AULA DE CIÈNCIA
I CULTURA, 20

Matemàtiques del segle XXI: dels fonaments a la tecnologia

Cicle Ferran Sunyer i Balaguer

Manuel Castellet (ed.)

Joan Bagaria
Manuel Castellet
Xavier Messeguer
Carles Puente
James Stirling
Sebastià Xambó

FUNDACIÓ Caixa Sabadell

AULA DE CIÈNCIA
I CULTURA, 20

Matemàtiques del segle XXI: dels fonaments a la tecnologia

Manuel Castellet (ed.)

Joan Bagaria
Manuel Castellet
Xavier Messeguer
Carles Puente
James Stirling
Sebastià Xambó

Catalogació en publicació de la
FUNDACIÓ CAIXA SABADELL

Matemàtiques del segle XXI: dels fonaments a la tecnologia / Manuel Castellet (ed.) .-- Sabadell: Fundació Caixa Sabadell, 2003.- (Aula de ciència i cultura; 20)

Conté: De Hilbert als "problemes del mil·lenni" / Manuel Castellet; Una petita excursió al paradís de Cantor / Joan Bagaria / Recreacions computacionals de la teoria de codis correctors d'errors / Sebastià Xambó Descamps; Mathematical modelling in sport / J.R. Stirling, M. S. Zakynthinaki; La tecnologia fractal: de les matemàtiques a l'empresa de tecnologia / Carles Puente i Baliarda; Genòmica computacional, una ciència interdisciplinar / Xavier Messeguer

I. Fundació Caixa Sabadell
II. Castellet, Manuel, ed.

1. Fundació Caixa Sabadell ? Conferències
2. Matemàtiques ? S. XXI ? Conferències