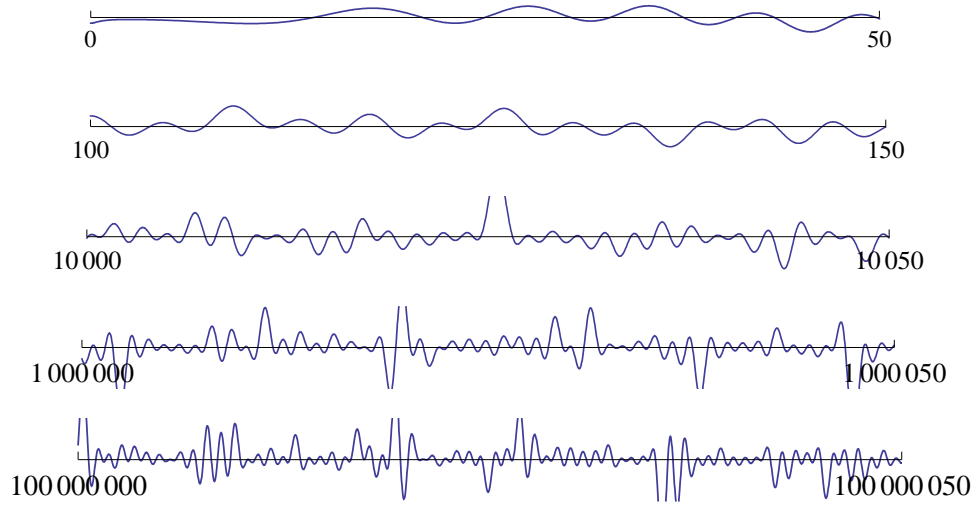


FIGURA 10. La funció  $Z$  de Riemann-Siegel en diversos intervals de mida 50.

el qual garanteix que el hi ha almenys  $n$  zeros sobre la recta crítica. Si s'aconsegueix arribar a  $n = N(T)$  es dedueix que la hipòtesi de Riemann és certa en la regió considerada. Naturalment, aquest mètode donarà resultat només si tots els zeros en la regió són de multiplicitat 1, propietat que es conjectura certa, i de fet ha funcionat sempre en tots els càlculs fets fins ara.

Backlund l'any 1915 i Hutchinson l'any 1925 van ampliar els càlculs fets per Gram fins a  $T = 200$  (regió on hi ha 79 zeros) i  $T = 300$  (138 zeros), respectivament.

Un avenç important en les tècniques per avaluar numèricament la funció zeta a la recta crítica és una fórmula asimptòtica per al càlcul de la funció  $Z(t)$ , coneguda pel nom de *fórmula de Riemann-Siegel*, que Siegel va trobar en els papers de Riemann, la qual va permetre avaluar numèricament aquesta funció de manera molt més eficient i precisa, i ampliar considerablement els resultats numèrics obtinguts fins aleshores.

La revolució informàtica de les darreres dècades i l'increment continuat de la potència de càlcul consegüent han permès estendre els resultats,



tal i com es pot observar a la taula 3. El rècord actual és de l'any 2004: s'han calculat els primers 10 bilions de zeros de la funció zeta de Riemann i cap d'ells s'ha desviat el més mínim de la recta crítica!

Any	Nombre de zeros	Autors
1903	15	J. P. Gram
1914	79	R. J. Backlund
1925	138	J. I. Hutchinson
1935	1 041	E. C. Titchmarsh
1953	1 104	A. M. Turing
1956	15 000	D. H. Lehmer
1956	25 000	D. H. Lehmer
1958	35 337	N. A. Meller
1966	250 000	R. S. Lehman
1968	3 502 500	J. B. Rosser, J. M. Yohe, L. Schoenfeld
1977	$4 \times 10^7$	R. P. Brent
1979	$8.1 \times 10^7$	R. P. Brent
1982	$2 \times 10^8$	R. P. Brent, J. van de Lune, H. J. J. te Riele, D. T. Winter
1983	$3 \times 10^8$	J. van de Lune, H. J. J. te Riele
1986	$1.5 \times 10^9$	J. van de Lune, H. J. J. te Riele, D. T. Winter
2001	$10^{10}$	J. van de Lune
2003	$2.5 \times 10^{11}$	S. Wedeniwski
2004	$10^{13}$	X. Gourdon

TAULA 3. Nombre de zeros calculats

**Aproximacions a la hipòtesi de Riemann.** En el darrer segle i mig, a través de les investigacions dels continuadors de l'obra de Riemann, han proliferat els problemes i les teories directament relacionades amb els seus resultats, i en especial amb la conjectura encara per demostrar.

El més significatiu és sens dubte el fet que la hipòtesi de Riemann és equivalent al fet que l'error del teorema dels nombres primers sigui en un cert sentit el més petit possible (no pot ser menor degut a la magnitud de les oscil·lacions predites per Littlewood). L'any 1901, Helge von Koch va demostrar que la hipòtesi de Riemann és equivalent a que

$$\text{Li}(x) = \pi(x) + O(\sqrt{x} \log x).$$

Mertens estudià la funció  $M(x) = \sum_{n \leq x} \mu(x)$  que dona la suma dels valors de la funció de Möbius en els enters menors que  $x$ ; tenint en compte que aquesta funció multiplicativa proporciona els coeficients de

la sèrie de Dirichlet que dóna el recíproc de  $\zeta$ ,

$$\frac{1}{\zeta(s)} = \prod_p \left(1 - \frac{1}{p^s}\right) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}, \quad \Re(s) > 1,$$

i com a conseqüència de condicions ben conegudes per a la convergència d'aquest tipus de sèries, no és difícil concloure que la hipòtesi de Riemann és equivalent al fet que es tingui  $M(x) = O(\sqrt{x}^{1+\varepsilon})$  per a tot  $\varepsilon > 0$ . Mertens enuncia, el 1897, la conjectura  $|M(x)| \leq \sqrt{x}$  (que implica la hipòtesi de Riemann) però l'any 1985 Odlyzko i te Riele van refutar-la: emprant una ingent quantitat de zeros de la funció zeta calculats amb prou precisió es pot veure que hi ha algun nombre menor que  $e^{3.21 \times 10^{64}}$  en què la suma dels valors de la funció de Möbius supera l'arrel quadrada. Els experts opinen que la desigualtat  $|M(x)| \leq C\sqrt{x}$  serà també eventualment falsa per a tota constant  $C$ , tot i que això no s'ha aconseguit provar fins ara.

Una altra via d'aproximació a la hipòtesi de Riemann és l'estudi del creixement de la funció  $\zeta$ . Per a cada nombre real  $\sigma$  la funció  $\zeta$  té ordre de creixement finit a la recta vertical d'abscissa  $\sigma$ , o sigui  $|\zeta(\sigma+it)| \leq t^\alpha$  per algun  $\alpha > 0$  i  $t$  prou gran. Definint  $w(\sigma)$  com l'ínfim d'aquests nombres  $\alpha$  s'obté una funció contínua monòtona decreixent a tot  $\mathbb{R}$ . És fàcil calcular els valors d'aquesta funció per als valors fora de l'interval  $0 < \sigma < 1$  corresponent a la regió crítica i la gran incògnita és el valor en  $\sigma = 1/2$ . La *hipòtesi de Lindelöf* prediu que  $w(1/2) = 0$  o, de manera equivalent, que

$$w(\sigma) = \begin{cases} \frac{1}{2} - \sigma, & \text{si } \sigma < \frac{1}{2}, \\ 0, & \text{si } \sigma \geq \frac{1}{2}, \end{cases}$$

i es pot demostrar que equival a la hipòtesi de Riemann. El 1916 Hardy i Littlewood van demostrar que  $w(1/2) \leq 0.166\dots$ ; el rècord actual és de Bombieri i Iwaniec, que donen la fita  $9/56$ .

Una aproximació deguda a Beurling i Nyman des de l'anàlisi funcional obté la hipòtesi de Riemann com a conseqüència que l'espai generat per les funcions

$$\eta_\alpha(t) := \left\{ \frac{\alpha}{t} \right\} - \alpha \left\{ \frac{1}{t} \right\},$$

on  $\{\cdot\}$  indica la part fraccionària d'un nombre real, per als valors de l'interval  $0 < \alpha < 1$ , generen tot l'espai  $L^2(0, 1)$ . Una descripció detallada d'aquesta teoria (i molt més) es pot consultar a [5].

Per acabar, tot i que la llista podria seguir, una equivalència curiosa: la hipòtesi de Riemann és equivalent a que  $\det(A(n)) = O(\sqrt{n}^{1+\varepsilon})$ , on  $A(n)$  és la matriu binària que té un 1 en el lloc  $(i, j)$  si  $j = 1$  o bé  $i \mid j$ , i zeros a la resta (vegeu [4]).

**Operador de Riemann i matrius aleatòries.** Una de les aproximacions a la hipòtesi de Riemann que ha originat més discussió i controvèrsia és la coneguda amb el nom de conjectura de Hilbert-Pólya. Es tracta de l'especulació sobre la possibilitat que els zeros de la funció  $\xi(\frac{1}{2} + it)$  siguin valors propis d'un operador hermític auto-adjunt en un espai de Hilbert de dimensió infinita adequat, l'operador de Riemann. Com que aquests zeros són reals, la hipòtesi de Riemann queda reduïda «només» a la construcció d'un tal operador.

L'analogia amb el cas de les funcions  $L$  locals, en que la hipòtesi de Riemann es va demostrar a partir de la construcció de certs espais (la cohomologia  $\ell$ -àdica) i de les propietats d'un automorfisme que hi opera (l'automorfisme de Frobenius), ha estat de vegades presentada com a argument en favor d'aquesta via, tot i que els experts són molt cautes en les seves opinions al respecte (veure [1, 3, 14]).

La conjectura de Hilbert-Pólya rebé un fort impuls durant els anys 70 quan Montgomery i Dyson observaren que l'espaiat entre zeros consecutius de la funció zeta, convenientment normalitzat, es distribueix de manera increïblement idèntica a com ho fa l'espaiat entre els valors propis de matrius aleatòries hermítiques. La verificació numèrica d'aquesta «correlació de parells» ha estat l'objectiu de molts projectes de càlcul intensiu durant les darreres dècades, essent aquest un dels principals esperons per al desenvolupament de tècniques numèriques relacionades amb la recerca de zeros de la funció  $\zeta$ . A més, la teoria s'ha estès amb èxit a altres famílies de funcions  $L$ , fent-les correspondre a matrius aleatòries de diversos tipus. En tots els casos les dades confirmen el comportament previst, tot i que fins ara no s'ha obtingut cap resultat teòric rellevant sobre la funció  $\zeta$  com a conseqüència d'aquesta teoria.

## REFERÈNCIES

- [1] P. Bayer. *La hipòtesi de Riemann (IV Cicle Ferran Sunyer i Balaguer)*. Aula de ciència i cultura, **25** (2007). Fundació Caixa de Sabadell.
- [2] C. Bays, R.H. Hudson. *A new bound for the smallest  $x$  with  $\pi(x) > Li(x)$* . Math. Comp. **69** (1999), no. 231, pp. 1285–1296.
- [3] E. Bombieri. *Problems of the Millennium: The Riemann Hypothesis*. Pàgina web de la fundació Clay.
- [4] E. Bombieri. *The new book of prime number records*. Springer-Verlag (1995).
- [5] Bruna, J. *Euler, sèries i funció zeta de Riemann*. Conferències FME, volum iv (Curs Euler 2006-2007), 2007.
- [6] J. B. Conrey. *The Riemann Hypothesis*. Notices of the AMS **50** (2003), no. 3, pp. 341–353.
- [7] A. Granville, G. Martin. *Prime number races*. Amer. Math. Monthly **113** (2006), no. 1, pp. 1–33.
- [8] M. du Sautoy. *The Music of Primes. Why an Unsolved Problem in Mathematics Matters*. Harper Perennial (2004). Traducció al castellà publicada per Acanalado (2007).
- [9] M. Deléglise, J. Rivat. *Computing  $\pi(x)$ : the Meissel, Lehmer, Lagarias, Miller, Odlyzko, method*. Math. Comp. **65** (1996), no. 213, pp. 235–245.
- [10] H. M. Edwards. *Riemann's Zeta function*. Academic Press (1974).
- [11] X. Gourdon. *The  $10^{13}$  first zeros of the Riemann zeta function, and zeros computation at very large height*. Preprint. 24 Octubre 2004.
- [12] G. F. B Riemann. *Über die Anzahl der Primzahlen unter einer gegebenen Größe*. Monatsberichte der Berliner Akademie (1859), pp. 671–680.
- [13] M. Rubinstein, P. Sarnak. *Chebyshev's bias*. Experiment. Math. **3** (1994), pp. 173–197.
- [14] P. Sarnak. *Problems of the Millennium: The Riemann Hypothesis (2004)*. Pàgina web de la fundació Clay.
- [15] G. Tenenbaum, M. Mendès France. *The prime numbers and their distribution*. AMS Student Mathematical Library vol. 6 (2000).
- [16] D. Zagier. *The first 50 million prime numbers*. Math. Intelligencer vol. 0 (1977), pp. 7–19.

# RIEMANN'S INFLUENCE IN GEOMETRY, ANALYSIS AND NUMBER THEORY

SIR MICHAEL F. ATIYAH

ABSTRACT. The purpose of this article is to describe some ways in which Zeta functions enter geometry, and their relation to the theory of Riemann surfaces.

Riemann's collected works take one small volume, but every contribution to this volume was very original work that supplied foundations for the mathematics of the next century.

To give an idea of his range of interests, and of his influence up to recent times, I will first describe briefly three areas which bear his name:

- (1) Riemann Zeta function (number theory)
- (2) Riemann surfaces (algebraic geometry-topology)
- (3) Riemannian metric (fundamental in differential geometry)

Then I will review three examples where these areas interact. They are three examples of work in the 20th century which have their roots in the ideas of Riemann. The first one will be a theorem of the 1930's which involves the interaction of (2) and (3) on this list. Next one will be a theorem of the 1950's, also involving the interaction of algebraic topology/geometry with differential geometry. And the last one will be a theorem of the 1980's which involves in fact all three, including number theory.

In summary, the main points will be:

- Review the three topics (1), (2) and (3) above.
- A theorem of 1930's involving (2) and (3).
- A theorem of 1950's involving (2) and (3).
- A theorem of 1980's involving (1), (2) and (3).

## RIEMANN ZETA FUNCTION

It is the function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad s \in \mathbb{C}.$$

This function converges for  $\operatorname{Re}(s) > 1$  and has analytic continuation to whole complex  $s$ -plane with a simple pole at  $s = 1$ . For  $s = 0$  we have  $\zeta(0) = 1/2$ .

**Functional equation.** If we write

$$\xi(s) = \frac{s(s-1)}{2} \Gamma(s/2) \pi^{-s/2} \zeta(s)$$

then

$$\xi(s) = \xi(1-s).$$

**Euler product**

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1}$$

where the product ranges over the primes  $p$ .

**Riemann Hypothesis (10<sup>6</sup> dollar prize).** It is the most famous unsolved problem since Riemann's time and it is about the zeros of the Riemann zeta function.

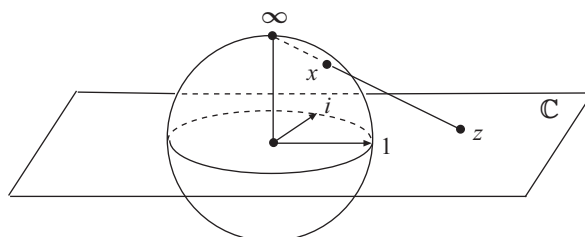
The "trivial zeros" of  $\zeta(s)$  occur at  $s = -2, -4, \dots$  (even negative integers).

The Riemann hypothesis states that *all non-trivial zeros of  $\zeta(s)$  lie on the line  $\operatorname{Re}(s) = 1/2$ .*

It has very important consequences for the detailed distribution of prime numbers.

## RIEMANN SURFACES

Adding the point at infinity ( $\infty$ ) to the complex plane  $\mathbb{C}$  we get the *Riemann sphere*.



The complex solutions  $(x, y)$  of the equation

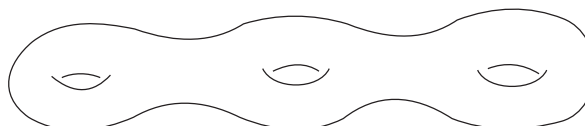
$$y^2 = f(x), \quad f \text{ quartic polynomial}$$

(including  $\infty$ ) form a *complex torus*. It is a double covering of the sphere with 4 branch points at the 4 roots of  $f(x) = 0$  (assumed distinct).



The sphere leads to the theory of *rational functions* and the torus to the theory of *elliptic functions*.

Higher degree polynomial equations  $p(x, y) = 0$  lead to a theory of functions involving surfaces of higher genus.<sup>1</sup> These are called *Riemann surfaces*. Here is a picture of a genus 3 surface:



The genus is a topological invariant. It is  $B_1/2$ , where  $B_1$  is the first Betti number, or the rank of the first homology group, which agrees with the number of independent 1-cycles that can be drawn on the surface. For a torus, for example, two independent cycles are one going

---

<sup>1</sup>The genus  $g$  of a compact orientable surface is the number of its handles or holes. Thus  $g = 0$  for the sphere and  $g = 1$  for the torus.

around the hole and one around the tube. In general, for each hole we have one more pair, which means  $2g$  in all.

On the Riemann surface we have a relationship with analysis. We can write down what are called *holomorphic differentials*. If  $z$  is a local coordinate on the surface, a holomorphic differential has (locally) the form

$$f(z)dz,$$

where  $f$  is a holomorphic function. If we make a change of variable from one region to another one, say  $z = z(u)$ , then  $f(z)dz$  becomes  $f(z(u))z'(u)du$ . Now *the genus  $g$  is the dimension of the space of holomorphic differentials*. For example, if  $g = 0$  there are no non-zero holomorphic differentials. The key fact is that if  $z$  is the usual coordinate on the complex plane, then  $u = 1/z$  is a local coordinate at  $\infty$  and  $dz = -du/u^2$  has a pole there. On the other hand on the torus, thought of as the quotient  $\mathbb{C}/L$  of  $\mathbb{C}$  by a lattice  $L$ ,  $dz$  is a holomorphic differential, and this differential is unique up to scalar factor. For higher genus there are more complicated formulas giving the  $g$  independent holomorphic differentials.

Thus we have a fundamental link between the topology (the number of holes) and the analysis (the description of the holomorphic differentials). Holomorphic differentials occur as integrands of integrals along curves. In fact, that is how the theory of elliptic integrals first arose.

### THE RIEMANNIAN METRIC

The Riemannian metric is the foundation of modern differential geometry. It has an interesting beginning, because when Riemann was to submit his thesis for a doctoral degree the custom in Germany of the time was that the examiners would ask him some minor topic outside of the main topic of interest. Gauss asked Riemann to investigate the foundations of geometry. I think he knew that Riemann had ideas, and Riemann's answer was to produce the foundations of differential geometry –as a side issue to the main thesis.

On an  $n$ -dimensional manifold, which locally looks like an Euclidean space with  $n$  real coordinates  $x_1, \dots, x_n$ , a *Riemannian metric*, which is a way of measuring distances, is a symmetric positive definite quadratic



form

$$ds^2 = \sum_{i,j} g_{ij} dx_i dx_j,$$

with the  $g_{ij}$  varying smoothly. That gives the element of length,  $ds$ , and the length of a curve is then given by integrating  $ds$  along it.

What Riemann discovered was that the most essential and important object that can be associated with the Riemannian metric is what is called the *Riemann curvature tensor*: An expression  $R_{ijkl}$ , given by a quite simple formula involving the  $g_{ij}$  and their derivatives, which describes all the fundamental ways in which geometry is curved.

Note that a Riemann surface is at the same time, a 1-dimensional complex manifold and a 2-dimensional Riemannian manifold, where the real coordinates are the real and imaginary part of the complex coordinate.

### Examples of Riemannian manifolds

- (1) Take an ordinary smooth surface in 3-dimensional space, as for instance a sphere, and restrict to it the ordinary Euclidean metric. This gives a way of measuring lengths of curves on the surface.
- (2) We can do similar things in higher dimensions. Consider a non-singular algebraic curve in the the complex projective plane  $\mathbb{CP}_2$ , which has real dimension 4. It is given by an equation

$$f_d(x, y, z) = 0,$$

where  $f_d$  is a homogeneous polynomial of degree  $d$ , and its genus  $g = \frac{1}{2}(d-1)(d-2)$ . This curve acquires a natural Riemannian metric, namely, the metric induced on it by the standard metric in  $\mathbb{CP}_2$ .

- (3) Similarly, a non-singular algebraic surface  $f_d(x, y, z, t) = 0$  in  $\mathbb{CP}_3$  with the induced metric.

**Note.** The Riemannian curvature tensor has four indices  $i, j, k, l$  in all dimensions, but it reduces to a scalar function in dimension 2, the Gauss scalar curvature, and it is essentially a 2-tensor in dimension 3 (the Ricci tensor). Dimension 4 is the first dimension in which the most general form of the curvature tensor is required. It is the dimension

of Einstein's space-time, in which Riemannian ideas, and especially Riemann's curvature tensor, played a very important role. It is also the dimension of Simon Donaldson's theory, in which he applied ideas of Riemannian geometry and got many fantastic results showing that 4 real dimensions in some ways are quite unique amongst all dimensions.

### HODGE THEORY (1930's)

Consider an  $n$ -dimensional real manifold, compact and oriented, with a Riemannian metric. Hodge considered the exterior differential forms  $\Omega^q$ , the natural integrands of  $q$ -dimensional integrals ( $q = 0, 1, \dots, n$ ). They are skew-symmetric  $q$ -tensors. Thus  $\Omega^0$  are the scalar functions  $f(x)$ , and  $\Omega^1$  the forms of degree 1 (locally,  $\sum_i f_i(x) dx_i$ ). Higher degree forms  $\Omega^q$  involve wedge products (skew symmetric) of any  $q$  of the  $dx_1, \dots, dx_n$ . For  $q = n$  there is only one such product (a volume form). There is also the exterior differentiation operator,  $d : \Omega^{q-1} \rightarrow \Omega^q$ , which gives rise to the *De Rham complex*:

$$\Omega^0 \xrightarrow{d} \Omega^1 \xrightarrow{d} \Omega^2 \xrightarrow{d} \Omega^3 \xrightarrow{d} \dots \xrightarrow{d} \Omega^n$$

The metric defines a dual or orthogonal form  $*\omega \in \Omega^{n-q}$  associated to a  $q$ -form  $\omega \in \Omega^q$ . In particular, if  $f$  is a function,  $*f$  is a volume form.

In Hodge theory, a form  $\omega$  is called harmonic if  $d\omega = 0$  and  $d*\omega = 0$ . One of Hodge's motivations for introducing this concept was the fact that the electromagnetic field can be represented by a 2-form  $\omega$  in Minkowski space and that Maxwell's equations are then equivalent to  $d\omega = 0$  and  $d*\omega = 0$ .

Another equivalent way of defining harmonic forms is the following. Define

$$d^* : \Omega^q \rightarrow \Omega^{q-1}$$

as the adjoint of  $d$  (if  $n$  is even,  $d^* = -*d*$ ) and let

$$\Delta = dd^* + d^*d : \Omega^q \rightarrow \Omega^q.$$

This operator is the *Hodge Laplacian* and  $\Delta\varphi = 0$  is equivalent to saying that  $\varphi$  is harmonic. This fact explains why they are called 'harmonic', as  $\Delta\varphi$  is the ordinary Laplacian for a function  $\varphi$  in Euclidean space.

The fundamental theorem of Hodge was a simple relationship between the harmonic forms and the topology:

**Theorem.** *The space of harmonic  $q$ -forms is isomorphic the real cohomology space  $H^q$  (the map is given  $\omega \mapsto \int_\gamma \omega$ ,  $\gamma$  any  $q$ -cycle).*

The expression  $\int_\gamma \omega$  is called the *period* of  $\omega$  along  $\gamma$ , and the theorem says that every harmonic form is determined by its periods and that every ‘period’ can actually occur.

For example, on a Riemann surface the harmonic 1-forms, for an appropriate metric, are the real and imaginary parts of the holomorphic forms. Note that this implies that  $\dim H^1 = 2g$ .<sup>2</sup>

Now a big problem was going from algebraic curves to higher dimensions (algebraic surfaces, for example) and replacing the holomorphic differentials. Hodge’s great idea was using a Riemannian metric and taking harmonic forms. That was the beginning of the foundations of modern algebraic geometry.

#### SIGNATURE OF A 4-MANIFOLD

For  $n = 2$  (surface) the only topological invariant is the genus  $g$  (or  $B_1 = 2g$ ).

For  $n = 4$  we have

$$B_1 = B_3 \quad \text{and} \quad B_2.$$

But there is a further invariant on  $H^2$  (or  $H_2$ ): the *intersection form* (or intersection matrix),  $H^2 \times H^2 \rightarrow \mathbb{R}$ . This pairing is a non-degenerate quadratic form. Now a quadratic form over the real numbers has another invariant, aside from the rank, namely the signature. When diagonalized, it has  $B_2^+$  positive terms and  $B_2^-$  negative terms, where  $B_2 = B_2^+ + B_2^-$  (no zeros, as it is non-degenerate), and the signature is  $B_2^+ - B_2^-$ . This signature is called the *signature* of the manifold, and it is another topological invariant.

---

<sup>2</sup>Another way of saying this is that the complex valued harmonic forms on the surface are the holomorphic and antiholomorphic forms.

**Note.** There is no analogue of the signature in dimension 2, since the intersection pairing is skew-symmetric: (21) is an odd permutation while (3412) is an even permutation.

This signature invariant for quadratic forms was known for a long time, but its applications to manifolds topology were first pointed out by Hermann Weyl, another of my heroes in mathematics. Interestingly enough for this audience, it was published in a Spanish journal, in Spanish, in an old paper that is not very well known.

Let's connect this with the work of Hodge in differential geometry. On a 4-manifold  $M$ , the  $*$  operator acts on  $\Omega^2$  and  $*^2 = 1$ . The corresponding eigenspaces  $H_+^2$  and  $H_-^2$  have dimensions  $B_2^+$  and  $B_2^-$ , respectively. Hence

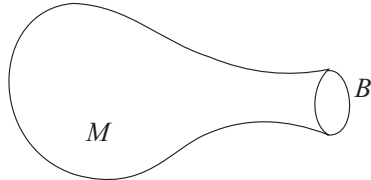
$$\text{signature} = \dim H_+^2 - \dim H_-^2.$$

By a famous theorem of Hirzebruch (1950's),

$$\text{signature} = \int_M f(R),$$

where  $f$  is a polynomial in the curvature. This beautiful theorem connects topology on one side with differential geometry on the other and it is the first of a large family of theorems of this kind. It is part of a big development in the 1950's. Hirzebruch proved that the theorem also holds in dimensions  $n = 4k$  (with  $f$  depending on  $k$ ).

Hirzebruch's result was for closed manifolds. So the natural question now is to ask what happens for manifolds with boundary.



Choose a Riemannian metric on  $M$  which is isometric to the product  $B \times \mathbb{R}$  near the boundary:

$$ds_M^2 = ds_B^2 + dr^2.$$

Then  $f(R) = 0$  on  $B \times \mathbb{R}$  and therefore  $\int_M f(R)$  is independent of the length of the boundary region.

The signature is still defined on  $H^2(M, B)$ , but the (diagonal) matrix has some zeros. Count only non-zero diagonal elements.

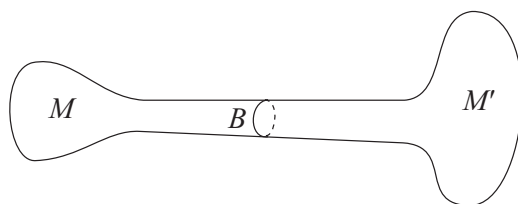
**Question.** Is  $\text{signature}(M) = \int_M f(R)$ ?

**Answer.** No.

But we have:

**Proposition.** *The difference between the signature and  $\int_M f(R)$ , which is called the signature defect, depends only on the boundary.*

*Proof.* If  $B = \partial M$  and  $B = \partial M'$ , stick  $M$  and  $M'$  together along  $B$  to get a manifold  $X$  without boundary.



Then we have

$$\text{signature}(X) = \text{signature}(M) + \text{signature}(M'),$$

$$\int_X f(R) = \int_M f(R) + \int_{M'} f(R),$$

and it is enough to subtract and use Hirzebruch's theorem for  $X$  (a sign change occurs for  $M'$  since its boundary is  $B$  with the opposite orientation).  $\square$

**Problem.** What kind of invariant of the Riemannian 3-manifold  $B$  is the signature defect?

**Note.** The signature defect changes sign when we reverse the orientation of  $B$ .

To solve the problem, introduce the 1st order elliptic differential operator  $A$  acting on

$$\begin{aligned} \Omega^{\text{ev}} &= \Omega^0 \oplus \Omega^2, \\ A &= \pm(*d - d*) \quad [- \text{ on } \Omega^2, \quad + \text{ on } \Omega^0]. \end{aligned}$$

It is self-adjoint ( $A^* = A$ ) and<sup>3</sup>

$$A^2 = \Delta^0 \oplus \Delta^2 \quad (\text{Hodge Laplacian}).$$

The eigenvalues  $\lambda$  of  $A$  are discrete and  $\lambda^2$  are the eigenvalues of  $\Delta^0 \oplus \Delta^2$ . The eigenvalue  $\lambda = 0$  gives the harmonic forms  $H^0 \oplus H^2$ . Define

$$\eta(s) = \sum_{\lambda \neq 0} \frac{\text{sign}(\lambda)}{|\lambda|^s}.$$

The function  $\eta(s)$  converges for  $\text{Re}(s) > 3$  (because  $3 = \dim(B)$ ) and has an analytical continuation with no pole at  $s = 0$  (so  $\eta(0)$  is well defined).

**Theorem** (Atiyah–Patodi–Singer, 1970’s).

$$\text{signature}(M) = \int_M f(R) - \eta(0).$$

See [APS] and [Atiyah].

**Note.** 1) Reversing orientation changes

$$\begin{array}{lcl} * & \text{to} & -* \\ A & \text{to} & -A \\ \lambda & \text{to} & -\lambda \end{array}$$

It follows that  $\eta(0)$  changes to  $-\eta(0)$ , and this behaviour checks with the theorem.

2) In classical notation for  $\mathbb{R}^3$ , and identifying a 2-form with a vector field,

$$A = \begin{pmatrix} 0 & \text{div} \\ \text{grad} & \text{curl} \end{pmatrix}$$

The theorem shows that the signature defect is a spectral invariant. It measures the degree to which the positive eigenvalues and the negative eigenvalues differ. If they are the same (for example if the manifold has an orientation reversing isometry), then the signature defect is 0, as the positive eigenvalues pair up one by one with the negative eigenvalues. In general they are not the same, and we say that the signature defect measures the *spectral asymmetry*.

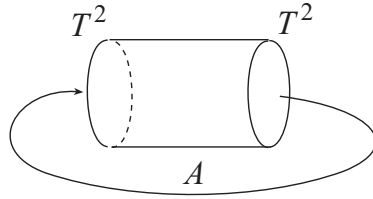
---

<sup>3</sup> $A^2$  is the square of the operator  $A$ , while  $\Delta^2$  simply refers to the Hodge Laplacian on 2-forms.

## EXAMPLE INVOLVING NUMBER THEORY

Here we will consider an example of an interesting application of the APS theorem.

Let  $B$  be the 2-torus bundle over a circle defined by an automorphism of  $T^2 = \mathbb{R}^2/\mathbb{Z}^2$ , i.e., an element  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  of  $SL(2, \mathbb{Z})$ .



Note that  $SL(2, \mathbb{Z})$  gives linear automorphisms of  $\mathbb{R}^2$  that map the lattice  $\mathbb{Z}^2$  onto itself. We have  $\det(A) = ad - bc = 1$  and we will assume that  $a + d = \text{trace}(A) > 2$ . This implies that the eigenvalues of  $A$ , namely the roots  $\lambda$  and  $\lambda' = 1/\lambda$  of

$$\lambda^2 - (a + d)\lambda + 1 = 0,$$

are real and positive, as its discriminant  $\Delta = (a + d)^2 - 4 > 0$ .

This situation arises in number theory from the real quadratic field

$$K = \mathbb{Q}(\sqrt{\Delta}).$$

Let us examine more carefully the matrix  $A$ . Over  $\mathbb{R}$ , it is conjugate to the diagonal matrix  $\begin{pmatrix} \lambda & \\ & \lambda' \end{pmatrix}$  and we have a 1-parameter group  $A^t \sim \begin{pmatrix} \lambda^t & \\ & (\lambda')^t \end{pmatrix}$ . The orbits of  $A^t$  in the  $(x, y)$  plane are (branches of) hyperbolas whose asymptotes are given by the equation of the eigenvectors ( $N$  is the norm):

$$N(y, x) = 0,$$

where

$$N(y, x) = cy^2 + (d - a)xy - bx^2 = c(y - \alpha x)(y - \beta x),$$

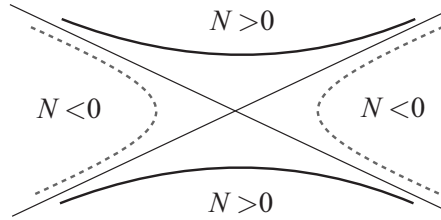
with

$$\alpha = \frac{(a-d) + \sqrt{\Delta}}{2c}, \quad \beta = \frac{(a-d) - \sqrt{\Delta}}{2c}$$

and

$$\Delta = (d-a)^2 + 4bc = (a+d)^2 - 4.$$

Assume  $c < 0$ .



Since  $A \in SL(2, \mathbb{Z})$ , it preserves each branch and maps integer points to integer points, so branches that have one integer point actually have infinitely many. Since the norm is constant along a branch,  $N$  is an integer for branches containing an integer point.



**Zeta function of  $K = \mathbb{Q}(\sqrt{\Delta})$ :**

$$\zeta_A(s) = \sum_{(\gamma)} \frac{1}{|N(\gamma)|^s},$$

where  $\gamma = (m, n)$  is an integer lattice point and we sum over the non-zero orbits  $(\gamma)$  of  $\{A\}$  (the group generated by  $A$ ).

$\zeta_A$  converges for  $\text{Re}(s) > 1$ , has analytic continuation, with no pole at  $s = 0$ , and has an Euler product expansion.

**$L$ -function of  $K = \mathbb{Q}(\sqrt{\Delta})$ :**

$$L_A(s) = \sum \frac{\text{sign}(N(\gamma))}{|N(\gamma)|^s}$$

over the non-zero integer orbits of  $\{A\}$ . It has similar properties to  $\zeta_A(s)$ .

$\zeta_A(0)$  is well-defined (actually it is a rational number).



**Theorem** (Atiyah–Donnelly–Singer, 1980's).

$$L_A(0) = \eta_A(0)$$

where  $A \in SL(2, \mathbb{Z})$ .

**Notes.** 1)  $L(s)$  and  $\eta(s)$  are very different analytic functions:

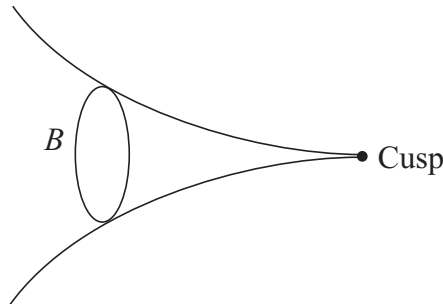
- $L(s)$  converges for  $\operatorname{Re}(s) > 1$
- $\eta(s)$  converges for  $\operatorname{Re}(s) > 3$ .

But the theorem can be proved by using Fourier series on  $T^2$ , an approach that was initiated by Carl L. Siegel.

2) The theorem generalizes to any totally real field  $K$  of any degree:

- $B$  becomes a  $T^r$  bundle over  $T^{r-1}$ , where  $r$  is the rank of the integers in  $K$  (and  $r - 1$  the rank of the group of units of  $K$ ).

3) The theorem is also related to algebraic geometry. Hirzebruch proved [Hirz] that the signature defect of a 3-manifold  $B$  is  $L_A(0)$  by a direct method related to the periodic continued fraction of  $\sqrt{\Delta}$  and the resolution of the cusp singularities of an algebraic surface. This provided



one motivation for the ADS theorem, as it led to a conjectured generalization to higher degree fields. See [ADS].

#### REFERENCES

- [Hirz] F. Hirzebruch: *The Hilbert modular group, resolution of the singularities at the cusps and related problems*. Springer Lecture Notes in Mathematics **244**, 275-288.  
 [Also: Collected Works, vol. II, p. 192-205]

- [APS] M.F. Atiyah, V.J. Patodi, and I.M. Singer: *Spectral asymmetry and Riemannian Geometry, I*. Math. Proc. Camb. Phil. Soc. **77** (1975), 43-69. [Also: MFA Collected Works, vol. 4, p. 81-107]
- [ADS] M.F. Atiyah, H. Donnelly, and I.M. Singer: *Eta invariants, signature defects of cusps and values of L-functions*. Ann. of Math. **118** (1983) 131-177 Math. Proc. Camb. Phil. Soc. **71** (1975), 43-69. [Also: MFA Collected Works, vol. 4, p. 563-612]
- [Atiyah] M.F. Atiyah: *The logarithm of the Dedekind  $\eta$ -function*. Math. Ann. **278** (1987), 335-380. [Also: Collected Works, vol. 6, p. 151-199]

# DUALITY IN MATHEMATICS AND PHYSICS\*

SIR MICHAEL F. ATIYAH

ABSTRACT. Duality is one of the oldest and most fruitful ideas in Mathematics. I will survey its history, showing how it has constantly been generalized and has guided the development of Mathematics. I will bring it up to date by discussing some of the most recent ideas and conjectures in both Mathematics and Physics.

## INTRODUCTORY REMARKS

Duality in mathematics is not a theorem, but a “principle”. It has a simple origin, it is very powerful and useful, and has a long history going back hundreds of years. Over time it has been adapted and modified and so we can still use it in novel situations. It appears in many subjects in mathematics (geometry, algebra, analysis) and in physics. Fundamentally, duality gives *two different points of view of looking at the same object*. There are many things that have two different points of view and in principle they are all dualities.

**Linear duality in the plane.** It starts off classically in geometry with linear duality in the plane.

In the plane we have points and lines. Two different points can be joined by a unique line. Two different lines meet in one point unless

---

\* Editorial note. On December 18, 2007, Atiyah delivered the lecture *Riemann's Influence in Geometry, Analysis and Number Theory* at the Facultat de Matemàtiques i Estadística of the Universitat Politècnica de Catalunya and the lecture *Duality in Mathematics and Physics*, this one sponsored by the Centre de Recerca Matemàtica (CRM), at the Institut de Matemàtica de la Universitat de Barcelona (IMUB). The present text has been produced from a transcription of the second lecture and can be downloaded from the web pages [www.imub.ub.es](http://www.imub.ub.es) or [www-fme.upc.edu](http://www-fme.upc.edu). The text of the first lecture can also be downloaded from the latter web page.



they are parallel. People did not like this exception and so they worked hard and realized that if they added some points at infinity then they got what is called the *projective plane* in which the duality is perfect: the relationship between points and lines is perfectly symmetrical. This led to the classical principle of projective duality, which says that the “dual statement” of a theorem is also a theorem, so that we can talk about the dual theorem.

**Linear algebra.** In linear algebra duality involves the pairing

$$\langle \boldsymbol{\xi}, \boldsymbol{x} \rangle = \sum_i \xi_i x_i$$

of *dual vector spaces*.

For example, let  $n = 3$ . If we fix a  $\boldsymbol{\xi}$ , then the equation  $\langle \boldsymbol{\xi}, \boldsymbol{x} \rangle = 0$  gives a linear condition on  $\boldsymbol{x}$  which determines a plane in 3-dimensional vector space. On the other hand if we fix  $\boldsymbol{x}$  and let  $\boldsymbol{\xi}$  vary, then we get a plane in the dual space. These are vector spaces and if we factor out by homogeneous coordinates (this leads to one dimension lower) we get the lines of the projective plane. So the relationship between projective geometry and linear algebra is very simple: we write equations and we identify those that differ by a non-zero scalar factor.

## LINEAR ANALYSIS

When we go from linear algebra to linear analysis, in principle we make  $n = \infty$  (infinite dimension linear spaces), and we have to be careful about questions of convergence. Formal infinite series do not make sense and so we have to have suitable continuity conditions.

**Example.** If we have functions of one variable  $x$ , then we can define the  $\langle f, g \rangle$  of two functions  $f$  and  $g$  by

$$\langle f, g \rangle = \int f(x)g(x)dx.$$

Assuming that the integral converges, this defines a pairing which is an  $\infty$ -dimensional version of the finite pairing of linear algebra. So immediately the ideas of duality go (provided we are careful about convergence) into linear analysis.

**Fourier transform.** Now in linear analysis one of the most powerful tools is that of Fourier theory. Let us recall what is the *Fourier transform*. Let  $\mathbf{x} = (x_1, \dots, x_n)$  be  $n$  variables and  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)$  corresponding ‘dual’ variables. Then given a function  $f(\mathbf{x})$ , its *Fourier transform*  $\hat{f}(\boldsymbol{\xi})$  is defined by

$$\hat{f}(\boldsymbol{\xi}) = \int f(\mathbf{x}) \exp(2\pi i \langle \mathbf{x}, \boldsymbol{\xi} \rangle) d\mathbf{x}.$$

**Fourier series.** We can also have Fourier series, which in fact are more elementary. If we have a function  $f(x)$  of the variable  $x$  which is *periodic* (it is a function defined on the circle  $\cong \mathbb{R}/\mathbb{Z}$ ), then we have Fourier *coefficients*  $a_n$ :

$$a_n = \int f(x) \exp(2\pi i n x) dx.$$

This transforms functions on the circle  $\mathbb{R}/\mathbb{Z}$  to functions on the integers  $\mathbb{Z}$ . The *inversion formula* says how to go backwards.<sup>1</sup> In this example we see that the two things that are dual need not be of the same kind. In linear duality usually the two sides are the same sort of thing, but here one side is a circle and the other side are the integers, but it is still an example of duality.

**Poisson summation formula.** This is a very useful formula in connection with the theory of Fourier series. It uses Fourier theory on  $\mathbb{R}$  and on  $\mathbb{R}/\mathbb{Z}$  and it says, given a function  $f$ , that

$$\sum f(n) = c \sum \hat{f}(n),$$

$c$  a suitable constant. In the particular case when the function is the Gaussian  $f(x) = \exp(-x^2)$ , and taking into account that its Fourier

---

<sup>1</sup>  $f(x) = \sum_n a_n \exp(-2\pi i n x)$

transform  $\hat{f}$  is also Gaussian, we get the famous transformation law<sup>2</sup>

$$\Theta(-1/z) = (-iz)^{1/2}\Theta(z)$$

of the *theta function*

$$\Theta(z) = \sum_n \exp(\pi i n^2 z).$$

Besides this formula we have another obvious relation:

$$\Theta(z+2) = \Theta(z).$$

The two operations  $z \mapsto z+2$  and  $z \mapsto -1/z$  generate a subgroup of index two the modular group  $SL(2, \mathbb{Z})$  (we would get the whole group  $SL(2, \mathbb{Z})$  if we had the condition  $\Theta(z+1) = \Theta(z)$ ). A function which transforms according to such laws under this group is called a *modular form*. They are very important things in arithmetic and the  $\Theta$  *modular form* is one of the most classical examples. As we see, modular forms turn up in connection with dualities in Fourier theory. This is of course all very classical, as it was known to Riemann and to people before him.

#### DUALITY FOR ABELIAN GROUPS

The real line, the circle and the integers are different-looking things. But in fact they are all abelian groups, and for these groups there is a duality theory. Given an abelian group  $G$ , the *dual group*  $\hat{G}$  is the group of characters of  $G$  (homomorphisms of  $G$  to the circle group  $U(1)$ ):

$$\hat{G} = \text{Hom}(G, U(1)).$$

Examples. 1) If  $G = \mathbb{Z}$ ,  $\hat{G} = S^1$  (circle of radius 1), where  $x \in S^1 \cong \mathbb{R}/\mathbb{Z}$  corresponds to the homomorphism  $n \mapsto \exp(2\pi i n x)$ .

2) If  $G = \mathbb{R}$ ,  $\hat{G} = \mathbb{R}$ .

3) If  $G$  is discrete, then  $\hat{G}$  is a compact group (like  $S^1 = \hat{G}$  for  $G = \mathbb{Z}$ ).

The general class of abelian groups for which duality theory works are the *locally compact topological abelian groups*. This class includes the groups  $\mathbb{R}$ ,  $S^1$  and  $\mathbb{Z}$ , and also others. If  $G$  is finite (hence it is both

---

<sup>2</sup>This relation gives one of the ways in which the functional equation for the Riemann  $\zeta$  function can be proved.

discrete and compact), then  $\hat{G}$  is finite. Besides the examples given, there are also important examples in number theory, as the group of  $p$ -adic integers:

$$\mathbb{Z}_p = \varprojlim \mathbb{Z}/p^n\mathbb{Z}.$$

Fourier theory works also in this case and is important in applications in number theory.

## NON-ABELIAN GROUPS

Abelian groups give a unifying framework for these classical dualities. We have a good theorem, but it is natural to try to go beyond that. In that regard a reasonable aim is to look at non-abelian groups.

(1) Let us start with a (non-abelian) finite group  $G$ . Introduce the set  $\hat{G}$  of (isomorphism classes) of irreducible *unitary representations*  $\text{Hom}(G, U(n))$ , which is called the *dual* of  $G$ . It is not a group anymore, it is only a (finite) set.

(2) Consider now a compact Lie group  $G$ , like for example the orthogonal group  $O(n)$ . Then the dual  $\hat{G}$ , defined as the set of (isomorphism classes of) irreducible unitary representations, is an *infinite discrete set*.

**Fourier theory.** For any class function  $f$  on  $G$ , there is a Fourier expansion

$$f(g) = \sum_{\chi \in \hat{G}} a_\chi \chi(g), \quad a_\chi = \int_G f(g) \bar{\chi}(g) dg.$$

The notation  $\chi \in \hat{G}$  indicates that the sum is extended over all irreducible characters  $\chi$  of  $G$ .<sup>3</sup> It may be finite or an infinite series depending on the group  $G$ .

---

<sup>3</sup>The character  $\chi = \chi_\rho$  of a unitary representation  $\rho : G \rightarrow U(n)$  of  $G$  is the function  $\chi : G \rightarrow \mathbb{C}$  defined by  $\chi(g) = \text{trace}(\rho(g))$ . It only depends on the isomorphism class of the representation. The character is said to be *irreducible* if the representation is irreducible. The mapping of isomorphism classes of irreducible representations to irreducible characters is one-to-one.

**(3) Non-compact Lie groups.** In this case the theory gets more difficult. Let  $G$  be a non-compact Lie group, as for example the general linear group  $GL(n)$ . Then  $\hat{G}$  should be the set of isomorphism classes of irreducible representations, including those that are infinite dimensional (i.e., representations in Hilbert space). If  $G$  satisfies some conditions,  $\hat{G}$  is a set with a measure on it (the Plancherel measure  $d\mu$ ) and for any class function  $f$  on  $G$  the Fourier inversion formula is

$$f(g) = \int_{\hat{G}} \hat{f}(\chi)\chi(g)d\mu, \quad \hat{f}(g) = \int_G f(g)\bar{\chi}(g)dg.$$

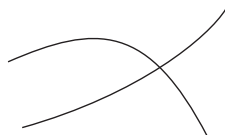
This theory, valid for semisimple Lie groups, is due to Harish-Chandra. It involves infinite-dimensional spaces and continuous parameters, but it is very satisfactory as it generalizes all the other cases.

#### NON-LINEAR GEOMETRY

After the preceding excursions to algebra and to groups, let us go back to geometry.

Suppose we have an  $n$ -dimensional manifold which is compact and oriented. Then we can define the *homology groups*, which are given by cycles of various dimensions up to the maximum dimension  $n$ . If  $q$  is an integer,  $0 \leq q \leq n$ , we can define the homology group in dimension  $q$ ,  $H_q$ . It is a finite dimensional vector space if we use real coefficients. Then we also have cycles in the complementary dimension,  $n - q$ , which give the homology group (space)  $H_{n-q}$ .

**Intersection pairing.** If we have a cycle of dimension  $q$  and a cycle of dimension  $n - q$  in general position, then they intersect in finitely many points and so we get an intersection number.



From this it is possible to construct a pairing of real vector spaces

$$H_q \otimes H_{n-q} \rightarrow \mathbb{R}$$

(the proof uses the local Euclidean nature of the manifold).



**Poincaré duality.** The pairing above is non-degenerate. Therefore each of the homology spaces  $H_q$  and  $H_{n-q}$  becomes the dual of the other. Thus we obtain a duality between finite-dimensional vector spaces from a curved manifold. In particular,  $\dim H_{n-q} = \dim H_q$ .

Poincaré duality actually works at the global level of topology and it is a very important tool in for the study of the topology of higher dimensional manifolds.

### HODGE THEORY

Another way of examining this, which is a combination of geometry and analysis, is to use Hodge theory. Given an  $n$ -dimensional manifold  $M$ , let us consider, instead of cycles, differential forms. The differential forms of degree  $q$ , which form a space that we will denote  $\Omega^q$ , are the natural integrands of integrals on cycles of dimension  $q$ .

We will also use the *derivative operator*

$$d : \Omega^q \rightarrow \Omega^{q+1}.$$

The  $\alpha \in \Omega^q$  such that  $d\alpha = 0$  are said to be *closed forms*.

Now if we have  $\alpha \in \Omega^q$  and  $\beta \in \Omega^{n-q}$ , then the wedge product  $\alpha \wedge \beta$  is an  $n$ -form (also called a *volume form*). This form can be integrated over  $M$ ,

$$\int_M \alpha \wedge \beta,$$

and this gives a bilinear pairing

$$\Omega^q \times \Omega^{n-q} \rightarrow \mathbb{R}$$

under which  $\Omega^q$  and  $\Omega^{n-q}$  are dual  $\infty$ -dimensional spaces.

Given a Riemannian metric on  $M$ , we also have the Hodge duality operator  $*$  :  $\Omega^q \rightarrow \Omega^{n-q}$  that is defined by the relation ( $\alpha, \beta \in \Omega^q$ )

$$\int_M \alpha \wedge * \beta = (\alpha, \beta),$$

where  $(\alpha, \beta)$  denotes the pairing of  $\alpha$  and  $\beta$  by the natural extension of the Riemannian metric to  $\Omega^q$ .

Hodge theory is concerned with the solutions (called *harmonic forms*) of the two equations

$$d\omega = 0, \quad d*\omega = 0.$$

Thus  $\omega$  is harmonic if and only if  $\omega$  and  $*\omega$  are closed.

**Hodge theorem.** The space  $H^q$  of harmonic forms of degree  $q$  has finite dimension and is isomorphic to the cohomology space of dimension  $q$ . Moreover,  $*$  :  $H^q \rightarrow H^{n-q}$  is an isomorphism. In particular,  $H^q$  and  $H^{n-q}$  are dual finite dimensional spaces (Poincaré duality) and  $\dim H^q = \dim H^{n-q}$ .

This theorem provides a reinterpretation of the geometrical duality using harmonic forms.

## PHYSICS

Now we go to (classical) physics.

**1. Position-momentum.** Here is an elementary situation. Let  $x$  be a variable in ordinary space (it may describe the position of a particle) and a dual variable  $\xi$  (it may represent momentum of the same particle). The the Fourier transform gives rise to “dual pictures”, which is what physicists call *spectral analysis*.

**2. Quantum mechanics.** In quantum mechanics, there is the famous particle-wave duality. Duality appears here as two ways of looking at the same thing: either a particle behaves like a point going around or as (quantum) wave. These two aspects are part of the mystery of quantum mechanics.

**3. Electromagnetism.** In ordinary Minkowski space  $\mathbb{R}^{3,1}$ , the electromagnetic force is described by a 2-form (skew-symmetric 2-tensor)  $\omega$ . In this notation, Maxwell’s equations *in vacuo* are

$$d\omega = 0, \quad d*\omega = 0,$$

where now  $*$  is defined using the Lorentz metric in  $\mathbb{R}^{3,1}$ . Formally they are the same as Hodge equations for forms of degree 2 on a 4-dimensional Riemannian manifold, but here the space is ordinary Minkowski space, not the 4-dimensional Euclidean space. From this it

appears that Maxwell's equations, which unified electricity and magnetism, also encode a duality between electricity and magnetism in the sense that the  $*$  operator interchanges both aspects. Physically, this is a very fundamental fact of the universe.

Maxwell's equations actually motivated Hodge for his work on harmonic forms in general. As indicated, Maxwell's equations are about forms of degree 2 in 4 dimensions and Hodge went to forms of any degree  $q$  in any dimension  $n$ . He also worked in Riemannian geometry, not in Minkowski space.

### MODERN PHYSICS: GAUGE-THEORY

A lot of modern physics is concerned about *gauge theory*. This is what physicists use to describe elementary particles. In some naive sense it is just a (non-abelian, non-linear) matrix generalization of Maxwell's theory.

A *Yang-Mills field*  $F$  is the curvature of a connection<sup>4</sup> on a fibre bundle ( $F$  may be thought as a as a 2-form with matrix coefficients, or as a matrix of 2-forms). Then the Yang-Mills equations take the same form again:

$$dF = 0, \quad d * F = 0.$$

These are now matrix equations. When defined on Lorentz manifolds, they are the fundamental equations that physicists use for elementary particle physics. They can also be defined on a Riemannian 4-manifold, and this leads to geometry, as in Donaldson's celebrated theory, one of the most exciting developments in the last quarter of the 20th century. In particular it produced new invariants of 4-manifolds. The physics inspired the equations and then Donaldson used the equations, in the context of Riemannian geometry, to develop his theory.

This was just mathematics. We might say that mathematicians took an equation that the physicists had written down and studied it in the Riemannian context, just as Hodge did with Maxwell's equations. But subsequently Witten showed ([W94]) that what Donaldson did could be understood in the language of physics. In fact, he interpreted

---

<sup>4</sup>In gauge theory, connections play the role of potentials.

Donaldson's theory as a Topological Quantum Field Theory, which can be thought of as a "non-abelian Hodge theory".

Going beyond ordinary quantum mechanics, which describes the quantum behaviour of particles, quantum field theory (QFT) describes the quantum behaviour of fields, like the electromagnetic field. Usually quantum field theory supplies information about real particles and their interactions.

In the special case of a topological quantum field theories, the only things that can be extracted are topological invariants. These theories supply discrete quantities that have a topological meaning, not real numbers that can vary continuously. Donaldson's invariants, for example, turned out to be, as shown by Witten, instances of such topological invariants of a QFT. This physical interpretation of Donaldson's theory made the link between physics and mathematics much easier to understand.

One of the most exciting things that physicists have discovered is that in QFT there are many dualities which are not at all easy to understand geometrically. These theories are non-linear, and so they are not trivial. An important observation here is that if we have a given classical geometrical picture, then there is some kind of procedure, called 'quantization', by which we replace classical variables by operators to produce, with a bit of guesswork, a 'quantization' of the original theory. In this quantized theory there is a parameter that plays the role of Planck's constant and which allows us to recover the classical theory by letting it go to 0. Given a quantum theory, however, it may turn out that there are several ways in which it can be realized as the quantization of a classical geometry picture. For example, for a single classical particle we can use the position observable or the momentum observable (they are actually symmetrical), and so there are two ways of reaching the same theory from two geometrical points of view. Then these two points of view are called a duality. This is, of course, a very simple example, but it turns out that physicists have found many much deeper dualities which exist in complicated QFTs which link two very different-looking geometric pictures.

**Example: Donaldson’s theory.** A first example of such a duality is that Donaldson’s theory, which has been very roughly described above, is “dual” to Seiberg–Witten theory. Physicists predicted that these two theories should be equivalent, in the sense that one quantum theory has two different descriptions: one in terms of Donaldson’s geometry and the other in terms of Seiberg–Witten geometry.

In Donaldson’s theory we just take  $2 \times 2$  matrices and write down equations analogous to Maxwell’s equations. In Seiberg–Witten theory, instead of  $2 \times 2$  matrices, the ordinary circle group is used, and Maxwell’s equations for the (classical) electro-magnetic field, but coupled to a spinor field.

Spinors are very important things in physics (they describe particles like the electron) and they turn out to be mysteriously important in geometry. Their geometrical meaning is much more difficult to understand than differential forms, but they exist and they are beautiful geometrical objects. They satisfy the famous equation introduced by Dirac (the Dirac equation), which is a linear equation for spinors like Maxwell’s equations are for the electro-magnetic field, but now the spinors and the electro-magnetic field can be coupled together. The reason that they can be coupled together is fairly simple algebra: spinors can be multiplied together and the results are differential forms (not a single degree, but all degrees simultaneously). So spinors are like square roots of differential forms, and that explains why they are so mysterious.<sup>5</sup>

So Spinors $\times$ Spinors=Forms. If the form obtained as the product of two spinors is related to Maxwell’s equations, we get equations that are quadratic, and therefore non-linear. So there is a non-linear coupling between spinor fields and the electro-magnetic field and this coupling gives rise to a system of equations that are called the *Seiberg–Witten equations*. They involve on one hand spinors, which satisfy Dirac’s

---

<sup>5</sup>One of the most fantastic discoveries in mathematics was the invention of the square root of  $-1$ . It took 200 years to understand, and even when it was understood Gauss said that “the real metaphysics of the square root of  $-1$  is not simple”. In a similar way, spinors are something like trying to understand the square root of ‘area’, or the square root of ‘volume’. What does that mean? Very mysterious notions, but nevertheless spinors exist in a formal way and can be used in formulas just as  $i$  is used all the time in mathematics and physics.

equation, and involve also the electromagnetic field, which satisfies not just the pure Maxwell equations, but the Maxwell equation with a right hand side which is constructed from a spinor roughly speaking by squaring.

The Seiberg–Witten equations are, in some ways, very elementary equations, more elementary than the matrices that Donaldson used, but they are not linear—otherwise they would be rather trivial. What was predicted, on the grounds of physics intuition, was that Donaldson theory and the Seiberg–Witten theory should be equivalent: they should give two different ways of realizing the same QFT. So the invariants that Donaldson had written down and the similar invariants that we can get by solving the Seiberg–Witten equations, are really equivalent information, and what this means is that we have some kind of Poisson summation formula to relate the two. This does not at all say that the invariants on one side are equal to the invariants on the other side. What it says is that the sum of all the Donaldson invariants (they depend on an integer degree) is equal to the sum of the Seiberg–Witten invariants (they also depend on an integer degree), or that the two series give the same information. The equality (Poisson summation formula) is in the sense of generating functions. Symbolically:

$$\sum_{SW} = \sum_D.$$

It is a very deep kind of relationship, a non-linear  $\infty$ -dimensional version of a Poisson summation formula—that’s why this formula was mentioned at the beginning. It is a very important way of getting interesting formulae in ordinary classical mathematics. But it also points to much deeper things that link non-linear dualities in geometry and in physics. This is very much part of modern day physics, for it was physicists who discovered it, not mathematicians.

Of course, we might say that we ought to prove theorems, in this case this prediction of the physicists. It is a very hard question and *no direct mathematical proof is known yet*. Ideas for a proof could come from the fact that the same geometric results are obtained by applying Donaldson’s theory or Seiberg–Witten’s theory, so that in practical terms they are equally useful. But mathematically, strictly speaking, we don’t have a proof because we do not yet have a framework in which

this new kind of duality involving infinite dimensional spaces has been formally set up. We are still in the early stages of a theory which will probably be developed in the future. In any case the question is obviously very important.

**Example: Mirror symmetry.** Another example, also coming out of physics, was equally spectacular. It was called *mirror symmetry*. It does not have the name ‘duality’, but ‘symmetry’ here means essentially the same thing. Mirror symmetry can be described in many different ways. Here I will consider one of them.

In geometry there are different kind of geometries:

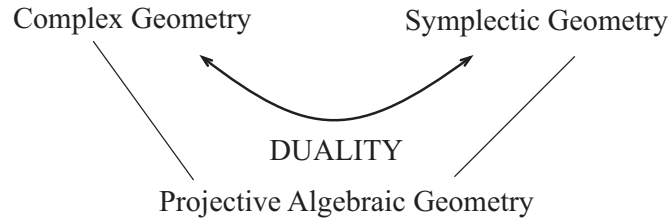
- *Complex geometry.* The coordinates are complex and the admissible changes of variables are holomorphic. Examples: Riemann surfaces and algebraic varieties over the complex numbers.
- *Symplectic geometry.* Real manifolds with a non degenerate skew-symmetric 2-form. They generalize the phase space of classical Hamiltonian mechanics. Algebraic varieties in complex projective space are also symplectic, because the latter has a natural symplectic structure (given by the Kähler metric) and this structure induces a symplectic structure on subvarieties.
- *Riemannian geometry.* Real manifolds with a positive definite metric.

These geometries may be thought as the non-linear versions of the three classical (lie) groups:

- Complex general linear group,  $GL(n, \mathbb{C})$ .
- Symplectic group,  $Sp(n)$  (preserves a skew-symmetric form).
- Orthogonal group,  $O(n)$  (preserves a metric).

Seen the other way around, each of these linear theories has a non-linear generalization to a geometry, and this is how complex geometry, symplectic geometry and Riemannian geometry arise.

What the physicists discovered is that there is a certain duality, or symmetry, between the complex geometry and the symplectic geometry. This duality relates in a rather precise way the information on one side to the information on the other side. In particular, *when this*



*symmetry is applied to complex projective algebraic varieties we obtain again complex projective algebraic varieties.*

There are many concrete examples of this, which were first discovered by physicists. In one class of examples, the algebraic varieties  $M$  and  $M'$  have (complex) dimension 3 (so 6 real dimensions) and the duality exchanges information on one side with information on the other (very subtle information, as we will see below). Moreover,

$$\dim H^{\text{ev}}(M) = \dim H^{\text{odd}}(M')$$

The dimensions of the homology are the same, but even homology on one side corresponds to the odd homology on the other, which is very hard to understand. It means, for example, that the Euler characteristic (the dimension of the even part minus the dimension of the odd part) changes sign when we go from one to the other. It is a very precise relationship, but very mysterious. Moreover, the details of the geometry give the following: On the  $M$  side the information is about *algebraic curves*, for example enumerative questions about rational curves of any degree, and on the  $M'$  side it is related to *periodic matrices*.<sup>6</sup>

So this marvellous theorem tells us that easy information on one side (periodic matrices, that can be calculated by classical means) is equivalent to difficult information on the other side (algebraic curves, for whose determinations there is very little information). In physics language, the easy information is what is called classical and the difficult one is what is called quantum. We are thus getting information of a quantum character on one side out of a classical calculations on the other side.

---

<sup>6</sup>Period matrices started with Riemann and its entries are integrals of forms on cycles.



Using this fantastic duality the physicists were able to calculate the numbers of rational curves of any degree on very simple examples of algebraic varieties of dimension 3. The formulae they got were so spectacular to algebraic geometers that at first they did not believe them, but eventually they were converted. Then they began a big industry that has produced many books on mirror symmetry. It is a whole new area in algebraic geometry that arises out of this particular simple example, just one example of duality in quantum theory.

The applications of physics to mathematics have been numerous over the years. Usually the mathematics is quite close to the physics, but complex algebraic varieties seem as far away removed from physics as one can imagine. Why should physics have anything to do with these concepts in pure geometry?

[So its a spectacular coup: physicists go up into the sky, they land by parachute in the middle of algebraic geometers and they capture immediately the whole city.]

The discovery of mirror symmetry is certainly one of the most remarkable developments of the last part of the 20th century. It provided an example of two different classical theories, two different algebraic varieties, giving rise to the same quantum theory, and with spectacular applications. The mirror symmetry and Donaldson's theory, two of the most interesting developments in recent times, had their origin in physics and in terms of the dualities we have been considering. We see, therefore, that the payoffs of these dualities is very great.

It should be emphasized that the physics, even if it gives the exact answers, does not necessarily give proofs. It gives answers we can believe if we believe that the physics intuition is correct. But mathematicians like to have a proof, and then they have to work very hard to prove the theorems by other methods. Now lots of this has been done, and when they have succeeded, essentially the physicists have been correct. This is an interesting interaction: the physicists generate results that the mathematicians can prove by other means and then they can exchange information.

## NUMBER THEORY

The preceding two examples were from geometry. Now we will move to number theory and to begin with we will consider a nice beautiful simple example about prime numbers.

Let  $p$  be an odd prime. Then we can ask whether it can be represented as the sum of two squares of two integers. The answer is that it can be done if and only if  $p \equiv 1 \pmod{4}$ . In other words, if  $p \equiv 3 \pmod{4}$ , the answer is no, and if  $p \equiv 1 \pmod{4}$ , the answer is yes.

Writing  $p = x^2 + y^2$  ( $x, y$  integers), which is equivalent to

$$p = (x + iy)(x - iy),$$

we see that we are factorizing a prime number, not in the rational numbers, but in the field we get by adjoining  $i$  to the rational numbers.<sup>7</sup> This is an example of algebraic extension of the field of rational numbers.

This is the beginning of what is called *class field theory*, which is trying to give information which relates algebraic extensions of a number field to properties of primes in that field (in the example, information on  $\mathbb{Q}(i)$  relates to properties of primes in  $\mathbb{Q}$ ). The theorem above is one of the facts at the origin of class field theory, which essentially worked for all abelian field extensions.

Abelian class field theory, a very fine theory finished in the 20th century, was founded by Gauss, Kronecker and others, and gives a complete answer to the story about abelian extensions of the rational numbers, and also about abelian extensions of number fields. The important thing that comes in is the Galois group of the extension. If the Galois group is assumed to be abelian, the answer is that we can relate this to constructions that can be performed inside the field concerning primes, whereas the extensions are outside the field.

This was a great triumph of classical number theory and ended the questions that Gauss and Hilbert posed of developing a theory in general for all number fields (the original works were about quadratic fields). But in the 20th century people became more ambitious and

---

<sup>7</sup>If we do this with the integers, we obtain the *Gaussian integers*.

asked about what could be done for non-abelian extensions. Would there still be some kind of duality?

Taking into account the role of representations of groups in linear duality, it is natural to ask whether this idea will also apply to number theory. In other words, is there a way of using representations of non-abelian groups to extend the classical class field theory?

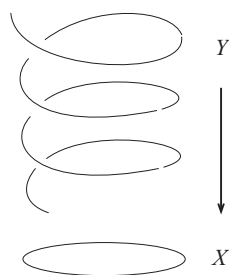
The search for non-abelian class field theory has been going on for over forty years and nowadays it goes under the name of the Langlands programme, because Langlands was the person who put this in a rather precise form—he made many conjectures and proved many special cases.

**Langlands programme.** The Langlands programme is a large programme which, if eventually completed, will give a very satisfactory answer to the question of how to extend abelian class field theory to the general case of non-abelian extensions. It is a very big programme, that in many ways has just started, and in it a fundamental role is played by complex linear representations of the Galois groups of algebraic extensions.

**Simple geometric analogue.** Let us consider a simple example or analogue between number theory and ordinary geometry.

In ordinary geometry, or topology, if  $X$  is a space, for example the circle  $S^1$ , there are two things that we can do.

One is that we can look at the covering spaces that project down onto it as the picture indicates.



For example, if we take  $X$  to be a circle, then  $Y$  can be the real line, which covers the circle infinitely many times (that is why we take the real numbers modulo 1). When we go to a covering space that cannot go any further, that is called the *universal covering space*.

Consider the fundamental group of the space  $X$ ,  $\Gamma = \pi_1(X)$ , It is constructed inside the space, by geometry.<sup>8</sup> Then the fact is that this is related to the covering space situation: if we have the universal covering space  $Y$ , then the fundamental group  $\Gamma$  acts freely on  $Y$  and the quotient of  $Y$  by this group action is  $X$ . This is an example of the sort of duality we are talking about, the fundamental group and the universal covering being the ‘inside’ and ‘outside’ aspects of the situation.

For example, given a representation of  $\Gamma$  into a matrix group (a homomorphism  $\Gamma \rightarrow GL(n, \mathbb{C})$ ), then this gives rise over the space  $X$  to what is called a *vector bundle* (the quotient of  $Y \times \mathbb{C}^n$  by the action of  $\Gamma$  induced by the representation). This bundle is *flat*.<sup>9</sup> In the case where  $n = 1$ , we get a *flat line bundle* (a line bundle is a vector bundle of dimension 1). Note that in this case  $GL(1, \mathbb{C}) = \mathbb{C}^*$  is abelian.

If  $X$  is an algebraic curve or, in other words, a Riemann surface, then it is well known that a Riemann surface is a close analogue of a number field. Not the same thing, of course, but only an analogue. Points of the Riemann surface are the analogues of prime numbers and the number field itself is the analogue of the rational functions (meromorphic functions) on the curve (Riemann surface).

Rational (meromorphic) functions on  $X \leftrightarrow$  Number field

The analogy is stressed by the fact that many theorems just look the same on both sides. Even more, people can guess what a theorem

---

<sup>8</sup>This group is obtained by taken closed paths in  $X$  and identifying two paths if they are deformable one into the other.

<sup>9</sup>Any closed loop yields a linear endomorphism of the vector bundle, which only depends on the homotopy class of the loop. The bundle is said to be flat because loops that can be deformed to a point produce the identity endomorphism. If we had a non-trivial endomorphisms even for loops in small regions of  $X$ , then the bundle is said to have ‘curvature’. Flat vector bundles are vector bundles with trivial curvature. For flat bundles, therefore, the only interesting endomorphisms come from going around big paths.

should be on the number field side by using the corresponding piece of geometry and making the translation of languages.<sup>10</sup> There is a whole dictionary. For example, the fundamental group of the Riemann surface corresponds to the Galois group of the number field and coverings correspond to field extensions.

We also remark that there is some analogy between these pictures and what we have been saying about class field theory. On both sides we have the abelian case and the non-abelian case, depending to which we want to take.

### GEOMETRIC LANGLANDS PROGRAMME

The Langlands programme has, along these lines, something called the geometric Langlands programme, which replaces the number fields by Riemann surfaces. It is a very interesting theory which is much easier than the number field case, but not trivial and still quite big. It is developed by using the theory of vector bundles on Riemann surfaces. This theory has been going on for quite a long time, has nice results and in it there are geometrical analogues of the Langlands conjectures. There is also the intermediate case of the Langlands program on algebraic curves over finite fields from which we can transfer either to Riemann surfaces or to a number fields.

Very recently, in the last two years, Witten and his collaborators (mainly Kapustin and Gukov, two young Russians) have managed to deduce what is required for the geometric Langlands program from non-abelian dualities in physics. The kind of dualities they use are close to the dualities in Donaldson's theory and to the dualities in the mirror symmetry, and are based (at least) on the electric-magnetic duality.

The original dualities in physics are those between electricity and magnetism. They are very well understood and known, but it is a linear theory. When people started to play with non-linear theories, and tried to look for something analogous to duality, they had vague ideas that perhaps it would be possible to have some generalization of the electric-magnetic duality to the non-abelian case.

---

<sup>10</sup>The Weil conjectures are of this kind.

In the dualities it often happens that there is a pair of corresponding parameters, one on each side, that are inverse.<sup>11</sup> That means that if the parameter is large on one side, then it is small on the other side. In physics there are many theories that have a parameter called the *coupling constant*. If this parameter is small, then we can expand in a power series, and one gets a good approximation of the theory (this approach is called *perturbation theory*). If the parameter is large, this is hopeless. But if the theory can somehow be related to a dual theory with a small coupling constant, then of course we are in good business. We expand on the other side, then translate backwards and we get information about the theory with a large coupling constant, which is very difficult to get by other means. This was the dream that physicists had.

Could it be that the strong force, with a large coupling constant, and the weak force, with a weak coupling constant, are in some sense dual? There was lot of speculation along these lines and eventually, for example, the theories that Seiberg and Witten produced were of that kind.

Now the question of what Witten and his collaborators showed is that the geometric Langlands program for a Riemann surface can be reconstructed from one of the non-abelian dualities that involve basically electric-magnetic duality. They take a duality in 4 dimensions, where electricity and magnetism are given by a 2-form on each side, then they develop a non-abelian theory which is applied in the special case where the 4 dimensions is the product of the (curved) Riemann surface by 2 flat dimensions. In this situation they reconstruct interesting results about the Riemann surface and out of it they get the geometric Langlands program. Mirror symmetry of particular algebraic varieties plays a key role in this programme. The particular algebraic varieties they use are (Hitchin) moduli spaces of certain kinds of bundles on curves (bundles with some extra information called a Higgs field).

Altogether it is a great tour de force (see [K-W, G-W, F-W]) and one of the most exciting things happening in the last few years. It is a non trivial, very deep operation, and it is quite spectacular that the

---

<sup>11</sup> This phenomenon appear in ordinary Fourier theory, where the Fourier transform of the Gaussian  $e^{-a^2}$  is the Gaussian  $e^{-1/a^2}$ .

electric-magnetic duality should be behind the duality which is used in this Langlands programme, itself born out of number theory.

I must admit that many years ago, when I knew just a little bit of these ingredients, I made some speculations that perhaps the Langlands programme could be understood in terms of duality of electricity and magnetism. It was a good guess: 20 years later we are there.<sup>12</sup>

### MODULAR FORMS

To finish with, let us go back to modular forms (at the beginning we mentioned the modular form of the  $\Theta$  function, which corresponds to  $SL(2, \mathbb{Z})$ ). Modular forms arise all over the place in mathematics, and in physics, and basically they are objects which transform nicely under certain arithmetic groups (there are large numbers of these groups; they are like groups of integer matrices).

Modular forms are the things which number theorists really play with. Much of the Langlands programme is a reinterpretation of questions about modular forms, which are a very key part of that theory. The framework of representations and so on that Langlands developed is a framework in order to handle modular forms. Results about modular forms have immediate applications and so they are really at the heart of the matter.

Interestingly enough, modular forms also frequently turn up in quantum field theories, usually in the form of what is called a *partition*

---

<sup>12</sup>Editorial note. “For a gauge group  $G$ , the GNO dual group is actually the same as the Langlands dual group  ${}^L G$ , which plays an important role in formulating the Langlands conjectures. [...] This was observed by Atiyah, who suggested to the second author [Witten] at the end of 1977 that the Langlands program is related to quantum field theory and recommended the two papers [10,12].” (Quoted from [K-W], p. 3). The two papers cited in this quote are: [10] P. Goddard, J. Nuyts, and D. I. Olive, *Gauge Theories And Magnetic Charge*, Nucl. Phys. B125 (1977) 1-28; and [12] C. Montonen and D. I. Olive, *Magnetic Monopoles As Gauge Particles?*, Phys. Lett. B72 (1977) 117-120.

*function*.<sup>13</sup> There are many examples of modular forms in the literature, but the reason why they exist, where they come from, is often not at all clear.

One of the big unsolved problems will be to understand the connection between number theory and physics which involves modular forms. That will be related to what I have been talking about, maybe much beyond that, and might need new ideas, but it is very concrete: we see modular forms created on both sides, one asks why, and the answer is that we don't know. But the evidence is there, it is very spectacular, and large parts of mathematics have been swept up by this kind of band-wagon of duality.

Duality is an old topic, but it is still very much alive and kicking. We have seen how it relates to many things that everyone is familiar with in mathematics (group theory, topology, analysis, Fourier theory), and so it is not surprising that it also arises in physics, where one can use the same sort of ideas. It is a very exciting story that hopefully the younger people here will continue by proving some theorems on modular forms for the future.

#### REFERENCES

- [FGV] E. Frenkel, D. Gaitsgory, and K. Vilonen: *On the geometric Langlands conjecture*. J. AMS, **15** (2001), 367-417.
- [F-W] E. Frenkel, E. Witten: *Geometric endoscopy and mirror symmetry*. Preprint [arXiv:0710.5939v2](https://arxiv.org/abs/0710.5939v2) [Math AG] 14 Mar 2008. 130 pp.
- [G-W] S. Gukov, W. Witten: *Gauge theory, ramification, and the geometric Langlands program* Preprint [arXiv:hep-th/0612173v2](https://arxiv.org/abs/hep-th/0612173v2) 8 Oct 2007. 160 pp.
- [K-W] A. Kapustin and E. Witten: *Electric-Magnetic duality and the geometric Langlands program*. Preprint [arXiv:hep-th/0604151](https://arxiv.org/abs/hep-th/0604151) 10 Jan 2007. 231 pp.
- [Mor] John w. Morgan: *The Seiberg–Witten equations and applications to the topology of smooth four-manifolds*. Mathematical Notes 44, Princeton University Press, 1996.

---

<sup>13</sup>The notion of partition function comes from statistical mechanics, but it is basically the simplest invariant quantity that one can construct from quantum field theory. In particular, topological quantum field theories supply topological invariants.



- [S-W] N. Seiberg and E. Witten: *Monopoles, duality and chiral symmetry breaking in  $N = 2$  supersymmetric QCD*. Nuclear Phys. B.
- [W94] E. Witten: *Monopoles and four-manifolds*. Math. Res. Lett. **1** (1994), 769-796.
- [W08] E. Witten: *Mirror symmetry, Hitchin's equations, and Langlands duality*. Preprint [arXiv:0802.0999v1](https://arxiv.org/abs/0802.0999v1) [Math.RT] 7 Feb 2008.



## AN OVERVIEW OF RIEMANN'S LIFE AND WORK

ROSSANA TAZZIOLI

ABSTRACT. Riemann made fundamental contributions to mathematics –number theory, differential geometry, real and complex analysis, Abelian functions, differential equations, and topology– and also carried out research in physics and natural philosophy.

The aim of this note is to show that his works can be interpreted as a unitary programme where mathematics, physics and natural philosophy are strictly connected with each other.

### A BRIEF BIOGRAPHY

Bernhard Riemann was born in Breselenz –in the Kingdom of Hanover– in 1826. He was of humble origin; his father was a Lutheran minister. From 1840 he attended the Gymnasium in Hanover, where he lived with his grandmother; in 1842, when his grandmother died, he moved to the Gymnasium of Lüneburg, very close to Quickborn, where in the meantime his family had moved. He often went to school on foot and, at that time, had his first health problems –which eventually were to lead to his death from tuberculosis.

In 1846, in agreement with his father's wishes, he began to study the Faculty of Philology and Theology of the University of Göttingen; however, very soon he preferred to attend the Faculty of Philosophy, which also included mathematics. Among his teachers, I shall mention Carl Friedrich Gauss (1777-1855) and Johann Benedict Listing (1808-1882), who is well known for his contributions to topology.

In 1847 Riemann moved to Berlin, where the teaching of mathematics was more stimulating, thanks to the presence of Carl Gustav Jacob Jacobi (1804-1851), Johann Peter Gustav Lejeune Dirichlet (1805-1859), Jakob Steiner (1796-1863), and Gotthold Eisenstein (1823-1852). Therefore, in Berlin classical mechanics and the theory of elliptic and Abelian

functions (with Jacobi), number theory (with Dirichlet and Eisenstein), geometry (with Steiner), real analysis and series of functions (with Dirichlet) were all well represented and taught; and in fact Riemann's stay in Berlin strongly influenced his future research.

In 1848 in Berlin –as well as in all Europe– there was a democratic uprising; Riemann was among the conservatory students who opposed it. This attitude was in contrast with the ideas of his teachers; Dirichlet, Eisenstein and Jacobi had indeed an active role in the fight against the King of Prussia.

In 1849 Riemann followed his father's wishes and came back to the University of Göttingen, where he began to attend the seminar on mathematical physics held by Gauss and Wilhelm Eduard Weber (1804-1891) who was appointed professor of physics in 1831. Gauss was also conservative; when Victoria became Queen of Britain in 1837 her uncle became ruler of Hanover and revoked the liberal constitution. Weber was one of seven professors at Göttingen to sign a protest and all were dismissed. Gauss did not support them, even though Weber –who was very close to him– and his son-in-law (the orientalist G.H.A. von Ewald) were among the seven professors who lost their positions at the University. Weber remained at Göttingen without a position until 1843, when he became professor of physics at Leipzig; but in 1848 he came back to his old position in Göttingen.

Riemann held his *Inauguraldissertation* (*Grundlagen für eine allgemeine Theorie der Funktionen einer veränderlichen complexen Grösse*) on complex analysis –supervised by Gauss– in 1851; his *Habilitationschrift* (*Über die Darstellbarkeit einer Function durch eine trigonometrische Reihe*) on real analysis and his *Habilitationsvortrag* (*Über die Hypothesen, welche die Geometrie zu Grunde liegen*) on the foundations of geometry were held in 1853 and in 1854 respectively, and published posthumously in 1868.

In 1855 Gauss died and Dirichlet took his place. Two years later, in 1857, Riemann was appointed extraordinary professor at the University of Göttingen. In the same year he published his celebrated paper on the theory of Abelian functions (“Theorie der Abel'schen Functionen”).

In 1859 he visited Paris, where he knew some important mathematicians –such as Joseph Bertrand (1822-1900), Jean-Baptiste Biot (1774-1862), Jean Claude Bouquet (1819-1885), Charles Hermite (1822-1901), Victor Puiseux (1820-1883) and Joseph Serret (1819-1885). In the same year he was appointed ordinary professor at the University of Göttingen and published his paper on number theory (“Über die Anzahl der Primzahlen unter einer gegebenen Grösse”), where he stated the famous Riemann hypothesis on the so-called zeta function.

In 1862 Riemann married Elise Koch, a friend of his sisters.

From 1863 to 1865 he was at the University of Pisa, invited by Enrico Betti (1823-1892). In fact, in this period Riemann's health problem worsened and he tried to recover by staying for a time in a mild climate.

In 1866 he went again to Italy because of his illness, but with no positive effect; he died on July 20 in Selasca.

#### ANALYSIS, MATHEMATICAL PHYSICS, TOPOLOGY AND DIFFERENTIAL EQUATIONS

In Riemann's work, research fields seem to be closely connected; as an explanatory example I show the case of Riemann's *Inauguraldissertation*. His starting point is given by the equations

$$(1) \quad \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

which were supposed to be satisfied by a function  $w = u + iv$  of a variable  $z = x + iy$ ; they lead to the conditions  $\Delta u = 0$ ,  $\Delta v = 0$  – that is to say,  $u$  and  $v$  satisfy the Laplace equations and, therefore, are called harmonic functions.

It is well known that Riemann's complex function theory is strongly connected with potential theory in two dimensions, which Riemann knew very well. As a student he had indeed participated in the physics seminar by Gauss and Weber. In potential theory, an important task concerned the so-called “Dirichlet problem”, which many mathematicians of the 19th century dealt with. It can be stated as follows:

To find a function  $u$  with continuous first partial derivatives on a given bounded domain, which satisfies the Laplace equation within the domain and has given values on the boundary.

Both Gauss and Dirichlet had also confronted such a problem. Gauss (1839) was led to it by studying the distribution of masses or electric charges on a closed surface  $S$ , assuming the potential constant on  $S$ ; while Dirichlet published a treatise on potential theory (*Vorlesungen über die im umgekehrten Verhältnisse des Quadrats der Entfernung wirkenden Kräfte*) in 1876 where this problem has an important role.

The Dirichlet problem plays a fundamental role also in Riemann's theorem concerning the existence of a conformal mapping between two simply connected plane regions. Riemann proved this theorem by solving a special Dirichlet problem and remarked that the assumption "simply connected" can be removed and that, as a consequence, the theorem is valid for Riemann surfaces as well. A "Riemann surface" is introduced for the first time in his *Inauguraldissertation* in order to study multi-valued functions –such as algebraic functions and their integrals. The Riemann surface associated with a function is composed of as many sheets as there are branches of the function, connected in a particular way –so that continuity is preserved and a single-valued function on the surface is obtained. Therefore, he gave an abstract conception of the space of complex variables by using a geometrical formulation.

The concept of "transversal cut" (*Querschnitte*) probably come from mathematical physics too. This definition allowed him to make the surface simply connected with suitable transversal cuts and to study the behaviour of the function in the neighbourhood of the singularities. On the basis of some letters by Riemann himself, it seems that the idea of a transversal cut on a surface struck Riemann after a long discussion with Gauss on a mathematical-physical problem. The origin of the ideas of Riemann surfaces and "Querschnitte" can indeed be found in his note on a problem of electrostatic or thermic equilibrium on the surface of a cylinder with transversal cuts ("Gleichgewicht der Electricität auf Cylindern mit kreisförmigen Querschnitt und parallelen Axen", 1876). In this unpublished note, he was led to consider a Dirichlet problem on a simply connected and simple sheeted surface.

Finally, I note that the Laplace equation –which appears in Dirichlet's problem– is a special partial differential equation of the elliptic kind. Green (1828) showed that –at least theoretically– a Dirichlet problem can be solved by using the so-called Green function. Unfortunately the solution can be explicitly deduced only for special cases. Many mathematicians of the 19th century –such as Hermann von Helmholtz (1821-1894), Rudolph Lipschitz (1832-1903), Betti, Carl Neumann (1832-1925), Franz Neumann (1798-1895), and Riemann himself– deduced functions similar to Green's function in order to solve problems in acoustics, electrodynamics, magnetism, theory of heat, and elasticity.

In a paper (“Ueber die Fortpflanzung ebener Luftwelle von endlicher Schwingungsweite”) published in 1860, Riemann applied the method of Green's function in order to integrate the differential equation of hyperbolic type describing the diffusion of acoustic waves. He introduced a function which plays the same role as Green's function did for the Laplace equation and is today called “Green's function for the hyperbolic problem”.

#### NATURPHILOSOPHIE

In a undated note, written after the completion of his *Inauguraldissertation*, Riemann wrote that his “main work” involved “a new interpretation of the known laws of nature –whereby the use of experimental data concerning the interaction between heat, light, magnetism and electricity would make possible an investigation of their interrelationships. I was led to this primarily through the study of the works of Newton, Euler and, on the other side, Herbart”. The note remained unpublished and appeared only in 1876, posthumously.

Riemann agreed “almost completely” with Herbart's psychology, which inspired both Riemann's model of the ether (the elastic fluid which was supposed to fill all the universe) and his principles of *Naturphilosophie*. According to Herbart, the “psychic act” (or “representation”) is an act of self-preservation with which the “ego” opposes the perturbations coming from the external world. A continuous flow of representations go from the ego to the conscious and back. Herbart studied the connections between different representations in mechanical terms as compositions of forces.

Riemann followed Herbart's psychology and tried to apply Herbart's theory to his conception of the universe in a paper drafted in March 1853 ("Neue mathematische Principien der Naturphilosophie") and in other notes on *Naturphilosophie* which he intended to publish, as he wrote to the brother Wilhelm in December 1853. However, he never published his writings on this subject and all of them only appeared in the edition of his collected works (1876).

In his "Neue mathematische Principien der Naturphilosophie" Riemann followed Herbart's ideas and supposed that the universe is filled with a substance (*Stoff*) flowing continually through atoms and there disappearing from the material world (*Körperwelt*). From this obscure assumption, Riemann tried to build a mathematical model of the space surrounding two interacting particles of substance: if a single particle of substance is concentrated at the point  $O(x_1, x_2, x_3)$  at time  $t$  and at the point  $O'(x'_1, x'_2, x'_3)$  at time  $t'$ , then he considered the two homogeneous forms:

$$(2) \quad ds^2 = dx_1^2 + dx_2^2 + dx_3^2, \quad ds'^2 = dx_1'^2 + dx_2'^2 + dx_3'^2.$$

Riemann considered an appropriate new basis and compared the two forms associated to the "particle of substance" at times  $t$  and  $t'$ . The difference between the two forms is given by:

$$(3) \quad \delta(ds) = ds'^2 - ds^2 = (G_1^2 - 1)ds_1^2 + (G_2^2 - 1)ds_2^2 + (G_3^2 - 1)ds_3^2.$$

If  $\delta(ds) = 0$ , then the particle will not change its form from the time  $t$  to the time  $t'$ ; in this case, the particle does not propagate any force since space is not submitted to deformation by a force. On the contrary, if  $\delta(ds)$  is different from zero, a physical phenomenon is propagated through space.

Riemann tried to connect this formula with the different forces, such as gravity, and heat and light propagation. Riemann could not explicitly show these connections. He limited himself to state that gravity, and light and heat can be explained by assuming that every particle of the homogeneous substance filling space has a direct effect only on its neighbourhood and the mathematical law according to which this happens is due to:

- 1) The resistance with which a particle opposes a change of its volume,  
and



2) The resistance with which a physical line element opposes a change of length.

Gravity and electric attraction and repulsion are founded on the first part, light, heat propagation, electrodynamic and magnetic attraction and repulsion on the second part.

Riemann's *Naturphilosophie* is connected both to some of his physical concepts on electricity and electromagnetism and to his ideas on differential geometry. In fact, as regards physics, I mention a paper on Kohlrausch's experiment written by Riemann in September 1854. According to Kohlrausch's experiment, in a Leyden jar - the first "capacitor" in the history of physics - which had been charged, then discharged and left for some time, a residual charge appeared. Riemann tried to explain "the residual charge" by using the model of ether developed in his *Naturphilosophie* and developed a physical explanation of the electromotive force and of electric propagation. Riemann addressed this paper to the famous journal *Annalen der Physik und Chemie*, but Kohlrausch - the editor of the journal - asked for so many changes that Riemann retracted the paper.

In this paper Riemann proposed a new theory of electricity by assuming that the electric current was caused by a reaction of the body opposed to the change of its own electric state. This reaction is proportional to the charge density, and it decreases or increases the electric density depending on whether the body contains positive or negative electricity. Therefore the transmission of electricity could not be instantaneous but electricity moves "against ponderable bodies" with a certain speed.

In the paper, "Ein Beitrag zur Electrodynamik" (1858), he developed a new theory of electromagnetism, by assuming that electric phenomena travel with the velocity of light and that the differential equations for the electric force are the same as those valid for light and heat propagation. Riemann's research on electromagnetic forces was also influenced by Gauss and Weber. In an unpublished note - dated July 1835 - Gauss had already suggested a new theory of electrodynamics. According to Gauss, two elements of electricity attract and repulse each other with a force depending on their moving state. In a letter to Weber written in 1845, Gauss supposed that electricity propagated from one point to another not instantaneously, but in time, as in the case of light.

He never published his ideas on electrodynamics during his lifetime; therefore, Weber's theory of electrodynamics, published in *Electrodynamische Maasbestimmungen* (1846), was the first result of this kind—where electricity propagates in time—known to the scientific world.

In his lectures on partial differential equations (*Partielle Differentialgleichungen und deren Anwendungen auf physikalische Fragen*, ed. by Hattendorff), published in 1876, Riemann tried to describe the ether surrounding two interacting electric particles. For this purpose, he assumed that the ether possesses physical properties which guarantee electric propagation. He actually deduced a differential equation expressing the flux of ether in space, by using classical Lagrangian mechanics.

His model of the ether and, more generally, his ideas contained in the notes on *Naturphilosophie* are not only connected to physics, but also to his geometrical concepts expressed in the *Habilitationsvortrag*. Riemann indeed extended the “local” investigation of particles of ether—developed in his writings on natural philosophy—to the “global” analysis of  $n$ -dimensional spaces. In fact, if one considers an ether filling all space, then a deformation of space will be closely linked to the force which has to be propagated. Force and curvature of space are then deeply connected; one could say that it is space which propagates forces by changing its curvature.

A deep analysis on the connection between  $n$ -dimensional manifolds and their curvatures is widely developed in a paper on mathematical physics (“*Commentatio Mathematica...*”), which Riemann wrote in 1861, when trying to answer a question proposed by the Paris Academy of Sciences on heat conduction in homogeneous solid bodies.

Riemann's ideas on space and curvature, which seem to anticipate the theory of relativity, were shared by other mathematicians of the 19th century—such as Eugenio Beltrami (1835-1900), Nicolai Lobachevskij (1792-1856), and William Kingdon Clifford (1845-1879). In this connection, in *The common sense of exact sciences* (1885) Clifford asked the question “whether physicists might not find it simpler to assume that space is capable of a varying of curvature, and of a resistance to that variation...”, and that this resistance was the responsible of the propagation of phenomena.

The idea of being in a curved space which, thanks to his changes of curvature, transmitted physical force was at the basis of many mathematical and physical reflections during the 19th century and, in any case, long before Einstein's theory of relativity.

## REFERENCES

- [1] U. Bottazzini & R. Tazzioli, "Naturphilosophie and its role in Riemann's mathematics", *Revue d'histoire des mathématiques*, 1, 1995, p. 3-38.
- [2] D. Laugwitz, *Bernhard Riemann (1826-1866)*, Boston/Berlin, Birkhäuser, 1996.
- [3] B. Riemann, *Gesammelte mathematische Werke und wissenschaftlicher Nachlass* (H. Weber, ed.), Leipzig, Teubner, 1876; repr. New York, Springer, 1991.
- [4] R. Tazzioli, "Fisica e 'filosofia naturale' in Riemann", *Nuncius*, 8, 1993, p. 105-120.
- [5] R. Tazzioli, *Riemann. Le géomètre de la nature. Les génies de la science*, Pour la Science, Paris, 2002



# GEOMETRIA DE RIEMANN

JOAN PORTI

RESUM. Expliquem la contribució de Riemann a la geometria al voltant de la conferència pronunciada a Göttingen el 1854, que suposà el naixement de la geometria de Riemann i obrí nous horitzons a la física.

## 1. INTRODUCCIÓ

El 10 de juny de 1854, Bernhard Riemann pronuncià, a la facultat de filosofia de Göttingen, la conferència d'habilitació “*Über die Hypothesen, welche der Geometrie zu Grunde liegen*”, que podem traduir com “Sobre les hipòtesis en què reposen els fonaments de la geometria”. Tot i haver estat publicada pòstumament, aquesta conferència fou extremadament influent en l'evolució de la geometria i la física. Se la considera una mena de manifest fundacional de la geometria de Riemann —geometria riemanniana— que ha tingut un gran desenvolupament posterior.

Els canvis conceptuals i tècniques que procurà la nova geometria de Riemann són essencials per la teoria de la relativitat general, 60 anys després. Les idees introduïdes en aquesta conferència ajudaren a dir més endavant que la matèria deforma la geometria de l'espai, que deixava de ser inamovible. Riemann creà els fonaments que permetrien desenvolupar les matemàtiques necessàries per explicar com l'espai queda afectat per la matèria, segons la relativitat general.

---

Aquest article està basat en una conferència a la jornada Riemann de l'UPC, el 20 de febrer de 2008. Vull agrair a Sebastià Xambó i a tota la “Comissió Riemann” per la invitació. També vull agrair a Eva Miranda l'ajut en la preparació de la conferència, en particular en el guió dels *Riemann cartoons*, cf. la figura 1. Finalment, el projecte FEDER/MEC MTM2006-04353 financia la meua recerca.



FIGURA 1. Diàleg apòcrif entre Gauss (1777-1855) i Riemann (1826-1866).

Entre les nocions més importants, la primera que Riemann planteja és la de varietat de dimensió  $n$ , ja que la majoria d'exemples fins aleshores eren subvarietats de  $\mathbf{R}^m$ , i Riemann els dóna entitat i valor intrínsec.

Després del tensor mètric, una altra noció clau que apareix és la de curvatura. Gauss ja l'havia introduïda per superfícies a l'espai, aquí Riemann la desenvolupa en qualsevol dimensió i en mostra la importància. A més ho fa amb una simplicitat i elegància sorprenents. Actualment, en els cursos i llibres de text s'arriba a la curvatura mitjançant un procés força més complicat, que inclou les connexions. Val la pena llegir la conferència de Riemann i, tot i que no sabem com s'ho va fer Riemann per arribar a aquestes conclusions, podem consultar l'aproximació que proposa Spivak [38].

Després de la introducció, aquest article està organitzat en sis capítols més. En el 2 situem el context en què es va realitzar la conferència de Riemann, en el 3 fem un breu recordatori de què entenem actualment per geometria de Riemann, que ens servira per entendre i valorar el capítol 4, el més extens, on s'expliquen els continguts de la conferència. Al capítol 5 expliquem els desenvolupaments principals realitzats posteriorment, que consolidaren la geometria de Riemann. Al capítol 6 citem alguns resultats que relacionen la topologia amb la curvatura de les varietats. Finalment, al capítol 7 expliquem com s'ha

resolt la conjectura de Poincaré mitjançant el flux de Ricci, una eina d'origen riemannià.

## 2. EL CONTEXT

**2.1. La conferència d'habilitació.** La conferència pronunciada el 10 de juny de 1854 completava el procés d'habilitació de Bernhard Riemann, per això se l'anomena *Habilitationsvortrag*. Al desembre de 1953 havia presentat la memòria d'habilitació, *Habilitationsschrift*, amb el títol: *Über die Darstellbarkeit einer Function durch eine trigonometrische Reihe* (Sobre la representabilitat d'una funció en sèrie trigonomètrica).

Tal com encara es fa a Alemanya en l'actualitat, després de presentar l'*Habilitationsschrift* Riemann va proposar tres títols i un tribunal encapçalat per Gauss en va escollir un. El primer títol proposat per Riemann era "*Gesichte der Frage über die Darstellbarkeit einer Function durch eine trigonometrische Reihe*" ("Història de la qüestió de representabilitat d'una funció com a sèrie trigonomètrica"); el segon, "*Über die Auflösung zweier Gleichungen zweiten Grades mit zwei unbekanntem Grössen*" ("Sobre la solució de dues equacions quadràtiques amb dues variables"); el tercer, que fou l'escollit, era "*Über die Hypothesen, welche der Geometrie zu Grunde liegen*". Tot i que tradicionalment mai no s'escollia el tercer, cal dir que el primer títol era molt proper al tema de la memòria d'habilitació, i que el segon, la intersecció de dues quàdriques, podia haver-se considerat com a trivial per Gauss. No s'han trobat manuscrits sobre l'elaboració de cap dels tres temes.

La conferència estava adreçada als professors de la facultat de filosofia, i per tant hi ha molt poques fórmules i cap càlcul en el text. Queda com un misteri esbrinar quins càlculs havia fet Riemann per arribar a les conclusions i fórmules que exposa.

**2.2. Antecedents.** Està fora del meu abast establir amb precisió quins treballs previs tingueren influència en la conferència de Riemann. Probablement els treballs de Johannes Bernouilli, d'Euler i de Meusnier sobre superfícies eren coneguts per Riemann, donat que tingueren continuïtat en l'obra de Gauss. En canvi, és molt més difícil de saber si Riemann coneixia dels treballs recents sobre geometria no euclidiana,

Ueber  
die Hypothesen, welche der Geometrie zu Grunde liegen.  
Von  
B. R i e m a n n.

---

FIGURA 2. El títol del text de la conferència, publicat pòstumament.

de Bolyai i Lobachevsky. Gauss hi estava interessat [35], però Riemann no en diu res en cap moment.

Riemann sí que menciona Gauss, i no només perquè estava en el tribunal, sinó perquè generalitza alguns dels seus treballs. En particular, les coordenades intrínseques utilitzades per Gauss, la mètrica intrínseca de la varietat i la curvatura.

**2.3. Les reaccions.** En les obres completes de Riemann [37], Dedekind descriu la conferència com una obra mestra d'exposició, tot i que sembla difícil que la majoria de professors de la facultat de filosofia de Göttingen la poguessin comprendre suficientment. Qui sense cap dubte s'adonà de la importància de la conferència fou Gauss, que —segons escriu Dedekind—

*amb gran apreciació i una animació rara per ell, parlà a  
Wilhem Weber sobre la profunditat de les idees presen-  
tades per Riemann.*

El text de la conferència fou publicat pòstumament i les seves idees encara van trigar temps a ser assimilades: almenys fins a la teoria de la relativitat general, no se li va donar tota la rellevància. Per citar-ne un exemple, Henri Poincaré escriu en el seu assaig de 1902 “La science et l'hypothèse” [34]:

*Les geometries de Riemann, tan interessants en molts  
camps, mai no seran, però, altra cosa que ens purament  
analítics, i no conduiran a demostracions anàlogues a les  
d'Euclides.*

Encara faltaven uns anys per la teoria de la relativitat general.



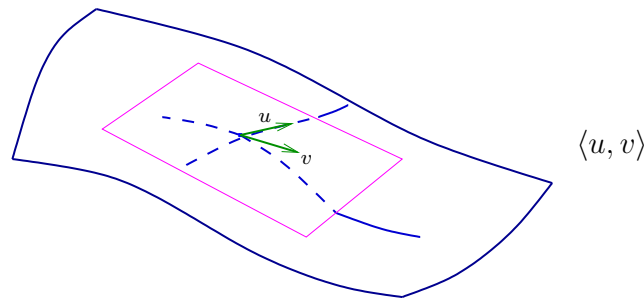


FIGURA 3. Producte escalar.

### 3. GEOMETRIA DE RIEMANN ACTUALMENT

Abans d'entrar en el text de la conferència, repassem breument el que entenem avui en dia per geometria de Riemann. De moment donem les construccions bàsiques per entendre el que Riemann deia, després ja anirem introduïnt més eines.

Una *varietat de Riemann* és una varietat diferenciable  $M$  amb un producte escalar en l'espai tangent de cada punt, que escrivim  $g$ . Que la varietat sigui diferenciable vol dir que els canvis de coordenades són diferenciables (normalment demanem de classe  $C^2$  per poder fer segones derivades), i es demana que el producte escalar  $g$  variï diferenciablement.

La manera de descriure-ho és utilitzar coordenades d'un obert  $U$  de  $\mathbf{R}^n$ :

$$x = (x^1, \dots, x^n) \in U \subset \mathbf{R}^n.$$

Els coeficients de la matriu del producte escalar són funcions de les coordenades

$$g_{ij}(x) = \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle.$$

Per tant, donats dos vectors  $u$  i  $v$ , si la seva expressió en coordenades és

$$\left. \begin{aligned} u &= u^i \frac{\partial}{\partial x^i} \\ v &= v^j \frac{\partial}{\partial x^j} \end{aligned} \right\},$$

aleshores el producte escalar de dos vectors  $u$  i  $v$  tangents al punt de coordenades  $x$  s'expressa com

$$\langle u, v \rangle = \sum u^i g_{ij}(x) v^j = (u^1 \cdots u^n) \begin{pmatrix} g_{11}(x) & \cdots & g_{1n}(x) \\ \vdots & & \vdots \\ g_{n1}(x) & \cdots & g_{nn}(x) \end{pmatrix} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix},$$

cf. la figura 3. Per bilinearitat, és equivalent donar el producte escalar o la forma quadràtica corresponent, és a dir la norma dels vectors.

**3.1. Tensors.** L'assignació d'una forma bilineal a cada punt del tangent és un exemple de *tensor*, en aquest cas el tensor mètric o tensor de Riemann. Si canviem bilineal per multilinear i fem intervenir el dual, tenim la noció general de *tensor*.

Per exemple, els camps vectorials, que a cada punt de la varietat associen un vector, són tensors, i les formes diferencials també, perquè assignen a cada punt de la varietat un element del dual del tangent. Les funcions són el cas més senzill de tensor.

La notació dels tensors pot arribar a ser força ferragosa, i per simplificar s'utilitza *el conveni de sumació d'Einstein*, segons el qual, quan tenim *superíndexs* i *subíndexs coincidents* en expressions multiplicades, suposem que se sumen i ja no escrivim el sumatori. Per exemple, un camp  $V$  i una forma  $\omega$  es poden escriure com:

$$V = \sum_i v^i \partial_i = v^i \partial_i \quad \text{i} \quad \omega = \sum_i w_i dx^i = w_i dx^i.$$

Aquí hem utilitzat l'abreviació dels camps de coordenades

$$\partial_i = \frac{\partial}{\partial x^i}.$$

Per simplificar, moltes vegades s'ometen els termes  $\partial_i$  i  $dx^i$ , que se sobreentenen en funció dels índexs i superíndexs. Així el camp  $V$  i la forma  $\omega$  es poden escriure:

$$V = v^i \quad \text{i} \quad \omega = w_i.$$

Per exemple, una mètrica, sense cap convenció prèvia s'escriuria com

$$\sum_{ij} g_{ij} dx^i \otimes dx^j,$$

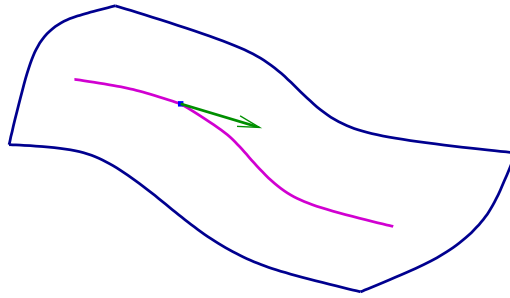


FIGURA 4. La longitud de corbes i l'element de línia.

però amb la convenció d'Einstein i ometent les bases canòniques de formes, escrivim

$$g_{ij}.$$

3.2. **Longituds.** Per mesurar longituds utilitzem l'element de línia:

$$ds = \sqrt{g_{ij} dx^i dx^j}.$$

Riemann ja utilitzava la notació  $ds$ , i l'expressió seria la mateixa si no fos perquè ell utilitzava sumatoris i no la convenció d'Einstein.

L'element de línia es fa servir per calcular la longitud  $L(\gamma)$  d'un corba diferenciable  $\gamma$ :

$$L(\gamma) = \int_{\gamma} ds.$$

Més precisament, si la corba s'escriu en coordenades com

$$\gamma(t) = (x^1(t), \dots, x^n(t)) \quad t \in [a, b],$$

aleshores la seva longitud és

$$L(\gamma) = \int_{\gamma} ds = \int_a^b \left\| \frac{d\gamma}{dt} \right\| dt = \int_a^b \sqrt{g_{ij} \frac{dx^i}{dt} \frac{dx^j}{dt}} dt.$$

Això ens permet calcular longituds de corbes contínues diferenciables a trossos.

La *distància* entre dos punts es defineix com l'ínfim de la longitud dels camins que uneixen dos punts. Aquest ínfim és en realitat un mínim, i les corbes que el realitzen s'anomena *geodèsiques*.

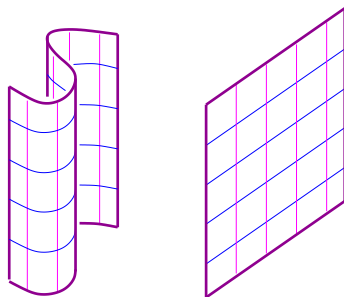


FIGURA 5. Exemple de superfícies a l'espai *isomètriques.*, amb la mateixa mètrica. La identificació (o *isometria*) preserva les línies paral·leles de la figura.

La distància determina l'element de línia, i per tant el tensor mètric de Riemann  $g$  i totes les altres propietats com ara els angles i les nocions que veurem més endavant: transport paral·lel, curvatura, etc. Per tant les varietats de Riemann estan caracteritzades per les propietats mètriques.

**3.3. Curvatura de Gauss.** Un cop sabem mesurar longituds, les nocions més importants són les diferents curvatures: la de Riemann, la de Ricci, l'escalar i la seccional o de Gauss, però ja les veurem més endavant. De moment tornem a Riemann i parlem de la curvatura de Gauss.

En la seva memòria de 1827 "*Disquisitiones generales circa superficies curvas*", Gauss ens parla de superfícies a l'espai i de les diferents nocions intrínseques. És a dir, Gauss considera superfícies a l'espai amb la mètrica induïda, però pot passar que la mateixa mètrica es realitza per diferents immersions de la varietat a l'espai, cf. la figura 5. Les nocions intrínseques són les que només depenen de la mètrica, no de com la superfície està situada a l'espai.

Gauss demostra el Teorema Egregi que porta el seu nom. Els treballs d'Euler i Meusnier havien introduït les anomenades curvatures principals, Gauss definí una nova noció de curvatura, que és el producte de les curvatures principals, actualment anomenada curvatura de Gauss (cf. figura 6). Observem que la curvatura de Gauss de les superfícies de la figura 5 és zero.

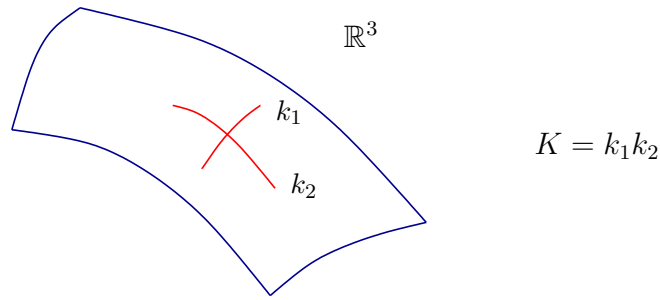


FIGURA 6. La cruvatura de Gauss  $K$  és el producte de les principals  $k_1$  i  $k_2$ .

**Teorema 3.1** (Teorema Egredi de Gauss). *La curvatura Gaussiana és un invariant intrínsec.*

És a dir la curvatura de Gauss només depèn de la mètrica intrínseca de la superfície i no de com està situada a l'espai. Riemann no només generalitzarà la noció de curvatura a qualsevol dimensió, sinó que la noció d'invariant intrínsec és una porta per la qual Riemann accedeix a un nou concepte de geometria.

#### 4. ÜBER DIE HYPOTHESEN, WELCHE DER GEOMETRIE ZU GRUNDE LIEGEN

En aquesta secció comentarem alguns aspectes de la conferència. Spivak [38] conté una discussió més detallada a nivell matemàtic, i altres autors l'analitzen des d'un punt de vista més històric [12, 22, 27].

La introducció es titula “*Plan der Untersuchung*”, és a dir “Pla de la recerca”. Riemann comença parlant de la confusió sobre l'estat de la geometria no euclidiana, que en aquest temps encara no estava completament acceptada. De manera independent, el 1829 Lobachevsky i Bolyai havien construït geometries que començaven suposant que per un punt hi ha més d'una línia paral·lela a una línia donada, però encara s'esperava que es trobaria una contradicció a aquestes construccions. Riemann atribueix les dificultats trobades en l'estudi de la geometria euclidiana al fet que els geòmetres mai no havien separat el que anomenem propietats topològiques de l'espai de les seves propietats mètriques.

En el desenvolupament axiomàtic de la geometria, fins i tot la noció d'espai no està definida, i les seves propietats es dedueixen mitjançant els axiomes. Riemann ens proposa primer el plantejament de totes les mètriques possibles i després la determinació empírica de la mètrica a l'espai. De manera natural sorgeix la qüestió de trobar les dades més simples que ens permetin distingir l'espai d'altres possibles varietats i mètriques de dimensió tres. Amb una gran visió física, qüestiona la legitimitat d'extendre, a les escales massa petites o massa grans, les propietats que poguem determinar experimentalment per l'espai.

Després de la introducció, la memòria està estructurada en tres parts, que analitzem a continuació amb diferents nivells de detall.

**4.1. Part I. Varietats de dimensió  $n$ .** La primera part es titula "*Begriff einer  $n$ -fach ausgedehnten Größe*" i s'hi introdueix la noció de varietat de dimensió  $n$ .

Tot i que no podem determinar fins a quin punt havia trobat la noció precisa de varietat, Riemann tenia clar que una varietat és localment homeomorfa a l'espai euclidià  $\mathbf{R}^n$ . Riemann exemplifica com es passa d'una varietat a una altra de dimensió superior o inferior: les varietats de dimensió  $n + 1$  s'obtenen variant en un paràmetre les de dimensió  $n$ , i els conjunts de nivell de certes funcions són varietats de dimensió inferior. Entre altres exemples, posa els espais de paràmetres, els continus dels colors.

Riemann també ens descriu la noció de *coordenades*, en una varietat de dimensió  $n$  calen  $n$  valors numèrics per determinar un punt:

*Mitjançant  $n$  iteracions d'aquest procés [fent referència a les funcions de nivell], la determinació de la posició en una  $n$ -varietat queda determinada per  $n$  determinacions numèriques...*

Per acabar aquesta primera part, també fa referència a les varietats de dimensió infinita i, sense entrar en detalls, en posa exemples:

*Aquestes varietats estan formades, per exemple, per les possibles funcions en un determinat domini, o per les possibles formes d'una figura sòlida, etc...*

**4.2. Part II. Secció 1. Element de línia.** Un cop introduïdes les varietats de dimensió  $N$ , Riemann passa a les nocions mètriques, i comença definit l'element de línia, que ens permet calcular longituds de corbes. L'element de línia només dependrà del punt de l'espai i del vector tangent de la corba, és a dir de la direcció cap on va la corba. Aquesta és la manera que té Riemann de donar la mètrica com una propietat intrínseca de l'espai, i la notació  $ds$  s'ha mantingut fins a l'actualitat. A continuació ho descrivim en detall.

Riemann comença la secció dient de quines corbes es cuida. En termes moderns, parlem de corbes que són diferenciables amb diferencial continua, és a dir de classe  $C^1$ . En el seu llenguatge, Riemann ens parla de corbes de les quals...

*...les raons de les quantitats  $dx$  varien continuament.*

Per aquestes corbes, la longitud s'obté integrant l'element de línia  $ds$ :

$$L(\gamma) = \int_{\gamma} ds.$$

La fórmula que ens proposa Riemann per a l'element de línia és la següent:

*$ds$  = l'arrel quadrada d'una funció sempre positiva, homogènia de segon grau en les quantitats  $dx$  i amb coeficients funcions de  $x$ .*

És a dir

$$ds = \sqrt{\sum g_{ij}(x) dx^i dx^j},$$

on  $g_{ij}(x)$  són els coeficients d'una matriu d'una forma quadràtica definida positiva, que és l'expressió que hem donat a la secció 3.2.

Riemann també ens diu que per a l'espai amb les coordenades canòniques l'expressió és

$$ds = \sqrt{\sum dx_i^2}.$$

L'arrel quadrada queda justificada per homogenitat, donat que tenim una forma quadràtica. Per què una forma quadràtica i no una forma de grau quatre? Riemann ens diu que...

*...el següent cas en ordre de complexitat seria el de varietats en que l'element de línia es pot expressar com l'arrel quarta d'una expressió diferencial de grau quatre.*

Això demanaria molt esforç de càlcul que difícilment es justificaria ja que, segons Riemann...

*...els resultats no es poden expressar geomètricament.*

L'explicació però no sembla gaire convincent, i potser una altra justificació és que, amb una forma quadràtica, es generalitza sense problemes la geometria intrínseca que Gauss havia desenvolupat per les superfícies a l'espai.

**4.3. Canvi de coordenades.** Un cop definit l'element de línia mitjançant coordenades locals, Riemann planteja la qüestió del *canvi de coordenades*. És a dir, donades dues eleccions de coordenades d'una mateixa regió de la varietat, com podem saber si corresponen a la mateixa mètrica?

Diem que dues expressions d'una mètrica són *equivalents* si podem passar d'una a l'altra mitjançant un canvi de coordenades. El problema de decidir si dues mètriques arbitràries són equivalents és molt complicat i no es pot respondre en general, ja que tal com Riemann ens diu, hi intervenen massa indeterminades, les  $n(n-1)/2$  funcions  $g_{ij}$ . Malgrat això, es poden trobar invariants per distingir classes d'equivalència de mètriques, per exemple la curvatura.

Aquesta secció acaba amb les varietats que Riemann anomena *planes*, les que l'element de línia és equivalent al de l'espai Euclidià:

$$\sqrt{\sum dx^2}.$$

En la secció següent Riemann introduirà la curvatura i la veurà com una obstrucció a la plitud.

**4.4. Part II. Secció 2. La curvatura.** Abans de la curvatura, Riemann tria un sistema de coordenades adequat. Una de les primeres coses que aprenem en geometria de Riemann és que l'elecció d'un bon sistema de coordenades pot simplificar substancialment els càlculs. Les



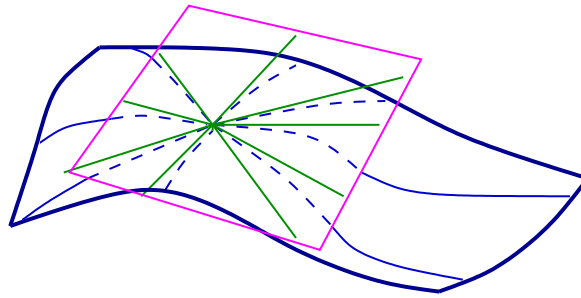


FIGURA 7. Aplicació exponencial. La figura representa les rectes del tangent que passen per l'origen i la seva imatge per l'aplicació exponencial, les geodèsiques que surten del punt base.

coordenades escollides per Riemann són les que avui anomenem *coordenades geodèsiques* o coordenades normals.

L'*aplicació exponencial* identifica rectes radials que surten de l'origen amb geodèsiques (minimitzants) de la varietat que surten del punt. La identificació es fa de manera que la distància a l'origen de coordenades correspon a la distància al punt: és a dir cada vector és enviat a una distància del punt base igual a la seva longitud. Aquesta aplicació està ben definida en un entorn de l'origen, i utilitza el fet que donat un vector existeix una única geodèsica tangent a aquest vector, llevat parametritzacions, cf. la figura 7.

En altres paraules, si  $T_pM$  denota l'espai tangent a un punt  $p \in M$ , existeix  $U \subset T_pM$ , un entorn de 0, tal que l'exponencial està ben definida i compleix:

$$\begin{aligned} \exp_p : U \subset T_pM &\rightarrow M \\ v &\mapsto \gamma(1) \end{aligned} ,$$

on  $\gamma(t)$  és l'única corba geodèsica determinada per les condicions inicials  $\gamma(0) = p$  i  $\gamma'(0) = v$ .

A més tenim el resultat següent.

**Teorema 4.1.** *L'aplicació exponencial indueix un difeomorfisme entre un entorn de l'origen a l'espai tangent i un entorn del punt base a la varietat.*

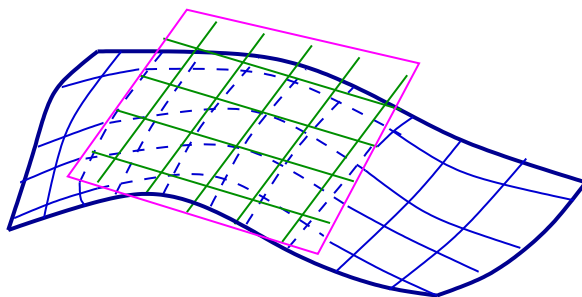


FIGURA 8. Coordenades normals o exponencials. En el pla tangent representem les coordenades rectangulars, i a la varietat la seva imatge per l'exponencial, és a dir les coordenades normals.

En conseqüència podem definir:

**Definició 4.2.** *Les coordenades normals són les coordenades cartesianes induïdes per l'aplicació exponencial i les coordenades canòniques o rectilínies de l'espai tangent (és a dir per les quals la forma quadràtica és la standard), cf. la figura 8.*

Riemann afirma que en coordenades geodèsiques:

- A l'origen  $ds^2$  val  $\sum(dx^i)^2$  (és a dir,  $g_{ij}(0) = \delta_{ij}$ ).
- No hi ha terme de primer ordre en la sèrie de Laurent.
- El de segon ordre és combinació de  $(x^1 dx^2 - x^2 dx^1)$ ,  $(x^1 dx^3 - x^3 dx^1)$ .

En llenguatge actual, això és equivalent al teorema 4.3 i a les simetries del tensor de Riemann.

**Teorema 4.3.** *En coordenades geodèsiques:*

$$g_{ij}(x) = \delta_{ij} + \frac{1}{3}R_{i\alpha\beta j}x^\alpha x^\beta + O(|x|^3).$$

Les simetries són:

$$\begin{aligned} R_{i\alpha\beta j} &= -R_{i\alpha j\beta} = -R_{\alpha i\beta j} = R_{\beta j i\alpha}, \\ R_{i\alpha\beta j} + R_{i\beta j\alpha} + R_{ij\alpha\beta} &= 0. \end{aligned}$$

Aquestes simetries són equivalents a la frase “El de segon ordre és combinació de  $(x^1 dx^2 - x^2 dx^1)$ ,  $(x^1 dx^3 - x^3 dx^1)$ ”.

A partir d'aquí podem definir:

**Definició 4.4.** *El tensor  $R_{i\alpha\beta j}$  s'anomena curvatura de Riemann.*

És a dir, utilitzant la multilinealitat definim un tensor pel seu valor en una base:  $R$  és l'únic tensor que satisfà  $R(\partial_i, \partial_\alpha, \partial_\beta, \partial_j) = R_{i\alpha\beta j}$ . Alternativament:

$$R(u^i \partial_i, v^\alpha \partial_\alpha, w^\beta \partial_\beta, z^j \partial_j) = u^i v^\alpha w^\beta z^j R_{i\alpha\beta j}.$$

Actualment aquesta no és la definició que se sol donar als llibres de text ni als cursos de geometria. Més endavant parlarem de connexions i de la definició actual, però val la pena veure com Riemann la dóna amb molt pocs prerequisits, tot i que quedi pendent de saber quins càlculs va fer per arribar a aquesta conclusió.

Un cop introduïda la seva curvatura, Riemann ens diu que recupera la curvatura de Gauss per superfícies. En llenguatge modern tindriem que si  $K$  és la curvatura de Gauss d'una superfície, aleshores

$$K = R_{1212} = -R_{1221} = -R_{2112} = R_{2121}.$$

Riemann no introdueix símbols pel tensor que més endavant portarà el seu nom, però que calcularà en un article que comentarem al paràgraf 5.1.

En el text ens diu:

*Multiplicirt mit  $-3/4$  wird sie der Grösse gleich, welche Herr Geheimer Hofrath Gauss das Krümmungsmass einer Fläche ganannt hat.*

Aquí el factor  $1/4$  surt d'una normalització diferent a la nostra que Riemann fa, el signe es deu a una convenció diferent, mentre que el 3 prové del teorema 4.3.

**4.5. Part II. Secció 3. Curvatura seccional.** En la resta de seccions d'aquesta part II, Riemann desenvolupa la idea de curvatura i en dóna exemples.

En el cas bidimensional, Riemann ja ha explicat que la seva curvatura no és altra que la de Gauss. Per entendre-la en dimensió superior, Riemann explica a la secció tres el que actualment s'anomena *curvatura seccional*. Donat l'espai tangent a un punt base i un pla de l'espai tangent, és a dir un subespai vectorial de dimensió dos, podem considerar la superfície que conté totes les geodèsiques que passen pel punt base i són tangents a aquest pla. Obtenim d'aquesta manera una superfície, i la seva curvatura de Gauss és la curvatura de Riemann de les coordenades corresponents (una certa quantitat multiplicada per un factor  $-3/4$ ), de la mateixa manera que el cas purament bidimensional. Riemann ens diu que quan ens restringim a plans lineals que passen per l'origen, en coordenades geodèsiques tornem a recuperar la curvatura de Gauss de les superfícies.

**4.6. Part II. Secció 4. Curvatura seccional constant.** Un cop introduïda la curvatura seccional, Riemann dóna l'expressió d'una mètrica de curvatura constant  $\alpha$ . Amb una bona elecció de coordenades, aquesta mètrica s'expressa com:

$$ds = \frac{1}{1 + \frac{\alpha}{4}\sum x_i^2} \sqrt{\sum_i dx_i^2}.$$

Aquest càlcul és un altre exemple de bona elecció de coordenades, car ara ja no són les geodèsiques o normals que ha fet servir en la definició, sinó unes altres que li permeten tenir una expressió molt més senzilla.

Encara que Riemann no ho digui, aquesta fórmula porta implícit el resultat que, quan la curvatura seccional és constant, la mètrica queda determinada pel valor  $\alpha$  de la curvatura, ja sigui localment, ja sigui globalment en el cas simplement connex.

Observem que dóna models de curvatura constant negativa, com els que havien construït Bolyai i Lobachevsky. Per exemple, per  $\alpha = -1$  i amb el canvi de variable  $x = 2y$ , obtenim el que actualment s'anomena model de Poincaré.

A més, aquesta fórmula ens permet tenir una intuïció sobre la curvatura. Per exemple, si mirem les esferes  $S_r$  de radi  $r > 0$  (el conjunt de punts a distància  $r$  d'un punt donat) la distància a  $S_r$  és igual a la de l'esfera unitat de  $\mathbf{R}^n$  multiplicada per un factor homotètic. Aquest

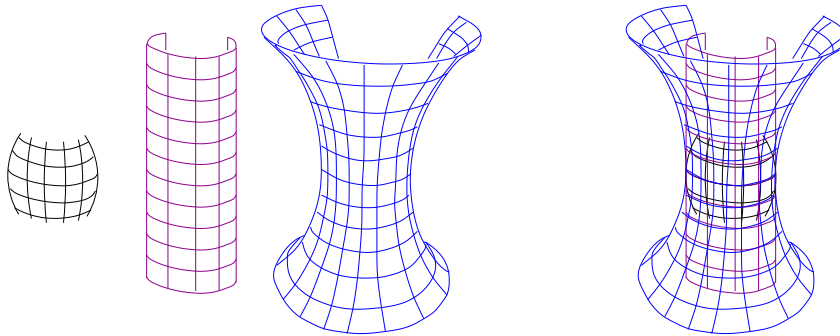


FIGURA 9. Superfícies de rotació de curvatura constant que són paral·leles a l'equador.

factor depén de la curvatura i el seu valor és

$$\begin{cases} \frac{1}{\sqrt{\alpha}} \sin(r\sqrt{\alpha}) & \text{si } \alpha > 0; \\ r & \text{si } \alpha = 0; \\ \frac{1}{\sqrt{-\alpha}} \sinh(r\sqrt{-\alpha}) & \text{si } \alpha < 0. \end{cases}$$

Aquest factor fa referència a les distàncies, el tensor mètric s'hauria de multiplicar pel seu quadrat. Per tant en curvatura positiva les esferes “són menors” que les corresponents a l'espai euclidià, i en curvatura negativa, “són majors”. En particular, per  $\alpha > 0$  aquestes esferes col·lapsen quan  $r = \pi/\sqrt{\alpha}$ , que és el diàmetre de l'esferes de curvatura constant  $\alpha > 0$ .

**4.7. Part II. Secció 5. Superfícies de rotació.** Per acabar la part II, a la secció 5 Riemann descriu superfícies de rotació de curvatura constant, sense ni un dibuix, és clar. Si col·loquéssim superfícies de rotació amb un mateix paral·lel (òrbita de rotació d'un punt) que fossin tangents al cilindre corresponent, les superfícies de més curvatura estarien a dins de les menys curvades. En particular, les de curvatura positiva estarien a dins del cilindre, i les de curvatura negativa a fora. Vegeu la figura 9.

**4.8. Part III. Aplicacions a l'espai.** A la tercera part, Riemann discuteix com es pot intentar determinar quina és la mètrica de l'espai físic que ens envolta. En altres paraules, quina és la relació de la geometria amb el món real.

Les qüestions que planteja són: primer quina és la dimensió de l'espai i segon quina geometria descriu l'espai físic. En la primera qüestió, Riemann observa que la noció de dimensió assumeix una noció de continuïtat, que per conjunts finits o discrets no té cap sentit. Suposant que l'espai físic es pugui veure com una varietat de dimensió finita, té límits o frontera? És acotat? Per exemple, si la curvatura és constant i positiva, necessàriament serà acotat.

La geometria just definida i més en general el càlcul infinitesimal li serveix a una certa escala, però es planteja la seva validesa a escala infinitesimal. En aquest sentit la visió de Riemann per la física és gairebé profètica:

*“La qüestió de la validesa de la geometria en l'infinitament petit està relacionada amb la qüestió de la base interna de la relació mètrica de l'espai. Per a aquesta qüestió, que encara pot ser considerada part de la teoria de l'espai, l'observació prèvia és aplicable, és a dir que una varietat discreta té una relació mètrica inherent, mentre que en el cas continu ha de venir d'un altre lloc. Per tant o bé la realitat que està a la base de l'espai ha de formar una varietat discreta, o bé hem de buscar les bases per les relacions mètriques a fora, en forces de lligadura que hi actuen.”*

*“La resposta a aquesta qüestió només es pot trobar començant des de la concepció dels fenòmens justificats previament per l'experimentació, en els quals Newton es fonamentà, i per canvis successius deguts a fenòmens no explicats. Investigacions com la que hem portat a terme aquí i que provenen de nocions generals només poden servir per garantir que aquest treball no està limitat per conceptes massa restringits, i que el progrés en el reconeixement de la relació entre les coses no està obstaculitzat per prejudicis tradicionals.”*

*“Això ens porta als dominis d'una altra ciència, al reialme de la física, en el qual la natura de la present ocasió no ens permet d'entrar.”*

## 5. DESENVOLUPAMENT POSTERIOR

A l'apartat 5.1 discutim una construcció de Riemann en què calcula el tensor de curvatura en altres coordenades i el veu efectivament

com una obstrucció a la platitud. Això donà lloc als treballs de Christoffel i Lipschitz. A la resta del capítol comentarem treballs de Ricci-Curbastro i Levi-Civita i mencionarem breument Einstein i Grossmann.

**5.1. Tensor de curvatura en altres coordenades.** En la conferència d'habilitació, Riemann deia que la curvatura era una obstrucció a la platitud. Això no ho havia demostrat, ni tampoc no havia donat l'expressió de  $R_{i\alpha\beta j}$  en un sistema de coordenades arbitrari. En un article [37] sotmés a l'academia de ciències el dia 1 de juliol de 1861, Riemann dóna la fórmula per  $R_{i\alpha\beta j}$  en funció de  $g_{ij}$  i les seves derivades. L'article estableix un criteri de platitud per a una mètrica qualsevol.

Explicuem el càlcul fent servir el formalisme de Christoffel, introduït posteriorment en un article [8] que tracta sobre aquest treball de Riemann. Definim els símbols de Christoffel com

$$\Gamma_{ab|c} = \frac{1}{2}(\partial_c g_{ab} - \partial_a g_{bc} + \partial_b g_{ca}),$$

si fem servir la notació  $\partial_a = \frac{\partial}{\partial x_a}$ . Aleshores

$$-\frac{1}{2}R_{i\alpha\beta j} = \partial_\alpha \partial_j g_{i\beta} + \partial_\beta \partial_i g_{j\alpha} - \partial_\alpha \partial_\beta g_{ij} - \partial_i \partial_j g_{\alpha\beta} + g^{lm}(\Gamma_{l\alpha|j} \Gamma_{mi|\beta} - \Gamma_{li|j} \Gamma_{m\alpha|\beta}),$$

on  $g^{lm}$  denota la matriu inversa de  $g_{ij}$ , i per tant

$$\sum_k g^{ik} g_{kj} = \delta_j^i = \begin{cases} 1 & \text{si } i = j; \\ 0 & \text{si } i \neq j. \end{cases}$$

La notació que Riemann utilitzava per a l'expressió  $-\frac{1}{2}R_{i\alpha\beta j}$  era

$$(i\alpha, \beta j).$$

L'article va romandre gairebé desconegut fins a la publicació de les obres completes el 1876.

Riemann donava la condició necessària

$$R_{i\alpha\beta j} = 0$$

per tal que una varietat fos plana i va indicar vagament que era suficient. La demostració de la suficiència la donaren Christoffel (l'introduïdor dels símbols  $\Gamma$ ) el 1867 [8] i Lipschitz el 1870 [24].



FIGURA 10. Christoffel (1829-1900) i Lipschitz (1832-1903).

**5.2. Derivada covariant.** A principis de 1887 [36], Gregorio Ricci-Curbastro va desenvolupar la derivació covariant de tensors. Comencem amb la derivada covariant de camps. A partir dels símbols  $\Gamma_{jk|r}$  de l'apartat anterior, definim el que actualment es coneix com a símbols de Christoffel:

$$\Gamma_{jk}^i = \sum g^{ir} \Gamma_{jk|r}.$$

Aleshores la *derivada covariant* del camp  $V = \sum v^i \partial_i$  en la direcció  $\partial_k$  és:

$$\nabla_{\partial_k} V = \partial_k v^i \partial_i + \Gamma_{jk}^i v^j \partial_i.$$

La derivada covariant és lineal respecte a la direcció de derivació:

$$\nabla_{u^k \partial_k} V = u^k \nabla_{\partial_k} V.$$

Per a les funcions, la derivada covariant s'entén que és la derivada direccional.

Aquesta derivada covariant de camps s'estén de manera natural a tensors, sempre utilitzant la regla de Leibnitz i la compatibilitat amb contraccions. És a dir, si  $X$  i  $Y$  són dos tensors, aleshores

$$\nabla_v (X \otimes Y) = \nabla_v X \otimes Y + X \otimes \nabla_v Y.$$

A més,  $\nabla_v$  commuta amb les contraccions. Per posar un exemple de contracció, si  $X$  és un camp i  $\omega$  una forma, aleshores  $\omega(X)$  és una funció obtinguda per contracció del tensor  $\omega \otimes X$ . Per tant la regla de



Leibnitz i la compatibilitat amb les contraccions ens diu:

$$\nabla_v(\omega(X)) = (\nabla_v\omega)(X) + \omega(\nabla_vX).$$

Això ens permet calcular la derivada covariant de formes:

$$\begin{aligned} \nabla_{\partial_k}\omega &= (\nabla_{\partial_k}\omega)(\partial_j)dx^j = (\partial_k(\omega(\partial_j)) - \omega(\nabla_{\partial_k}\partial_j))dx^j \\ &= (\partial_k\omega_j - \Gamma_{jk}^i\omega_i)dx^j. \end{aligned}$$

Un altre exemple: si  $X$  i  $Y$  són camps, aleshores  $g(X, Y)$  és una funció, i tenim:

$$v(g(X, Y)) = \nabla_v(g)(X, Y) + g(\nabla_vX, Y) + g(X, \nabla_vY).$$

Per tant

$$(1) \quad \nabla_v(g)(X, Y) = v(g(X, Y)) - g(\nabla_vX, Y) - g(X, \nabla_vY).$$

Pel fet de ser lineal respecte a  $v$ , complir la regla de Leibnitz i ser compatibles per contraccions, diem que  $\nabla$  és una derivada covariant o una *connexió*. Tal com l'hem definida, compleix dues propietats més que la caracteritzen entre totes les connexions possibles:

$$\begin{aligned} \nabla g &= 0, \\ \nabla_X Y - \nabla_Y X &= [X, Y]. \end{aligned}$$

De la primera propietat diem que el tensor de Riemann  $g$  és constant, i de la segona que és compatible amb el parèntesi de Lie. En particular  $\nabla g = 0$  es llegeix, mitjançant l'equació 1, com:

$$v(g(X, Y)) = g(\nabla_vX, Y) + g(X, \nabla_vY).$$

Aquesta connexió s'anomena precisament connexió de Levi Civita, de qui parlem al proper paràgraf.

**5.3. Transport paral·lel.** Levi Civita introduí el 1917 [23] la noció de transport paral·lel. El transport paral·lel d'un vector al llarg d'una corba consisteix a demanar que la derivada covariant respecte al vector tangent sigui zero. La teoria d'equacions diferencials lineals ens diu que el transport paral·lel existeix i és únic, perquè per construir el transport paral·lel cal resoldre una equació diferencial lineal.

La noció de transport paral·lel serveix per derivar, car dóna la manera de transportar vectors i tensors al llarg dels espais tangents a una corba, i un cop en el mateix espai tangent té sentit parlar de derivada.



FIGURA 11. Ricci-Curbastro (1853-1925) i Levi-Civita (1873-1941).

El transport paral·lel i la derivada covariant són equivalents, i la paraula paral·lel s'ha d'entendre com constant.

Les geodèsiques poden definir-se com les corbes tals que el seu vector tangent és paral·lel, de la mateixa manera que les rectes a l'espai euclidià tenen vector tangent constant.

En general el transport paral·lel depèn de la corba, llevat del cas pla, de curvatura constant zero.

Podem veure la curvatura de Riemman com una obstrucció a la commutativitat de les derivades covariants (per tant dels transports paral·lels). Per camps de coordenades  $\partial_i$ ,  $\partial_j$ ,  $\partial_k$  i  $\partial_l$  tenim:

$$R(\partial_i, \partial_j, \partial_k, \partial_l) = g(\nabla_{\partial_i} \nabla_{\partial_j} \partial_k - \nabla_{\partial_j} \nabla_{\partial_i} \partial_k, \partial_l).$$

Per camps en general, el tensor de curvatura es defineix com

$$R(X, Y, Z, T) = g(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z, T).$$

A partir d'aquí es pot veure com una mesura de la no commutativitat de  $\nabla$ .

**5.4. Contraccions o traces.** El tensor  $R_{i\alpha\beta j}$  és força complicat de manipular i moltes vegades es pensa com un endomorfisme simètric del fibrat de 2-formes. A vegades se'n consideren contraccions, que enumerem a continuació.

- *Curvatura de Ricci.* Té el mateix ordre que la mètrica  $g_{ij}$ , és a dir és una forma bilineal simètrica:

$$R_{ij} = R_{\alpha i \beta j} g^{\alpha \beta} = -R_{i \alpha \beta j} g^{\alpha \beta}.$$

Recordem que  $g^{\alpha \beta}$  denota la matriu inversa de  $g_{ij}$  i que utilitzem la convenció d'Einstein, per tant estem fent la suma  $\sum_{\alpha, \beta}$ .

Sense coordenades és un tensor  $Ric(X, Y)$ , on  $X$  i  $Y$  són camps. És a dir,  $R_{ij} = Ric(\partial_i, \partial_j)$  i donats dos camps  $u^i \partial_i$  i  $v^j \partial_j$ ,

$$Ric(u^i \partial_i, v^j \partial_j) = u^i v^j R_{ij}.$$

Com que és un tensor simètric,  $Ric(X, Y) = Ric(Y, X)$ , podem pensar en la forma quadràtica corresponent, de manera equivalent:

$$Ric(X, X).$$

- *Curvatura escalar.* Aquesta ja és una funció, perquè és una nova contracció de la curvatura de Ricci

$$s = R_{ij} g^{ij}.$$

Aquestes contraccions poden veure's com traces: per una base ortonormal  $g^{ij} = \delta_j^i$ , i així  $s = \sum R_{ii}$  és la traça.

- *Curvatura seccional d'un pla.* Si escollim  $\{e_1, e_2\}$  una base ortonormal per un pla, aleshores la curvatura seccional del pla que generen  $\langle e_1, e_2 \rangle$  es defineix com:

$$K(\langle e_1, e_2 \rangle) = R(e_1, e_2, e_1, e_2).$$

Si la base no és ortonormal, aleshores

$$K(\langle e_1, e_2 \rangle) = \frac{R(e_1, e_2, e_1, e_2)}{g(e_1, e_1)g(e_2, e_2) - g(e_1, e_2)^2}.$$

La curvatura seccional ens permet recuperar la curvatura de Ricci. De fet si  $v$  és un vector unitari i el completem en una base ortonormal del tangent  $\{v, e_2, \dots, e_n\}$ , aleshores

$$Ric(v, v) = \sum_{i=2}^n K(\langle v, e_i \rangle).$$

La curvatura de Ricci es relaciona amb el volum. En coordenades  $(x^1, \dots, x^n)$ , la forma de volum s'escriu com

$$d \text{vol} = \sqrt{\det(g_{ij})} dx^1 \wedge \dots \wedge dx^n.$$

L'anàleg del teorema 4.3 és:

**Teorema 5.1.** *En coordenades geodèsiques:*

$$d \text{vol}(x) = \left( 1 - \frac{1}{6} R_{ij} x^i x^j + O(|x|^3) \right) dx^1 \wedge \dots \wedge dx^n.$$

**5.5. L'Equació d'Einstein.** Com que se situa en l'espai temps quadridimensional, per l'equació d'Einstein no parlem d'una mètrica de Riemann sinó d'una pseudomètrica o mètrica de Lorentz. L'única diferència és que el tensor  $g_{ij}$  no és una mètrica definida positiva, sinó que té signatura  $(3, 1)$ , és definida positiva a les direccions espaials i negativa en les temporals.

L'equació d'Einstein és la següent:

$$R_{ij} - \frac{1}{2} s g_{ij} = T_{ij}.$$

A part del tensor mètric  $g_{ij}$ , els termes que hi intervenen són la curvatura de Ricci  $R_{ij}$ , la curvatura escalar  $s$  i el tensor impusió-energia  $T_{ij}$ . Aquest darrer descriu el flux d'energia i moment en l'espai temps, generalitzant el tensor d'impulsió en física de Newton. És la font del camp gravitacional en relativitat general, i l'equació d'Einstein ens descriu com es veu afectada la geometria per aquest tensor.

Més tard Einstein escrigué sobre aquell període:

...gairebé mai en la meua vida no he treballat tant durament, i m'he imbuït de gran respecte per les matemàtiques, la part més subtil de la qual havia mirat des de la meua ingeniutat com una simple luxúria fins ara.

La teoria de la relativitat general va suposar la consolidació definitiva de la geometria de Riemann i del càlcul tensorial. Per fer el pas de la relativitat restringida a la general, Riemann va comptar amb l'ajut d'un matemàtic de Zurich, Marcel Grossmann [11].

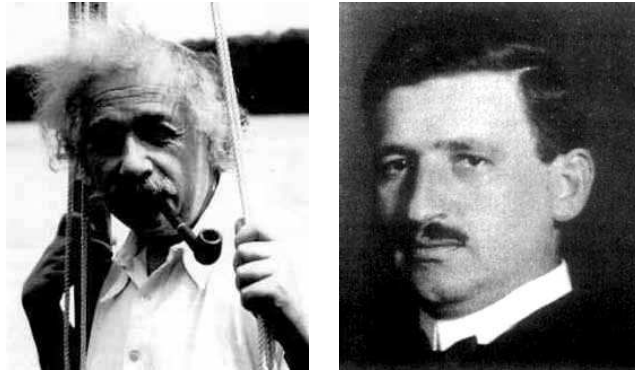


FIGURA 12. Einstein (1879-1955) i Grossmann (1878-1936).

A l'equació d'Einstein se li pot afegir un terme  $cg_{ij}$ , on  $c$  és una constant cosmològica i el valor de la qual els cosmòlegs no es posen d'acord a determinar.

## 6. CURVATURA I TOPOLOGIA

Comentem breument alguns resultats que ens relacionen la curvatura amb la topologia. És a dir, tenir la curvatura d'un signe o un altre ens dona informació sobre la topologia. Els resultats d'aquest capítol es poden trobar a molts llibres de geometria de Riemann, com per exemple [3, 6, 17, 33].

Comencem per un exemple de curvatura positiva, pel teorema de Myers, quan la curvatura de Ricci està acotada inferiorment. El tensor de Ricci  $R_{ij}$  i el de Riemann  $g_{ij}$  es poden comparar com a formes quadràtiques.

Una varietat de Riemann  $(M^n, g)$  s'anomena completa si és un espai mètric complet (les successions de Cauchy convergeixen). Equivalentment, pel teorema de Hopf-Rinow, les geodèsiques es poden allargar indefinidament.

**Teorema 6.1** (Myers). *Sigui  $(M^n, g)$  una varietat de Riemann completa. Si  $R_{ij} \geq (n-1)c g_{ij}$  per cert  $c > 0$ , aleshores  $\text{diam}(M^n) \leq \frac{\pi}{\sqrt{c}}$*

El teorema ens diu que una cota inferior positiva a la curvatura de Ricci ens dóna una cota superior del diàmetre. En particular, aplicada al recobridor universal:

**Corol·lari 6.2.** *Sigui  $(M^n, g)$  una varietat de Riemann completa. Si  $R_{ij} \geq (n-1)c g_{ij}$  per cert  $c > 0$ , aleshores  $\pi_1(M^n)$  és finit.*

La curvatura de Ricci es pot veure com una suma de curvatures seccionals, per tant el teorema de Myers s'aplica a varietats de curvatura seccional positiva.

**Corol·lari 6.3.** *Sigui  $(M^n, g)$  una varietat de Riemann completa. Si la curvatura seccional és  $\geq c > 0$  aleshores  $\text{diam}(M^n) \leq \frac{\pi}{\sqrt{c}}$  i  $\pi_1(M^n)$  és finit.*

**Teorema 6.4** (Cartan-Hadamard). *Sigui  $(M^n, g)$  una varietat de Riemann completa. Si la seva curvatura seccional és negativa o nul·la, aleshores per tot punt  $p \in M^n$ , l'aplicació exponencial  $T_p M^n \rightarrow M^n$  és un recobriments. En particular el recobridor universal de  $M^n$  és  $\mathbf{R}^n$ .*

**Corol·lari 6.5.** *Si  $(M^n, g)$  és una varietat de Riemann compacta de curvatura seccional negativa o nul·la, aleshores  $\pi_1(M^n)$  és infinit.*

Per tant, és difícilment compatible la curvatura de Ricci positiva (o curvatura seccional positiva) amb la curvatura seccional negativa o nul·la.

Cal anar amb compte perquè la condició de curvatura de Ricci negativa no és cap restricció, de fet un teorema de Lohkamp [25] afirma que en dimensió superior o igual a tres tota varietat de Riemann admet una mètrica de curvatura de Ricci negativa.

Els dos teoremes següents ens diuen que curvatura seccional estrictament negativa i curvatura zero són incompatibles.

**Teorema 6.6** (Preissman). *Sigui  $(M^n, g)$  una varietat de Riemann completa. Si la seva curvatura seccional és estrictament negativa, aleshores  $\pi_1(M^n)$  no conté cap subgrup isomorf a  $\mathbf{Z} \oplus \mathbf{Z}$ .*

**Teorema 6.7** ( Bieberbach ). *Sigui  $(M^n, g)$  una varietat de Riemann plana de dimensió  $n$ . Aleshores  $\pi_1(M^n)$  és una extensió finita del grup abelià lliure  $\mathbf{Z}^m$ , amb  $m \leq n$ . A més  $m = n$  si i només si  $M^n$  és compacta.*

Les idees de geometria de comparació s'han extés a espais mètrics (espais d'Alexandrov, espais CAT) i s'han aplicat a teoria de grups, donant lloc a l'anomenada teoria geomètrica de grups. La idea de curvatura s'aplica doncs a diversos camps de les matemàtiques, més enllà de la geometria.

## 7. CURVATURA I VARIETATS TRIDIMENSIONALS

En aquesta darera secció volem explicar un desenvolupament recent, la demostració de la conjectura de Poincaré mitjançant la geometria de Riemann, en particular el flux de Ricci, introduït per R. Hamilton el 1982. Els fluxos per curvatura tenen moltes aplicacions i estan molt extesos, però aquí només ens ocuparem del flux de la curvatura de Ricci.

Recordem que la conjectura de geometrització fou enunciada per Thurston a la dècada dels 1970 [39] i afirma el següent:

**Conjectura 7.1.** *Tota varietat tridimensional compacta es descompon de manera canònica en trossos geomètrics.*

Per tros *geomètric* entenem una varietat amb interior equipat amb una mètrica localment homogènia, és a dir, que dos punts qualsevol de l'interior de la varietat tenen entorns isomètrics. Dit d'una altra manera, les propietats locals no ens permeten distingir els punts.

La descomposició canònica de la conjectura té dues etapes. Fa referència a la descomposició en summa connexa, deguda a Kneser l'any 1929 [21], i a la descomposició en tors de obtinguda el 1979 de manera independent per Jaco-Shalen i Johansson [18, 19].

Aquesta conjectura implica la de Poincaré: tota varietat tancada de dimensió tres simplement connexa és homeomorfa a l'esfera  $S^3$ . Cal insistir en que Thurston va revolucionar la topologia de les varietats tridimensionals al fer intervenir la geometria. Sense la seva capacitat de visió és poc probable que en aquests moments la conjectura de Poincaré estigués demostrada.

Perelman conclogué la demostració de la conjectura de geometrització a partir del flux de Ricci construït per Hamilton. En particular demostrà la conjectura de Poincaré. Podem veure-ho com una altra aplicació de

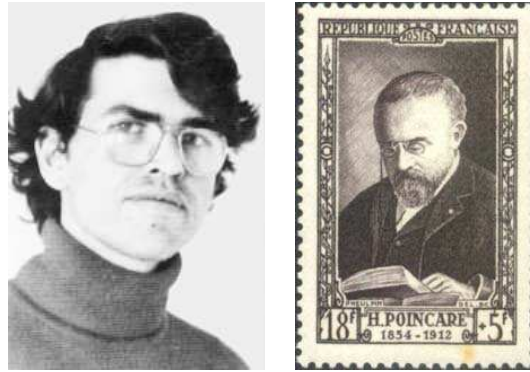


FIGURA 13. Thurston (1946) i Poincaré (1854-1912).

les idees de Riemann, que exemplifica el fet que no es poden deslligar les propietats topològiques de les varietats de les seves mètriques.

**7.1. Flux de Ricci.** Definim el flux de Ricci com la solució de l'equació d'evolució

$$\frac{\partial g}{\partial t} = -2 \text{ Ric},$$

on  $g$  denota la mètrica de Riemann i Ric el seu tensor de Ricci corresponent. Es tracta d'una equació en derivades parcials en l'espai de tensors de la varietat que són dues vegades covariants i simètrics. Aquest tipus d'equacions s'anomenen d'evolució, ja que la mètrica de la varietat anirà canviant al llarg del temps segons aquesta equació.

En coordenades  $(x^1, \dots, x^n)$ , l'equació s'escriu:

$$\frac{\partial g_{ij}}{\partial t} = -2 R_{ij}.$$

Podriem intentar escriure el terme  $R_{ij}$  en funció dels  $g_{ij}$  i les seves derivades parcials, però ens quedaria un sistema massa complicat.

**7.2. Existència.** L'existència de solucions en temps positius no és gens senzilla de demostrar. L'existència de solucions en temps negatius es falsa, com per l'equació de la calor.



**Teorema 7.2** (Existència i unicitat en temps curt). *Si  $M$  és una varietat compacta amb mètrica  $g_0$ , aleshores l'equació del flux de Ricci amb condició inicial  $g_0$  té una única solució definida en temps  $t \in [0, T)$  per cert  $T > 0$ .*

Aquest teorema fou provat per Hamilton [13], mitjançant un teorema de la funció inversa de Nash-Moser, i poc després DeTurck [10] va donar-ne una demostració molt més senzilla trencant la invariància per difeomorfismes.

Si la varietat no és compacta, l'existència i unicitat no tenen perquè ser certes. Es compleixen quan la varietat és completa i la curvatura està acotada [7], però es poden construir exemples no complets amb infinites solucions, o exemples amb curvatura no acotada sense solució.

**7.3. Exemples.** Comencem amb l'exemple més senzill possible: suposem que la mètrica inicial  $g_0$  té curvatura seccional constant  $K$ , la qual cosa vol dir que el tensor de Ricci és un múltiple constant del tensor mètric:

$$\text{Ric}_{g_0} = (n - 1)K g_0,$$

on  $n$  és la dimensió. Aleshores, sabent l'existència i unicitat de solucions, ens restringim a mètriques homotètiques

$$g_t = f(t)g_0$$

i trobem solucions d'aquest tipus. Com que el tensor de Ricci és invariant per homotècies,  $\text{Ric}_{g_t} = \text{Ric}_{g_0} = (n - 1)K g_0$ , l'equació esdevé:

$$f' = -2(n - 1)K,$$

i la solució és

$$g_t = (1 - 2K(n - 1)t) g_0.$$

Tenim tres tipus de comportament segons el signe de la curvatura seccional  $K$  inicial:

- Per  $K = 0$ , la solució és constant.
- Per  $K < 0$  la solució s'expandeix durant un temps infinit, i la curvatura s'acosta a zero.
- Per  $K > 0$  la solució es contrau fins a col·lapsar en un temps finit  $T = \frac{1}{2K(n-1)}$ , i la curvatura tendeix a infinit.

Essencialment veiem que, en el cas de curvatura constant, el flux es limita a fer homotècies, a diferents velocitats.

**7.4. Primers resultats.** A [13, 14], a més de l'existència, Hamilton provà que el flux de Ricci servia per demostrar la conjectura de geometrització quan la varietat tenia una mètrica amb curvatura de Ricci no negativa (primer quan  $Ric > 0$  i després quan  $Ric \geq 0$ ).

**Teorema 7.3** (Hamilton 1982). *Si una varietat compacta tridimensional  $M^3$  admet una mètrica amb  $Ric > 0$ , aleshores el flux de Hamilton-Ricci convergeix, després d'homotècia, cap a una mètrica de curvatura seccional positiva.*

Hamilton va desenvolupar una sèrie de principis del màxim per a tensors, que li permeteren demostrar el resultat següent.

**Teorema 7.4** (Hamilton 1984). *Si una varietat compacta tridimensional  $M^3$  admet una mètrica amb  $Ric \geq 0$ , aleshores tenim tres possibilitats:*

- (1) *La mètrica és plana (curvatura constant zero).*
- (2)  *$Ric > 0$  per  $t > 0$ , i per tant el flux convergeix a una mètrica de curvatura constant positiva, després de renormalitzar-la.*
- (3) *La mètrica és localment un producte  $g = g_1 + dx^2$ . En aquest cas la varietat és  $S^2 \times S^1$  o un quocient seu compatible amb el producte.*

El flux té molt bon comportament en dimensió tres per a varietats amb curvatura de Ricci positiva o nul·la, però ara veurem que es poden crear singularitats.

**7.5. Pinçament.** Veiem ara un exemple amb una singularitat. Apliquem el flux a l'esfera  $S^3$  amb una mètrica que té un coll. Això vol dir que és una mètrica en la qual l'equador és molt més estret que els tròpics, cf. la figura 14.

Veiem el coll com  $S^2 \times I$ , de manera que per cada valor de  $x \in I$  tenim una mètrica a l'esfera  $S^2 \times \{x\}$ . Escollim una mètrica tal que l'esfera del centre tingui un diàmetre molt més petit que les esferes de la vora del coll. Si a més escollim la longitud de l'interval molt llarga, això es pot fer de tal manera que el diàmetre de l'esfera del centre tendeixi a

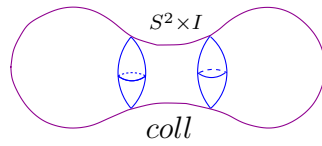


FIGURA 14.  $S^3$  amb un coll.

zero abans, ja que un diàmetre molt petit correspon a molta curvatura seccional, i com que l'interval és molt llarg, en les altres direccions gairebé no hi ha curvatura. Això s'anomena una punxada (figura 15).

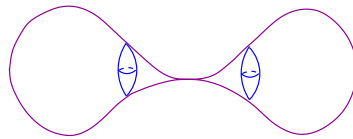


FIGURA 15.  $S^3$  punxada.

Tot i que l'exemple sigui força intuïtiu, no és gens fàcil de demostrar que es crea un pinçament [2].

**7.6. Heurística.** Per entendre perquè aquesta equació pot ser útil per la conjectura de geometrització, observem que en coordenades harmòniques (i.e.  $\Delta(x^i) = 0$ ) l'equació del flux de Ricci s'escriu:

$$\frac{\partial g_{ij}}{\partial t} = \Delta(g_{ij}) + Q_{ij}(g^{-1}, \frac{\partial g}{\partial x})$$

on  $\Delta(g_{ij})$  és el laplacià de la funció escalar  $g_{ij}$  (no el tensor) i  $Q_{ij}$  és una expressió quadràtica. Aquesta és una equació de *reacció-difusió*. El terme del Laplacià és el de *difusió*, la seva contribució a l'equació s'interpreta que tendeix a repartir la mètrica de manera uniforme. És un comportament semblant al de l'equació de la calor: en un cos on hi ha zones fredes i calentes, la temperatura flueix de les parts calentes a les fredes per tendir a uniformitzar-se. El terme quadràtic  $Q_{ij}$  és el de *reacció*, perquè s'entén que contribueix a crear singularitats, per analogia amb les equacions que descriuen certes reaccions químiques. En conseqüència podríem pensar que l'heurística del programa de Hamilton és la següent:

*“O bé  $g(t)$  convergeix cap a una mètrica localment homogènia o bé crea singularitats que corresponen a una descomposició canònica.”*

Evidentment les coses no seran tan senzilles. Hi haurà singularitats que donaran sumes connexes però potser sumes topològicament trivials (amb esferes), com de fet passa al pinçament descrit al paràgraf 7.5.

Els tors de Jaco-Shalen i Johansson no surten de singularitats, sinó que en el comportament a llarg terme del flux, hi ha parts que convergeixen en una mètrica hiperbòlica (de curvatura constant  $-1$ ) i parts que s'enfonsen (part en què el radi d'injectivitat tendeix a zero després d'una homotècia que ens normalitza la curvatura 1). En les parts que s'enfonsa, no hi ha convergència de la mètrica, i és mitjançant l'estudi topològic de les varietats enfosades que sabem que la varietat és geomètrica [1].

Cal tenir en compte que la dimensió també compta en si guanya la part de difusió o de reacció. En dimensió dos guanya la part de difusió [15], perquè el flux sempre convergeix. De fet, el flux redemuestra el teorema d'uniformització per a superfícies de Poincaré-Koebe dintre de la classe conforme, perquè el flux en superfícies no canvia la classe conforme. En dimensió quatre, guanya clarament la part de reacció, perquè no hi ha uniformització possible. Per tant és en dimensió tres quan es produeix un equilibri delicat.

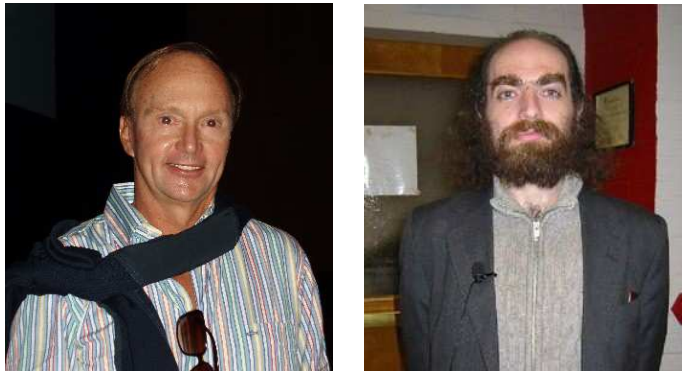


FIGURA 16. Hamilton (1943) i Perelman (1966).

Per acabar, direm molt breument que, a partir dels resultats que ja havia aconseguit, Hamilton desenvolupà un programa per demostrar la conjetura de geometrització [16], que no pogué portar a terme perquè no controlava les singularitats del flux. Fou Perelman, ben conegut per les seves habilitats tècniques en geometria de Riemann, qui aconseguí entendre les singularitats i completà el programa de Hamilton en tres preprints [30, 32, 31]. La comprovació que els preprints de Perelman donaven la demostració ha generat molta activitat, [20, 28, 5, 29, 4, 26, 9], i tot i que moltes coses no hi eren, tothom està d'acord en que la demostració final és de Perelman, i en conseqüència li donaren la medalla Fields el 2006. No la recollí, però això ja és una altra història.

#### REFERÈNCIES

- [1] L. Bessières, G. Besson, M. Boileau, S. Maillot, and J. Porti. *Weak collapsing and geometrisation of aspherical 3-manifolds*. Preprint. math.GT/0706.2065.
- [2] Sigurd Angenent and Dan Knopf. An example of neckpinching for Ricci flow on  $S^{n+1}$ . *Math. Res. Lett.*, 11(4):493–518, 2004.
- [3] Arthur L. Besse. *Einstein manifolds*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge, Bd. 10. Berlin etc.: Springer-Verlag. XII, 510 p. DM 198.00, 1987.
- [4] Gérard Besson. Preuve de la conjecture de Poincaré en déformant la métrique par la courbure de Ricci [d'après G. Perelman]. *Astérisque*, (307):Exp. No. 947, 309–347, 2006. Séminaire Bourbaki, Vol. 2004/2005.
- [5] Huai-Dong Cao and Xi-Ping Zhu. A complete proof of the Poincaré and geometrization conjectures—application of the Hamilton-Perelman theory of the Ricci flow. *Asian J. Math.*, 10(2):165–492, 2006.
- [6] Jeff Cheeger and David G. Ebin. *Comparison theorems in Riemannian geometry*. North-Holland Publishing Co., Amsterdam, 1975. North-Holland Mathematical Library, Vol. 9.
- [7] Bing-Long Chen and Xi-Ping Zhu. Uniqueness of the Ricci flow on complete noncompact manifolds. *J. Differential Geom.*, 74(1):119–154, 2006.
- [8] E. B. Christoffel. Ueber die Transformation der homogenen Differentialausdrücke 2<sup>ten</sup> Grades. *J. Reine Angew. Math.*, 70:46–70, 1869.
- [9] Tobias H. Colding and William P. Minicozzi, II. Estimates for the extinction time for the Ricci flow on certain 3-manifolds and a question of Perelman. *J. Amer. Math. Soc.*, 18(3):561–569 (electronic), 2005.
- [10] Dennis M. DeTurck. Deforming metrics in the direction of their Ricci tensors. *J. Differential Geom.*, 18(1):157–162, 1983.
- [11] A. Einstein and M. Grossmann. Entwurf einer verallgemeinerten Relativitätstheorie und einer Theorie der Gravitation. *Zs. f. Math u. Phys.*, 62:225–261, 1914.

- [12] Joan Girbau i Badó. La geometria diferencial, de Gauss a Riemann. In *El desenvolupament de les matemàtiques al segle XIX*, Arx. Sec. Cièn., LXXV, pages 41–53+photo on p. 40. Inst. Estudis Cat., Barcelona, 1984.
- [13] Richard S. Hamilton. Three-manifolds with positive Ricci curvature. *J. Differential Geom.*, 17(2):255–306, 1982.
- [14] Richard S. Hamilton. Four-manifolds with positive curvature operator. *J. Differential Geom.*, 24(2):153–179, 1986.
- [15] Richard S. Hamilton. The Ricci flow on surfaces. In *Mathematics and general relativity (Santa Cruz, CA, 1986)*, volume 71 of *Contemp. Math.*, pages 237–262. Amer. Math. Soc., Providence, RI, 1988.
- [16] Richard S. Hamilton. The formation of singularities in the Ricci flow. In *Surveys in differential geometry, Vol. II (Cambridge, MA, 1993)*, pages 7–136. Int. Press, Cambridge, MA, 1995.
- [17] Emmanuel Hebey. *Introduction à l'analyse non linéaire sur les variétés*. Fondations. Paris: Diderot Editeur. viii, 406 p. FF 195.00 , 1997.
- [18] William H. Jaco and Peter B. Shalen. Seifert fibered spaces in 3-manifolds. *Mem. Amer. Math. Soc.*, 21(220):viii+192, 1979.
- [19] Klaus Johannson. *Homotopy equivalences of 3-manifolds with boundaries*, volume 761 of *Lecture Notes in Mathematics*. Springer, Berlin, 1979.
- [20] Bruce Kleiner and John Lott. *Notes on Perelman's papers*. Preprint. math.DG/0605667.
- [21] H. Kneser. Geschlossene Flächen in dreidimensionalen Mannigfaltigkeiten. *Jahresbericht D. M. V.*, 38:248–260, 1929.
- [22] Detlef Laugwitz. *Bernhard Riemann. 1826-1866. Turning points in the perception of mathematics. (Bernhard Riemann. 1826-1866. Wendepunkte in der Auffassung der Mathematik.)*. Vita Mathematica. 10. Basel: Birkhäuser. 346 p. , 1996.
- [23] T. Levi-Civita. Nozione di parallelismo in una varietà qualunque e conseguente specificazione geometrica della curvatura *riemanniana*. *Palermo Rend.*, 42:172–205, 1917.
- [24] R. Lipschitz. Fortgesetzte Untersuchungen in Betreff der ganzen homogenen Functionen von n Differentialen. *J. Reine Angew. Math.*, 72:1–56, 1869.
- [25] Joachim Lohkamp. Metrics of negative Ricci curvature. *Ann. of Math. (2)*, 140(3):655–683, 1994.
- [26] John Milnor. Towards the Poincaré conjecture and the classification of 3-manifolds. *Notices Amer. Math. Soc.*, 50(10):1226–1233, 2003.
- [27] Mikhail Monastyrskij. *Riemann, topology, and physics. Transl. from the Russian by James King and Victoria King, ed. by R. O. Wells jun. With a foreword by Freeman J. Dyson*. Boston-Basel-Stuttgart: Birkhäuser. XIII, 158 p.; DM 82.00 , 1987.
- [28] John Morgan and Gang Tian. *Ricci flow and the Poincaré conjecture*, volume 3 of *Clay Mathematics Monographs*. American Mathematical Society, Providence, RI, 2007.

- [29] John W. Morgan. Recent progress on the Poincaré conjecture and the classification of 3-manifolds. *Bull. Amer. Math. Soc. (N.S.)*, 42(1):57–78 (electronic), 2005.
- [30] Grisha Perelman. *The entropy formula for the Ricci flow and its geometric applications*. Preprint. math.DG/0211159.
- [31] Grisha Perelman. *Finite extinction time for the solutions to the Ricci flow on certain three-manifolds*. Preprint. math.DG/0307245.
- [32] Grisha Perelman. *Ricci flow with surgery on three-manifolds*. Preprint. math.DG/0303109.
- [33] Peter Petersen. *Riemannian geometry*, volume 171 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.
- [34] H. Poincaré. *La science et l'hypothèse*. Paris: Flammarion. 284 S. 12<sup>mo</sup>, 1902.
- [35] A. Reventós and J.C. Rodríguez. Gauss i la geometria. *Conferències FME*, III:155–214, 2006.
- [36] G. Ricci. Sulla derivazione covariante ad una forma quadratica differenziale. *Rom. Acc. L. Rend. (4)*, III:15–18, 1887.
- [37] Bernhard Riemann. *Gesammelte mathematische Werke und wissenschaftlicher Nachlass. Herausgegeben unter Mitwirkung von Richard Dedekind von Heinrich Weber. Zweite Auflage bearbeitet von Heinrich Weber*. Leipzig. B. G. Teubner. X u. 558 S. gr. 8°, 1892.
- [38] Michael Spivak. *A comprehensive introduction to differential geometry. Vol. II. 2nd ed.* Berkeley, California: Publish or Perish, Inc. XIII, 423 p., 1979.
- [39] William P. Thurston. Three-dimensional manifolds, Kleinian groups and hyperbolic geometry. *Bull. Amer. Math. Soc. (N.S.)*, 6(3):357–381, 1982.





**RIEMANN AND PARTIAL DIFFERENTIAL  
EQUATIONS.  
A ROAD TO GEOMETRY AND PHYSICS**

JUAN LUIS VÁZQUEZ

1. INTRODUCTION

This is the edited text of a talk given at the Facultat de Matemàtiques i Estadística of UPC on February 20th, 2008, in the framework of the “Jornada Riemann”.

In preparing the text I have endeavored to avoid or minimize repetitions of material that appears in other contributions. I have also tried to keep the text as simple as possible, in accordance with the style of the talk and the open and lively discussions held during the event.

Before entering the subject proper of this work, let me summarize a few reflections on Riemann’s life and work that I find relevant for a deeper understanding of the significance of his legacy.



Left: the scientist as a young man; Right: a classical picture.

Bernhard Riemann died very young, like other geniuses of the 19th Century, but he left an impressive legacy to Mathematics, pure and applied. He had a permanent love for Italy, where he traveled for health reasons, but where he also had friends that continued his deep ideas. Besides Mathematics, Riemann had a sustained interest in Philosophy and in Physics.

An easy source for his mathematics is the collection of his works, [12], which has translations to several languages. His life is described in Detlef Laugwitz's book [7], and I have used M. Monastyrsky's [10]. I have found very useful Ferreirós' "Riemanniana Selecta" [5], in Spanish. Besides that, Internet is a very convenient source of details, and I have consulted Wikipedia, MacTutor, Encyclopaedia Britannica, and other Internet archives.

Riemann had a philosophical formation, and as such he sustains that *the essence of Reality lies in a hidden world*. That is not really new; medieval scholars would say that *Videmus in aenigmata, et per speculum*.

In his century Riemann was not alone in seeing the key to understanding the hidden reality of the World in the *Concepts and Formulas of Abstract Mathematics*, or in the original German, in *die Begriffe und Formeln der Höhere Mathematik*.

So maybe we can see here a new look for old ideas. Indeed, the look is not only new, it will prove to be revolutionary.

In keeping with his philosophical frame of mind, and also because of his short life, his work is deep in concepts and ideas, but rather scarce in details. A century and a half of research has provided answers and details to a large number of the topics treated in his complete works, but not to all: remember the Riemann hypothesis!

## 2. MATHEMATICS, PHYSICS AND PDES

**2.1. PDEs and the origins of differential calculus.** The Differential World, i.e, the world of derivatives, was invented, or discovered as you may prefer to see it, in the 17th Century, almost at the same time that Modern Science (then called *Natural Philosophy*), was born. We

owe it to the great Founding Fathers: Galileo, Descartes, Leibnitz and Newton (mainly). Motivation came from the desire to understand the World around us, more specifically Motion, Mechanics and Geometry.

Newton formulated Mechanics in terms of ODEs, by concentrating on the movement of particles. The main magic formula is

$$m \frac{d^2 \mathbf{x}}{dt^2} = F(t, \mathbf{x}, \frac{d\mathbf{x}}{dt})$$

though he would write derivatives with dots, and not as quotients, which is Leibnitz style. Here are the magical words, to which we are now so used: mass ( $m$ ), force at a distance ( $F$ ), and acceleration, and here is where the (second) derivative enters the picture.

Newton thought about fluids, in fact he invented *Newtonian fluids*, and there you need dependence on space and time simultaneously,  $x$  as well as  $t$ . It involves the partial derivatives, which means that we get Partial Differential Equations (PDEs for short). But his progress was really small if you compare it with the rest of *Principia Mathematica Philosophiae Naturalis*, 1687, and other works of the early Calculus time.

*We conclude that there was not much time for PDEs from the Big Bang to 1700 AD.*

In the 18th Century, PDEs appear in the work of Jean Le Rond D'Alembert about string oscillations: there a *set of particles* moves together due to *elastic forces*, but every one of the infinitely many *solid elements* has a different motion,  $u = u(x, t)$ . This is one of the first instances of *continuous collective dynamics*. *PDEs are the mode of expression of such CCD.*

Johann and Daniel Bernoulli and then Leonhard Euler lay the foundations of Ideal Fluid Mechanics (1730 to 1750), in Basel and St Petersburg. This is PDEs of the highest caliber:

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p = 0, \quad \nabla \cdot \mathbf{u} = 0.$$

The system is *nonlinear*; it does not fit into one of the three main types that we know today (elliptic, parabolic, hyperbolic); the main pure-mathematics problem is still unsolved (existence of classical solutions for good data; Clay Problems, year 2000).

**2.2. Modern times: PDEs in the 19th Century.** The 19th Century confronts revolutions in the concept of heat and energy, electricity and magnetism, and also in the very concept of space. The Newtonian edifice begins to shake. You may add a lesser revolution, *real fluids*.

Mathematically, all of these fields take up the form of PDEs:

(i) Heat leads to the heat equation,  $u_t = \Delta u$ , and the merit goes primarily to J. Fourier.

(ii) Electricity leads to the Coulomb equation in the Laplace-Poisson form:  $-\Delta V = \rho$ . Surprisingly, this equation also represents gravitation! (with a difference, for electricity the right-hand side may have two signs).

(iii) Electromagnetic fields are represented by the Maxwell system. The vector potential satisfies a wave equation, the same as D'Alembert's (but now it is vector-valued and in several dimensions).

(iv) Real fluids are represented by the Navier-Stokes equations. Sound waves follow wave equations, but they can create discontinuous solutions called *shocks* (and here Riemann appears as we will see).

**2.3. PDEs continued in the 19th Century.** Geometry was transformed from the Euclid tradition plus Cartesian Algebra to the spirit of PDEs by G. Gauss and B. Riemann. The new spirit is condensed in a number of key words. Space is determined by its *metric* which is a local object that has tensor structure. The connection from point to point is a new object called *covariant derivative*, the *curvature* is a second order operator, a nonlinear relative of the Laplace operator.

After the work of these people, in particular Riemann, *Reality* is seen as mainly continuous, and its essence lies in the *physical law*, that is a

law about a *field* or a number of fields. In symbols, we have  $\Phi(x, y, z, t)$  and an equation (or system)

$$L\Phi = \mathbf{F},$$

where  $F$  is the force field (a tensor).

**2.4. 20th Century. Summing Up.** In the 20th Century General Relativity and Quantum Mechanics take this same form. *Space, matter and interactions become fields.*

A main variant from the scheme is Statistical Mechanics, a thread that leads to Brownian motion (Einstein, Smoluchowski), abstract probability (Kolmogorov, Levy, Wiener), stochastic calculus and stochastic differential equations (Itô).

Summing up in a rather succinct form: *the main (technical) task of the Mathematician working in Mathematical Physics is to understand the world of Partial Differential Equations, linear and nonlinear.*

The same is true nowadays for geometers (you only need to travel to the Universitat Autònoma of Barcelona and attend the now running CRM semester on Ricci flows!).

The main abstract tool is Functional Analysis. The *combination* of Functional Analysis, PDES and ODEs, Geometry, Physics and Stochastic Calculus is one of the Great Machines of today's research, a child of the 20th Century.

### 3. RIEMANN, COMPLEX VARIABLES AND 2-D FLUIDS

In the sequel we will try to convey some of the mathematics of B. Riemann that had a profound impact on PDEs, with attention to specific concepts and calculations. In other words, let us do some math!

**3.1. Complex Variables (Euler, Cauchy, Gauss, Riemann, Weierstrass).** We start with a function

$$u(x, y) = u(z), \quad z = x + iy = (x, y)$$

that is supposed to be a good function of two real variables.

- A good function of two real variables means (could mean)  $u \in C^1(\Omega)$  for some subdomain  $\Omega$  of  $\mathbb{R}^2$ .
- Therefore,  $u$  has a gradient:  $\nabla u(z_0) = (u_x, u_y)$ .
- But, what is a good function of one complex variable?
- First of all, to keep the symmetry, there must be two real functions of two real variables:

$$u = u(x, y), \quad v = v(x, y)$$

which we write as  $f = f(z)$  with  $f = u + iv$ , and  $z = x + iy$ .

- The question is: Do we ask that  $f \in C^1$  and that is all? The answer is no and this is a consequence of algebra.
- Let us explain why: very nice real functions of one variable are polynomials, and very nice complex functions of one complex variable should also be polynomials.

- Now, polynomials are easy to define, for instance  $f(z) = z^2$  means

$$u = x^2 - y^2, \quad v = 2xy$$

while  $f(z) = z^3$  means

$$u = x^3 - 3xy^2, \quad v = 3x^2y - y^3.$$

- Can the reader do  $f(z) = z^n$  by heart? Euler could! In fact, Euler and Moivre could see the whole trigonometry (by using the polar form  $z = re^{i\theta}$  and expanding the power  $z^n$ ).
- Can you see something special in these pairs of functions,  $u$  and  $v$ ? Cauchy and Riemann could! They saw the whole theory of *complex holomorphic functions*.

**3.1.1. The PDE code for complex variables.** What they saw is this hidden symmetry:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

These equations are called Cauchy-Riemann (CR) equations for complex variables. They are one of the most important examples of a PDE system with extraordinary geometric and analytic consequences.

- We see that  $\nabla u = (a, -b)$  is orthogonal to  $\nabla v = (b, a)$ . Consequence: the level lines  $u = c_1$  and  $v = c_2$  are orthogonal sets of curves.
- The linear algebra of infinitesimal calculus at every point is not 4-dimensional but two-dimensional. In fact, the system

$$du \sim adx - bdy, \quad dv \sim bdx + ady,$$

can be written together in the complex form  $df \sim Jf(z) dz$ , where  $Jf$  is the Jacobian matrix that we begin to call  $f'(z) = a + ib$ .

- *We will assume from such glorious moment on that this is the correct derivative of a 2-function of 2 variables that is a candidate to be a good complex differentiable function.*

**3.1.2. Complex Variables, analysis and geometry.** Hence we know some magic formulas:

$$f'(z) = a + bi = f_x, \quad f_y = -b + ai = if'(z).$$

Thus  $df = f'dx$  along the  $x$ -axis and  $df = if'dy$  along the  $y$  axis.

Comming back to the Jacobian,

$$Jf(z) = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = aE + bJ.$$

This is a similarity matrix with determinant

$$\det(Jf) = a^2 + b^2 = u_x^2 + u_y^2 = u_x^2 + v_x^2 = v_x^2 + v_y^2 = u_y^2 + v_y^2,$$

which can be written as

$$|Jf| = |f'(z)|^2 = \|f_x\|^2 = \|f_y\|^2 = \|\nabla u\|^2 = \|\nabla v\|^2.$$

The infinitesimal transformation preserves the angles (of tangent curves) and scales the size by  $Jf = |f'(z)|^2$ .

If the 2-2 function  $f$  is CR, then it defines a conformal transformation of the part of the plane where  $f'(z) \neq 0$ . Riemann's geometric theory of one complex variable is based on this idea!

**3.1.3. Complex Variables and PDEs. Solving the equations. Potential theory.** Once we enter that framework, the question is: how to find pairs of functions satisfying CR?

Of course, real and imaginary parts of algebraic complex functions satisfy CR. The Taylor Series, typical of the Cauchy–Weierstrass approach, also satisfies CR.

But PDE people want their way. Here is the wonderful trick:

$$\Delta u = u_{xx} + u_{yy} = v_{yx} + (-v_x)_y = 0.$$

Idem  $\Delta v = 0$ . Solutions of this equation are *harmonic functions*, and they count among the most beautiful  $C^\infty$  functions in analysis and among the most important in physics, where solving  $\Delta u = -\rho$  means finding the *potential* of  $\rho$ , where  $\rho$  is a volume distribution of mass or of electric charge.

Note a novelty full of promise: harmonic functions live in all dimensions, not just two. But in  $d = 2$  they produce complex holomorphic functions. Given  $u$ , we find  $v$ , its conjugate partner, by integration of the differential form

$$dv = Pdx + Qdy, \quad \text{with } P = v_x = -u_y, Q = v_y = u_x.$$

This is an exact differential thanks to CR.

**3.2. Complex Variables and ideal fluids in  $d = 2$ .** Recall that Riemann was a friend of Weber, the famous physicist.

The velocity of a 2-D fluid is a field  $\mathbf{v} = (v_1(x, y), v_2(x, y))$ . Irrotational means that  $\nabla \times \mathbf{v} = 0$ . Incompressible means  $\nabla \cdot \mathbf{v} = 0$ . For the PDE person this is easy:

$$v_{2,x} - v_{1,y} = 0, \quad v_{1,x} + v_{2,y} = 0.$$

Does this look like what we saw in the previous subsection? Yes, combining both we get  $\Delta v_1 = \Delta v_2 = 0$ .

Is  $v_2$  harmonic conjugate to  $v_1$ ? No, but  $-v_2$  is.

*Idea to eliminate sign problems.* Go to the scalar potential of the vector field  $\mathbf{v}$ :

$$d\Phi = v_1 dx + v_2 dy$$

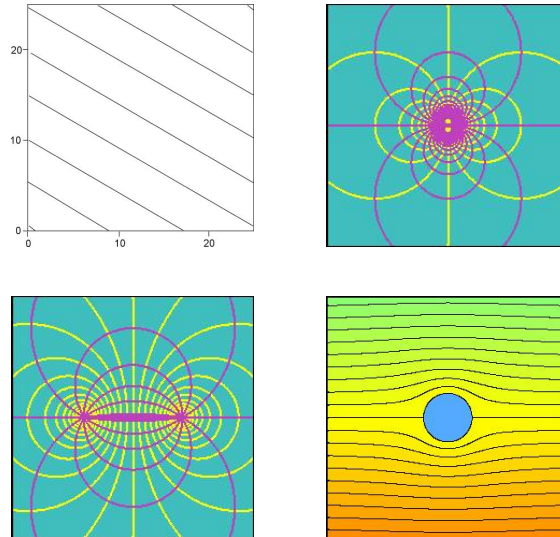


(it is exact by irrotationality). Take the harmonic conjugate  $\Psi$  and define the complex potential of the flow as  $F = \Phi + i\Psi$ , a complex holomorphic function. In that case

$$F'(z) = F_x = \Phi_x + i\Psi_x = \Phi_x - i\Phi_y = \bar{\mathbf{v}}.$$

Consequence:  $\bar{\mathbf{v}} = v_1 - iv_2$  is a complex holomorphic function.

**3.3. Some Pictures of 2D glory.** The following pictures come in every book about two-dimensional perfect fluids and conformal transformations. We ask the reader to identify them as a linear flow, a dipole configuration, a source-sink combination, and lastly the famous streamlines for the planar flow around an obstacle. This complex variable theory stands at the core of the science of aerodynamics.



#### 3.4. Summary. The big picture in 2D.

- There is an equivalence between holomorphic complex variable theory  $\Leftrightarrow$  conformal geometry  $\Leftrightarrow$  harmonic functions  $\Leftrightarrow$  ideal fluids.
- Any two dimensional ideal fluid generates an analytic function, and conversely, and it is a conformal mapping, and viceversa.
- The complex derivative of the complex potential is just the conjugate of the velocity field.

- The stream function  $\Psi$  indicates the lines of current via the formula  $\Psi = c$ .
- What happens when  $F'(z) = 0$ , i.e., when  $\mathbf{v} = \mathbf{0}$ ? These are singular points, called in physics the *stagnation points*. Many things can happen on a singularity, but essentially only one thing may happen on a regular point, where the implicit function theorem is valid. Riemann was an expert in singular points.

#### 4. RIEMANN AND THE EQUATIONS OF GEOMETRY

4.1. **From 2D to 3D.** Riemann was able to understand very well the Two-Dimensional Space with its functions, analysis, geometry and physics.

It is not as easy as it seems because complex holomorphic functions *try to follow their name and be globally defined*, actually they have analytic continuation. But they may have singularities blocking their way to global (global is called here *entire*).

Riemann's main contribution to 2D analysis+geometry is the concept of *Riemann surface* (RS) with the curious branching points. A simple Riemann surface may be a part of  $\mathbb{R}^3$  but more complicated RS live in a very strange situation, *a different world*.

But we want now to forget 2D and remember that we live in 3D. Thinking about the geometry of 3D is an old pastime, masterfully encoded by Euclid of Alexandria (325 BC-265 BC).

The 3D world is much more complicated than 2D and no part of the equivalence between analysis, Taylor series, elementary PDEs, conformal geometry and ideal Physics survives.

4.1.1. **What is Geometry according to Riemann.** Let us follow the Encyclopædia Britannica article on B. Riemann.

In 1854 Riemann presented his ideas on geometry for the official post-doctoral qualification at Göttingen; the elderly Gauss was an examiner and was greatly impressed.

Riemann argued that the fundamental ingredients for geometry are a *space of points* (called today a manifold) and a *way of measuring distances* along curves in the space.

He argued that the space *need not be ordinary Euclidean space* and that it could have *any dimension* (he even contemplated spaces of infinite dimension). Nor is it necessary that the surface be drawn in its entirety in three-dimensional space. According to Riemann, *many spaces are possible*. This happened more than four decades before Relativity!

It seems that Riemann was led to these ideas partly by his dislike of the (Newton's) concept of *action at a distance* in contemporary physics and by his wish to endow space with the ability to transmit forces such as electromagnetism and gravitation.

A few years later this inspired the Italian mathematician Eugenio Beltrami to produce just such a description of non-Euclidean geometry, the first physically plausible alternative to Euclidean geometry. More italians influenced by B. Riemann: Ricci, Levi-Civita, Bianchi.

**4.1.2. Habilitationsvortrag, 1854. Riemannian Geometry.** This is one of the most famous and influential habilitation documents in the history of Mathematics.

Space around only has a definite sense locally around the place. The basic tool to do geometry is the *metric*, which is given by

$$ds^2 = \sum g_{ij} dx^i dx^j$$

It is local since it works on local entities, tangent vectors. Forget Pithagoras but remember  $ds^2 = d\theta^2 + \sin^2 \theta d\phi^2$  on the sphere.

The metric field changes from point to point,  $g_{ij}(x)$ ,  $x$  is locally a set like  $\mathbb{R}^d$ . He even says that for a space of functions  $d$  can be infinite.

So there is no sense in principle of parallel vectors (at least if we do not work more). We can instead define the derivative of a tangent vector

$X = \sum_i a_i \mathbf{e}_i$  when we move along another vector  $Y = \sum_j b_j \mathbf{e}_j$ . This is the famous covariant derivative  $\nabla$ :

$$\nabla_Y X = \sum_i Y(a_i) \mathbf{e}_i + \sum_{ijk} a_i b_j \Gamma_{ij}^k \mathbf{e}_k.$$

which depends on a set of functions  $\Gamma_{ij}^k$ , the so-called Christoffel symbols. For the correct covariant derivative, called Levi-Civita connection, the Christoffel symbols are given by

$$\Gamma_{ij}^k = \frac{1}{2} g^{kl} \left( \frac{\partial g_{il}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^l} \right).$$

Objects with several indices are usually tensors. Note that although the Christoffel symbols have three indices on them, they are not tensors. Sorry, local coordinates are intuitive but messy to work with, this really cumbersome aspect of modern geometry is also part Riemann's inheritance!!

**4.1.3. Curvatures at the center of geometry.** The covariant derivative opens the way to a whole new Differential Calculus, where curvature tensors and Laplacians play a key role.

**Curvature tensor.** The Riemann curvature tensor is given by

$$R(X, Y, Z) = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X, Y]} Z.$$

In that case we have

$$R_{jkl}^i = \frac{\partial \Gamma_{jk}^i}{\partial x^l} - \frac{\partial \Gamma_{jl}^i}{\partial x^k} + \Gamma_{jk}^s \Gamma_{sl}^i - \Gamma_{jl}^s \Gamma_{sk}^i.$$

Contraction gives the low tensor  $R_{ijkl} = g_{im} R_{jkl}^m$ . Wikipedia gives

$$R_{iklm} = \frac{1}{2} (\partial_{kl}^2 g_{im} + \partial_{im}^2 g_{kl} - \partial_{km}^2 g_{il} - \partial_{il}^2 g_{km}) + g_{np} (\Gamma_{kl}^n \Gamma_{im}^p + \Gamma_{km}^n \Gamma_{il}^p).$$

**Ricci curvature.** The Ricci curvature of  $g$  is a contraction of the general curvature tensor:

$$R_{ij} = \sum_j R_{ijs}^s = \sum_{s,m} g^{sm} R_{isjm}$$

The Ricci tensor has the same type  $(0, 2)$  (twice covariant) of the metric tensor. In coordinates we have (Nirenberg's sign, Wikipedia)

$$R_{ij} = \frac{\partial \Gamma_{ij}^l}{\partial x^l} - \frac{\partial \Gamma_{il}^j}{\partial x^j} + \Gamma_{ij}^l \Gamma_{lm}^m - \Gamma_{il}^m \Gamma_{jm}^l.$$

4.1.4. **The Laplacian operator in such geometries.** Here is the definition of the geometer's Laplacian (Laplace–Beltrami operator):

$$\Delta_g(u) = -g^{ij}(\partial_{ij}u - \Gamma_{ij}^k \partial_k u) = -\frac{1}{|g|^{1/2}} \partial_i(|g|^{1/2} g^{ij} \partial_j u).$$

This is minus the contraction of the second covariant derivative tensor

$$(\nabla^2 u)_{ij} = \partial_{ij}u - \Gamma_{ij}^k \partial_k u.$$

A coordinate chart  $(x^k)$  is called **harmonic chart** if and only if  $\Delta_g x^k = 0$  for all  $i$ . Note that

$$\Delta_g(x^k) = -g^{ij} \Gamma_{ij}^k.$$

Therefore,  $(x^i)$  is harmonic iff  $g^{ij} \Gamma_{ij}^k = 0$  for all  $k$ .

The Laplacian is convenient for doing analysis and PDEs on manifolds because the basic integration by parts formula

$$\int_M (\Delta_g u) v \, d\mu + \int_M \langle \nabla_g u, \nabla_g v \rangle \, d\mu = 0$$

makes sense if you use the correct definitions.

4.1.5. **Yamabe problem. Ricci flow.** Like in the rest of disciplines of Mathematics, the combination of differential geometry and PDEs has given rise to classical problems that have focused the interest of generations. We will comment here on two problems that have attracted attention in the last decades.

• **Yamabe Problem.** Let  $g$  be a metric in the conformal class of a metric  $g_0$ . Let  $D$  denote the Levi-Civita connection of  $g$ . We denote by  $R = R_g$  and  $R_0$  the scalar curvatures of the metrics  $g, g_0$ , respectively. Write  $\Delta_0$  for the Laplacian operator of  $g_0$ . Then we can write

$$g = u^{4/(n-2)} g_0$$

locally on  $M$  for some positive smooth function  $u$ . Moreover, we have the formula

$$R = -u^{-N}Lu \quad \text{on } M,$$

with  $N = (n + 2)/(n - 2)$  and

$$Lu = \kappa\Delta_0u - R_0u, \quad \kappa = \frac{4(n - 1)}{n + 2}.$$

Note that  $\Delta_0 - \frac{n-2}{4(n-1)}R_0$  is the conformal Laplacian relative to the background metric. Write equivalently,  $R_gu^N = R_0u - \kappa\Delta_0u$ .

The standard Yamabe Problem can be stated thus: *given  $g_0$ ,  $R_0$  and  $R_g$ , find  $u$* . This is a nonlinear elliptic equation for  $u$ .

An evolution version of the Yamabe problem leads to the so-called Fast Diffusion Equation,  $u_t = \Delta u^m$  with exponent  $m = (n - 2)/(n + 2) < 1$ . The problem is described for instance in [17].

• **Ricci flow.** The Ricci curvature features prominently in R. Hamilton's program, 1982, to classify three dimensional manifolds by continuous deformation of the original metric. This is a remarkable idea to try to solve by PDE methods the old *Poincaré conjecture* (one of the 7 problems of the Clay list). The proposed flow is

$$(4.1) \quad \partial_t g_{ij} = -R_{ij}.$$

In view of the expression of  $R_{ij}$  in terms of  $g_{ij}$  and its partial derivatives, this turns out to be a system of nonlinear partial differential equations for the evolution of the metric tensor  $g_{ij}$ . It is formally of parabolic type, so Hamilton established for it maximum principles and Harnack inequalities. But the program faced difficulties related to the blow-up of solutions in finite time

In 2002-03 G. Perelman posted three papers with a complete solution of that problem, and this seems to be one of the main mathematical events of the running century. In this way Riemann's legacy is more present than ever before for the mathematical research community.

More details on this topic can be found in J. Porti's contribution in this volume.

**4.2. General relativity. Einstein's equation.** Riemann's ideas went further and they turned out to provide the mathematical foundation for the four-dimensional geometry of space-time in Einstein's theory of general relativity.

The Einstein tensor  $\mathbf{G}$  is a 2-tensor on pseudo-Riemannian manifolds which is defined in index-free notation as

$$(4.2) \quad \mathbf{G} = \mathbf{R} - \frac{1}{2}R\mathbf{g},$$

where we use the following notations:  $\mathbf{R}$  is the Ricci tensor,  $\mathbf{g}$  is the metric tensor and  $R$  is the Ricci scalar (or scalar curvature). In components, the above equation reads

$$G_{ij} = R_{ij} - \frac{1}{2}Rg_{ij},$$

and Einstein's field equations (EFE's) are:

$$G_{ij} = \frac{8\pi G}{c^4}T_{ij},$$

a system of second order partial differential equations in 4 variables. The quantities  $T_{ij}$  are the components of the stress-energy tensor, so-called because it describes the flow of energy and momentum.

## 5. RIEMANN AND THE PDES OF PHYSICS

**5.1. Riemann's interest in Physics.** The influence of the famous experimental physicist W. Weber was important in Riemann's view of mathematics. Apart from his contributions to the mechanics of air waves, he was keenly interested in the contemporary developments, and in teaching. Let us mention here the famous book [15]: Riemann's lectures on the partial differential equations of mathematical physics and their application to heat conduction, elasticity, and hydrodynamics were published after his death by his former student, Hattendorff. Three editions appeared, the last in 1882; and few books have proved so useful to the student of theoretical physics. The object of Riemann's lectures was twofold: first, to formulate the differential equations which are based on the results of physical experiments or hypotheses; second, to integrate these equations and explain their limitations and their application to special cases.

**5.2. Paper “Über die Fortpflanzung...”, 1860. The equations of gas dynamics.** Let us go back to creative science. Though Riemann’s fame is usually associated among mathematicians with pure mathematics (Riemann’s Hypotheses, Riemann surfaces, Riemannian geometry), his contribution to applied science is fundamental in the area of aerodynamics.

Let us present his contribution in a brief form. One-dimensional isentropic gas flow is a mathematical abstraction described by the system of differential equations

$$(5.1) \quad \begin{cases} u_t + u u_x + p_x/\rho = 0, \\ \rho_t + (\rho u)_x = 0 \end{cases}$$

plus the algebraic equation  $p = p(\rho)$ .

In applications  $x$  is interpreted as length along a tube, whose transversal dimensions are supposed to be irrelevant,  $u$  is interpreted as *fluid particle speed* and  $\rho$  as *density*. The equation  $p = p(\rho)$  is called *equation of state* and for ideal gases it takes the form  $p = C\rho^\gamma$  where  $\gamma = 1, 4$ . Evidence on the determination of this  $\gamma$  really worried Riemann as he says at the beginning of his paper (he was not an absent-minded theoretician!).

In a more modern style, we may write the equations in a compact way

$$(5.2) \quad \mathbf{U}_t + A(\mathbf{U})\mathbf{U}_x = \mathbf{0},$$

which encodes the system:

$$(5.3) \quad \begin{pmatrix} u_t \\ \rho_t \end{pmatrix} + \begin{pmatrix} u & p'(\rho)/\rho \\ \rho & u \end{pmatrix} \begin{pmatrix} u_x \\ \rho_x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

**5.2.1. Hyperbolic systems.** In order to continue we do linear algebra, calculating the eigenvectors and eigenvalues of the matrix  $A$ . We obtain

$$\lambda_1 = u + c, \quad \lambda_2 = u - c$$

where  $c^2 = p'(\rho)$  ( $c$  is called the speed of sound). Note that  $\lambda = \lambda(u, \rho)$ , hence it changes with  $(x, t)$  depending on the flow you are solving at this time.



If  $\rho \neq 0$  then  $c \neq 0$  and we have two different eigenvalues and we are entering with Riemann into the theory of Nonlinear Hyperbolic Differential Systems, still frightening today. Peter Lax, Courant Institute, Abel Prize winner, is a world leader in the topic. See his monograph [8].

We now get a map from  $(x, t)$  into  $(u, \rho)$ , with two nice directions for the linearization of the evolution equation,

$$\mathbf{U}_t + A(\mathbf{U}_0)\mathbf{U}_x = \mathbf{0}.$$

If you are Riemann, or you are able to follow his train of thought, this allows you to construct some magical local coordinates where the flow is not complicated. Correct coordinates are Riemann's specialty.

**5.2.2. Riemann invariants.** The eigenvectors of the system are

$$\mathbf{U}_1 = (c/\rho, 1), \quad \mathbf{U}_2 = (c/\rho, -1)$$

Now Riemann tells us to find the characteristic lines: if we think that the solution is known, then solve the ODE Systems

$$\frac{dx}{dt} = \lambda(x, t, \mathbf{U})$$

He tells you then to find functions  $F_1, F_2$ , called the *Riemann invariants*, which are independent and constant along the corresponding characteristics. In the gas example they are

$$F_{\pm} = u \pm \int \frac{c(\rho)}{\rho} d\rho.$$

Since these functions are constant on the characteristics, they allow to see what the characteristics do and this says what the flow does at any moment. Replace  $(u, \rho)$  by  $F_1, F_2$  and try to see something. The reader can follow the detailed development in Chorin-Marsden [3] or Smoller [16].

**5.2.3. Shocks.** The theory Riemann develops allows to solve the system in a classical way if and only if the characteristics of the same type for different points do not cross. In that case the invariant takes two values, a shock appears.

Shocks appear in the examples even in  $d = 1$ , which is the Burger's equation

$$u_t + u u_x = 0.$$

Since it happens in the simplest nontrivial mathematics, Riemann concludes that you cannot avoid shock formation, and that a theory of solutions with discontinuities that propagate in some magical way is needed. This is today the theory of shocks and discontinuous solutions of conservation laws.

Very soon the physical community recognized this work as a fundamental new insight into the complexity inherent to compressible fluids.

Rankine and Hugoniot completed the work of Riemann when the gas is not isentropic and the system is three dimensional. They even found that Riemann made an error in that general case! (cf. [6], [11]).

**Aftermath.** The story of how *discontinuous functions can be correct solutions of a partial differential equation of mathematical physics*, and even more, *how important is what happens at the point where classical analysis breaks down*, is one of the deepest and most beautiful aspects of PDEs in the 20th century. The catch word is *entropy solutions*, a theory that counts famous names in the last decades like P. Lax, O. Oleinik, S. Kruzhkov, J. Glimm, and continues for instance with the recent work by A. Bressan.

*Before Riemann nobody really dared to face those problems, after him all of us must!*

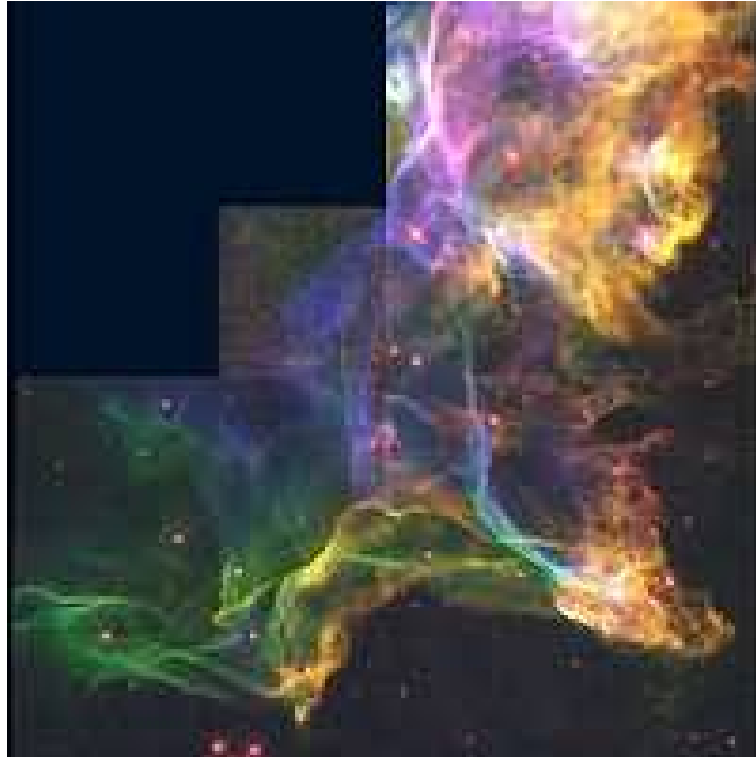
Follow the whole “shocking story” in references like [8], [3], [16], and more recently, [1], [2]. Connections of shocks and general relativity are described in [4].

## 6. PICTURE GALLERY. SOME SHOCK WAVES IN NATURE



**Schlieren Image – Convection Currents and Shock Waves**, *Steve Butcher, Alex Crouse, and Loren Winters – August, 2001.*

The projectiles were 0.222 calibre bullets fired with a muzzle velocity of 1000 m/s (Mach 3). The Schlieren lighting technique used for these images makes density gradients in fluids visible. Color filtration provides false color images in which the colors provide information about density changes.



This Hubble telescope image shows a small portion of a nebula called the "Cygnus Loop."

This nebula is an expanding blast wave from a stellar cataclysm, a supernova explosion, which occurred about 15,000 years ago. The supernova blast wave, which is moving from left to right across the picture, has recently hit a cloud of denser-than-average interstellar gas. This collision drives shock waves into the cloud that heats interstellar gas, causing it to glow.

Sandia Releases New Version of Shock Wave Physics Program; it can be found in:

<http://composite.about.com/library/PR/2001/blsandia1.htm>

As a conclusion of this work, as practitioners of PDEs interested in understanding how the real world works, and above all as mathematicians, we would like to say

**Danke schön, Herr Riemann!**

#### REFERENCES

- [1] A. BRESSAN, "Hyperbolic Systems of Conservation Laws: The One-Dimensional Cauchy Problem". Oxford University Press, 2000.
- [2] A. BRESSAN, D. SERRE, M. WILLIAMS, K. ZUMBRUN, "Hyperbolic Systems of Balance Laws": Lectures given at the C.I.M.E. Summer School held in Cetraro, Italy, July 14-21, 2003 (Lecture Notes in Mathematics / Fondazione C.I.M.E., Firenze). Springer, 2007.
- [3] A. CHORIN, J. MARSDEN, "A Mathematical Introduction to Fluid Mechanics". Springer, 1990.
- [4] J. GROAH, J. SMOLLER, B. TEMPLE, "Shock Wave Interactions in General Relativity". Springer Science + Business Media, LLC, New York, NY USA, 2007.
- [5] JOSÉ FERREIRÓS, "Riemann, Bernhard. Riemanniana selecta" (Spanish). Edited and with an introductory study by José Ferreirós. Clásicos del Pensamiento. CSIC. Consejo Superior de Investigaciones Científicas, Madrid, 2000.
- [6] H. HUGONOT, *Propagation des Mouvements dans les Corps et spécialement dans les Gaz Parfaits*, Journal de l'École Polytechnique, **57**, (1887), p. 3; **58** (1889), p. 1.
- [7] D. LAUGWITZ, "Bernhard Riemann (1826-1866): Turning Points in the Conception of Mathematics". Birkhäuser (1999).
- [8] P. D. LAX, "Hyperbolic partial differential equations". With an appendix by Cathleen S. Morawetz. Courant Lecture Notes in Mathematics, 14. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2006.
- [9] T. LEVI-CIVITA, *Nozione di parallelismo in una varietà qualunque e conseguente specificazione geometrica della curvatura riemanniana*, Rend. Cir. Mat. Palermo, 42 (1917), 172–205.
- [10] M. MONASTYRSKY, "Riemann, Topology, and Physics". Birkhäuser, 2nd ed., 1999.

- [11] W. J. M. RANKINE, *On the thermodynamic theory of waves of finite longitudinal disturbances*, Phil. Trans. Roy. Soc. London, 160 (1870), p. 277.
- [12] B. RIEMANN, “Gesammelte mathematische Werke, wissenschaftlicher Nachlass und Nachträge” [Collected mathematical works, scientific Nachlass and addenda] Based on the edition by Heinrich Weber and Richard Dedekind. Edited and with a preface by Raghavan Narasimhan. Springer-Verlag, Berlin; BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1990.
- [13] B. RIEMANN, *Ueber die Fortpflanzung ebener Luftwellen von endlicher Schwingungsweite* [*Propagation of plane air waves of finite oscillation amplitude*], Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, 8 (1860), 43-65. Classic papers in shock compression science, 109–128, High-pressure Shock Compression Condensed Matter, Springer, New York, 1998.
- [14] B. RIEMANN, *Ueber die Hypothesen, welche der Geometrie zu Grunde liegen* [*On the Hypotheses which lie at the Bases of Geometry*], Habilitationsschrift (1854), Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, 13 (1868).
- [15] B. RIEMANN, “Die Partiellen Differentialgleichungen der Mathematischen Physik” [Partial Differential Equations Of Mathematical Physics], Herausgeg. von K. Hattendorf. Braunschweig: F. Vieweg, 1938. XII, 325.
- [16] J. SMOLLER, “Shock Waves and Reaction-Diffusion Equations”, Springer, 2nd ed., 1994.
- [17] J. L. VÁZQUEZ, “Smoothing and decay estimates for nonlinear diffusion equations”, vol. 33 of Oxford Lecture Notes in Maths. and its Applications. Oxford Univ. Press, 2006.

# RIEMANN & PHYSICS

EMILI ELIZALDE

ABSTRACT. Riemann's interest in physics is clear from his legacy, which is discussed in detail in all the contributions to this volume. Here, after providing a very concise review of the main publications of Bernhard Riemann on physical problems, we will turn to consider some rather less known (but not less interesting) connections between Riemann's papers and contemporary Physics. More specifically, we will address, among other aspects: (a) the influence of Riemann's work on the zeta function, its functional formula, and related extensions of those concepts, to the regularization of quantum field theories in curved space-time (in particular, that of the quantum vacuum fluctuations); and (b) the uses of the Riemann tensor in general relativity and in very recent generalizations of this celebrated theory, which aim at understanding the presently observed acceleration of the universe expansion (the dark energy issue). We shall argue that the importance of the influence in Physics of Riemann's *purely mathematical* works exceeds by far that of his papers which were directly devoted to *physical* issues.

## 1. INTRODUCTION

The presentation at the meeting, on which this paper on Riemann's work is based, took place at the very end of a long day of dense and interesting discussions. It is in this framework that the contents which follow have to be pondered. The author is somehow afraid he tried to present in the talk rather deep concepts in a light, almost casual way. This was however maybe not too bad, since it gave rise to a number of clever questions from the audience. They can be found, hopefully, in the recordings of the lecture, to which the reader is addressed for additional information, when needed. We will try here to avoid repeating concepts and arguments already contained in the other contributions

to this volume. For all these reasons—and also for lack of space—the present article will not be self-contained, but the reader will be addressed to the relevant references at the appropriate places. Even then, some repetition will be unavoidable, but hopefully the viewpoints will be different and maybe enriching.

It is quite clear that Riemann was definitely interested in physics. This may sound to more than one a weird statement nowadays, when he is considered to be a pure mathematician, who gave name to so many concepts in different fields of mathematics, as the Riemann integral, the Riemann surface, the Cauchy-Riemann equations, the Riemann-Roch theorem, the Riemannian manifolds, the Riemann curvature tensor and, most notably the Riemann zeta function, with its associated conjecture—the only one of Hilbert's problems that after the turn of the XXth Century has entered the new list of *Million Dollar Problems* of the XXIst (awarded by the Clay Foundation). However, historians of science assure that during his life and till as late as 60 to 80 years after his death, Riemann was counted among the list of important physicists, whose ideas on the unification of all known forms of energy preceded the ground-breaking work of Hilbert and Einstein (see later, and also the other contributions in this volume). Even more surprising is to learn that Riemann was not only a theoretical physicist, but also an experimental one, and that he made use of physical proofs with charged surfaces in order to establish supplementary checks of the validity of some mathematical theorems (as boundary problems involving partial differential equations).

Let me here just recall that, as a student at Göttingen university, Riemann worked with Weber on electromagnetism, which happened starting around 1849. Like Riemann, Weber was also a student of Gauss, but at that time Weber had already a faculty position. He had proposed a theory of electromagnetism which gained him a name in history, as every physicist knows, although not through his theory in fact, that was eventually superseded by Maxwell's one, the real landmark in classical electromagnetism. Gauss himself is also famous for his important work on this subject.

Riemann publications include some fifteen papers, four of which were published after his death. Needless to say, this does not include a



number of important notes, letters, books and other writings that also form part of his written scientific production. In the first section of the present paper a brief summary will be provided (the reader is again addressed to the other contributions in this issue for more detailed discussions) of the six papers (among the mentioned fifteen) which are devoted to physical problems. Then, sort of a panoramic view will be presented of the enormous influence of Riemann's work on pure mathematics to past and present Physics. In the last part of the paper I will concentrate more specifically on a couple of issues of my own speciality, namely, on the one hand, the use of zeta functions as a very elegant regularization tool in quantum field theory, including a brief description of its uses for the calculation of quantum vacuum fluctuations, the Casimir effect, and the related cosmological constant problem. The other issue to be addressed is the very well known applications of the Riemann curvature tensor and all his geometrical formalism in general relativity and the, much less known but very important nowadays, proposed modifications of the Einstein-Hilbert Lagrangian with additional terms—a function of the curvature scalar, the so-called  $f(R)$  theories. Only ca. hundred years after the formulation of general relativity, on response to the demand of the observed acceleration of the universe expansion (the crucial dark energy issue), have some attempts at a modification of Einstein's equations started to appear. But again, notably, in terms of its basic Riemannian building blocks, as we shall later see.

## 2. SIX RIEMANN PIECES ON PHYSICAL PROBLEMS

The starting reference list of works by Riemann, previously mentioned, consists of *The Mathematical Papers of Georg Friedrich Bernhard Riemann (1826-1866)*, a collection which contains scientific papers of Bernhard Riemann as transcribed and edited by David R. Wilkins [1]. These texts are based on the second edition of the *Gesammelte Mathematische Werke* and, in the case of some of the papers, the original printed text in the *Journal für die reine und angewandte Mathematik*, *Annalen der Physik und Chemie* and *Annali di Matematica*. Included in Ref. [1] are all papers published in Riemann's lifetime, papers and correspondence published after Riemann's death by Dirichlet and others prior

to the publication of the first edition of the *Gesammelte Mathematische Werke* (with the exception of the fragment *Mechanik des Ohres*, which is non-mathematical in character), and one of the papers from his *Nachlass*, first published in the *Gesammelte Mathematische Werke*. There is also a translation by W. K. Clifford of Riemann's inaugural lecture on the foundations of geometry, and a biographical sketch by Richard Dedekind that was included in the *Gesammelte Mathematische Werke*.

However, I will not go here through *all* these works. I will restrict my attention to a subset which, although not complete as viewed by a historian of science, I think it is fair enough in order to establish my point that Riemann's physical production was actually a good part of his complete scientific work. I will reduced the whole sample in Ref. [1] to that of the published papers—during Riemann's lifetime and posthumously—and limit my study to the papers on physical issues among them. The list of these published articles is as follows.

## 2.1 Papers published in Riemann's lifetime

- (1) *Grundlagen für eine allgemeine Theorie der Functionen einer veränderlichen complexen Grösse*, Inauguraldissertation, Göttingen (1851).
- (2) *Ueber die Gesetze der Vertheilung von Spannungselectricität in ponderabeln Körpern, wenn diese nicht als vollkommene Leiter oder Nichtleiter, sondern als dem Enthalten von Spannungselectricität mit endlicher Kraft widerstrebend betrachtet werden*, Amtlicher Bericht über die 31. Versammlung deutscher Naturforscher und Aerzte zu Göttingen (im September 1854).
- (3) *Zur Theorie der Nobili'schen Farbenringe*, *Annalen der Physik und Chemie*, **95** (1855) 130-139.
- (4) *Beiträge zur Theorie der durch die Gauss'sche Reihe  $F(\alpha, \beta, \gamma, x)$  darstellbaren Functionen*, *Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen*, **7** (1857) 3-32.
- (5) *Selbstanzeige: Beiträge zur Theorie der durch die Gauss'sche Reihe darstellbaren Functionen*, *Göttinger Nachrichten* (1857) 6-8.

- (6) *Theorie der Abel'schen Functionen*, Journal für die reine und angewandte Mathematik, **54** (1857) 101-155.
- (7) *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse*, Monatsberichte der Berliner Akademie (November, 1859) 671-680.
- (8) *Ueber die Fortpflanzung ebener Luftwellen von endlicher Schwingungsweite*, Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, **8** (1860) 43-65.
- (9) *Selbstanzeige: Ueber die Fortpflanzung ebener Luftwellen von endlicher Schwingungsweite*, Göttinger Nachrichten (1859) 192-197.
- (10) *Ein Beitrag zu den Untersuchungen über die Bewegung eines flüssigen gleichartigen Ellipsoides*, Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, **9** (1860) 3-36.
- (11) *Ueber das Verschwinden der Theta-Functionen*, Journal für die reine und angewandte Mathematik, **65** (1866) 161-172.

## 2.2 Posthumously published papers of Riemann

- (12) *Ueber die Darstellbarkeit einer Function durch eine trigonometrische Reihe*, Habilitationsschrift, 1854, Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, **13** (1868).
- (13) *Ueber die Hypothesen, welche der Geometrie zu Grunde liegen*, Habilitationsschrift (1854), Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, **13** (1868).
- (14) *Ein Beitrag zur Elektrodynamik* (1858), Annalen der Physik und Chemie, **131** (1867) 237-243.
- (15) *Ueber die Fläche vom kleinsten Inhalt bei gegebener Begrenzung*, Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, **13** (1868).

Six among these fifteen papers (namely, those with numbers 2, 3, 8, 9, 10, 14) are the ones that I have selected because they directly address issues of theoretical and experimental physics. I now provide a free translation of their titles, together with a short summary of each of

them (indeed very brief, since they are also described in some detail in the other contributions to this volume).

### 2.3 Riemann papers on Physics

2. *About the distribution laws of electric tension in ponderable bodies, when these cannot be considered as absolutely conductors or non-conductors, but as opposing with a finite force to the electric tension they contain*, Official Report at the 31st Meeting of German Scientists and Physicians at Göttingen (September, 1854).

Riemann considers in this paper Leyden jars, where an electric charge is kept, and studies in particular how, once the bottle has been emptied, a certain amount of charge still remains, which gradually disappears with time. He studies in detail the corresponding law describing this phenomenon. Riemann deals, in particular, as the title clearly indicates, with bodies that are neither perfect conductors nor perfect isolators. He elaborates on previous work by Ohm, Weber, Kirchhoff and Kohlrausch. An important point in the whole development is the contact with the corresponding experimental results. The mathematical basis of the paper are partial differential equations, as is also the case in the ones to follow.

3. *On the theory of noble color rings*, Annals of Physics and Chemistry, **95** (1855) 130-139.

Here an experimental study of the propagation and of the distribution of an electrical current in a conductor is presented. The rings mentioned in the title are generated when one covers a plate of a noble metal, as platinum, gold plated silver, or similar, with a solution of lead oxide. Then, an electric current, produced by a battery, is connected to the plate. In this way, the so-called Newton color rings are produced. Riemann elaborates here on previous results by Becquerel, Du-Bois-Reymond and Beetz, improving their calculation and discussing about the hypothesis previously considered by these authors.

8. *About the propagation of plane airwaves of finite oscillation amplitude*, Sessions of the Royal Science Society at Göttingen, **8** (1860) 43-65.

Riemann integrates in this paper the differential equations corresponding to the movement of gases, under different conditions of pressure and temperature. He remarks that he can bring his calculations further away in the order of approximation, with respect to those previously carried out by Helmholtz, for instance, who only got to the second order in the perturbative expansion. He refers to previous results by Helmholtz, Regnault, Joule and Thomson, improving their calculations, discussing the set up and improving the hypothesis in the works by these authors. With 22 pages, this is quite a long paper as compared with other papers of Riemann.

9. *Self-announcement: About the propagation of plane airwaves of finite oscillation amplitude*, Göttingen Notices (1859) 192-197.

This is a very short compendium of the main mathematical formalism that is used in the former paper, of the same title, in order to obtain the results. In spite of its title, this one could be considered as a mathematical article. Indeed, it deals with the theory of propagation of a gas, but the only physical input in the whole paper is the mathematical equation giving the behavior of gas pressure as a function of the density (that is, its equation of state), in the absence of any heat exchange. He develops the mathematical formalism in detail and compares with previous results by other mathematical physicists as Challis, Airy, Stokes, Petzval, Doppler and von Ettinghausen (most of them have given names to quite famous equations).

10. *A contribution to the investigation of the movement of a uniform fluid ellipsoid*, Sessions of the Royal Science Society at Göttingen, **9** (1860) 3-36.

Again as clearly indicated in the title, Riemann deals here with the movement of a uniform fluid ellipsoid, which is considered to be constituted by isolated points that attract themselves under the influence of gravity. This is considered one of the finest papers by Riemann within the class of those considered here, i.e. the ones dealing with actual physical problems. In the paper, the equilibrium configurations of the ellipsoid are identified, what has many and important applications, e.g., to the study of the possible forms of celestial bodies as galaxies or clusters. Riemann studies in particular the evolution of the

principal axis of the ellipsoid and the relative movement of its components. As the one before, this is also a rather mathematical paper, since the only physics it contains is practically reduced to the initial conditions and Newton's law. Previous results of Dirichlet and Dedekind on this problem are extended.

14. *A contribution to Electrodynamics*, Annals of Physics and Chemistry, **131** (1867) 237-243.

This paper is generally considered to incorporate the main results of Riemann's physical (and also philosophical) ideas on the 'unification' of gravity, electricity, magnetism, and heat. It contains indeed his observation on how a theory of electricity and magnetism is closely related with those for the propagation of light and heat radiation. He presents in the paper a complete mathematical theory, with "*an action that does not differentiate*" the already mentioned four cases of "*gravity, electricity, magnetism, and temperature*". The finite velocity of propagation of the interaction (as opposed to the predominant concept, at the epoch, of action at a distance) is clearly presented, identifying such velocity with that of light, which has been considered by many to be a really remarkable achievement of Riemann's genius. The paper, which with only six pages is in fact quite short, relies on experimental results by Weber and Kohlrausch, Busch, and by Bradley and Fizeau.

## 2.4 Some additional considerations

- (1) Once more, those above are not all the works on physical issues Riemann wrote, but just the ones extracted from a uniform sample, namely his published articles.
- (2) A good example of a work not in the list is the well-known book by H. Weber and B. Riemann, *Die partiellen Differential-Gleichungen der mathematischen Physik nach Riemanns Vorlesungen*, 6. unveränderte Aufl., 2 vols (Vieweg, Braunschweig, 1919), that was used for many years as a textbook in different universities, together with several other papers by Riemann.
- (3) An interesting biography of Riemann is the book by Monastyrsky [2]. A lot of emphasis is made there on the importance of the

contribution to physics of paper 14 of the list above. In particular, it is underlined how Riemann was searching for “... *a completely self-contained mathematical theory ... , which leads from the elementary laws up to the actions in an actually given filled space, without making a difference between gravity, electricity, magnetism or the equilibrium of temperature.*”

- (4) In the celebrated biography of David Hilbert by Constance Reid [3] we can read that Hilbert sustained the opinion (referring to what is nowadays known as the Einstein-Hilbert action) that “... *the invariance of the action integral unifies electromagnetism with gravity ...*”, yielding in this way a solution to a problem that, as he recognizes, “... *was already posed by Riemann: the connection between gravitation and light.*” Hilbert goes on to observe that, since then, many investigators had tried to arrive at a deeper understanding of this connection by merging the gravitational and electromagnetic potentials into a unity. The one example Hilbert mentions explicitly is Weyl’s unification of the two fields in a “*unified world metric,*” as he calls it, by means of Weyl’s notion of gauge invariance.
- (5) Remarkably enough, in what is probably one the most exhaustive biographies of Riemann ever written, Laugwitz [4] forgets almost completely about Riemann’s work on physical issues. This is, in my view, to push to an extreme the opinion that I maintain here, which is much more moderate and doublefaced.

### 3. INFLUENCE IN PHYSICS OF RIEMANN’S PURELY MATHEMATICAL PAPERS

It is the opinion of the author, shared also by others (see, e.g., Ref. [4]), that the influence in Physics of Riemann’s purely mathematical papers exceeds by far, in its manifest importance, the above mentioned contributions on actual physical problems; even if the interest of the last attains already, as we have pointed out in the preceding section, a fair high level.

I would need more space and time than I have at disposal in order to describe all such intertwining influences. In the following, to start, a rather short list of items will provide some basic ideas about those

influences. Then we will elaborate on some of them in more detail, not only because of their importance, but also because of the fact that they have to do with my own scientific expertise and published record — mostly joint works with a number of different colleagues— in the last few years.

**3.1. On the concept of space.** One reason why the discovery of non-Euclidean geometry took so long might have been the fact that there was universal belief that Euclidean geometry was special because it described the space we live in. Stemming from this uncritical acceptance of the view that the geometry of space is Euclidean was the conviction that there was *no other* geometry. Philosophers like Emmanuel Kant argued that the Euclidean nature of space was a fact of nature, and the weight of their authority was very powerful. From our perspective, we know of course that the question of the geometry of space is in fact entirely different from the question of the existence of geometries which are non-Euclidean. Gauss was the first who clearly understood the difference between these two issues. In Gauss' *Nachlass* one can find his computations of the sums of angles of each of the triangles that occurred in his triangulation of the Hanover region. His conclusion was that the sum was always two right angles, within the limits of observational errors.

Nevertheless, quite early in his scientific career Gauss became convinced of the possibility of constructing non-Euclidean geometries, and in fact came up with the '*theory of parallels*,' but because of the fact that the general belief in Euclidean geometry was deeply ingrained, Gauss decided not to publish his researches in the 'theory of parallels' and the construction of non-Euclidean geometries for fear that there would arise criticisms of such investigations by people who did not understand those things (*'the outcry of the Boeotians'*).

Riemann took this entire circle of ideas to a higher, completely different level. In his famous inaugural lecture of 1854, written under the advice (or, better, compulsory choice) of Gauss himself, he touched upon all of the aspects that his thesis advisor had considered. He pointed out, to start, the very crucial idea that a space does not have any structure except that it is a continuum in which points are specified by the values of  $n$  coordinates,  $n$  being the *dimension* of the space. On such a space



one can then impose many geometrical structures. His great insight was that a geometry should be built from these infinitesimal parts. He treated in depth geometries where the distance between pairs of infinitely near points is pythagorean, formulated also central questions about such geometries, and discovered the set of functions—the sectional curvatures—whose vanishing characterized the geometries which are Euclidean, namely those whose distance function is pythagorean not only for infinitely near points, but even for points which are a finite but small distance apart.

If the space is the one we live in, he formulated the principle that its geometrical structure could only be determined empirically. In fact he stated explicitly that the question of the geometry of physical space does not make sense independently of physical phenomena, i.e., that space has no geometrical structure until we take into account the physical properties of matter in it, and that this structure can be determined by measurement only. Indeed, he went so far as to say that “*the physical matter determines the geometrical structure of space*”. This groundbreaking idea took definite form some half a century later with Einstein equations.

Indeed, it is also important to remark that Riemann’s ideas constituted a profound departure from the perceptions that had prevailed until his time. No less an authority than Newton had asserted that space by itself is an “*absolute entity endowed with Euclidean geometric structure*”, and had built his entire theory of motion and celestial gravitation on that premise. Riemann went completely *away* from this point of view. Thus, for Riemann, space *derived its properties from the matter that occupied it*, and he asserted that the only question that could be studied was *whether* the physics of the world made its geometry Euclidean. It followed from this idea that only a mixture of geometry and physics could be tested against experience. For instance, measurements of the distance between remote points clearly depended on the assumption that a light ray would travel along shortest paths. This merging of geometry and physics, which is a central and dominating theme of modern physics, since Einstein’s work, may be thus traced back to Riemann’s inaugural lecture [5].

**3.2. Linear algebra, the concept of  $n$ -dimensional space (linear, or trivial ‘variety’).** It has been often reported (and this seems indeed to be the case) that linear algebra was a ‘trivial matter’ for Riemann. However, in Laugwitz’s book [4] (p. 242) we can read that the early developments of Riemannian geometry were ‘prolix and opaque’ because ‘the development of linear algebra failed for a long time to keep pace with the progress of analysis.’ This may be true, in fact: although nowadays  $n$ -dimensional linear spaces and their algebraic properties are considered to be one of the simplest theories in Mathematics, and its uses in classical and quantum physics are so basic and widespread (including infinite dimensional spaces, topological spaces, Banach and Hilbert spaces, etc.), that even the most basic issues of modern physics would not be possible without such concepts. One cannot simply translate this view to Riemann’s time. But it was already clear that these abstract linear spaces had nothing to do with the space we live in, and were not even called ‘spaces’ by Riemann or Gauss, but rather ‘varieties’ or ‘manifolds’ again.

**3.3. Riemann’s integral.** Riemann may have arrived at his notion of an ‘integral’ in answer to the question of whether the Fourier coefficients,  $c_n$ , of a given function tend to 0 (as  $n$  goes to infinity). Yet Laugwitz [4] characterizes Riemann’s introduction of his integral as *ad hoc* and remarks that “History would have been different if he had asked himself the question: what kind of integral implies the equality

$$\lim \int_a^b f_n = \int_a^b f,$$

where  $f_n$  is a monotonically increasing sequence of integrable functions that converge pointwise to the limit  $f$ ?”

It is of course true that Lebesgue’s integral is the ultimate extension of Riemann’s one, a most fundamental tool in measure theory. But it is not less certain that, concerning the subject at discussion here, Riemann’s integral is much more intuitive for a physicist. This I know well since I have been teaching both kind of integrals to physicists during many years. The Riemann integral is such a physically fashionable object, in particular the incremental version before the limit is taken, before an ‘increment’ is transformed into a ‘differential’, which is a far more elusive concept indeed! (I know well from my students). The

corresponding upper and lower finite sums<sup>1</sup>, ... no realization of an integral could be more suited to the physicist's mind.

**3.4. Complex Variables, Cauchy-Riemann equations.** Riemann's Thesis studied the theory of complex variables and, in particular, what we now call Riemann surfaces. It therefore introduced topological methods into complex function theory. The work elaborates on Cauchy's foundations of the theory of complex variables built up over many years and also on Puiseux's ideas of branch points. However, Riemann's Thesis is a strikingly original piece of work which examined geometric properties of analytic functions, conformal mappings and the connectivity of surfaces. Riemann's work was always based on intuitive reasoning which would fall at instances a little below the rigor required to make the conclusions watertight. However, the brilliant ideas which his works contain are so much clearer because his papers are not overly filled with lengthy computations [6], and this is why they were so frequently used in lecture courses (in special in Italy) afterwards. Again, one recognizes the physicist's approach in many of his discussions, but more important than this is the enormous use that both classical and quantum physics has made of the complex calculus that Riemann (among others) contributed to expand and popularize in a very efficient way.

**3.5. Riemann surface, sphere, manifold.** In principle, those are very abstract concepts, but which have been applied, e.g., by engineers to the study of aerodynamics and hydrodynamics. At a different level, theoretical physicists have more recently drawn upon them very heavily in their formulations of string theory.

String theory is the modern version of a Theory of Everything (TOE). It suggests replacing pointlike particles with infinitesimal vibrating strings as the basic units of the physical world. Some ten to fifteen years ago, when string theory was overwhelmingly dominating the landscape in theoretical physics<sup>2</sup>, there have even been jokes about the typical

---

<sup>1</sup>With references to the determination of areas of real fields in ancient Egypt and Mesopotamia.

<sup>2</sup>There is a fashionable string theory *landscape* right now which contains an enormous amount (maybe  $10^{500}$ ) of possible vacuum solutions of the corresponding theory. Choosing one among them seems hopeless, for the moment, and it is one of the main problems of M theories (M stands for 'Mother', or 'Mysterious').

theoretical physicist always carrying Farkas and Kra's book (*Riemann Surfaces*) under his arm, everywhere from place to place [7]. Edward Witten, from the Institute for Advanced Study at Princeton, has been, and continues to be, one of the main architects of string theory. He has given talks from time to time on Riemann's work, when discussing some of the relations between physics and mathematics in the 20th and 21st centuries, to which the reader is addressed for material that complements a lot, from a different, much more ambitious perspective, what I will discuss below.

**3.6. Analytic continuation, complex power series.** Most of Riemann's predecessors concentrated on a power series expansion rather than on the function that it represents. By shifting emphasis to the latter, Riemann could eliminate superfluous information, determining a complex function from its singularities. Riemann's work used simple concepts in place of the lengthy and sometimes obscure computations typical of his predecessors and contemporaries. The steady decrease in the amount of attention Riemann seems to have paid to power series between 1856 and 1861 indicates how Riemann's thought matured, shifting further away from computation. Even when using his great computational abilities, Riemann still focused upon concepts rather than the computation itself. Since relations obtained from series expansions of functions retain their validity outside their regions of convergence, he asked himself *what actually continues functions from region to region?* For example, Riemann constructed a function that has simple zeros at  $z = 0, 1, 2, \dots$  and is finite for all finite  $z$  (see Laugwitz [4]). The road to his function  $g(z)$  was heuristic, but this was of no consequence to Riemann. All he wanted was to find some function with the prescribed zeros. By contrast, Weierstrass always aimed to obtain formula representations of given functions. The Riemann approach to this issue is one of the main starting points in a big part of the present author's work, as it will be commented later in more detail.

**3.7. Curvature tensor. Differential Geometry.** In his general theory of relativity, Einstein used Riemann's concept of curved space as the basis for his elegant explanation of gravitation. Massive objects put a dimple in space in their vicinity. So when other physical objects, including photons, which do not have any mass, wander into the object's

vicinity, they encounter this curved space. Such curvature determines the path the objects follow, in a way that was formerly attributed to the force of attraction that we call gravity.

In much the same way that Riemann conceived of curving and twisting space in innovative ways, he also described a set of abstract surfaces that were created by cutting and pasting together normal surfaces in ways that cannot be employed with real surfaces, but can be thought abstractly. You can do a lot of mathematics on those abstract surfaces. So this has been an amazingly important idea for many parts of mathematics, and now for physics.

General relativity, quantum field theory in curved spaces, string theories, gravitation, modern cosmology, would had been impossible without those basic concepts introduced by Riemann.

**3.8. The Riemann zeta function.** This is known to be of extreme importance in analytic number theory. But also, through its analytical continuation (the so called functional equation or reflection formula of the zeta function), and extrapolating the concept of the zeta function to the domain of pseudodifferential operators (the spectral values of the operator replace the natural numbers in the zeta function definition), as a regularization tool in quantum field theory (notably in curved space-time), for dynamical systems (classical and quantum), the concept of chaos (also present nowadays in the issue of the distribution of non-trivial zeros, or Riemann conjecture), etc. The interconnections between pure mathematics and physical uses here is becoming more on more profound as decades advance.

#### 4. SELECTED HOT SUBJECTS: CONCEPT OF SPACE, ZETA REGULARIZATION, MODIFIED GRAVITY THEORIES

##### 4.1. The concept of space.

4.1.1. *Historical evolution of the concept of space.* A summary of the evolution of the concept of space, from the very remote times of its inception, could be as follows.

- (1) The introduction of the concept of space seems to go back to the pre-Socratic philosophers, who already had coined this notion, together with some other very important ones as those of substance, number, power, infinity, movement, being, atom, and of course time, among others.
- (2) The Pythagorean school should be mentioned as another important step, in its attempt at bringing all these concepts, in particular the one of space to the domain of numbers (*“all things are numbers”*). Just recall the importance of Pythagoras theorem, that has so much to do with space and with Gauss’ search to check if the space we live in was or not Euclidean.
- (3) Euclid’s *“Elements”*, this goes without saying as one of the most important pieces of work in the History of Mankind. It was so influential, for generations, that departing from the concept of Euclidean space was absolutely impossible for many centuries to come.
- (4) Indeed, still for Isaac Newton *“space is, by itself, an absolute entity embedded with a Euclidean geometrical structure”*.
- (5) On the side of the philosophers, for Immanuel Kant *“that space is Euclidean is a property of nature itself”*.
- (6) Now Bernhard Riemann came to clearly say, as we have advanced before, the following: *“many spaces are possible; it is the physical matter that determines the geometrical structure of space”*.
- (7) And Albert Einstein gave a precise mathematical formulation of this concept, with the important help of Grossmann and making use of Riemann’s manifolds and tensors: *space-time is curved by matter*, as prescribed by Einstein’s equations (in terms of the Riemann curvature tensor).
- (8) Finally, an embracing reflection by Eugene Wigner, which can be extended to the whole development of the concept of space, is that of *“the unreasonable effectiveness of Mathematics in the Natural Sciences”*.

4.1.2. *On the topology and curvature of our universe.* Let us now connect, briefly, these philosophical ideas about space with recent precise determinations of the topology and geometrical curvature of the universe we live in—what can be considered as the modern version of

the pioneering attempts by Gauss, already mentioned, to determine its possible curvature.

The Friedmann-Robertson-Walker (FRW) model, which can be obtained as the *only* family of solutions to the Einstein's equations compatible with the assumptions of *homogeneity* and *isotropy* of space, is the generally accepted model of the cosmos nowadays. But, as the reader surely knows, the FRW is a family with a free parameter,  $k$ , the curvature, that can be either positive, negative or zero (the flat or Euclidean case). This curvature, or equivalently the curvature radius,  $R$ , is not fixed by the theory and should be matched with cosmological observations. Moreover, the FRW model, and Einstein's equations themselves, can only provide local properties, not global ones, so they cannot tell about the overall topology of our world: is it closed or open? is it finite or infinite? Those questions are very appealing to any human being. All this discussion will only concern three dimensional space curvature and topology, time not being for the moment involved.

Serious attempts to measure the possible curvature of the space we live in go back to Gauss, who measured the sum of the three angles of a big triangle with vertices on the peaks of three far away mountains (Brocken, Inselberg, and Hohenhagen). He was looking for evidence that the geometry of space is non-Euclidean. The idea was brilliant, but condemned to failure: one needs a much bigger triangle to try to find the possible non-zero curvature of space. Now cosmologists have recently measured the curvature radius  $R$  by using the largest triangle available, namely one with us at one vertex and with the other two on the hot opaque surface of the ionized hydrogen that delimits our visible universe and emits the cosmic microwave background radiation (CMB, some  $3$  to  $4 \times 10^5$  years after the Big Bang) [8]. The CMB maps exhibit hot and cold spots. It can be shown that the characteristic spot angular size corresponds to the first peak of the temperature power spectrum, which is reached for an angular size of  $.5^\circ$  (approximately the one subtended by the Moon) if space is flat. If it has a positive curvature, spots should be larger (with a corresponding displacement of the position of the peak), and correspondingly smaller for negative curvature.

The joint analysis of the considerable amount of data obtained during the last years by balloon experiments (BOOMERanG, MAXIMA, DASI), combined also with galaxy clustering data, have produced a lower bound for  $|R| > 20h^{-1}\text{Gpc}$ , that is, twice as large as the radius of the observable universe, of about  $R_U \simeq 9h^{-1}\text{Gpc}$ .

General Relativity does not prescribe the topology of the universe, or its being finite or not. The universe could perfectly be flat and finite. The simplest non-trivial model from the theoretical viewpoint is the toroidal topology (that of a tyre or a donut, but in one dimension more). Traces for the toroidal topology (and more elaborated ones, as negatively curved but compact spaces) have been profusely investigated, and some circles in the sky with near identical temperature patterns were identified [9]. And yet more papers appear, from time to time, proposing a new topology [10]. However, to summarize all these efforts and the observational situation, and once the numerical data are interpreted without bias (what sometimes was not the case, and led to erroneous conclusions), it seems at present that available data still point towards a very large (we may call it *infinite*) flat space.

**4.2. On zeta-function regularization and its uses in quantum field theory.** The fact that the infinite series

$$s = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots + \frac{1}{2^n} + \cdots$$

has the sum  $s = 1$  is nowadays clear to any school student. It was not so, even to well learned persons, for many centuries, as we can recall from *Zeno of Elea's paradox* (or Zeno's paradox of the tortoise and Achilles), transmitted by Aristotle and based on the pretended impossibility to do an *infinite* number of summations (or recurrent 'jumps' or steps of any kind, in a *finite* amount of time). In fact there are still modern versions of the Zeno paradox (e.g. the *quantum Zeno paradox*) which pop up now and then [11, 12].

In a modern version of this paradox, extrapolated to its more far reaching consequences, Krauss and Dent affirm, in a recent paper appeared in *The New Scientist* that, incredible as it may seem, our detection of dark energy could have reduced the life-expectancy of the universe (!). This idea is a worrying new variant of the *quantum Zeno paradox*,



as these cosmologists claim that astronomers may have accidentally nudged the universe closer to its death by observing dark energy, the anti-gravity force which is thought to be accelerating the expansion of the cosmos. These allegations suggest that by making this observation in 1998 we may have caused the cosmos to revert to an earlier state when it was more likely to end. Krauss and Dent came to such astonishing conclusion by calculating how the energy state of our universe—a summation of all its particles and all their energies—has evolved since the big bang of creation some 13.7 billion years ago.

The *quantum Zeno effect* is a well known phenomenon in quantum physics, with sufficiently precise experimental proofs. It says that, whenever we observe or measure a quantum system repeatedly, we make its evolution slower and slower, until it could stop decaying. That is, if an observer makes repeated, quick observations of a microscopic object undergoing change, the object can stop changing (just as, according to common lore, a watched kettle never boils).

A couple of months ago, under the request of some journalists, I was asked to report on this issue in the scientific sections of a couple of Spanish newspapers. What I said, in short, is that even being the quantum Zeno effect a widely accepted phenomenon, the extrapolation made by Krauss and Dent is far from clear. Actually, I was able to find some loopholes in the mathematical derivation, what points out to the conclusion that, even in the best of cases, the computed result would be many orders of magnitude smaller than the one reported and, therefore, negligible.

I shall not discuss on this philosophical point here any further, but rather concentrate on the beautiful mathematics behind the Zeno paradox. Let us continue with the very simple example above. It is quite clear that, by taking the first term,  $1/2$ , to the left what remains on the r.h.s. is just one half of the original series (extracting  $1/2$  as a common factor), so that

$$s - \frac{1}{2} = \frac{s}{2} \quad \implies \quad s = 1.$$

Thus the conclusion follows that when to one half of an apple pie we add a quarter of it and then an eighth, and so on, what we get in the end is the whole pie.

Now, something more difficult: what is the sum of the following series?

$$s = 1 + 1 + 1 + \cdots + 1 + \cdots$$

Again, any of us will answer immediately:  $s = \infty$ . In fact, whatever  $\infty$  is, everybody recognizes in this last expression the definition itself of the concept of infinity, e.g. the piling of one and the same object, once and again, without an end. Of course, this idea is absolutely true, but it is at the same time of little use to modern Physics. To be more precise, since the advent of Quantum Field Theory (QFT). In fact, calculations there are plagued with divergent series, and it is of no use to say that: look, this series here is divergent, and this other one is also divergent, and the other there too, and so on. One gets non-false but also non-useful information in this way, and actually we *do not observe* these many infinities in Nature. Thus it was discovered in the 30's and 40's that something very important was missing from the formulation or mathematical modelization of quantum physical processes.

Within the mathematical community, for years there was the suspicion that one could indeed give sense to divergent series. This has now been proven experimentally (with  $10^{-14}$  accuracy in some cases) to be true in physics, but many years earlier mathematicians were the first to realise that it was possible. In fact, Leonard Euler (1707-1783) was convinced that "*To every series one could assign a number*" [13] (that is, in a reasonable, consistent, and possibly useful way, of course). Euler was unable to prove this statement in full, but he devised a technique (Euler's summation criterion) in order to 'sum' a large family of divergent series. His statement was however controverted by some other great mathematicians, as Abel, who said that "*The divergent series are the invention of the devil, and it is a shame to base on them any demonstration whatsoever*". [14] There is a classical treatise due to G.H. Hardy and entitled simply *Divergent Series* [15] that can be highly recommended to the reader.

Actually, regularization and renormalization procedures are essential in present day Physics. Among the different techniques at hand in order to implement these processes, zeta function regularization is one of the most beautiful. Use of this method yields, for instance, the vacuum energy corresponding to a quantum physical system, which could, e.g., contribute to the cosmic force leading to the present acceleration of the

expansion of our universe. The zeta function method is unchallenged at the one-loop level, where it is rigorously defined and where many calculations of QFT reduce basically (from a mathematical point of view) to the computation of determinants of elliptic pseudodifferential operators ( $\Psi$ DOs) [16]. It is thus no surprise that the preferred definition of determinant for such operators is obtained through the corresponding zeta function (see, e.g., [17, 18]).

4.2.1. *The zeta function as a summation method.* The method of zeta regularization evolved from the consideration of the Riemann zeta function as a ‘series summation method’. The zeta function, on its turn, was actually introduced by Euler, from considerations of the harmonic series

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n} + \cdots ,$$

which is logarithmically divergent, and of the fact that, putting a real exponent  $s$  over each term,

$$1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \cdots + \frac{1}{n^s} + \cdots ,$$

then for  $s > 1$  the series is convergent, while for  $s \leq 1$  it is divergent. Euler called this expression, as a function of  $s$ , the  $\zeta$ -function,  $\zeta(s)$ , and found the following important relation

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right)^{-1} ,$$

which is crucial for the applications of this function in Number Theory. By allowing the variable  $s$  to be complex, Riemann saw the relevance of this function (that now bears his name) for the proof of the prime number theorem<sup>3</sup>, and formulated thereby the *Riemann hypothesis*, which is one of the most important problems (if not *the* most important one) in the history of Mathematics. More of that in the excellent review by Gelbart and Miller [19].

For the Riemann  $\zeta(s)$ , the corresponding complex series converges absolutely on the open half of the complex plane to the right of the

---

<sup>3</sup>Which states that the number  $\Pi(x)$  of primes which are less than or equal to a given natural number  $x$  behaves as  $x/\log x$ , as  $x \rightarrow \infty$ . It was finally proven, using Riemann’s work, by Hadamard and de la Vallée-Poussin.

abscissa of convergence  $\operatorname{Re} s = 1$ , while it diverges on the other side, but it turns out that it can be analytically continued to that part of the plane, being then everywhere analytic and finite except for the only, simple pole at  $s = 1$  (Fig. 1).<sup>4</sup> In more general cases, namely

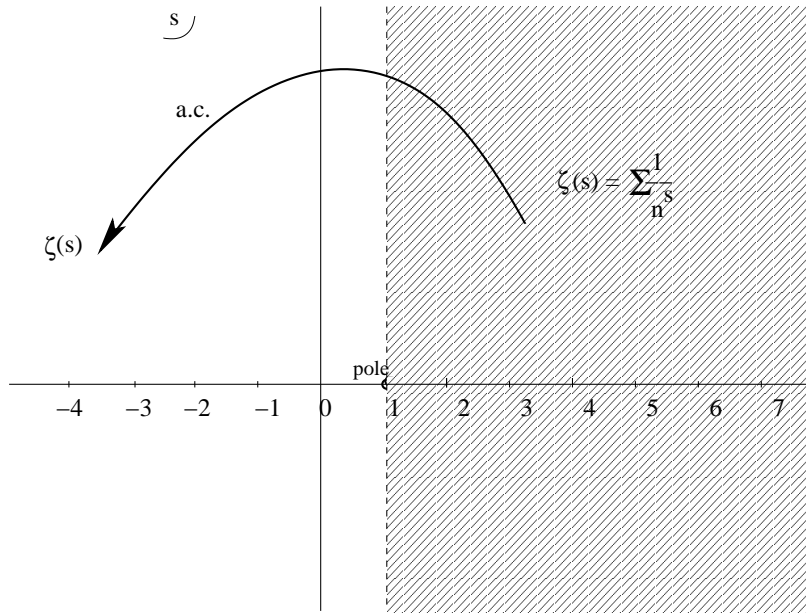


FIGURE 1. The zeta function  $\zeta(s)$  is defined in the following way, on the whole complex plane,  $s \in \mathbf{C}$ . To start, on the open half of the complex plane which is on the r.h.s of the abscissa of convergence  $\operatorname{Re} s = 1$ ,  $\zeta$  is defined as the absolutely convergent series:  $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ . In the rest of the  $s$ -complex plane,  $\zeta(s)$  is defined as the (unique) analytic continuation of the preceding function, which turns out to be meromorphic. Specifically, it is analytic everywhere on the complex plane except for one simple pole with residue equal to 1, which is at the point  $s = 1$  (notice that it corresponds to the logarithmically divergent harmonic series, as already discussed).

corresponding to the Hamiltonians which are relevant in physical applications, [20, 21, 22] the situation is in essence quite similar, albeit in

<sup>4</sup>Where it yields the harmonic series: there is no way out for this divergence.

practice it can be rather more involved. A mathematical theorem exists, which assures that under very general conditions the zeta function corresponding to a Hamiltonian operator will be also meromorphic, with just a discrete number of possible poles, which are usually simple and extend to the negative side of the real axis.<sup>5</sup>

The above picture already hints towards the use of the zeta function as a summation method. Let us consider two examples.

- (1) We interpret our starting series

$$s_1 = 1 + 1 + 1 + \cdots + 1 + \cdots$$

as a particular case of the Riemann zeta function, e.g. for the value  $s = 0$ . This value is on the left hand side of the abscissa of convergence (Fig. 1), where the series as such diverges but where the analytic continuation of the zeta function provides a perfectly finite value:

$$s_1 = \zeta(0) = -\frac{1}{2}.$$

So this is the value to be attributed to the series  $1+1+1+1+\cdots$ .

- (2) The series

$$s_2 = 1 + 2 + 3 + 4 + \cdots + n + \cdots$$

corresponds to the exponent  $s = -1$ , so that

$$s_2 = \zeta(-1) = -\frac{1}{12}.$$

A couple of comments are in order.

- (1) In two following years, some time ago, two distinguished physicists, A. Slavnov from Moscow and F. Yndurain from Madrid, gave seminars in Barcelona, about different subjects. It was quite remarkable that, in both presentations, at some point the speaker addressed the audience with these (or equivalent) words: “*As everybody knows,  $1 + 1 + 1 + \cdots = -1/2$ ”*.<sup>6</sup>

<sup>5</sup>Although there are some exceptions to this general behavior, they correspond to rather twisted situations, and are outside the scope of this brief presentation.

<sup>6</sup>Implying maybe: *If you do not know this it will be no use for you to continue listening.* Remember by the way the *lemma* of the Pythagorean school: *Do not cross this door if you do not know Geometry.*

- (2) That positive series, as the ones above, can yield a *negative* result may seem utterly nonsensical. However, it turns out that the most precise experiments ever carried out in Physics do confirm such results. More precisely: models of regularization in QED built upon these techniques lead to final numbers which are in agreement with the experimental values up to the 14th figure [23]. In recent experimental proofs of the Casimir effect [24] the agreement is also quite remarkable (given the difficulties of the experimental setup) [25].
- (3) The method of zeta regularization is based on the analytic continuation of the zeta function in the complex plane. Now, how easy is to perform that continuation? Will we need to undertake a fashionable complex-plane computation every time? It turns out that this is not so. The result is immediate to obtain, in principle, once you know the appropriate functional equation (or reflection formula) that your zeta function obeys: in the case of the Riemann zeta  $\xi(s) = \xi(1-s)$ ,  $\xi(s) \equiv \pi^{-s/2}\Gamma(s/2)\zeta(s)$ . In practice these formulas are however not optimal for actual calculations, since they are ordinarily given in terms of power series expansions (as the Riemann zeta itself), which are very slowly convergent near the corresponding abscissa. Fortunately, sometimes there are more clever expressions, that can be found, which converge exponentially fast, as the celebrated Chowla-Selberg [26] formula and some others [27, 28]. Those formulas are an speciality of the author, and give enormous power to the method of zeta function regularization.

4.2.2. *Zeta regularization in physics.* As advanced already, the regularization and renormalization procedures are essential issues of contemporary physics—without which it would simply not exist, at least in the form we now know it. [29] Among the different methods, zeta function regularization—which is obtained by analytic continuation in the complex plane of the zeta function of the relevant physical operator in each case—is maybe the most beautiful of all. Use of this method yields, for instance, the vacuum energy corresponding to a quantum physical system (with constraints of any kind, in principle). Assume the corresponding Hamiltonian operator,  $H$ , has a spectral decomposition of the form (think, as simplest case, in a quantum harmonic

oscillator):  $\{\lambda_i, \varphi_i\}_{i \in I}$ , where  $I$  is a set of indices (which can be discrete, continuous, mixed, multiple, ...). Then, the quantum vacuum energy is obtained as follows [20]:

$$\begin{aligned} E/\mu &= \sum_{i \in I} \langle \varphi_i, (H/\mu) \varphi_i \rangle = \text{Tr}_\zeta H/\mu = \sum_{i \in I} \lambda_i/\mu \\ &= \sum_{i \in I} (\lambda_i/\mu)^{-s} \Big|_{s=-1} = \zeta_{H/\mu}(-1), \end{aligned}$$

where  $\zeta_A$  is the zeta function corresponding to the operator  $A$ , and the equalities are in the sense of analytic continuation (since, generically, the Hamiltonian operator will not be of the trace class).<sup>7</sup> Note that the formal sum over the eigenvalues is usually ill defined, and that the last step involves analytic continuation, inherent with the definition of the zeta function itself. Also, the unavoidable regularization parameter with dimensions of mass,  $\mu$ , appears in the process, in order to render the eigenvalues of the resulting operator dimensionless, so that the corresponding zeta function can indeed be defined. We shall not discuss these important details here, which are just at the starting point of the whole renormalization procedure. The mathematically simple-looking relations above involve very deep physical concepts (no wonder that understanding them took several decades in the recent history of quantum field theory).

4.2.3. *The Casimir energy.* In fact things do not turn out to be so simple. One cannot assign a meaning to the *absolute* value of the zero-point energy, and any physical effect is an energy difference between two situations, such as a quantum field in curved space as compared with the same field in flat space, or one satisfying BCs on some surface as compared with the same in its absence, etc. This difference is the Casimir energy:  $E_C = E_0^{BC} - E_0 = \frac{1}{2} (\text{tr } H^{BC} - \text{tr } H)$ .

But here a problem appears. Imposing mathematical boundary conditions (BCs) on physical quantum fields turns out to be a highly non-trivial act. This was discussed in much detail in a paper by Deutsch

---

<sup>7</sup>The reader should be warned that this  $\zeta$ -trace is actually no trace in the usual sense. In particular, it is highly non-linear, as often explained by the author [30]. Some colleagues are unaware of this fact, which has led to important mistakes and erroneous conclusions too often.

and Candelas a quarter of a century ago [31]. These authors quantized em and scalar fields in the region near an arbitrary smooth boundary, and calculated the renormalized vacuum expectation value of the stress-energy tensor, to find out that the energy density diverges as the boundary is approached. Therefore, regularization and renormalization did not seem to cure the problem with infinities in this case and an infinite *physical* energy was obtained if the mathematical BCs were to be fulfilled. However, the authors argued that surfaces have non-zero depth, and its value could be taken as a handy (dimensional) cutoff in order to regularize the infinities. Just two years after Deutsch and Candelas' work, Kurt Symanzik carried out a rigorous analysis of QFT in the presence of boundaries [32]. Prescribing the value of the quantum field on a boundary means using the Schrödinger representation, and Symanzik was able to show rigorously that such representation exists to all orders in the perturbative expansion. He showed also that the field operator being diagonalized in a smooth hypersurface differs from the usual renormalized one by a factor that diverges logarithmically when the distance to the hypersurface goes to zero. This requires a precise limiting procedure and point splitting to be applied. In any case, the issue was proven to be perfectly meaningful within the domains of renormalized QFT. In this case the BCs and the hypersurfaces themselves were treated at a pure mathematical level (zero depth) by using (Dirac) delta functions.

Recently, a new approach to the problem has been postulated [33]. BCs on a field,  $\phi$ , are enforced on a surface,  $S$ , by introducing a scalar potential,  $\sigma$ , of Gaussian shape living on and near the surface. When the Gaussian becomes a delta function, the BCs (Dirichlet here) are enforced: the delta-shaped potential kills *all* the modes of  $\phi$  at the surface. For the rest, the quantum system undergoes a full-fledged QFT renormalization, as in the case of Symanzik's approach. The results obtained confirm those of [31] in the several models studied albeit they do not seem to agree with those of [32]. They are also in clear contradiction with the ones quoted in the usual textbooks and review articles dealing with the Casimir effect [34], where no infinite energy density when approaching the Casimir plates has been reported. This issue is also of importance at the cosmological level, in braneworld models.



### 4.3. Present day cosmology from modified theories of gravity.

4.3.1. *Uses of the Riemann tensor in cosmology.* As was mentioned before, Riemann's revolutionary ideas about the concept of physical space were given a definite form by Albert Einstein when he formulated the *Theory of General Relativity*, with the help of his, more mathematically minded, classmate and friend Marcel Grossmann<sup>8</sup>. The community of relativists celebrates Grossmann's contributions to physics by organizing the very important Marcel Grossman meetings, every three years (MG12 will take place in Paris, in 2009). Let us summarize the main points of the so called "*curved-space-time physics*" (excellent references are the books by Robert Wald [35]):

- (1) Space-time, the set of all events, is a four-dimensional manifold endowed with a metric  $(M, g)$ .
- (2) The metric is physically measurable by rods and clocks.
- (3) The metric of space-time can be put in the Lorentz form momentarily at any particular event by an appropriate choice of coordinates.
- (4) Freely-falling particles, unaffected by other forces, move on time-like geodesics of the space-time.
- (5) Any physical law that can be expressed in tensor notation in special relativity has exactly the same form in a locally-inertial frame of a curved space-time.

We cannot go into much detail in the standard theory of General Relativity, since we here aim at putting our emphasis on the very recent developments concerning its application to modern cosmology. Let us just recall Einstein's equations

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi GT_{\mu\nu},$$

where on the lhs we have the curvature, the geometry of space-time, under the form of contractions of the Riemann curvature tensor:

$$R_{\mu\nu\rho}{}^{\sigma} = \Gamma_{\mu\rho,\nu}^{\sigma} - \Gamma_{\nu\rho,\mu}^{\sigma} + \Gamma_{\mu\rho}^{\alpha}\Gamma_{\alpha\nu}^{\sigma} - \Gamma_{\nu\rho}^{\alpha}\Gamma_{\alpha\mu}^{\sigma},$$

---

<sup>8</sup>Who later became a Professor of Mathematics at the Federal Polytechnic Institute in Zurich, today the ETH Zurich, specializing in descriptive geometry.

the  $\Gamma$ 's being, as usual, Christoffel symbols of the Riemannian connection, and

$$R_{\mu\rho} = R_{\mu\sigma\rho}{}^{\sigma}, \quad R = R_{\mu}{}^{\mu}.$$

Einstein observed that the solution of these equations, subject to the constraints of the cosmological principle, led to a universe that was not static. He was disappointed because at that time (1915-20) the expansion of the Universe had not yet been discovered (Hubble, 1925-30) and the universe was considered by everybody to be in a stationary state. This led Einstein to introduce (almost against his actual will) a constant term in his equations (known now as the cosmological constant,  $\Lambda$ ), that was perfectly compatible with all of the principles of his gravity theory (but otherwise unnecessary):

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi GT_{\mu\nu} - \Lambda g_{\mu\nu}.$$

When a few years later Hubble discovered that the universe was in fact expanding, Einstein said the introduction of the cosmological constant had been the greatest blunder of his life. He was right to be upset since, to get a static Universe, he had added an artificial term to his field equations that stabilized the Universe against expansion or contraction. Had he possessed sufficient confidence in his original equations, he could have predicted that either his theory was wrong or the Universe was expanding or contracting, well before there was any experimental evidence of the expansion!<sup>9</sup>

An important historical issue (also for what will follow) was the derivation of Einstein's equations from a variational principle, starting from what is now called the Einstein-Hilbert action<sup>10</sup>:

$$S = \int d^4x \sqrt{-g} (L_G + L_m - \lambda),$$

where  $\lambda = \Lambda/8\pi G$ . Here the first two terms within the brackets are the Lagrangians corresponding to gravity and matter, and the last one is the cosmological constant term. By variation in the Euler-Lagrange

---

<sup>9</sup>What would have been an enormous accomplishment. This explains why Einstein got so angry.

<sup>10</sup>In fact Hilbert preceded Einstein by one day in the submission of his results for publication, in 1915.

sense, one obtains

$$\delta S_m = \int d^4x \sqrt{-g} \left( \frac{\partial L_m}{\partial g^{\mu\nu}} - \frac{1}{2} g_{\mu\nu} L_m \right) \delta g^{\mu\nu},$$

$$T_{\mu\nu} = -2 \frac{\partial L_m}{\partial g^{\mu\nu}} + g_{\mu\nu} L_m,$$

wherefrom Einstein's equations follow.

4.3.2. *Cosmological constant and the quantum vacuum energy.* However, this was not the end of the story. Any attempt at a unification of all fundamental interactions—already envisaged by Riemann and to which Einstein devoted an important part of his entire life and scientific effort—that is, a physical theory describing the gravitational interactions of matter and energy in which matter and energy are described by quantum theory, has failed. In most theories that aim at doing this, gravity itself is quantized. Since the contemporary theory of gravity, general relativity, describes gravitation as the curvature of space-time by matter and energy, a quantization of gravity implies some sort of quantization of space-time itself. As all existing physical theories rely on a classical space-time background, this presents profound methodological and ontological challenges, in fact it is considered to be maybe the most difficult problem in physics. However, new theories must always contain the successful previous ones, that have proven already to be perfectly valid in their corresponding domains of applicability.

Thus, special relativity reduces to classical Newtonian mechanics when the velocities  $v$  involved are  $v \ll c$ , and corrections to the classical formulas start with terms of the form  $v/c$  and higher powers (post-Newtonian, post-post-Newtonian approaches, etc.). In this sense, some successful semi-classical approaches to quantum gravity have been constructed. Summing up, even if we do not have a quantum theory of gravity, it is by now clear that the quantum correction to the Einstein equations corresponding to the fluctuations of the quantum vacuum will show out as an additional term in the energy-momentum tensor  $T_{\mu\nu}$ , side by side with Einstein's cosmological constant contribution, namely

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G (T_{\mu\nu} - E g_{\mu\nu} - \lambda g_{\mu\nu}),$$

where  $E$  denotes this vacuum energy density (and remember  $\lambda = \Lambda/8\pi G$ ). More precisely, the combination of this two terms (including all fundamental physical constants) reads

$$\frac{\Lambda c^2}{8\pi G} + \frac{1}{\text{Vol}} \frac{\hbar c}{2} \sum_i \omega_i,$$

$\omega_i$  being the energy modes (spectrum) of the Hamiltonian operator of the quantum theory. This fact will remain true in any quantum theory of gravitation, as far as vacuum fluctuations behave as an ordinary form of energy (e.g., they satisfy the equivalence principle), what seems indeed to be the case [36].

The dramatic consequence of this issue (already pointed out by Zel'dovich in the sixties) is that we cannot get rid any more of the cosmological constant as Einstein finally did. It will pop up, under this new form, as fluctuations of the quantum vacuum, that are allowed by the fundamental Heisenberg's uncertainty principle (unless, of course, all quantum vacuum fluctuations add up to zero, which is very difficult to realize; this is known as *the cosmological constant problem*).

4.3.3. *Cosmic acceleration.* Astrophysical observations clearly indicate that huge amounts of 'dark matter' and 'dark energy' are needed to explain the observed large scale structures and cosmic accelerating expansion of our universe. Up to now, no experimental evidence has been found, at the fundamental level, to convincingly explain such weird components. In particular, concerning the problem of the accelerating expansion, the only possibility to solve it within the domains of Einsteinian gravity is, again, through the cosmological constant term, that with the convenient sign provides the contribution needed to produce the observed acceleration (very similar to the way how Einstein tried to stabilize the universe against gravitational collapse, when he thought it should be static). However, this is not easy to do. First, when computed with care, the contribution of the vacuum energy density is many orders of magnitude *larger* than the value needed to explain the small acceleration rate of the universe expansion<sup>11</sup> (what is called the 'new' problem of the cosmological constant, which is even worse

---

<sup>11</sup>It is of the order of  $10^{123}$ , one of the largest discrepancies between theory and observation in the history of Physics.

than the older one). Second, it is not even clear (very specific models must be involved) whether the *sign* of the contribution of the vacuum fluctuations is the correct one in order to obtain expansion (and not contraction!). Making the story short, there are models where these two problems could be understood, but always with the help of some tailored hypothesis, and the general consensus is that the problem is far from having been solved yet.

This has led to consider completely different approaches (see, for instance, [37]). One of the most successful is the so-called  $f(R)$  gravity, which is a deviation from Einstein's General Relativity in the way we are going to see (note that the  $R$  stands here again for Riemann: the Riemann tensor contraction). This is an alternative theory of gravity in which dark energy and dark matter could be effects—illusions, in a sense—created by the curvature of space-time (the same bending of space and time as in General Relativity, caused by extremely massive objects, like galaxies, but now a bit modified). This theory does not require the existence of dark energy and dark matter. The problem then could be completely reversed considering dark matter and dark energy as 'shortcomings' of General Relativity and claiming for a more 'correct' theory of gravity as derived phenomenologically by matching the largest number of observational data available. As a result, accelerating behavior of cosmic fluid and rotation curves of spiral galaxies have been reported to be reproduceable by means of 'curvature effects' [38].

4.3.4.  $f(R)$  gravity. Modified gravity models constitute an interesting dynamical alternative to the  $\Lambda$ CDM cosmology—which is the standard approach nowadays—in that they are able to describe with success, and in a rather natural way, the current acceleration in the expansion of our Universe, the so called dark energy epoch (and even perhaps the initial de Sitter phase and inflation). As the name itself indicates, the modification in the action of the modified gravitational models consists of changing the  $R$  contribution by adding to it a term which is a (in principle arbitrary) function of  $R$  only. It thus reads

$$S = \frac{1}{\kappa^2} \int d^4x \sqrt{-g} [R + f(R)] + S_{(m)}.$$

Usually, one calls  $F(R) = R + f(R)$  and sometimes the theory itself is named  $F(R)$  theory (those are very recent concepts, and nomenclature is not yet completely fixed). The general equation of motion in  $F(R) \equiv R + f(R)$  gravity with matter is obtained as

$$\frac{1}{2}g_{\mu\nu}F(R) - R_{\mu\nu}F'(R) - g_{\mu\nu}\square F'(R) + \nabla_{\mu}\nabla_{\nu}F'(R) = -\frac{\kappa^2}{2}T_{(m)\mu\nu},$$

where  $T_{(m)\mu\nu}$  is the matter energy-momentum tensor.

Modified  $f(R)$  gravity has undergone already a number of studies which conclude that this gravitational alternative to dark energy is able to pass the solar system tests, that is, the very severe constraints imposed by the observational proofs that Einstein's gravity (with  $R$  only) is able to describe to extremely high precision the evolution of our solar system. Recently the importance of those modified gravity models has been reassessed, namely with the appearance of the so-called 'viable'  $f(R)$  models [39, 40]. Those are models which satisfy the stringent cosmological as well as the local gravity constraints, which had caused a number of serious problems to some of the first-generation theories of that kind. The final aim of all these phenomenological models is to describe a segment as large as possible of the whole history of our universe, as well as to recover all local predictions of Einsteinian gravity, which have been verified experimentally to very good accuracy, at the solar system scale.

In this last couple of years, we have investigated for two classes of 'viable' modified gravitational models what it means, roughly speaking, that they incorporate the vanishing (or fast decrease) of the cosmological constant in the flat ( $R \rightarrow 0$ ) limit, and that they exhibit a suitable constant asymptotic behavior for large values of  $R$ . A huge family of these models, which we term first class—and to which most of the models proposed in the literature belong—can be viewed as containing all possible smooth versions of the following sharp step-function model. To discuss this toy model, at the distribution level, proves to be very useful in order to grasp the essential features that *all* models in this large family are bound to satisfy. In other words, to extract the general properties of the whole family in a rather simple fashion (which involves, of course, standard distribution calculus). This simple model

(representative of the whole class) reads

$$f(R) = -2\Lambda_{\text{eff}} \theta(R - R_0),$$

where  $\theta(R - R_0)$  is Heaviside's step distribution. Models in this class are characterized by the existence of one or more transition scalar curvatures, an example being  $R_0$  in the above toy model.

The other class of modified gravitational models that has been considered contains a sort of 'switching on' of the cosmological constant as a function of the scalar curvature  $R$ . A simplest version of this kind reads

$$f(R) = 2\Lambda_{\text{eff}}(e^{-bR} - 1).$$

Here the transition is smooth. The two models above may be combined in a natural way, if one is also interested in the phenomenological description of the inflationary epoch. For example, a two-step model may be the smooth version of

$$f(R) = -2\Lambda_0 \theta(R - R_0) - 2\Lambda_I \theta(R - R_I),$$

with  $R_0 \ll R_I$ , the latter being the inflation scale curvature.

In a recent paper [41], we have developed a general approach to viable modified gravity in both the Jordan and the Einstein frames. We have focussed on the so-called step-class models mainly, since they seem to be most promising from the phenomenological viewpoint and, at the same time, they provide a natural possibility to classify all viable modified gravities. We have explicitly presented the cases of one- and two-step models, but a similar analysis can be extended to the case of an  $N$ -step model, with  $N$  being finite or countably infinite. No additional problems are expected to appear and the models can be adjusted, provided one can always find smooth solutions interpolating between the de Sitter solutions (what seems at this point a reasonable possibility), to repeat at each stage the same kind of de Sitter transition. We can thus obtain multi-step models which may lead to multiple inflation and multiple acceleration, in a way clearly reminiscent of braneworld inflation.

This looks quite promising, with the added bonus that the model's construction is rather simple. Use has been made of the simple but efficient tools provided by the corresponding toy model constructed with sharp distributions, a new technique that we have pioneered. It

is to be remarked that, for the infinite-step models, one can naturally expect to construct the classical gravity analog of the string-theory landscape realizations, as in the classical ideal fluid model.

The existence of viable (or “chameleon”)  $f(R)$  theories with a phase of early-time inflation is not known to us from the literature. The fact that we are able to provide several classes of models of this kind that are consistent also with the late-time accelerated expansion is thus a novelty, worth to be remarked.

Both inflation in the early universe and the recent accelerated expansion could be thus understood in these theories in a unified way. If we start with large curvature,  $f(R)$  becomes almost constant and plays the role of the effective cosmological constant, which would generate inflation. For a successful exit from the inflationary epoch we may need, in the end, more (say small non-local or small  $R^n$ ) terms. When curvature becomes smaller, matter could dominate, what would indeed lower the curvature values. Then, when the curvature  $R$  becomes small enough and  $R_0 \ll R \ll R_I$ ,  $f(R)$  becomes again an almost constant function, and plays the role of the small cosmological constant which generates the accelerated expansion of the universe, that started in the recent past. Moreover, the model naturally passes all local tests and can be considered as a true viable alternative to General Relativity. Some remark is however in order. On general grounds, one is dealing here with a highly non-linear system and one should investigate all possible critical points thereof (including other time-dependent cosmologies), within the dynamical approach method. Of course, the existence of other critical points is possible; anyhow, for viable  $f(R)$  models, to find them is not a simple task, and in Ref. [41] we have restricted our effort to the investigation of the de Sitter critical points. With regard to the stability of these points, the one associated with inflation should be unstable. In this way, the exit from inflation (what is always a very non-trivial issue) could be achieved in a quite natural way. In particular, for instance, this is in fact the case for the two-step model with an  $R^3$  term.

In conclusion, we are on the way to construct realistic modified gravities. Some of these models ultimately lead to the unification of the inflationary epoch with the late-time accelerating epoch, under quite



simple and rather natural conditions. What remains to be done is to study those models in further quantitative detail, by comparing their predictions with the accurate astrophysical data coming from ongoing and proposed sky observations. It is expected that this can be done rather soon, having in mind the possibility to slightly modify the early universe features of the theories here discussed, while still preserving all of their nice universal properties.

**4.4. Epilogue.** Let us finish this short overview of Riemann’s work and its uses in modern Physics—a clear example of the very fruitful relation between the worlds of Physics and Mathematics—with an extremely touching sentence that appears in a letter written by Albert Einstein and addressed to Arnold Sommerfeld, in the year 1912—this means, some 60 years after the celebrated *Habilitationschrift* of Bernhard Riemann—where Einstein comments on the efforts he is doing in trying to understand *Riemannian Geometry*:

*“Aber eines ist sicher, dass ich mich im Leben noch nicht annähernd so geplagt habe und dass ich große Hochachtung vor der Mathematik eingeflößt bekommen habe, die ich bis jetzt in ihren subtileren Teilen in meiner Einfalt für puren Luxus gehalten habe!”*

What means, in a free English translation: *“But one thing is sure, that never before in my life had I invested such an effort, and that I never had had such a high opinion of Mathematics, which I considered till very recently, in my boldness, and for what respects its most subtle parts, as a mere luxury!”*

I, ja per acabar, en català: *“Però una cosa és segura, que mai en la meua vida no m’havia afanyat ni de bon troç com ara, i que mai no havia dispensat tan alta consideració a la Matemàtica, la qual tenia fins fa poc, en la meua ingenuïtat, pel que fa a les seves parts més subtils, per un simple luxe!”*

**Acknowledgements.** The author is indebted with all his collaborators, that are listed as coauthors of the several joint works mentioned in the References. This investigation has been supported in part by MEC (Spain), project FIS2006-02842, and by AGAUR (Generalitat de Catalunya), contract 2005SGR-00790.

## REFERENCES

- [1] D.R. Wilkins, <http://www.maths.tcd.ie/pub/HistMath/People/Riemann/Papers.html>
- [2] M. Monastyrsky, *Riemann, Topology, and Physics*, 2nd ed. (Springer, NY, 1999).
- [3] C. Reid, *Hilbert* (Springer, Berlin, 1970) [Copernicus, NY, 1996].
- [4] D. Laugwitz, *Bernhard Riemann, 1826-1866: Turning Points in the Conception of Mathematics* (Birkhäuser, Boston, 1999).
- [5] [http://arxiv.org/PS\\_cache/physics/pdf/0405/0405110.pdf](http://arxiv.org/PS_cache/physics/pdf/0405/0405110.pdf),  
<http://www.math.ucla.edu/~vsv/papers/ch2.pdf>,  
<http://www.maths.tcd.ie/pub/HistMath/People/Riemann/Geom/WKCGeom.html>.
- [6] <http://www-groups.dcs.st-and.ac.uk/~history/Biographies/Riemann.html>
- [7] H. Farkas and I. Kra, *Riemann Surfaces* (Springer, Berlin-Heidelberg-New York, 1980); 2nd Ed. (1992).
- [8] M. Tegmark, *Measuring spacetime: from Big Bang to Black Holes*; Science, **296**, 1427 (2002).
- [9] N.J. Cornish et al., Phys. Rev. Lett. **92**, 201302 (2004).
- [10] J.-P. Luminet et al., Nature, **425**, 593 (2003).
- [11] B. Misra and E.C.G. Sudarshan, *J. Math. Phys.* **18**, 756 (1977); A. P. Balachandran and S. M. Roy, *Phys. Rev. Lett.* **84**, 4019 (2000).
- [12] L. Krauss and J. Dent, *The late time behavior of false vacuum decay: possible implications for cosmology and metastable inflating states*, arXiv:0711.1821 [hep-ph]; see also New Scientist, 21 November 2007.
- [13] H. W. Turnbull, *The Great Mathematicians* (New York University Press, NY, 1961); E. J. Alton, *Encyclopedia of World Biography* (McGraw-Hill, NY, 1973) vol. 4, pp. 30-31.
- [14] E. Maor, *To Infinity and Beyond: a Cultural History of the Infinite* (Princeton University Press, 1991).
- [15] G. H. Hardy, *Divergent Series* (Oxford University Press, 1949).
- [16] P. Ramond, *Field Theory: a Modern Primer* (Addison-Wesley, Redwood City, 1989).
- [17] E. Elizalde, *J. Phys.* **A22**, 931 (1989); E. Elizalde and A. Romeo, *Phys. Rev.* **D40**, 436 (1989); E. Elizalde, *J. Math. Phys.* **31**, 170 (1990).
- [18] E. Elizalde, *J. Phys.* **A18**, 1637 (1985); E. Elizalde, *Math. of Computation* **47**, 175 (1986); E. Elizalde and A. Romeo, *Int. J. Math. and Math. Sci.* **13**, 453 (1990); D. Fried, *Invent. Math.* **84**, 523 (1984).
- [19] S. S. Gelbart and S. D. Miller, *Bull. Amer. Math. Soc. (NS)* **41**, 59 (2004).
- [20] E. Elizalde, S.D. Odintsov, A. Romeo, A.A. Bytsenko and S. Zerbini, *Zeta Regularization Techniques with Applications* (World Scientific, Singapore, 1994); K. Kirsten, *Spectral Functions in Mathematics and Physics* (Chapman & Hall, London, 2001); M. Bordag, E. Elizalde and K. Kirsten, *J. Math. Phys.* **37**, 895 (1996); *J. Phys.* **A31**, 1743 (1998); M. Bordag, E. Elizalde, K. Kirsten

- and S. Leseduarte, Phys. Rev. **D56**, 4896 (1997); E. Elizalde, L. Vanzo and S. Zerbini, Commun. Math. Phys. **194**, 613 (1998); E. Elizalde, J. Comput. Appl. Math. **118**, 125 (2000); J. Phys. **A34**, 3025 (2001); E. Elizalde, M. Bordag and K. Kirsten, J. Phys. A **31**, 1743 (1998); E. Elizalde and K. Kirsten, J. Math. Phys. **35**, 1260 (1994); E. Elizalde and A. Romeo, J. Math. Phys. **30**, 1133 (1989).
- [21] E. Elizalde, *Ten Physical Applications of Spectral Zeta Functions* (Springer-Verlag, Berlin, 1995).
- [22] A.A. Bytsenko, G. Cognola, E. Elizalde, V. Moretti and S. Zerbini, *Analytic Aspects of Quantum Fields* (World Scientific, Singapore, 2004).
- [23] R. S. Van Dyck, Jr. *Anomalous Magnetic Moment of Single Electrons and Positrons: Experiment*, in Quantum Electrodynamics (Ed. T. Kinoshita) (World Scientific, Singapore, 1990) pp. 322-388; T. Kinoshita, *Rept. Prof. Phys.* **59**, 1459 (1996).
- [24] H.B.G. Casimir, Proc. K. Ned. Acad. Wet. **51**, 635 (1948).
- [25] S. K. Lamoreaux, *Phys. Rev. Lett.* **78**, 5 (1997); U. Mohideen and A. Roy, *Phys. Rev. Lett.* **81**, 4549 (1998); B. W. Harris, F. Chen, and U. Mohideen, *Phys. Rev.* **A62**, 052109 (2000).
- [26] S. Chowla and A. Selberg, Proc. Nat. Acad. Sci. U.S.A. **35**, 317 (1949).
- [27] E. Elizalde, J. Phys. **A27**, 3775 (1994).
- [28] E. Elizalde, Commun. Math. Phys. **198**, 83 (1998).
- [29] J. Collins, *Renormalization: An Introduction to Renormalization, the Renormalization Group and the Operator-Product Expansion* (Cambridge University Press, 1984); W. D. McComb, *Renormalization Methods: A Guide for Beginners* (Oxford University Press, 2004).
- [30] E. Elizalde, J. High Energy Phys. **9907**, 015 (1999).
- [31] D. Deutsch and P. Candelas, Phys. Rev. **D20**, 3063 (1979).
- [32] K. Symanzik, Nucl. Phys. **B190**, 1 (1981).
- [33] R. L. Jaffe, *Unnatural acts: unphysical consequences of imposing boundary conditions on quantum fields*, hep-th/0307014; N. Graham, R. L. Jaffe, V. Khe-  
mani, M. Quandt, M. Scandurra and H. Weigel, *Casimir energies in light of quantum field theory*, hep-th/0207205; N. Graham, R. L. Jaffe, V. Khe-  
mani, M. Quandt, M. Scandurra and H. Weigel, Nucl. Phys. **B645**, 49 (2002);  
N. Graham, R. L. Jaffe and H. Weigel, Int. J. Mod. Phys. **A17**, 846 (2002).
- [34] V.M. Mostepanenko and N.N. Trunov, *The Casimir Effect and its Application* (Clarendon Press, Oxford, 1997); K.A. Milton, *The Casimir Effect: Physical Manifestations of Zero-Point Energy* (World Scientific, Singapore, 2001); M. Bordag, U. Mohideen and V.M. Mostepanenko, Phys. Reports **353**, 1 (2001).
- [35] R.M. Wald, *General Relativity* (University of Chicago Press, Chicago, 1984); *Quantum Fields in Curved Spacetime and Black Hole Thermodynamics* (University of Chicago Press, Chicago, 1994).
- [36] S.A. Fulling et. al., *How Does Casimir Energy Fall?* arXiv:hep-th/070209.

- [37] E. Elizalde, S. Nojiri and S.D. Odintsov, Phys. Rev. **D70**, 043539 (2004); E. Elizalde, S. Nojiri, S.D. Odintsov and P. Wang, Phys. Rev. **D71**, 103504 (2005).
- [38] S. Capozziello, V.F. Cardone and A. Troisi, J. Cosmol. Astropart. Phys. **08**, 001 (2006).
- [39] W. Hu and I. Sawicki, arXiv:0705.1158[astro-ph].
- [40] S.A. Appleby and R.A. Battye, arXiv:0705.3199[astro-ph]; A. Starobinskii, JETP Lett.**86** 157 (2007).
- [41] G. Cognola, E. Elizalde, S. Nojiri, S. D. Odintsov, L. Sebastiani and S. Zerbini, Phys. Rev. **D77**, 046009 (2008); arXiv:0712.4141[hep-th]; E. Elizalde, S. Nojiri and S. D. Odintsov, Phys. Rev. D **70** (2004) 043539 [arXiv:hep-th/0405034]; G. Cognola, E. Elizalde, S. Nojiri, S. D. Odintsov and S. Zerbini, JCAP **0502** (2005) 010 [arXiv:hep-th/0501096]; G. Cognola, E. Elizalde, S. Nojiri, S. D. Odintsov and S. Zerbini, Phys. Rev. D **73** (2006) 084007 [arXiv:hep-th/0601008]; E. Elizalde, S. Nojiri, S. D. Odintsov and S. Ogushi, Phys. Rev. D **67** (2003) 063515 [arXiv:hep-th/0209242]; E. Elizalde and J. Quiroga Hurtado, Mod. Phys. Lett. **A19**, 29 (2004) [GR-QC 0310128]; E. Elizalde, S. Naftulin, S.D. Odintsov, Phys. Rev. **D49**, 2852 (1994).

# CÓDIGOS CONVOLUCIONALES Y GEOMETRÍA ALGEBRAICA

J.M. MUÑOZ PORRAS, J.A. DOMÍNGUEZ PÉREZ

RESUMEN. Las técnicas de Geometría Algebraica utilizadas para construir códigos de Goppa sobre una curva algebraica sobre un cuerpo finito  $\mathbb{F}_q$  pueden extenderse a la construcción de códigos convolucionales sobre el cuerpo infinito de funciones racionales en una variable  $\mathbb{F}_q(z)$ . De este modo es posible construir códigos convolucionales sobre la recta proyectiva que alcanzan la cota de Singleton generalizada.

## 1. INTRODUCCIÓN A LA TEORÍA DE CÓDIGOS CONVOLUCIONALES

Dado un cuerpo finito  $\mathbb{F}_q$  de  $q$  elementos, un  $[n, k]$ -codificador lineal es una aplicación lineal inyectiva

$$\begin{aligned} G: \mathbb{F}_q^k &\hookrightarrow \mathbb{F}_q^n \\ u &\mapsto x = u \cdot G, \end{aligned}$$

cuya imagen define el  $[n, k]$ -código lineal  $\mathcal{C}_k \equiv \text{Im } G \subset \mathbb{F}_q^n$ .

De este modo, la *codificación* es la transformación de una “palabra de información”  $u \in \mathbb{F}_q^k$  en una “palabras del código”  $x = u \cdot G \in \mathbb{F}_q^n$ , mediante el producto por una “matriz codificadora”  $G$  de  $k$  filas y  $n$  columnas.

Si lo que se considera es una sucesión finita de palabras de información  $u(1), u(2), \dots, u(r)$ , se obtiene a su vez una sucesión finita de palabras del código  $x(1), x(2), \dots, x(r)$ , donde

$$(1) \quad x(t) = u(t) \cdot G,$$

de modo que  $x(t)$  es una función lineal de  $u(t)$ .

Entendiendo  $u(t)$  como una *palabra de información en el instante de tiempo*  $t$ , la idea de la codificación convolucional es hacer que  $x(t)$  sea

una función lineal de  $u(t)$  y de un número fijo de palabras de información anteriores  $u(t), u(t-1), \dots, u(t-m)$ , donde  $m$  es un entero positivo denominado “memoria” del codificador (en particular, los códigos lineales son códigos convolucionales sin memoria,  $m = 0$ ).

Este concepto de *codificador convolucional* fue introducido por P. Elias [1] en 1955, y desde el punto de vista algebraico fue desarrollado por G.D. Forney Jr. [2] en una serie de papeles clásicos en la década de 1970, y posteriormente reformulado por R. McEliece [3] en 1998. Más recientemente se han introducido técnicas de Geometría Algebraica en la teoría de códigos convolucionales con los trabajos de J. Rosenthal y R. Smarandache [4] y V. Lomadze [5], así como los publicados por los autores en colaboración con G. Serrano Sotelo y J.I. Iglesias Curto [6], [7], en cuyos resultados está basada esta exposición.

Para dar una definición concreta de *código convolucional*, reexpresamos la fórmula (1) como un sumatorio formal (dependiente del tiempo  $t$ ) introduciendo un *operador de retardo*  $D$ ,

$$(2) \quad \sum_{t \geq 0} x(t)D^t = \sum_{t \geq 0} u(t)D^t \cdot G,$$

donde  $D(\sum_{t \geq 0} u(t)D^t) = \sum_{t \geq 0} u(t)D^{t+1} = \sum_{t \geq 1} u(t-1)$ . Si se denota

$$x(D) \equiv \sum_{t \geq 0} u(t)D^t \in \mathbb{F}_q[D]^n$$

$$u(D) \equiv \sum_{t \geq 0} u(t)D^t \in \mathbb{F}_q[D]^k$$

la expresión (2) se puede reescribir como

$$(3) \quad x(D) = u(D) \cdot G,$$

ecuación que resume el proceso de codificación lineal respecto a la matriz codificadora  $G$ . La codificación convolucional consiste entonces en hacer que la matriz codificadora que aparece en (3) dependa también del operador de retardo

$$x(D) = u(D) \cdot G(D).$$

De este modo se obtiene una de las posibles definiciones de código convolucional, para la que utilizaremos una notación polinómica, reemplazando  $D$  por la variable  $z$ .

**Definición 1.1.** *Un  $[n, k]$ -codificador convolucional (polinómico) es un homomorfismo inyectivo de  $\mathbb{F}_q[z]$ -módulos*

$$G(z) : \mathbb{F}_q[z]^k \hookrightarrow \mathbb{F}_q[z]^n$$

$$u(z) \mapsto x(z) = u(z) \cdot G(z).$$

*La imagen de este homomorfismo define (las palabras de) el  $[n, k]$ -código convolucional*

$$\mathcal{C}_k \equiv \text{Img } G(z).$$

Dado un  $[n, k]$ -codificador convolucional  $G(z)$ , se tiene una sucesión exacta

$$0 \rightarrow \mathbb{F}_q[z]^k \xrightarrow{G(z)} \mathbb{F}_q[z]^n \rightarrow M \rightarrow 0,$$

donde  $M$  es un  $\mathbb{F}_q[z]$ -módulo de rango  $n - k$ . Tomando entonces duales como  $\mathbb{F}_q[z]$ -módulos, resulta otra sucesión exacta

$$0 \rightarrow \widehat{M} \rightarrow \widehat{\mathbb{F}_q[z]^n} \xrightarrow{G(z)^t} \widehat{\mathcal{C}}_k \rightarrow \widehat{\mathcal{C}}_k / \text{Img } G(z)^t \rightarrow 0,$$

donde  $\widehat{\mathcal{C}}_k / \text{Img } G(z)^t$  es un  $\mathbb{F}_q[z]$ -módulo de torsión, cuyo anulador es precisamente

$$\text{Ann} \left( \widehat{\mathcal{C}}_k / \text{Img } G(z)^t \right) = \langle \{\text{menores } k \times k \text{ de } G(z)\} \rangle.$$

La aplicación  $G(z)^t$  permite distinguir diversos tipos de codificadores, e introducir una noción de “grado interno” del codificador

$$\text{Gr}_{in}(G(z)) \equiv \text{máximo grado de los menores } k \times k \text{ de } G(z),$$

en contraste con la noción de “grado externo”,

$$\text{Gr}_{ex}(G(z)) \equiv \sum_{i=1}^k \text{máximo grado de la fila } i\text{-ésima de } G(z).$$

Este grado externo se identifica con la idea de “memoria” que motivó la introducción de los codificadores convolucionales, por lo que denotaremos indistintamente

$$m_G \equiv \text{Gr}_{ex}(G(z)).$$

**Definición 1.2.** *Un codificador convolucional  $G(z)$  se dice*

- *básico*  $\Leftrightarrow G(z)^t$  es epiyectiva.
- *reducido*  $\Leftrightarrow \text{Gr}_{in}(G(z)) = \text{Gr}_{ex}(G(z))$ .

- *canónico*  $\Leftrightarrow \text{Gr}_{ex}(G(z))$  es mínimo entre todos los codificadores polinómicos que definen el mismo código  $\mathcal{C}_k = \text{Img } G(z)$ .

A su vez, los conceptos de “grado” y “memoria” pueden aplicarse al propio código convolucional  $\mathcal{C}_k$ : considerando el conjunto de los codificadores polinómicos  $G(z)$  tales que  $\mathcal{C}_k = \text{Img } G(z)$ , el grado del código convolucional es

$$\begin{aligned} \delta(\mathcal{C}_k) &\equiv \text{mínimo } \text{Gr}_{ex} \text{ de sus codificadores polinómicos} \equiv \\ &\equiv \text{memoria total de } \mathcal{C}_k . \end{aligned}$$

**Observación 1.3.** *Los códigos convolucionales se pueden definir a partir de codificadores más generales, considerando el cuerpo (infinito)  $\mathbb{F}_q(z)$  de funciones racionales en una variable: un  $[n, k]$ -codificador convolucional es un homomorfismo inyectivo de  $\mathbb{F}_q(z)$ -espacios vectoriales  $G(z): \mathbb{F}_q(z)^k \hookrightarrow \mathbb{F}_q(z)^n$  y su imagen define el  $[n, k]$ -código convolucional  $\mathcal{C}_k \equiv \text{Img } G(z)$ .*

## 2. REALIZACIÓN FÍSICA DE UN CODIFICADOR CONVOLUCIONAL

Los codificadores convolucionales pueden realizarse físicamente mediante *circuitos secuenciales lineales*, equivalentes a *sistemas lineales invariantes respecto del tiempo y con un número finito de variables de estado*.

Para ello, si  $G(z)$  es un  $[n, k]$ -codificador convolucional de memoria  $m$ , la ecuación

$$x(z) = u(z) \cdot G(z)$$

puede expresarse como un sistema lineal de ecuaciones, introduciendo las “variables de estado”  $s(t)$ , con  $s(0) = 0$ ,

$$\begin{cases} s(t+1) = s(t) \cdot A + u(t) \cdot B \\ x(t) = s(t) \cdot C + u(t) \cdot G \end{cases}$$

donde  $s(t)$ ,  $u(t)$  y  $x(t)$  son vectores de dimensiones  $m$ ,  $k$  y  $n$ , respectivamente, mientras que  $A$ ,  $B$ ,  $C$  y  $D$  son matrices con coeficientes en  $\mathbb{F}$  de dimensiones  $m \times m$ ,  $k \times m$ ,  $m \times n$  y  $k \times n$ , respectivamente.

**Ejemplo 2.1.** *Sea  $G(z) = (1 + z + z^2, 1 + z^2)$ . En este caso,  $k = 1$ ,  $n = 2$  y  $m = 2$ , y la codificación puede expresarse*

$$(x_1(t), x_2(t)) = (u_1(t)) \cdot (1 + t + t^2, 1 + t^2),$$



o equivalentemente

$$(x_1(t), x_2(t)) = (u_1(t) + u_1(t - 1) + u_1(t - 2), u_1(t) + u_1(t - 2)).$$

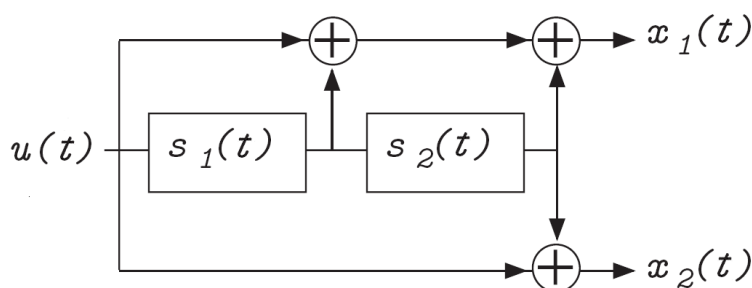
Introduciendo entonces las variables de estado,

$$\begin{cases} (s_1(t+1), s_2(t+1)) = (u_1(t), s_1(t)) \\ (x_1(t), x_2(t)) = (u_1(t) + s_1(t) + s_2(t), u_1(t) + s_2(t)) \end{cases}$$

resulta que la codificación es el sistema de ecuaciones lineales

$$\begin{cases} (s_1(t+1), s_2(t+1)) = (s_1(t), s_2(t)) \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + (u_1(t)) \cdot (1 \ 0) \\ (x_1(t), x_2(t)) = (s_1(t), s_2(t)) \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} + (u_1(t)) \cdot (1 \ 1) \end{cases}$$

cuya realización física es el circuito



### 3. CODIFICADORES POLINÓMICOS Y HACES SOBRE LA RECTA PROYECTIVA

Consideremos  $\mathbb{P}^1 = \text{Proj } \mathbb{F}_q[x_0, x_1]$  la recta proyectiva sobre el cuerpo  $\mathbb{F}_q$ , donde  $(x_0, x_1)$  son las coordenadas proyectivas, y su abierto afín  $U = \mathbb{P}^1 - \{P_\infty\}$ , con  $P_\infty \equiv \{x_0 = 0\}$  el punto del infinito de modo que la variable  $z = \frac{x_1}{x_0}$  sea la coordenada afín de esta recta proyectiva.

Dado entonces un codificador polinómico  $G(z)$ , es posible realizar su “homegenización”  $G(x_0, x_1)$ , haciendo un cambio de variable a las coordenadas proyectivas.

**Ejemplo 3.1.**

$$G(z) = (1 + z + z^2, 1 + z^2): \mathbb{F}_2[z] \hookrightarrow \mathbb{F}_2[z]^2$$

$$G(x_0, x_1) = (x_0^2 + x_0x_1 + x_1^2, x_0^2 + x_1^2): \mathbb{F}_2[x_0, x_1] \hookrightarrow \mathbb{F}_2[x_0, x_1]^2$$

El estudio de los codificadores polinómicos puede entonces plantearse en términos del haz de anillos sobre  $U$

$$\mathcal{O}_U \equiv \widetilde{\mathbb{F}[z]},$$

teniendo en cuenta las identificaciones:

$$\begin{aligned} \text{Hom}_{\mathbb{F}[z]}(\mathbb{F}[z]^k, \mathbb{F}[z]^n) &= \text{Hom}_{\mathcal{O}_U}(\mathcal{O}_U^k, \mathcal{O}_U^n) = \\ &= H^0(U, \mathcal{H}om_{\mathcal{O}_{\mathbb{P}^1}}(\mathcal{O}_{\mathbb{P}^1}^k, \mathcal{O}_{\mathbb{P}^1}^n)) = \\ &= \varinjlim_m H^0(\mathbb{P}^1, \mathcal{H}om_{\mathcal{O}_{\mathbb{P}^1}}(\mathcal{O}_{\mathbb{P}^1}^k, \mathcal{O}_{\mathbb{P}^1}^n) \otimes \mathcal{O}_{\mathbb{P}^1}(mP_\infty)) = \\ &= \varinjlim_m H^0(\mathbb{P}^1, \mathcal{H}om_{\mathcal{O}_{\mathbb{P}^1}}(\mathcal{O}_{\mathbb{P}^1}^k, \mathcal{O}_{\mathbb{P}^1}^n(mP_\infty))) = \\ &= \varinjlim_m \text{Hom}(\mathcal{O}_{\mathbb{P}^1}^k, \mathcal{O}_{\mathbb{P}^1}(mP_\infty)^n) = \\ &= \varinjlim_m \text{Hom}_{\mathbb{F}}(\mathbb{F}^k, H^0(\mathcal{O}_{\mathbb{P}^1}(mP_\infty)^n)). \end{aligned}$$

Siguiendo esta idea, los autores están trabajando conjuntamente con J.I. Iglesias Curto en la clasificación de los códigos convolucionales, en términos de esquemas Quot.

#### 4. LA NOCIÓN DE DISTANCIA PARA CÓDIGOS CONVOLUCIONALES

Dado un vector polinómico  $x(z) = (x_1(z), \dots, x_n(z)) \in \mathbb{F}[z]^n$ , es posible definir su *peso de Hamming* como

$$hwt(x(z)) = \#\{i \mid x_i(z) \neq 0\}.$$

Sin embargo, este peso no sirve para medir los errores de transmisión si se utiliza la codificación convolucional. En la teoría de códigos convolucionales, la noción natural de peso se obtiene considerando los vectores polinómicos  $x(z) \in \mathbb{F}[z]^n$  como  $x(z) = \sum_t x(t)z^t$ , con

$x(t) = (x_1(t), \dots, x_n(t)) \in \mathbb{F}^n$ , de modo que el *peso* de  $x(t)$  es

$$wt(x(z)) = \sum_t hwt(x(t)), \quad \text{donde } hwt(x(t)) = \#\{i \mid x_i(t) \neq 0\}.$$

De este modo, el concepto de *distancia libre* de un  $[n, k]$ -código convolucional  $\mathcal{C}_k \subseteq \mathbb{F}[z]^n$  es

$$d_{free}(\mathcal{C}_k) = \min \{wt(x(z)) \mid x(z) \in \mathcal{C}_k, x(z) \neq 0\}.$$

**Teorema 4.1.** (véase [4]) *Dado un  $[n, k]$ -código convolucional de grado  $\delta$ , su distancia libre  $d_{free}$  verifica la cota de Singleton generalizada*

$$d_{free} \leq (n - k) \left( \left\lfloor \frac{\delta}{k} \right\rfloor + 1 \right) + \delta + 1.$$

Un código convolucional cuya distancia libre alcanza el valor máximo de esta cota se denomina *MDS (Maximum Distance Separable)*.

La cuestión clave es entonces la construcción de códigos convolucionales MDS, para lo cual es posible introducir técnicas algebraicas, similares a las utilizadas en la teoría de códigos algebraico-geométricos de Goppa.

### 5. CÓDIGOS DE GOPPA CONVOLUCIONALES

Sea  $\mathbb{F}_q(z)$  el cuerpo (infinito) de funciones racionales en una variable, y  $X$  una curva proyectiva lisa sobre  $\mathbb{F}_q(z)$ . Consideremos  $D = P_1 + \dots + P_n$  un divisor de  $n$  puntos  $\mathbb{F}_q(z)$ -racionales distintos de  $X$  y  $G$  otro divisor con soporte disjunto a  $D$ , tal que

$$2g - 2 < \text{Gr}(G) < n.$$

Sobre el  $\mathbb{F}_q(z)$ -espacio de secciones globales  $L(G)$  se tiene una aplicación  $\mathbb{F}_q(z)$ -lineal inyectiva

$$\begin{aligned} \alpha: L(G) &\rightarrow \mathbb{F}_q(z) \times \overset{n}{\dots} \times \mathbb{F}_q(z) \\ f &\mapsto (f(P_1), \dots, f(P_n)) \end{aligned}$$

**Definición 5.1.** (véase [6] o [7]) *El código de Goppa convolucional de longitud  $n$  y dimensión  $k = \text{Gr}(G) + 1 - g$  asociado al par  $(D, G)$  es*

$$\mathcal{C}(D, G) \equiv \text{Img } \alpha \subset \mathbb{F}_q(z)^n.$$

Considerando sobre el  $\mathbb{F}_q(z)$ -espacio vectorial  $\mathbb{F}_q(z)^n$  el producto escalar  $\langle \cdot, \cdot \rangle$

$$\mathbb{F}_q(z)^n \times \mathbb{F}_q(z)^n \rightarrow \mathbb{F}_q(z)$$

$$(x(z), y(z)) \mapsto \langle x(z), y(z) \rangle = \sum_{i=1}^n x_i(z)y_i(z),$$

es posible construir el correspondiente *código de Goppa convolucional dual*,

$$\mathcal{C}(D, G)^\perp \equiv \{y(z) \in \mathbb{F}_q(z)^n \mid \langle x(z), y(z) \rangle = 0 \text{ para todo } x(z) \in \mathcal{C}(D, G)\}.$$

**Teorema 5.2.** (véase [6] o [7])  $\mathcal{C}^\perp(D, G)$  es también un código de Goppa convolucional: si  $K$  es el divisor canónico de formas diferenciales racionales sobre  $X$ , entonces  $\mathcal{C}^\perp(D, G)$  es la imagen de la aplicación  $\mathbb{F}_q(z)$ -lineal  $\beta: L(K + D - G) \rightarrow \mathbb{F}_q(z)^n$ , dada por

$$\beta(\eta) = (\text{Res}_{p_1}(\eta), \dots, \text{Res}_{p_n}(\eta)).$$

## 6. CÓDIGOS DE GOPPA CONVOLUCIONALES SOBRE LA RECTA PROYECTIVA

Consideremos  $X = \mathbb{P}_{\mathbb{F}_q(z)}^1 = \text{Proj } \mathbb{F}_q(z)[u_0, u_1]$  la recta proyectiva sobre el cuerpo  $\mathbb{F}_q(z)$ , donde  $w = u_1/u_0$  es la coordenada afín,  $P_0 = (1, 0)$  es el punto origen, y  $P_\infty = (0, 1)$  el punto del infinito. Sean  $P_1, \dots, P_n$  un conjunto de  $n$  puntos racionales distintos,  $P_i = (1, \alpha_i) \neq P_0, P_\infty$ .

Tomando los divisores  $D = P_1 + \dots + P_n$  y  $G = rP_\infty - sP_0$ , con  $0 \leq s \leq r < n$ , el  $\mathbb{F}_q(z)$ -espacio vectorial  $L(G)$  tiene como base

$$L(G) = \langle w^s, w^{s+1}, \dots, w^r \rangle$$

y el código de Goppa convolucional  $\mathcal{C}(D, G)$  es la imagen de la aplicación

$$\alpha: L(G) \rightarrow \mathbb{F}_q(z) \times \overset{n}{\dots} \times \mathbb{F}_q(z)$$

$$w^i \mapsto (\alpha_1^i, \dots, \alpha_n^i)$$

Se obtiene de este modo un código convolucional  $\mathcal{C}(D, G)$  de longitud  $n$  y dimensión  $k = r - s + 1$ , cuya matriz generadora es

$$G = \begin{pmatrix} \alpha_1^s & \alpha_2^s & \dots & \alpha_n^s \\ \alpha_1^{s+1} & \alpha_2^{s+1} & \dots & \alpha_n^{s+1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^r & \alpha_2^r & \dots & \alpha_n^r \end{pmatrix}.$$

En cuanto al código dual, como

$$\Omega(G-D) = \left\langle \frac{dw}{w^s \prod_{i=1}^n (w - \alpha_i)}, \frac{w dw}{w^s \prod_{i=1}^n (w - \alpha_i)}, \dots, \frac{w^{n-r+s-2} dw}{w^s \prod_{i=1}^n (w - \alpha_i)} \right\rangle$$

y calculando los residuos

$$\text{Res}_{P_j} \left( \frac{w^m dw}{w^s \prod_{i=1}^n (w - \alpha_i)} \right) = \frac{\alpha_j^m dw}{\alpha_j^s \prod_{i=1, i \neq j}^n (\alpha_j - \alpha_i)}$$

resulta que  $\mathcal{C}(D, G)^*$  es un código convolucional de longitud  $n$  y dimensión  $n - k = n - r + s - 1$  que tiene como matriz generadora (= matriz de control de  $\mathcal{C}(D, G)$ )

$$H = \begin{pmatrix} h_1 & h_2 & \dots & h_n \\ h_1 \alpha_1 & h_2 \alpha_2 & \dots & h_n \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ h_1 \alpha_1^{n-r+s-2} & h_2 \alpha_2^{n-r+s-2} & \dots & h_n \alpha_n^{n-r+s-2} \end{pmatrix},$$

donde  $h_j = \frac{1}{\alpha_j^s \prod_{i=1, i \neq j}^n (\alpha_j - \alpha_i)}$ .

Obsérvese que esa matriz generadora  $H$  de  $\mathcal{C}(D, G)^\perp$  tiene la forma de un “codificador alternante” sobre  $\mathbb{F}_q(z)$ , lo que sugiere la posibilidad de emplear en este contexto convolucional los algoritmos algebraicos de decodificación conocidos para los códigos lineales alternantes.

Como casos particulares, vamos a calcular las matrices  $G$  y  $H$  en el caso de puntos  $\mathbb{F}_q(z)$ -racionales  $P_i = (1, \alpha_i)$  cuando

$$\alpha_i = a^{i-1} z + b^{i-1}, \quad i = 1, \dots, n, \quad n < q,$$

comprobando además que en estos casos resultan códigos convolucionales MDS.

- Cuerpo  $\mathbb{F}_3(z) = \{0, 1, 2\}$

Caso  $a = 1, b = 2, G = P_\infty - P_0$ .

$$G = (z + 1 \quad z + 2)$$

$$H = \left( \frac{1}{2(z+1)} \quad \frac{1}{z+2} \right)$$

$$(n, k, \delta, d) = (2, 1, 1, 4)$$

- Cuerpo  $\mathbb{F}_4(z) = \{0, 1, \xi, \xi^2\}$ , con  $\xi^2 + \xi + 1 = 0$

Caso  $a = \xi, b = \xi^2, G = P_\infty$

$$G = \left( \frac{1}{z+1} \quad \frac{1}{\xi z + \xi^2} \quad \frac{1}{\xi^2 z + \xi} \right)$$

$$H = \left( \frac{1}{(\xi^2 z + \xi)(\xi z + \xi^2)} \quad \frac{1}{(\xi^2 z + \xi)(z+1)} \quad \frac{1}{(\xi z + \xi^2)(z+1)} \right)$$

$$(n, k, \delta, d) = (3, 2, 1, 3).$$

Caso  $a = 1, b = \xi, G = P_\infty - P_0$

$$G = (z+1 \quad z+\xi \quad z+\xi^2)$$

$$H = \begin{pmatrix} \frac{1}{z+1} & \frac{\xi}{z+\xi} & \frac{\xi^2}{z+\xi^2} \\ 1 & \xi & \xi^2 \end{pmatrix}$$

$$(n, k, \delta, d) = (3, 1, 1, 6).$$

- Cuerpo  $\mathbb{F}_5(z)$

Caso  $a = 1, b = 2, G = 2P_\infty - 2P_0$

$$G = ((z+1)^2 \quad (z+2)^2 \quad (z+4)^2)$$

$$H = \begin{pmatrix} \frac{2}{(z+1)^2} & \frac{2}{(z+2)^2} & \frac{1}{(z+4)^2} \\ \frac{2}{z+1} & \frac{2}{z+2} & \frac{1}{z+4} \end{pmatrix}$$

$$(n, k, \delta, d) = (3, 1, 2, 9).$$

Caso  $a = 2, b = 3, G = 2P_\infty - P_0$

$$G = \begin{pmatrix} z+1 & 2z+3 & 4z+4 & 3z+2 \\ (z+1)^2 & (2z+3)^2 & (4z+4)^2 & (3z+2)^2 \end{pmatrix}$$

$$H = \begin{pmatrix} \frac{4}{(z+1)^2(z+2)(z+3)} & \frac{4}{(z+2)(z+3)(z+4)^2} & \frac{4}{(z+1)^2(z+2)(z+3)} & \frac{4}{(z+2)(z+3)(z+4)^2} \\ \frac{4}{(z+1)(z+2)(z+3)} & \frac{3}{(z+2)(z+3)(z+4)} & \frac{1}{(z+1)(z+2)(z+3)} & \frac{2}{(z+2)(z+3)(z+4)} \end{pmatrix}$$

$$(n, k, \delta, d_{free}) = (4, 2, 3, 8).$$

**Agradecimientos.** Los autores manifiestan su gratitud a sus colaboradores del grupo de investigación *Geometría Algebraica, Aritmética y Teoría de Códigos* del Departamento de Matemáticas de la Universidad de Salamanca, especialmente a Gloria Serrano Sotelo y J.I. Iglesias Curto.

#### REFERENCIAS

- [1] P. Elias, Coding for noisy channels, *I.R.E. Nat. Conv. Record* **3**, 34–45, (1955).
- [2] G.D. Forney Jr, Convolutional codes I: Algebraic structure, *IEEE Trans. Inform. Theory*, **16** (3), 720–738, (1970).
- [3] R.J. McEliece, The algebraic theory of convolutional codes, in *Handbook of coding theory, Vol. I*, 1065–1138, North-Holland, Amsterdam, (1998).
- [4] J. Rosenthal and R.Smarandache, Maximum distance separable convolutional codes, *Appl. Algebra Engrg. Comm. Comput.*, **10** (1), 15–32, (1999).
- [5] V. Lomadze, Convolutional Codes and Coherent Sheaves, *Algebra Engrg. Comm. Comput.*, **12** (4), 273–326, (2001).
- [6] J.A. Domínguez Pérez, J.M. Muñoz Porras and G. Serrano Sotelo, Convolutional Codes of Goppa type, *Algebra Engrg. Comm. Comput.*, **15** (1), 51–61, (2004).
- [7] J.M. Muñoz Porras, J.A. Domínguez Pérez, J.I. Iglesias Curto and G. Serrano Sotelo, Convolutional Goppa Codes, *IEEE Trans. Inform. Theory*, **52** (1), 340–344, (2006).





# FUNCIONES THETA DE RIEMANN

ESTEBAN GÓMEZ GONZÁLEZ  
JOSÉ M. MUÑOZ PORRAS

## 1. INTRODUCCIÓN

El objetivo de este artículo es presentar al lector no especializado una introducción al tema de las funciones theta y sus aplicaciones geométricas y aritméticas. Las funciones theta fueron introducidas por primera vez por B. Riemann y su estudio ha sido desarrollado por diversos matemáticos como Frobenius, Weierstrass, Hecke, Artin, Tate, Weil, Langlands, Mumford y Fay. Actualmente el estudio de las funciones theta y las formas modulares desde un punto de vista aritmético es una parte esencial del programa de Langlands.

Esperamos que este artículo sea de utilidad para estudiantes de postgrado interesados en la conexión entre la Geometría Algebraica, la Aritmética y la Física Teórica.

Finalmente deseamos agradecer al profesor Sebastián Xambó la invitación para realizar la exposición de este artículo en el seminario organizado por la Facultad de Matemáticas de la Universitat Politècnica de Catalunya

## 2. FUNCIONES THETA DE UNA VARIABLE

Sea  $H$  el semiplano superior complejo:

$$H = \{\tau \in \mathbb{C} : \text{Im } \tau > 0\}$$

---

Este trabajo está parcialmente financiado por los proyectos de investigación MTM2006-07618 de D.G.I. y SA112A07 de la Junta de Castilla y León.

La función theta es la función analítica en  $\mathbb{C} \times H$  definida por la siguiente serie absolutamente convergente:

$$\theta(z, \tau) = \sum_{n \in \mathbb{Z}} \exp(\pi i n^2 \tau + 2\pi i n z) \quad \text{con } z \in \mathbb{C}, \tau \in H$$

Esta función verifica las siguientes propiedades:

- $\theta(z + 1, \tau) = \theta(z, \tau)$
- $\theta(z + \tau, \tau) = \exp(-\pi i \tau - 2\pi i z) \cdot \theta(z, \tau)$
- $\theta(z + a\tau + b, \tau) = \exp(-\pi i a^2 \tau - 2\pi i a z) \cdot \theta(z, \tau)$

Otra de las relaciones fundamentales que satisface la función theta es la ecuación de calor:

$$\frac{\partial}{\partial t} \theta(z, it) = \frac{1}{4\pi} \frac{\partial^2}{\partial z^2} \theta(z, it)$$

Se verifica que la función theta está caracterizada por estas propiedades, es decir, periodicidad y ecuación del calor.

**2.1. Funciones theta y grupo de Heisenberg.** El grupo de Heisenberg de dimensión 3 se define como:

$$\mathcal{G} = \mathbb{C}_1^* \times \mathbb{R} \times \mathbb{R}, \quad \text{donde } \mathbb{C}_1^* = \{z \in \mathbb{C} : |z| = 1\}$$

con la siguiente ley de grupo:

$$(\lambda, a, b) \cdot (\lambda', a', b') = (\lambda\lambda' \exp(2\pi i b a'), a + a', b + b')$$

Se verifica que el centro del grupo de Heisenberg es  $\mathbb{C}_1^*$  que coincide con el subgrupo conmutador.

El teorema de Stone-von Neumann afirma que existe una única representación unitaria e irreducible de  $\mathcal{G}$  en la cual  $\mathbb{C}_1^*$  actúa por la identidad. Esta representación se puede realizar de la siguiente forma. Sea  $\mathcal{E}$  el espacio de las funciones enteras (funciones holomorfas en todo el plano complejo) con la siguiente norma definida a partir de un elemento  $\tau \in H$ :

$$\|f\|^2 = \int_{\mathbb{C}} \exp(-2\pi y^2 / \text{Im } \tau) \cdot |f(x + iy)|^2 dx dy$$

Denotaremos por  $\mathcal{H}_\tau$  el subespacio de  $\mathcal{E}$  formado por las funciones con norma finita. Se define la siguiente representación unitaria de  $\mathcal{G}$  en  $\mathcal{H}_\tau$ :

$$U: \mathcal{G} \longrightarrow \text{Aut}(\mathcal{H}_\tau)$$

$$(\lambda, a, b) \longmapsto U_{\lambda, a, b}$$

con  $(U_{\lambda, a, b}f)(z) = \lambda \exp(\pi i a^2 \tau + 2\pi i a z) f(z + a\tau + b)$ . Se verifica que  $\mathcal{H}_\tau$  es un  $\mathcal{G}$ -módulo irreducible, por lo que es una realización de la única representación irreducible de  $\mathcal{G}$ .

La función theta admite la siguiente caracterización en términos de la acción del grupo de Heisenberg  $\mathcal{G}$ : *La función theta  $\theta(z, \tau)$  es la única función holomorfa en  $\mathcal{H}_\tau$  invariante por la acción del siguiente subgrupo de  $\mathcal{G}$ :*

$$\Gamma = \{(1, a, b) \in \mathcal{G} : a, b \in \mathbb{Z}\} \subset \mathcal{G}$$

A continuación vamos a introducir el análogo finito al grupo de Heisenberg y le teorema de Stone-von Neumann. Para cada número entero positivo  $l$  se define el siguiente subgrupo de  $\Gamma$ :

$$\Gamma_l = \{(1, la, lb) : a, b \in \mathbb{Z}\} \subset \Gamma$$

Sea  $V_l$  el subespacio vectorial de  $\mathcal{H}_\tau$  formado por las funciones enteras invariantes por la acción de  $\Gamma_l$ . Se verifica que una función  $f(z) \in V_l$  si y sólo si es de la forma

$$f(z) = \sum_{n \in \frac{1}{l}\mathbb{Z}} c_n \exp(\pi i n^2 \tau + 2\pi i n z)$$

siendo  $c_n = c_m$  si  $n - m \in l \cdot \mathbb{Z}$ . En particular se tiene que  $\dim_{\mathbb{C}} V_l = l^2$ .

Para cada entero  $l$ , sea  $\mu_l$  el subgrupo de  $\mathbb{C}_1^*$  formado por las raíces  $l$ -ésimas de la unidad. Se define el grupo de Heisenberg finito de nivel  $l$  como

$$\mathcal{G}_l = \mu_{l^2} \times \left(\frac{1}{l}\mathbb{Z}/(l \cdot \mathbb{Z})\right) \times \left(\frac{1}{l}\mathbb{Z}/(l \cdot \mathbb{Z})\right)$$

con la ley de grupo  $(\lambda, a, b) \cdot (\lambda', a', b') = (\lambda\lambda' \exp(2\pi i b a'), a + a', b + b')$ .

Se verifica que  $V_l$  es la única representación irreducible de  $\mathcal{G}_l$  en la que  $\mu_l$  actúa por homotecias (este enunciado es el análogo finito del teorema de Stone-von Neumann).

Una base de  $V_l$  está formada por las siguientes funciones:

$$\theta_{[b]}^{[a]}(z, \tau) = \sum_{n \in \mathbb{Z}} \exp(\pi i(a+n)^2 \tau + 2\pi i(n+a)(z+b))$$

con  $a, b \in \frac{1}{l}\mathbb{Z}$ . Las funciones  $\theta_{[b]}^{[a]}(z, \tau)$  se denominan funciones theta con características.

## 2.2. Función theta, función zeta y sus ecuaciones funcionales.

Dada una función  $f : \mathbb{R}^+ \rightarrow \mathbb{C}$  integrable y con apropiadas cotas en 0 y en el infinito, se define su transformada de Mellin como la siguiente función analítica:

$$(Mf)(s) = \int_0^\infty f(x) x^s \frac{dx}{x}, \text{ con } a < \operatorname{Re}(s) < b$$

La transformada de Mellin es simplemente una transformada de Fourier multiplicativa (haciendo el cambio de variable  $x = \exp(y)$  se convierte en una transformada de Fourier clásica). En particular, si consideramos la función  $f(x) = \theta(0, ix) - 1$  y  $\operatorname{Re}(s) > 1$ , se tiene que

$$\begin{aligned} (Mf)(1/2 s) &= 2 \sum_{n \in \mathbb{N}} \left( \int_0^\infty \exp(-\pi n^2 x) x^{\frac{1}{2}s} \frac{dx}{x} \right) = (y = \pi n^2 x) = \\ &= 2 \left( \sum_{n \in \mathbb{N}} (\pi n^2)^{-\frac{1}{2}s} \right) \int_0^\infty \exp(-y) y^{\frac{1}{2}s} \frac{dy}{y} = \\ &= 2\pi^{-\frac{1}{2}s} \left( \sum_{n \in \mathbb{N}} n^{-s} \right) \int_0^\infty \exp(-y) y^{\frac{1}{2}s} \frac{dy}{y} = \\ &= 2\pi^{-\frac{1}{2}s} \zeta(s) \Gamma(1/2 s) \end{aligned}$$

donde  $\zeta(s) = \sum_{n \in \mathbb{N}} n^{-s}$ , con  $\operatorname{Re}(s) < 1$ , es la función zeta de Riemann y  $\Gamma(1/2 s) = \int_0^\infty \exp(-y) y^{\frac{1}{2}s} \frac{dy}{y}$ . Por tanto, se tiene la fórmula fundamental:

$$(2.1) \quad 2\pi^{-\frac{1}{2}s} \zeta(s) \Gamma(1/2 s) = \int_0^\infty (\theta(0, ix) - 1) x^s \frac{dx}{x}$$

Ahora bien, si  $a, b, c, d \in \mathbb{Z}$  son números impares tales que  $ad - bc = 1$ , se tiene que

$$\theta\left(\frac{z}{c\tau + d}, \frac{a\tau + b}{c\tau + d}\right) = \mu(c\tau + d)^{1/2} \exp(\pi icz^2/(c\tau + d)) \theta(z, \tau)$$

siendo  $\mu$  una raíz octava de la unidad. En particular, se obtiene la ecuación funcional de la función theta:

$$(2.2) \quad \theta(0, i/x) = x^{1/2} \theta(0, ix) \quad \text{con } x \in \mathbb{R}^+$$

Por otro lado, si  $\xi(s) = \pi^{-\frac{1}{2}s} \zeta(s) \Gamma(1/2 s)$ , la ecuación funcional de la función zeta de Riemann es

$$\xi(s) = \xi(1 - s)$$

Esta igualdad se obtiene por la igualdad (2.1), aplicando la transformada de Mellin a la ecuación funcional (2.2) de la función theta, es decir, las ecuaciones funcionales de la función theta y de la función zeta son equivalentes.

**2.3. Función theta y toros complejos de dimensión uno.** Dado  $\tau \in H$ , se define el siguiente toro complejo:

$$E_\tau = \mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$$

Las funciones  $\theta\left[\frac{a}{b}\right](z, \tau)$  no definen funciones holomorfas en  $E_\tau$  porque no son periódicas respecto del retículo  $\mathbb{Z} + \tau\mathbb{Z}$ . Ahora bien las propiedades de periodicidad vistas para estas funciones se pueden interpretar geoméricamente en el sentido de que  $\theta\left[\frac{a}{b}\right](z, \tau)$  son secciones holomorfas de determinados fibrados de línea sobre el toro complejo  $E_\tau$ .

Por medio de las funciones theta se pueden construir explícitamente funciones meromorfas en  $E_\tau$  como productos de cocientes de las funciones theta y de sus trasladados, como por ejemplo, la función  $\wp$  de Weierstrass:

$$\wp(z) = \lambda + \frac{\theta\left[\frac{1}{1}\right]'(0, \tau)^2}{\theta\left[\frac{0}{1}\right](0, \tau)^2} \cdot \frac{\theta\left[\frac{0}{1}\right](z, \tau)^2}{\theta\left[\frac{1}{1}\right](z, \tau)^2} = -\frac{d^2}{dz^2} \log \theta\left[\frac{1}{1}\right](z, \tau) + \lambda$$

donde la constante  $\lambda$  se ajusta para que el desarrollo en serie de la función en  $z = 0$  no tenga término constante.

### 3. FUNCIONES THETA DE VARIAS VARIABLES. VARIEDADES ABELIANAS

Análogamente al semiplano superior, en varias variables se define el semiplano superior de Siegel:

$$\mathbb{H}_g = \{\tau \in \mathbb{M}(g \times g, \mathbb{C}) : \tau^t = \tau, \text{ Im } \tau \text{ es definida positiva}\}$$

Una variedad abeliana de dimensión  $g$  es un toro complejo de la forma

$$X(\tau) = \mathbb{C}^g / (\mathbb{Z}^g + \tau\mathbb{Z}^g) \text{ con } \tau \in \mathbb{H}_g$$

Además, para cada  $\tau \in \mathbb{H}_g$  se define la función theta en  $g$  variables como la función holomorfa en  $\mathbb{C}^g$  dada por la siguiente serie absolutamente convergente:

$$\theta(z, \tau) = \sum_{n \in \mathbb{Z}^g} \exp(\pi i n^t \tau n + 2\pi i n^t z)$$

con  $z = (z_1, \dots, z_g) \in \mathbb{C}^g$ . Análogamente, dado  $(a, b) \in \mathbb{R}^g \times \mathbb{R}^g$  se define la función theta con característica  $\begin{bmatrix} a \\ b \end{bmatrix}$  como:

$$\theta_{\begin{bmatrix} a \\ b \end{bmatrix}}(z, \tau) = \sum_{n \in \mathbb{Z}^g} \exp\left(\pi i (n + a)^t \tau (n + a) + 2\pi i (n + a)^t (z + b)\right)$$

Estas funciones theta también tienen propiedades de cuasi-periodicidad:

$$\theta_{\begin{bmatrix} a \\ b \end{bmatrix}}(z + m, \tau) = \exp(2\pi i a^t m) \theta_{\begin{bmatrix} a \\ b \end{bmatrix}}(z, \tau)$$

$$\theta_{\begin{bmatrix} a \\ b \end{bmatrix}}(z + \tau m, \tau) = \exp(-2\pi i b^t m) \exp(-\pi i m^t \tau m - 2\pi i m^t z) \theta_{\begin{bmatrix} a \\ b \end{bmatrix}}(z, \tau)$$

Por estas propiedades, las funciones theta definen secciones holomorfas de fibrados de línea sobre la variedad abeliana  $X(\tau)$ . En particular, la función theta  $\theta(z, \tau)$  es sección holomorfa de un fibrado de línea  $L_\Theta$  cuyas secciones holomorfas tienen dimensión compleja 1, por lo que define una polarización principal en  $X(\tau)$ .

El grupo  $\Gamma_g = \text{Sp}(2g, \mathbb{Z})$  actúa en  $\mathbb{H}_g$  del siguiente modo:

$$\gamma(\tau) = (A\tau + B) \cdot (C\tau + D)^{-1} \text{ con } \gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \text{Sp}(2g, \mathbb{Z})$$

Se verifica que dos variedades abelianas  $X(\tau)$  y  $X(\tau')$  son isomorfas si y sólo si existe  $\gamma \in \Gamma_g$  tal que  $\tau' = \gamma(\tau)$ . En consecuencia, la variedad que clasifica las variedades abelianas principalmente polarizadas

de dimensión  $g$ , es decir, la variedad de moduli de variedades abelianas principalmente polarizadas es:

$$\mathcal{A}_g = \mathbb{H}_g / \Gamma_g \quad \text{con } \dim \mathcal{A}_g = g(g+1)/2$$

**3.1. Funciones theta de superficies de Riemann compactas.**

**Variedad jacobiana.** Sea  $S$  una superficie de Riemann compacta de género  $g$ . Fijamos una base canónica  $\{\alpha_1, \dots, \alpha_g, \beta_1, \dots, \beta_g\}$  de  $H^1(S, \mathbb{Z}) \simeq \mathbb{Z}^{2g}$  tal que la intersección definida a través del producto cup es:

$$\alpha_i \cdot \alpha_j = \beta_i \cdot \beta_j = 0 \quad \alpha_i \cdot \beta_j = \delta_{ij}$$

Sea  $\mathcal{O}_S$  el haz de funciones holomorfas sobre  $S$ ,  $\mathcal{O}_S^*$  el haz de funciones holomorfas sin ceros en  $S$  y  $\Omega_S$  el haz de diferenciales holomorfas en  $S$ .

La sucesión exacta de grupos

$$0 \longrightarrow \mathbb{Z} \longrightarrow \mathbb{C} \longrightarrow \mathbb{C}^* \longrightarrow 1$$

$$z \longmapsto \exp(2\pi iz)$$

induce una sucesión exacta de grupos de cohomología:

$$0 \longrightarrow H^1(S, \mathbb{Z}) \xrightarrow{j} H^1(S, \mathcal{O}_S) \longrightarrow H^1(S, \mathcal{O}_S^*) \xrightarrow{deg} H^2(S, \mathbb{Z}) \longrightarrow 0$$

$$\begin{array}{ccccccc} & & \parallel & & \parallel & & \\ & & \mathbb{Z}^{2g} & & H^0(S, \Omega_S)^* & & \mathbb{Z} \end{array}$$

donde el morfismo  $j: H^1(S, \mathbb{Z}) \rightarrow H^1(S, \mathcal{O}_S)$  esta dado por  $j(\alpha)(\omega) = \int_{\alpha} \omega$ .

Por otro lado, se tiene que el grupo de Picard de  $S$  es:

$$\text{Pic}(S) = H^1(S, \mathcal{O}_S^*) \simeq \{\text{fibrados de línea en } S\} / \{\text{isomorfismos}\} \simeq \{\text{divisores en } S\} / \{\sim\}$$

donde dos divisores están relacionados si difieren en el divisor de ceros y polos de una función meromorfa. Teniendo en cuenta esta identificación, el morfismo  $deg$  de la sucesión exacta anterior viene dado por:

$$\text{Pic}(S) = H^1(S, \mathcal{O}_S^*) \xrightarrow{deg} \mathbb{Z}$$

$$\left[ D = \sum_{\text{finita}} n_s \cdot s \right] \longmapsto deg D = \sum n_s$$

Se define la variedad jacobiana de la superficie de Riemann  $S$  como:

$$J(S) = \text{Pic}^0(S) = \{\text{fibrados de línea de grado cero sobre } S\}$$

Sea  $\{\omega_1, \dots, \omega_g\}$  la base de diferenciales holomorfas de  $S$  normalizada por la base de ciclos elegida  $\{\alpha_1, \dots, \alpha_g, \beta_1, \dots, \beta_g\}$  por la condición:

$$\int_{\alpha_k} \omega_j = 2\pi i \delta_{kj}$$

Si  $\tau_{ij} = \int_{\beta_i} \omega_j \in \mathbb{C}$ , se verifica que la matriz  $\tau = (\tau_{ij})$  es una matriz simétrica de parte imaginaria definida positiva, es decir,  $\tau \in \mathbb{H}_g$ . Por tanto, por la sucesión exacta de cohomología anterior, se tiene que

$$J(S) \simeq H^0(S, \Omega)^* / H^1(S, \mathbb{Z}) = \mathbb{C}^g / (\mathbb{Z}^g + \tau \mathbb{Z}^g)$$

es decir, la jacobiana de  $S$  es una variedad abeliana principalmente polarizada. La matriz  $\tau$  se denomina matriz de periodos de la superficie de Riemann compacta  $S$ . Las funciones theta asociadas a esta matriz son conocidas como las funciones theta de la superficie de Riemann.

Sea  $\mathcal{M}_g$  la variedad de moduli de las superficies de Riemann compactas de género  $g$ , que es una variedad analítica de dimensión:

$$\dim \mathcal{M}_g = \begin{cases} 0 & \text{si } g = 0 \\ 1 & \text{si } g = 1 \\ 3g - 3 & \text{si } g > 1 \end{cases}$$

La construcción anterior de la matriz de periodos define una aplicación:

$$\begin{aligned} \mathcal{M}_g &\longrightarrow \mathcal{A}_g \\ S &\longmapsto J(S) \end{aligned}$$

que es inyectiva por el teorema de Torelli. Por dimensiones se verifica que  $\mathcal{M}_g = \mathcal{A}_g$  si  $g \leq 3$ , y  $\mathcal{M}_g \subsetneq \mathcal{A}_g$  para  $g \geq 4$ .

El problema clásico de Schottky consiste en calcular las ecuaciones de la imagen de  $\mathcal{M}_g$  en  $\mathcal{A}_g$ , y su estudio ha llevado a diferentes caracterizaciones de las jacobianas dentro del moduli de variedades abelianas principalmente polarizadas.



4. FÓRMULAS DE ADICIÓN DE FUNCIONES THETA

Para cualquier  $\tau \in \mathbb{H}_g$ , las funciones theta verifican unas relaciones (fórmulas) de adición (Riemann-Weierstrass) que describen analíticamente la ley de grupo de la variedad abeliana  $X(\tau)$ :

$$\theta(2x, 2\tau) \cdot \theta(2y, 2\tau) = 2^{-g} \sum_{2\sigma \in (\mathbb{Z}/2\mathbb{Z})^g} \theta_{[\sigma]}^0(x+y, \tau) \cdot \theta_{[\sigma]}^0(x-y, \tau)$$

y su inversión:

$$\theta(x+y, \tau) \cdot \theta(x-y, \tau) = \sum_{2\sigma \in (\mathbb{Z}/2\mathbb{Z})^g} \theta_{[\sigma]}^0(2x, 2\tau) \cdot \theta_{[\sigma]}^0(2y, 2\tau)$$

Si  $\tau$  es la matriz de periodos de una superficie de Riemann compacta, entonces sus funciones theta verifican ciertas ecuaciones que no satisfacen las funciones theta asociadas a elementos arbitrarios de  $\mathbb{H}_g$ . Una de ellas es la fórmula de adición de funciones theta de superficies de Riemann:

El morfismo de Abel definido por un punto  $p_0 \in S$  es:

$$A: S \longrightarrow J(S) = C^g / (\mathbb{Z}^g + \tau\mathbb{Z}^g)$$

$$p \longmapsto A(p) = \left( \int_{p_0}^p \omega_1, \dots, \int_{p_0}^p \omega_g \right)$$

que como clase de equivalencia de divisores de grado cero consiste en  $p \mapsto [p - p_0]$ . Por aditividad, se define el siguiente morfismo:

$$m: S^{2n} \longrightarrow J(S)$$

$$(x_1, y_1, \dots, x_n, y_n) \longmapsto \left[ \sum_{i=1}^n (x_i - y_i) \right]$$

Sea  $E(x, y)$  la prime-form de  $S$  construida a partir de las funciones theta de  $S$  y de sus diferenciales holomorfas, que es una sección holomorfa hemisimétrica del fibrado diagonal sobre  $S \times S$  con ceros simples a lo largo de la diagonal.

*Fórmula de adición para funciones theta de superficies de Riemann compactas:* Si  $e \in \mathbb{C}^g$  con  $\theta(e, \tau) \neq 0$  se verifica que:

$$\theta\left(\sum_{i=1}^n (x_i - y_i) - e, \tau\right) \theta(e, \tau)^{n-1} \frac{\prod_{i < j} E(x_i, x_j) E(y_j, y_i)}{\prod_{i, j} E(x_i, y_j)} = \det\left(\frac{\theta((x_i - y_j) - e, \tau)}{E(x_i, y_j)}\right)$$

para todo  $(x_1, y_1, \dots, x_n, y_n) \in S^{2n}$ , donde esta igualdad debe de ser entendido como igualdad de secciones holomorfas de ciertos fibrados de línea en  $S^{2n}$ . En el caso particular  $n = 2$ , esta fórmula es conocida como la *identidad trisecante de Fay*:

$$\begin{aligned} & \theta(x_1 + x_2 - y_1 - y_2 - e, \tau) \theta(e, \tau) E(x_1, x_2) E(y_1, y_2) = \\ & = \theta(x_1 - y_1 - e, \tau) \theta(x_2 - y_2 - e, \tau) E(x_1, y_2) E(y_1, x_2) + \\ & \quad + \theta(x_1 - y_2 - e, \tau) \theta(x_2 - y_1 - e, \tau) E(x_1, y_1) E(x_2, y_2) \end{aligned}$$

para todo  $x_1, x_2, y_1, y_2 \in S$ .

El nombre de esta fórmula se debe a su interpretación geométrica. Sea  $X(\tau)$  una variedad abeliana principalmente polarizada y consideramos la aplicación:

$$\begin{aligned} X(\tau) & \longrightarrow \mathbb{P}^N & (N = 2^g - 1) \\ z & \longmapsto \vec{\theta}_2(z, \tau) = \{\theta_0^\sigma(2z, 2\tau)\}_{\sigma \in (\mathbb{Z}/2\mathbb{Z})^g} \end{aligned}$$

De su definición se deduce que la imagen de un punto y su opuesto es la misma y dicha imagen es conocida con el nombre de la variedad de Kummer de  $X(\tau)$ . Si  $X(\tau)$  es la jacobiana de una superficie de Riemann compacta  $S$ , teniendo en cuenta las fórmulas de adición vistas al principio de esta sección para las funciones theta arbitrarias, la identidad trisecante de Fay implica que existen rectas en  $\mathbb{P}^N$  que cortan a la variedad de Kummer de  $J(S)$  en tres puntos o, equivalentemente, la variedad de Kummer de  $J(S)$  admite rectas trisecantes en  $\mathbb{P}^N$ . La conjetura trisecante afirma que esta propiedad caracteriza a las jacobianas y ha sido recientemente demostrada por I. Krichever.

Considerando el caso degenerado de la identidad trisecante de Fay (los tres puntos coinciden) se obtiene la *ecuación KP* (Kadomtsev-Petviashvili): sea  $\tau$  la matriz de periodos de una superficie de Riemann

compacta  $S$ , entonces existen tres derivaciones

$$D_i = \sum_{j=1}^g v_{ij} \frac{\partial}{\partial z_j} \quad i = 1, 2, 3$$

tales que:

$$\left( D_1^4 - D_1 D_3 + \frac{3}{4} D_2^2 + \lambda \right) \vec{\theta}_2(z, \tau)|_{z=0} = 0$$

siendo  $\lambda$  una constante. Shiota demostró que esta ecuación caracteriza a las funciones theta de superficies de Riemann compactas.

#### REFERENCIAS

- [FK] Farkas, H. M.; Kra, I., “Riemann surfaces”. Second edition. Graduate Texts in Mathematics 71. Springer-Verlag (1992)
- [F] Fay, J.D., “Theta Functions on Riemann Surfaces”, LNM 352, Springer-Verlag (1973)
- [G] Gunning, R. C. “Lectures on Riemann surfaces, Jacobi varieties”. Mathematical Notes 12. Princeton University Press (1972)
- [K] Krichever, I.M., “Characterizing Jacobians via trisecants of the Kummer Variety”, arXiv:math/0605625v4.
- [M] Mumford, D., “Tata Lectures on Theta I”, Birkhäuser (1983)
- [Sh] Shiota, T., “Characterization of Jacobian varieties in terms of soliton equations”, Invent. Math. **83** (1986), pp. 333–382.



## LA TESIS DE RIEMANN SOBRE LAS SERIES TRIGONOMÉTRICAS

ANTONIO CÓRDOBA BARBA

Bernhard Riemann, a pesar de su corta vida, está generalmente considerado como uno de los más universales, fecundos y originales creadores de todos los tiempos, ocupando, junto a los Arquímedes, Newton, Euler, Gauss, Hilbert y Poincaré, la posición más excelsa del Olimpo de las Matemáticas. Su obra, al mismo tiempo clásica y revolucionaria, ha tenido, y sigue aún teniendo, una influencia profunda en muchas y variadas áreas: en Geometría y en Teoría de los Números, por supuesto, pero también en Física y en el Análisis Matemático.

Si juzgamos por su tesis doctoral del año 1851 en la universidad de Göttingen: “Fundamentos para una teoría general de las funciones de una variable compleja” (*Grundlagen für eine allgemeine Theorie der Functionen einer veränderlichen complexen Grösse*), o por su primera tesis de Habilitación presentada en la misma universidad, tres años después, y titulada: “Sobre la representabilidad de una función mediante una serie trigonométrica” (*Ueber die Darstellbarkeit einer Function durch eine trigonometrische Reihe*), podríamos afirmar que, al menos durante los comienzos de su carrera, Riemann era, mayormente, un analista.

Cumpliendo con los requisitos de Göttingen presentó también una segunda Tesis de Habilitación titulada “Sobre las hipótesis en las que se funda la Geometría” (*Ueber die Hypothesen welche der Geometrie zu Grunde liegen*), siendo una anécdota muy conocida que Gauss, rompiendo con la tradición establecida en aquella Universidad, le pidió defender públicamente esta segunda opción, que tanta influencia ha tenido luego en el desarrollo y evolución posteriores de nuestras nociones de espacio y de estructura geométrica. Habiéndose convertido, con toda justicia, en un hito famoso de la historia de nuestra ciencia.

Pero la primera tesis, aunque sea menos popular y esté quizás un tanto oscurecida por la fama de la segunda, es también una maravilla. La preparación de este trabajo me ha brindado la oportunidad de volver a leerla con cuidado, siendo mi propósito compartir esa lectura con Uds. Pero a sabiendas de que mis comentarios no podrán nunca hacerle justicia, aunque consideraré un éxito que mis reflexiones puedan servir para animarles a estudiarla y aprender directamente de tan magnífica obra.

La tesis está estructurada en cuatro partes. En la primera describe de forma breve, pero precisa y amena, la historia y los antecedentes del problema de la representación de funciones “arbitrarias” por medio de series trigonométricas. Riemann detecta a los personajes fundamentales de la trama, a saber: D’Alembert, Euler, Bernoulli (Daniel), Lagrange, Fourier y Dirichlet. En un lenguaje muy claro nos ilustra de la cuestión en querella y de las contribuciones y puntos de vista de cada uno de esos artistas. En la parte segunda inicia su propio camino, que le conduce a la definición de la integral que ahora lleva su nombre y a la caracterización de las funciones que son integrables, presentando varios ejemplos que muestran la potencia, y las limitaciones, de la nueva definición. En ese empeño se demora en aclarar la diferencia entre continuidad y diferenciabilidad, ofreciéndonos los primeros ejemplos conocidos de funciones continuas que carecen de derivada en un conjunto denso de puntos, surgiendo naturalmente la pregunta: ¿habrá una función continua que no sea derivable en todos sus puntos?

En la tercera parte Riemann plantea el problema de la unicidad de los desarrollos trigonométricos y resuelve su caso más básico. Pero para hacerlo se ve obligado, entre otros logros, a generalizar el concepto de derivada segunda de una función. Luego Cantor prosiguió este análisis, lo que le llevó a crear la teoría de conjuntos y a preguntarse sobre la existencia de un conjunto de números reales cuyo cardinal esté estrictamente contenido entre el de los naturales y el de todos los reales. Es decir, a formular la hipótesis del continuo, cuya naturaleza independiente de los otros axiomas fue dilucidada por P. Cohen, en torno al año 1960. Por cierto que Cohen había realizado su tesis doctoral [10] en la Universidad de Chicago, dirigida por A. Zygmund, y ésta versó, precisamente, sobre el problema de la unicidad formulado por

Riemann. Finalmente, la parte cuarta de la tesis contiene varios ejemplos que nos ilustran sobre la complejidad y la diversidad de las series trigonométricas.

A lo largo de la memoria aparece repetidamente el empeño de tratar “funciones arbitrarias” e ir más allá de lo obtenido por Dirichlet y otros autores anteriores. Para entender el contexto y la presentación histórica, conviene pues considerar el concepto de función que se tenía en los primeros escritos de Euler y Lagrange (aunque en los postreros lo modificaron un poco, acercándose al actual) y su evolución hasta la época de Riemann, lo que queda patente en los textos siguientes:

• L. Euler (1748): *Institutiones calculi differentialis*.

1. *Quantitas constans est quantitas determinata perpetuo eundem valorem servans.*
2. *Quantitas variabilis est quantitas indeterminata seu universalis, quae omnes omnino valores determinatos in se complectitur.*
3. *Functio quantitatis variabilis est expressio analytica quomodunque composita ex illa quantitate variabili et numeris seu quantitibus constantibus.*

Que podemos traducir así:

1. Una cantidad constante es una cantidad determinada que mantiene permanentemente el mismo valor.
2. Una cantidad variable es una cantidad universal o indeterminada que contiene en sí misma todos los valores determinados.
3. Una función de una cantidad variable es una expresión analítica compuesta de cualquier manera a partir de esa cantidad variable y de números o constantes.

• J. L. Lagrange (1797): *Théorie des fonctions analytiques*.

*On appelle fonction d'une ou de plusieurs quantités toute expression de calcul dans laquelle ces quantités entrent d'une manière quelconque, mêlées ou non avec d'autres quantités qu'on regarde comme ayant des valeurs données et invariables, tandis que les quantités de la fonction peuvent recevoir toutes les valeurs possibles. Ainsi, dans les*

*fonctions, on ne considère que les quantités qu'on suppose variables, sans aucun égard aux constantes qui peuvent y être mêlées.*

Se llama función de una o varias cantidades a toda expresión de cálculo en la que estas cantidades entren de una manera arbitraria, mezcladas o no con otras cantidades a las que se considera teniendo valores dados e invariables, mientras que las cantidades de la función pueden tomar todos los valores posibles. Así, en las funciones, no se consideran más que las cantidades que se han supuesto variables, sin tener en cuenta a las constantes que puedan aparecer allí mezcladas.

- B. Riemann (1851): *Grundlagen für eine allgemeine Theorie der Functionen einer veränderlichen complexen Grösse.*

*Denkt man sich unter  $z$  eine veränderliche Grösse, welche nach und nach alle möglichen reellen Werthe annehmen kann, so wird, wenn jedem ihrer Werthe ein einziger Werth der unbestimmten Grösse  $w$  entspricht,  $w$  eine Function von  $z$  genannt, [...] Diese Definition setzt offenbar zwischen den einzelnen Werthen der Function durchaus kein Gesetz fest, indem, wenn über diese Function für ein bestimmtes Intervall verfügt ist, die Art ihrer Fortsetzung ausserhalb desselben ganz der Willkür überlassen bleibt. [...] Es ist daher einerlei, ob man die Abhängigkeit der Grösse  $w$  von der Grösse  $z$  als eine willkürlich gegebene oder als eine durch bestimmte Grössenoperationen bedingte definiert.*

Supongamos que  $z$  es una cantidad variable que puede asumir, gradualmente, todos los valores reales posibles, entonces, si a cada uno de sus valores corresponde un valor único de la cantidad indeterminada  $w$ , diremos que  $w$  es una función de  $z$ , [...] Esta definición no establece ninguna ley entre los distintos valores tomados por la función, con lo que si se conoce esta función en un cierto



intervalo, la forma de su continuación fuera de ese intervalo sigue siendo completamente arbitraria. [...] Es por lo tanto indiferente definir la dependencia de la cantidad  $w$  respecto de la cantidad  $z$  de una manera arbitraria o por medio de ciertas operaciones entre las cantidades involucradas.

1. PARTE I DE LA TESIS: *Historia de la cuestión de la representabilidad de una función arbitraria por una serie trigonométrica*

Esta parte es una síntesis, ágil, clara y precisa, del estado de la cuestión que va a abordar en el resto de la memoria. El mejor comentario que podemos hacer es simplemente una selección de sus frases más significativas, junto, una vez más, a la recomendación de leerla directamente:

Las series que Fourier llama trigonométricas, es decir las series de la forma

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos(nx) + b_n \sin(nx)),$$

desempeñan un papel importante en aquellas partes de las matemáticas en las que aparecen funciones arbitrarias. Incluso puede afirmarse con motivo que los avances más esenciales en estas partes de las matemáticas, tan importantes para la física, han sido consecuencia de un mejor conocimiento de la naturaleza de estas series. Ya en las primeras investigaciones que condujeron a la consideración de funciones arbitrarias, se formuló la pregunta de si una tal función totalmente arbitraria puede ser expresada por medio de una serie como la anterior. Esto sucedió a mediados del siglo pasado, con ocasión de las investigaciones sobre las cuerdas vibrantes, en las que se ocuparon entonces los matemáticos más afamados. Sus concepciones sobre nuestro tema no pueden ser expuestas adecuadamente sin entrar en ese problema.

Riemann continúa escribiendo la ecuación de la cuerda vibrante

$$\frac{\partial^2 y}{\partial t^2} = \alpha^2 \frac{\partial^2 y}{\partial x^2}$$

y presenta la solución obtenida por D'Alembert, publicada en las Memorias de la Academia de Berlín (1747), y que consiste en la suma de una onda progresiva,  $f(x - \alpha t)$ , y otra regresiva  $g(x + \alpha t)$ . Y no tiene inconveniente en informarnos de que la solución de D'Alembert se obtiene a través del sencillo cambio de variables ( $u = x - \alpha t, v = x + \alpha t$ ), ahora tan conocido:

$$\frac{\partial^2 y}{\partial u \partial v} = 0.$$

Constituye una muestra del estilo de exposición directo, pedagógico y nada pedante que encontramos en sus escritos.

Aparte de esta ecuación, que se deduce de las leyes generales del movimiento,  $y(x, t)$  debe satisfacer todavía la condición de anularse en los puntos de sujeción de la cuerda:  $y(0, t) = y(l, t) = 0$ .

$$\left. \begin{array}{l} f(-\alpha t) + g(\alpha t) = 0 \\ f(l - \alpha t) + g(l + \alpha t) = 0 \end{array} \right\} \\ \implies f(z) = -g(-z) = -g(l - (l + z)) = f(2l + z)$$

Luego

$$\left\{ \begin{array}{l} y = f(x - \alpha t) - f(-x - \alpha t) \\ f(x) = f(x + 2l) \end{array} \right.$$

Una vez que D'Alembert hubo establecido esto para la solución general del problema, se ocupó en una continuación de su memoria de la ecuación  $f(z + 2l) = f(z)$ ; es decir, buscó expresiones analíticas que permanezcan invariantes cuando  $z$  aumenta en  $2l$ .

El comentario siguiente se refiere a Euler, quien publicó, también en las Memorias de la Academia de Berlín (1748), un artículo en el que desentraña la relación entre la función  $f$  de D'Alembert y los datos, posición y velocidad, que tenía la cuerda inicialmente. Pero pocos años más tarde, en 1753, y en las mismas Memorias, apareció una solución distinta debida a Daniel Bernoulli basada en la observación de que la

función

$$y = \sin\left(\frac{\pi nx}{l}\right) \cos\left(\frac{\pi n\alpha t}{l}\right),$$

donde  $n$  es un entero, verifica la ecuación y las condiciones de contorno:  $y(0, t) = y(l, t) = 0$ .

Sobre esta base explicó el hecho físico de que una cuerda pueda dar, además de su tono fundamental, también el tono fundamental de una cuerda de igual constitución pero de longitud  $l/2, l/3, l/4, \dots$  y consideró que su solución particular era general.

La observación de que una cuerda pueda dar sus diferentes tonos simultáneamente, llevó a Bernoulli a considerar que la cuerda (según la teoría) también podría vibrar de acuerdo con la ecuación:

$$y = \sum_n c_n \sin\left(\frac{\pi nx}{l}\right) \cos\left(\frac{\pi n\alpha}{l}(t - \beta_n)\right)$$

Y como todas las modificaciones observadas del fenómeno se podían explicar partiendo de esta ecuación, la consideró como la más general.

La polémica suscitada está bien descrita en la tesis y en ella terciaron, además de D'Alembert y Bernoulli, también Euler y Lagrange, entre otros. ¿Es posible representar una función arbitraria por medio de una serie trigonométrica? D'Alembert lo negaba, mientras que Bernoulli creía que sí. Riemann lleva a cabo un análisis de las publicaciones de Euler y de Lagrange para concluir que estos dos grandes matemáticos de la Ilustración daban más la razón al primero que al segundo, al menos en el caso de las funciones arbitrarias, puesto que para las analíticas, o analíticas a trozos, Lagrange manifestó algunas dudas.

En el párrafo 2 de esta primera parte aparece Joseph Fourier, a quien Riemann rinde tributo en los términos siguientes:

Transcurrieron casi cincuenta años sin que se lograra ningún avance esencial en la cuestión de la representabilidad analítica de las funciones arbitrarias. Entonces, una consideración de Fourier arrojó nueva luz sobre este

tema. Fourier indicó que en la serie trigonométrica

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)),$$

los coeficientes pueden ser determinados mediante las fórmulas

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx, \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx.$$

El siguiente personaje destacado por Riemann es su maestro Gustav Lejeune Dirichlet, quien antes de ser profesor en Göttingen había estudiado con Fourier y con Poisson en la Universidad de París:

Sólo en enero de 1829 apareció en el Journal de Crelle una memoria de Dirichlet, donde la cuestión de la representabilidad mediante series trigonométricas se decidía con todo rigor para el caso de funciones que admiten integración en todo el recorrido y que no tienen una cantidad infinita de máximos y mínimos.

Resulta enternecedor leer hoy los comentarios del gran Riemann acerca de la diferencia radical entre la convergencia absoluta y condicional de las series, pero que, sin ningún género de duda, era uno de los puntos claves de la cuestión dilucidada por Dirichlet:

La idea del camino a seguir para la solución de este problema le vino al comprender que las series infinitas se dividen en dos clases absolutamente distintas, según que, al hacer positivos todos sus miembros, permanezcan convergentes o no. En las primeras los miembros pueden ser reordenados a voluntad, pero el valor de las últimas depende del orden que les demos. [...] Las leyes de las sumas finitas sólo son aplicables a las series de la primera clase; sólo ellas pueden realmente ser consideradas como la colección de todos sus miembros, no las series de la segunda clase; circunstancia que había pasado inadvertida a los matemáticos del pasado siglo, [...] Obviamente, la serie de Fourier no pertenece necesariamente a la primera clase.

Sea

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

la serie de Fourier de una función  $f$ . Dirichlet escribió sus sumas parciales en la forma:

$$\begin{aligned} S_N f(x) &= \frac{1}{2}a_0 + \sum_{n=1}^N (a_n \cos(nx) + b_n \sin(nx)) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\alpha) \frac{\sin\left(\frac{2n+1}{2}(x-\alpha)\right)}{\sin\left(\frac{1}{2}(x-\alpha)\right)} d\alpha \end{aligned}$$

y demostró que convergen al valor  $\frac{1}{2}(f(x+0) + f(x-0))$  cuando la función  $f$  es monótona y continua en un número finito de intervalos. Dice Riemann:

Dirichlet basa su demostración en dos teoremas:

1. Si  $0 < c \leq \frac{\pi}{2}$ , entonces

$$\int_0^c \varphi(\beta) \frac{\sin((2n+1)\beta)}{\sin(\beta)} d\beta$$

se aproxima infinitamente al valor  $\frac{\pi}{2}\varphi(0)$  con  $n$  creciente.

2. Si  $0 < b < c \leq \frac{\pi}{2}$ , entonces

$$\int_b^c \varphi(\beta) \frac{\sin((2n+1)\beta)}{\sin(\beta)} d\beta$$

se aproxima infinitamente al valor 0 con  $n$  creciente.

Supuesto que la función  $\varphi(\beta)$ , entre los límites de estas integrales, sea monótona, creciente o decreciente.

Con la ayuda de estos dos teoremas se puede, obviamente, si la función  $f$  no pasa infinitas veces de aumentar a disminuir o de disminuir a aumentar, descomponer la integral

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f(\alpha) \frac{\sin\left(\frac{2n+1}{2}(x-\alpha)\right)}{\sin\left(\frac{1}{2}(x-\alpha)\right)} d\alpha$$

en una cantidad finita de miembros, de los cuales uno converge a  $f(x + 0)/2$ , otro a  $f(x - 0)/2$ , y los restantes convergen a 0, cuando  $n$  crece a infinito. De aquí se deduce que es representable, por medio de una serie trigonométrica, toda función que se repite periódicamente según el intervalo  $2\pi$  y que:

1. admite integración en todo su recorrido;
2. no tiene infinitos máximos y mínimos;
3. y toma, donde su valor cambia por saltos, el valor medio entre los límites por ambos lados.

Con el trabajo de Dirichlet obtuvieron un fundamento firme una gran cantidad de investigaciones analíticas importantes [...] Le fue dado resolver una cuestión que ocupara a tantos matemáticos distinguidos desde hacía más de setenta años. En realidad, para todos los casos de la Naturaleza, los únicos de que se trataba, quedó plenamente resuelta; pues por grande que sea nuestra falta de conocimiento acerca de cómo las fuerzas y estados de la materia varían según lugar y tiempo en lo infinitesimal, sin duda podemos siempre suponer que las funciones a las que no se extiende la investigación de Dirichlet no se dan en la Naturaleza.

En esto, como ahora sabemos, Riemann se equivocaba. Pero su error se justifica quizás por la devoción sentida hacia su maestro, que le indujo a sobreestimar la universalidad de los resultados obtenidos por Dirichlet. Mas a renglón seguido se enmienda con la siguiente afirmación:

Los casos no resueltos por Dirichlet merecen atención por dos razones simples: en primer lugar, como Dirichlet mismo menciona al final de su memoria, este asunto está en la más estrecha conexión con los principios del cálculo infinitesimal, y puede servir para traer dichos principios a una mayor claridad y precisión. Pero en segundo lugar, la utilidad de las series de Fourier no se limita a investigaciones físicas; hoy se ha aplicado con éxito también a un campo de la matemática pura, la teoría de los números, y aquí parecen ser de importancia precisamente

aquellas funciones cuya representabilidad mediante series trigonométricas no ha investigado Dirichlet.

## 2. PARTE II: *Sobre la noción de integral definida y el ámbito de su validez*

Habiendo pues planteado el problema y presentados sus antecedentes, el primer paso para ir más allá de lo obtenido por Dirichlet era dar sentido a las fórmulas de Fourier para funciones “arbitrarias”. Escribe Riemann:

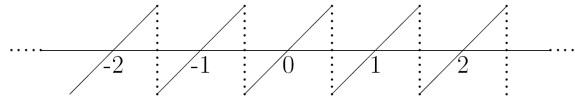
La incertidumbre que aún reina en algunos puntos fundamentales de la teoría de las integrales definidas, nos fuerza a anteponer algunas consideraciones sobre el concepto de integral definida y el ámbito de su validez: ¿qué hay que entender por  $\int_a^b f(x)dx$ ?

Para darle respuesta introduce la noción de integral de una función acotada a través de las sumas, ahora llamadas de Riemann, y da una condición necesaria y suficiente para que una tal función sea integrable, considerando la oscilación  $\omega_j$  de la función en un intervalo  $I_j$  de una partición dada del dominio de integración. La condición necesaria y suficiente para que las sumas converjan, y por tanto  $f$  sea integrable, es que la suma total de las longitudes de los intervalos de la partición donde la oscilación supera a un número positivo dado pueda hacerse arbitrariamente pequeña con el diámetro de la partición. Años después el criterio adquirió su versión actual equivalente:  $f$  es integrable (Riemann) si y sólo si el conjunto de sus puntos de discontinuidad tiene medida cero. Es un hecho notorio que todas las monografías que se han escrito desde entonces para explicar el cálculo diferencial introducen esta definición de integral.

A continuación, en la tesis se presenta un ejemplo de una función que es integrable en el nuevo sentido y que, sin embargo, es discontinua en un conjunto denso de puntos. Es decir, que se trata de una extensión genuina de la noción de Cauchy de integrales de funciones continuas, o continuas a trozos. Sea

$$(x) = \begin{cases} x - m, & \text{si } |x - m| < 1/2; \\ 0, & \text{si } x = m + 1/2 \end{cases} \quad m \in \mathbb{Z},$$

cuya gráfica es la función periódica en forma de dientes de sierra:



Entonces la función

$$f(x) = \sum_{n=1}^{\infty} \frac{(nx)}{n^2}$$

tiene sus discontinuidades en los puntos racionales de la forma  $a/2b$ ,  $\text{mcd}(a, 2b) = 1$ , que son densos en toda la recta real. Además, a partir de la fórmula de Euler

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6},$$

es un ejercicio sencillo comprobar que

$$f^+\left(\frac{a}{2b}\right) - f^-\left(\frac{a}{2b}\right) = -\frac{\pi^2}{8b^2}.$$

Pero  $f(x)$  está acotada por  $\frac{\pi^2}{6}$  y es integrable, por cuanto sus discontinuidades son numerables. La función

$$F(x) = \int_0^x f(t) dt$$

resulta ser continua:

$$|F(x+h) - F(x)| \leq \int_x^{x+h} |f(t)| dt \leq \frac{\pi^2}{6} h.$$

Sin embargo, carece de derivada precisamente en esos puntos  $a/2b$ ,  $\text{mcd}(a, 2b) = 1$ . Veamos:

$$\begin{aligned} \lim_{h \rightarrow 0, h > 0} \frac{F\left(\frac{a}{2b} + h\right) - F\left(\frac{a}{2b}\right)}{h} &= \lim_{h \rightarrow 0, h > 0} \frac{1}{h} \int_{a/2b}^{a/2b+h} f(t) dt = f^+\left(\frac{a}{2b}\right), \\ \lim_{h \rightarrow 0, h < 0} \frac{F\left(\frac{a}{2b} + h\right) - F\left(\frac{a}{2b}\right)}{h} &= \lim_{h \rightarrow 0, h < 0} \frac{1}{h} \int_{a/2b}^{a/2b+h} f(t) dt = f^-\left(\frac{a}{2b}\right). \end{aligned}$$

Surge la pregunta: ¿existirá una función continua que carezca de derivada en todos sus puntos?



3. PARTE III: *Sobre la posibilidad de representar una función por una serie trigonométrica, sin hacer ninguna hipótesis sobre la naturaleza de la función*

Mientras que los trabajos anteriores establecen proposiciones del tipo: “si una función goza de tal o cual propiedad, entonces puede ser desarrollada en serie trigonométrica”, nosotros nos proponemos la cuestión inversa: “si una función es desarrollable en una serie de Fourier, ¿qué podemos inferir sobre la función, sobre la variación de sus valores cuando el argumento cambia de forma continua?”

Las ideas y las técnicas introducidas por Riemann para abordar esta cuestión han tenido, y siguen teniendo, una gran influencia en el desarrollo del Análisis Matemático, muy por encima, quizás, de la importancia de los resultados concretos obtenidos en esta sección. Un ejemplo notable es la generalización de la noción de derivada (derivada segunda), para la que se aportan dos posibilidades.

La función continua  $F(x)$  tiene una derivada segunda en el punto  $x$  si existe el límite:

$$D^2F(x) = \lim_{h \rightarrow 0} \frac{F(x+h) + F(x-h) - 2F(x)}{h^2}.$$

Observemos que si  $F''(x)$  existe en el sentido de Newton y Leibniz, entonces tenemos que  $D^2F(x) = F''(x)$ . Pero el ejemplo

$$F(x) = \begin{cases} 0, & x = 0; \\ x^2 \sin \frac{1}{x}, & x \neq 0 \end{cases}$$

muestra que podemos tener  $D^2F(0)$ , mientras que  $F''(0)$  no está definida. Luego se trata de una genuina extensión de la noción de derivada segunda.

Riemann demuestra la proposición siguiente: si una función periódica  $f(x)$ , de periodo  $2\pi$ , puede ser representada por una serie trigonométrica, entonces existe una función continua  $F(x)$  tal que  $D^2F(x) = f(x)$  en todo punto.

Además, se verifica la siguiente identidad:

$$\int_a^b D^2 F(x) \varphi(x) dx = \int_a^b F(x) \varphi''(x) dx$$

para toda función  $\varphi$  con dos derivadas continuas y que se anule fuera de  $(a, b)$ .

Sería ocioso subrayar la importancia de esta noción, y su carácter precursor de las derivadas débiles de la teoría de distribuciones. La manera concreta en la que aparece la derivada débil es para obtener uno de los resultados notables de la tesis, el ahora llamado teorema de localización: la convergencia o divergencia de una serie trigonométrica

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

en un intervalo  $I$  depende sólo de la función

$$F(x) = \frac{1}{4}a_0 x^2 + \sum_{n=1}^{\infty} \frac{1}{n^2} (a_n \cos(nx) + b_n \sin(nx))$$

en ese intervalo.

Bajo la hipótesis  $|a_n| + |b_n| = o(1)$ , que Riemann deduce de la convergencia en todo  $x$ , demuestra que

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) - \frac{1}{2\pi} \int F(t) \varphi(t) \frac{d^2}{dt^2} \left( \frac{\sin\left(\frac{2N+1}{2}(x-t)\right)}{\sin\left(\frac{x-t}{2}\right)} \right) dt$$

tiende a 0 cuando  $N$  tiende a infinito, para toda función  $\varphi \in C_0^\infty(I)$  que sea idénticamente igual a 1 en un entorno del punto  $x$ .

Se debe a G. Cantor la detección del siguiente corolario del teorema de Riemann que es conocido como teorema de unicidad:

*Supongamos que la serie trigonométrica*

$$(*) \quad \frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

*converge al valor 0 en todo punto  $x \in [0, 2\pi]$ . Entonces todos los coeficientes son nulos.*

Observemos que si supiésemos de antemano que (\*) es la serie de Fourier de una función integrable ( $a_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos(nt) dt$ ,  $b_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin(nt) dt$ ), entonces la solución sería muy fácil. Lo que convierte al resultado de Riemann en algo delicado es el hincapié en que (\*) sea una serie trigonométrica general, de la que carecemos de información alguna sobre sus coeficientes.

Permitiéndonos tan solo la licencia de trastocar un poco el orden del razonamiento y describir algunos pasos con la notación contemporánea, la arquitectura de la demostración de Riemann sería la siguiente:

- Paso 1.- Se demuestra que la convergencia de la serie

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

en todo punto  $x \in E$  ( $E$  de medida positiva) implica que

$$\lim_{n \rightarrow \infty} (|a_n| + |b_n|) = 0.$$

- Paso 2.- Con una doble integración se obtiene la función continua

$$F(x) = \frac{1}{4}a_0x^2 - \sum_{n=1}^{\infty} \frac{1}{n^2} (a_n \cos(nx) + b_n \sin(nx)).$$

- Paso 3.- La hipótesis

$$\lim_{N \rightarrow \infty} S_N(x) = \lim_{N \rightarrow \infty} \left( \frac{1}{2}a_0 + \sum_{n=1}^N (a_n \cos(nx) + b_n \sin(nx)) \right) = 0$$

implica que  $D^2F(x) = 0$ , para todo  $x \in [0, 2\pi]$ , donde  $D^2F(x)$  designa a la derivada segunda generalizada de Riemann:

$$\begin{aligned} D^2F(x) &= \lim_{h \rightarrow 0} \frac{F(x+h) + F(x-h) - 2F(x)}{h^2} \\ &= \lim_{h \rightarrow 0} \left[ \frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) \left( \frac{\sin(nh/2)}{nh/2} \right)^2 \right]. \end{aligned}$$

- Paso 4.- Se comprueba que si  $D^2G(x)$  existe y es estrictamente positiva en un intervalo, entonces  $G$  es convexa.

Análogamente, si  $D^2G(x) < 0$  en un intervalo, entonces  $G$  es cóncava.

Finalmente, si  $D^2G(x) = 0$  en todo  $x$ , entonces  $G(x) + \varepsilon x^2$  es convexa para todo  $\varepsilon > 0$ , luego también lo es  $G$  por ser un límite uniforme de funciones convexas. De manera análoga,  $G(x) - \varepsilon x^2$  es cóncava y tomando límites cuando tiende a 0, obtenemos que  $G$  es cóncava.

Conclusión:  $D^2G \equiv 0$  en un intervalo implica que  $G$  es cóncava y convexa, luego es lineal.

- Paso 5.- De los pasos anteriores obtenemos que

$$F(x) = \frac{1}{4}a_0x^2 - \sum_{n=1}^{\infty} \frac{1}{n^2}(a_n \cos(nx) + b_n \sin(nx))$$

es lineal. En particular eso implica que  $a_0 = 0$  y

$$\sum_{n=1}^{\infty} \frac{1}{n^2}(a_n \cos(nx) + b_n \sin(nx)) \equiv 0$$

Pero esta última es la serie de Fourier de la función idénticamente nula y, por tanto,  $a_n = b_n = 0$  para todo  $n \geq 1$ .

Hasta la fecha no existe otra demostración, distinta de la dada por Riemann, del hecho fundamental de que si dos series trigonométricas convergen puntualmente al mismo valor, entonces son necesariamente idénticas. En el caso de varias variables existe la variante de tomar sumas parciales de diversos modos.

Cuando se consideran sumas esféricas, un resultado reciente de J. Bourgain [14] demuestra que el teorema de Riemann sigue siendo válido. Pero si consideramos las sumas en cubos la cuestión está todavía por decidir, concretamente:

En dimensión  $n \geq 2$ , tenemos series

$$f \sim \sum_{\nu \in \mathbb{Z}^n} a_{\nu} e^{i\nu \cdot x}, \quad \nu = (\nu_1, \dots, \nu_n).$$

Con la notación  $\|\nu\|_{\infty} = \max(|\nu_1|, \dots, |\nu_n|)$ , podemos escribir las sumas parciales cúbicas

$$S_N f(x) = \sum_{\|\nu\|_{\infty} \leq N} a_{\nu} e^{i\nu \cdot x}$$

**Problema abierto:** si existe  $\lim_{N \rightarrow \infty} S_N f(x) = 0$  en todo  $x \in [0, 2\pi]^n$ , ¿han de ser todos los coeficientes nulos necesariamente?

Como indicamos antes, Cantor se interesó por la tesis de Riemann y extendió el teorema de unicidad en el sentido siguiente: supongamos que la convergencia a cero de la serie trigonométrica la conocemos en todos los puntos salvo, quizás, por un conjunto finito para los que carecemos de información:

$$\lim_{N \rightarrow \infty} \left[ \frac{1}{2}a_0 + \sum_{n=1}^N (a_n \cos(nx) + b_n \sin(nx)) \right] = 0$$

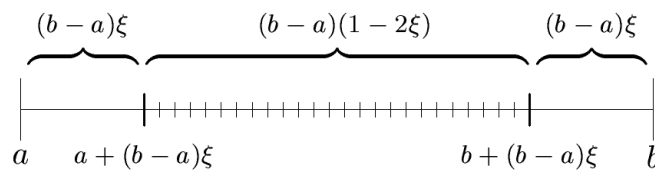
para todo  $x \in [0, 2\pi] - \{x_1, \dots, x_p\}$ . Cantor demostró que, también en este caso, la serie de partida ha de tener todos sus coeficientes nulos. ¿Qué ocurre si eliminamos un conjunto infinito? Se trata de una pregunta natural, pero muy difícil, que da lugar a una interesante definición. Diremos que  $U \subset [0, 2\pi]$  es un *conjunto de unicidad* si toda serie trigonométrica que converge puntualmente a 0 en el complemento de  $U$  ( $[0, 2\pi] - U$ ) ha de tener, necesariamente, todos sus coeficientes nulos. Con los métodos analíticos actuales resulta fácil comprobar que un conjunto de unicidad es de medida (Lebesgue) igual a cero. Pero el recíproco es falso: hay conjuntos de medida cero que no son de unicidad. Y esto es un hecho por lo menos inquietante, por cuanto implica la existencia de series trigonométricas que convergen en casi todo punto a una función integrable  $f$  sin coincidir con su serie de Fourier. En estos comienzos del siglo XXI sigue siendo un problema abierto caracterizar a los conjuntos de unicidad. No obstante Cantor demostró un resultado muy notable: una condición suficiente para que  $U$  sea de unicidad es que  $U_n$ , el conjunto derivado de orden  $n$ , sea vacío para algún entero positivo  $n$ .

Llama la atención que un problema tan concreto sobre las series trigonométricas llevase a Cantor a introducir conceptos tales como el de punto de acumulación y conjunto derivado. Y a crear la teoría de los cardinales transfinitos, de la que surgió, entre otros, el problema de la hipótesis del continuo. Un objeto importante es el conjunto de Cantor  $C_\xi$  de razón de disección  $\xi < 1/2$ , que no es numerable, puesto que su cardinal es el de todos los reales, pero que, sin embargo, tiene medida igual a cero.

A partir de un intervalo  $[a, b]$  obtenemos dos,

$$[a, a + (b - a)\xi] \cup [b - (b - a)\xi, b],$$

ambos de longitud  $(b - a)\xi$ .



Aplicado el proceso  $k$  veces a  $[0, 2\pi)$ , resultan  $2^k$  intervalos de longitudes  $2\pi\xi^k$  cuya unión designamos por  $C_k(\xi)$ . El conjunto de Cantor es la intersección de todos ellos:

$$C_\xi = \bigcap_{k=1}^{\infty} C_k(\xi).$$

En el año 1922, Alexander Rajchman demostró que el conjunto ternario de Cantor ( $\xi = 1/3$ ) es de unicidad. Su alumno, Antoni Zygmund, se doctoró en 1923 con una tesis sobre esta teoría, escribiendo posteriormente el libro *Trigonometric Series*, un clásico del Análisis Armónico del siglo XX, que está dedicado a Rajchman, su maestro, y a Marcinkiewicz, su discípulo, desaparecidos ambos trágicamente durante la Segunda Guerra Mundial. Del año 1955 es el siguiente resultado de R. Salem y A. Zygmund:

**Teorema.** *El conjunto de Cantor  $C_\xi$  es de unicidad si y sólo si  $\theta = 1/\xi$  es un número de Pisot.*

Un número de Pisot  $\theta$  es un entero algebraico cuyos conjugados algebraicos  $\theta = \theta_1, \theta_2, \dots, \theta_n$  verifican que  $\theta = \theta_1 > 1$ ,  $|\theta_2| < 1, \dots, |\theta_n| < 1$ . Estos números fueron definidos por su relación con los problemas de distribución uniforme módulo 1. Son ejemplos de números reales tales que las partes fraccionarias de sus potencias enteras no están uniformemente distribuidas en el intervalo unidad. La demostración de Salem y Zygmund es muy bella, puesto que conecta de forma precisa dos conceptos tan diferentes, *a priori*, como son la unicidad de las series y los números de Pisot. En la bibliografía sugerida ([7, 16, 17]) pueden encontrarse los detalles de la demostración.

4. PARTE IV. *Ejemplos y contraejemplos*

Riemann utiliza el concepto de valor principal de una integral, introducido por Cauchy, para ampliar el conjunto de funciones que son integrables, yendo más allá de las acotadas. En esta parte de la tesis considera el ejemplo siguiente:

$$f(x) = \frac{d}{dx} \left( x^\nu \cos \frac{1}{x} \right), \quad 0 < \nu < \frac{1}{2}, \quad \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^{2\pi} f(x) dx = (2\pi)^\nu \cos \frac{1}{2\pi}$$

Sin embargo, observa a renglón seguido que:

$$\int_0^{2\pi} f(x) \cos(nx) dx \sim \frac{1}{2} \sin\left(2\sqrt{n} + \frac{\pi}{4}\right) \sqrt{\pi n}^{(1-2\nu)/4}$$

Es decir, los coeficientes de su serie de Fourier se hacen arbitrariamente grandes y, por tanto, aquélla no puede ser convergente.

En sentido opuesto nos presenta a la función:

$$f(x) = \sum_{n=1}^{\infty} \frac{(nx)}{n},$$

donde, como antes,  $(x)$  representa la diferencia entre  $x$  y su entero más cercano. Luego escribe, sin dar la demostración, la identidad:

$$f(x) = \sum_{n=1}^{\infty} \frac{d^i(n) - d^p(n)}{\pi n} \sin(\pi nx),$$

siendo  $d^i(n)$  el número de divisores impares de  $n$  y  $d^p(n)$  el número de divisores pares de  $n$ .

Esta función está bien definida, es integrable Lebesgue, pero no es integrable a la Riemann, porque su oscilación se hace infinita en cualquier intervalo que consideremos.

Luego Riemann añade el comentario siguiente:

Se obtiene un ejemplo del mismo tipo con las series

$$\sum_{n=0}^{\infty} c_n \cos(n^2 x) \quad \text{y} \quad \sum_{n=0}^{\infty} c_n \sin(n^2 x)$$

cuando las cantidades positivas decrecientes  $c_0, c_1, c_2 \dots$  se hacen infinitamente pequeñas, pero para las cuales  $\sum_k c_k$  se hace infinitamente grande.

Resultaría ocioso mencionar a las funciones Theta para motivar el interés de Riemann por este tipo de series. Pero creo interesante resaltar que en la tesis se mencione un asunto que sigue siendo un problema abierto en el análisis armónico, con aplicaciones aritméticas notables:

Si una función integrable (Lebesgue) tiene una serie de Fourier de la forma  $\sum c_n e^{in^2 x}$ , ¿es cierto que  $\|f\|_p \ll \|f\|_1$ , para  $1 < p < 4$ ?

Cuando los coeficientes forman una sucesión monótona decreciente (como en los ejemplos de Riemann) sabemos que la respuesta a la pregunta anterior es afirmativa. Pero en el caso general, coeficientes arbitrarios, se ha mostrado hasta ahora muy difícil y elusiva.

De los ejemplos mostrados en la segunda parte de la tesis acerca de la relación entre continuidad y diferenciabilidad, se desprendía una pregunta natural a la que no se le dio respuesta: ¿existirá una función continua que carezca de derivadas en todos sus puntos?

K. Weierstrass encontró un ejemplo explícito de una función con esas características y lo presentó en una conferencia dada en la Academia de Ciencias de Berlín, el 8 de julio de 1872.

Consiste en una serie trigonométrica lacunar:

$$\sum_{n=0}^{\infty} b^n \cos(a^n x),$$

donde  $a$  es un entero impar,  $0 < b < 1$ , de manera que  $a \cdot b > 1 + \frac{3}{2}\pi$ .

Según parece, K. Weierstrass descubrió el ejemplo anterior como consecuencia de su fracaso en demostrar una conjetura de Riemann. En su carta a Du Bois-Raymond decía que “hasta donde yo alcanzo a saber, Riemann ha afirmado que las funciones

$$f(x) = \sum_{n=1}^{\infty} \frac{\sin(n^2 x)}{n^2} \quad \text{y} \quad g(x) = \sum_{n=1}^{\infty} \frac{\cos(n^2 x)}{n^2}$$



carecen de derivada en todos sus puntos”. Aunque Riemann no había comunicado la demostración, sino que, en una cierta ocasión, había indicado que la prueba se podía hacer usando las funciones elípticas.

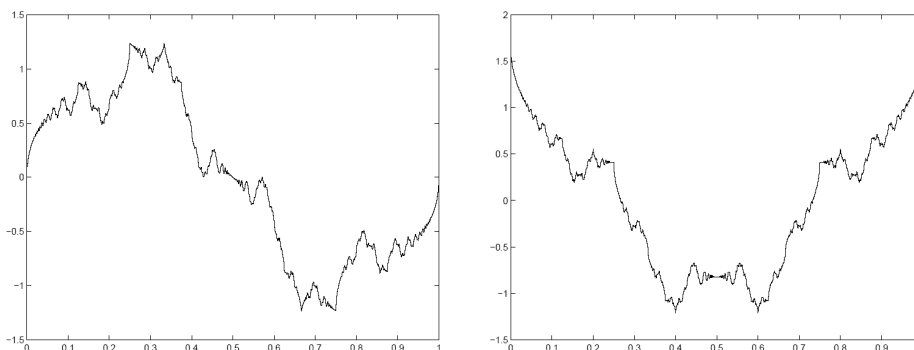
Cualquiera que se interese por la historia de esta notable función, puede comprobar con facilidad que se han publicado más de doscientos artículos sobre ella. En “Riemann’s example of a continuous non-differentiable function in the light of two letters of Christoffel to Prym”, publicado en 1986 en el Bulletin de la Société Mathématique de Belgique, los autores, P. Butzer y E. Stark, analizan hasta la saciedad la evidencia disponible acerca de si, en realidad, Riemann dijo, o no dijo, que la función carece de derivada en todos sus puntos. En mi opinión se trata de un ejemplo que muestra hasta qué extremos puede llegarse al hacer la historia de la ciencia. Seguramente es irrelevante discernir si Riemann afirmó, o no, tal cosa, excepto, quizás, por el posible morbo de encontrar una pifia menor en la obra de tan gran matemático. Habida cuenta de que muchos años después se encontraron puntos donde  $f$  es diferenciable. Lo cierto es que hallar una función continua carente de derivada en todos sus puntos era un problema natural e importante en esa época. También lo es que la función  $f$ , que está estrechamente relacionada con la función “theta”, era un objeto matemáticamente muy interesante para Riemann, y para muchos otros y durante bastante tiempo después, como es el caso de G. Hardy y J. Littlewood, quienes la trataron en [11, 12]. Demostraron que las funciones

$$\sum_{n=1}^{\infty} \frac{\sin(n^2x)}{n^2} \quad \text{y} \quad \sum_{n=1}^{\infty} \frac{\cos(n^2x)}{n^2}$$

carecen de derivada cuando  $x/\pi$  es irracional, o bien un racional cuya fracción reducida es de la forma  $a/4b$ ,  $a/(4b+1)$ . Sin embargo, J. Gerver [9] demostró que tanto  $f$  como  $g$  son derivables en los  $x = \pi a/q$ ,  $\text{mcd}(a, q) = 1$ , cuando  $q \equiv 2 \pmod{4}$ . Además el valor de las respectivas derivadas es  $-1$  y  $0$ , como se muestra en las figuras, obtenidas con la ayuda del ordenador, de la página siguiente.

Como dato curioso cabe reseñar que Gerver era un estudiante de primer curso en la universidad de Columbia (New York), y consiguió este resultado porque el profesor de Cálculo, S. Lang, había mencionado el problema en clase. Gerver lo resolvió y la demostración apareció en el American Journal of Mathematics (1970). De haber tenido acceso a

las gráficas que ahora nos proporcionan los ordenadores, es seguro que tanto Riemann primero, como Hardy y Littlewood después, hubiesen previsto esos puntos de diferenciabilidad. Aunque en honor a Gerver hay que añadir que tampoco en el año 1970 existían los excelentes programas para dibujar funciones de los que ahora disponemos.



Un capítulo notable de las matemáticas contemporáneas es el de las geometrías fractales, que aparecen en el estudio de los sistemas dinámicos caóticos y en los modelos creados para entender los regímenes turbulentos en la mecánica de fluidos. Existen diversas nociones de dimensión fractal, una de ellas es la llamada dimensión por cajas, o dimensión de Minkowski.

**Teorema** ([13]). *La dimensión de Minkowski de las gráficas de las funciones  $f$  y  $g$  es  $5/4$ .*

## 5. EPÍLOGO

En las obras completas de Riemann, publicadas después de su muerte, encontramos el siguiente texto de R. Dedekind refiriéndose a esta tesis:

Esta memoria fue presentada por su autor, en 1854, a la Facultad de Filosofía para obtener su Habilitación en la Universidad de Göttingen. Aunque el autor no parece haber tenido intención de publicarla, la impresión de este trabajo sin cambio alguno nos parece más que justificada, tanto por el considerable interés del tema en sí, cuanto

por la forma en la que son tratados los principios más importantes del análisis infinitesimal.

Brunswick, julio de 1867. R. Dedekind.

Habida cuenta de lo que hemos encontrado en su lectura (que incluye la formulación del problema de la unicidad de los desarrollos trigonométricos y su solución en un caso fundamental; la extensión de la noción de integral más allá de las definiciones de Newton, Leibniz y Cauchy; las generalizaciones de la noción de derivada, incluyendo una clara alusión al concepto de derivada débil del cálculo de distribuciones; los ejemplos de funciones continuas que carecen de derivada en conjuntos densos de puntos; y las diversas áreas futuras que supo entrever, como los fractales o la teoría de conjuntos de puntos), sorprende que Riemann no estuviese del todo satisfecho con su trabajo y que no se planteara su publicación.

Como no parece haber testimonio escrito del autor, sólo podemos especular acerca de sus motivos. Por un lado estaban los usos de aquella época, el “*pauca sed matura*” de Gauss, que debía imponer un tanto a un joven profesor de Göttingen. Pero eso no es todo y quizás sí podamos aventurar algunas de sus razones.

Por un lado Riemann había generalizado la noción de integral, pero para poder integrar funciones no acotadas tiene que hacer uso del “valor principal” y ahí aparece esa función  $\frac{d}{dx}(x^\nu \cos \frac{1}{x})$  que puede integrar entre 0 y  $2\pi$ , pero resulta que sus coeficientes de Fourier no tienden a cero. Por otro lado se encuentra con series trigonométricas que convergen a una función  $f$  que no es integrable según su definición. Es claro que a Riemann esta situación no podía satisfacerle y, además, tampoco pudo demostrar que las series de Fourier de sus funciones integrables convergiesen de manera razonable. Ahora sabemos que estas cuestiones eran muy difíciles y, quizás, imposibles para aquella época; pensemos en la integral de Lebesgue; en el ejemplo de Kolmogorov de una función integrable cuya serie de Fourier diverge en casi todo punto; en el teorema de Carleson, que es del año 1964; o en el problema todavía abierto de caracterizar los conjuntos de unicidad. No obstante, sí podemos afirmar que se trata de otra muestra de la profundidad y grandeza de Riemann: la insatisfacción por lo no realizado le impedía publicar los magníficos resultados que había conseguido.

**Agradecimientos.** Pablo Fernández Gallardo y Bernardo López Melero me ayudaron en la preparación de este escrito, leyendo su primera versión y haciendo sabias y oportunas sugerencias.

#### REFERENCIAS

- [1] B. RIEMANN: *Gesammelte mathematische Werke* (ed. H. Weber y R. Dedekind). Primer edición, Leipzig, Teubner, 1876. Existe una edición más reciente, del año 1990, por Springer Verlag.
- [2] B. RIEMANN: *Oeuvres mathématiques de Riemann*. Traducidas al francés por L. Laurel. Librairie scientifique et technique Albert Blanchard, 1968.
- [3] B. RIEMANN: *Riemanniana selecta*. Edición de José Ferreirós. CSIC, 2000.
- [4] D. MASCRÉ: *Bernhard Riemann, posthumous thesis on the representation of functions by trigonometric series (1867)*. Landmark writings in western mathematics, 1640-1940. I. Grattan-Guinness (Editor). Elsevier, 2005.
- [5] J. FOURIER: *Theorie analytique de la chaleur*, 1822.
- [6] G. CANTOR: *Über trigonometrische Reihen*. Ges. Abh. Berlin, 1871.
- [7] A. ZYGMUND: *Trigonometrics series*. Cambridge Univ. Press, 1959.
- [8] K. WEIERSTRASS: *Über continuierliche Functionen eines reellen Argumentes, die für keinen Werth des letzteren einen bestimmten Differentialquotient besitzen*. Math. Werke II, 1895.
- [9] J. GERVER: The differentiability of the Riemann function at certain rational multiples of  $\pi$ . *Amer. J. Math.* **92** (1970), 33–55.
- [10] P. COHEN: *Topics in the theory of trigonometrics series*. Ph. D. Thesis, University of Chicago, 1958.
- [11] G. H. HARDY, J. E. LITTLEWOOD: Some problems in diophantine approximation. *Acta Math.* **37** (1914), 193–238.
- [12] G. H. HARDY: Weierstrass non-differentiable function. *Trans. Amer. Math. Soc.* **17** (1916), 301–325.
- [13] F. CHAMIZO Y A. CÓRDOBA: The fractal dimension of a family of Riemann's graphs. *C. R. Acad. Sci. Paris Sér. I Math.* **317** (1993), no. 5, 455–460.
- [14] J. BOURGAIN: Spherical summation and uniqueness of multiple trigonometric series. *Internat. Math. Res. Notices* **1996**, no. 3, 93–107.
- [15] J. M. ASH AND G. WANG: A survey on uniqueness questions in multiple trigonometric series. In *Harmonic analysis and nonlinear differential equations (Riverside, CA, 1995)*, 35–71. Contemp. Math. **208**. Amer. Math. Soc., Providence, RI, 1997.
- [16] Y. MEYER: *Algebraic numbers and harmonic analysis*. North Holland, 1972.
- [17] R. SALEM: *Algebraic numbers and Fourier analysis*. D. C. Heath & Co., Boston, 1963.

## Annex: Presentation of Riemann's Academic Year (2007-2008)

The SCHOOL OF MATHEMATICS AND STATISTICS (FACULTAT DE MATEMÀTIQUES I ESTADÍSTICA, FME) of the Technical University of Catalonia (Universitat Politècnica de Catalunya) dedicated the academic year 2007-2008 to Bernhard Riemann (1826-1866).

The mission of the **Riemann Committee**, constituted by Víctor GONZÁLEZ, Eduard RECASENS, Oriol SERRA, Joan SOLÀ-MORALES and Sebastià XAMBÓ, was to design the academic content of that decision, including the definition of the lecture programme. It is the occasion to reaffirm the gratitude of the FME to its members for the time and effort dedicated to the initiative.

The diffusion of the lectures has been done by posting the presentations and corresponding articles in the **Butlletí Digital FME**, which can be accessed through the link with the same name at the top of the FME Web Page.\* Now, in addition to the materials that can be accessed via that page, the FME offers, with this volume, a collection of the lectures more directly related to the life and work of Riemann. In this way we add a “Riemann Volume” to the volumes dedicated to Poincaré, Einstein, Gauss and Euler. On behalf of the FME, it is a pleasing duty to thank very heartily the authors for the additional effort of transforming the lecture materials into the articles collected in this book.

Very special thanks are due to Professor Sir Michael Atiyah, who managed to find the time to visit the FME in December 2007. On that occasion he delivered a *keynote* lecture, *Riemann's influence in Geometry, Analysis and Number Theory*, and it is a great satisfaction that we could include it here. We also thank Professor Atiyah, the Centre

---

\* <http://www-fme.upc.es/>

de Recerca Matemàtica (CRM) and the Institut de Matemàtica de la Universitat de Barcelona (IMUB) for their kind permission to include the lecture *Duality in Mathematics and Physics*.

The papers are written in the language chosen by their authors: two in Catalan, three in Spanish and five in English.

We also have to say that we have used the same L<sup>A</sup>T<sub>E</sub>X document style (`amsart.sty`) to typeset the different works, even when the material received was not in L<sup>A</sup>T<sub>E</sub>X or it was in L<sup>A</sup>T<sub>E</sub>X but with a style that was not `amsart`. Nevertheless we have tried to respect as much as possible the original composition by the authors.

We will be satisfied if we contribute, with such endeavours, to have figures such as Riemann better known among the mathematicians, and above all if this knowledge reached, fresh and alive, the younger generations.

Sebastià Xambó Descamps  
Dean of the l'FME  
Barcelona, August 2008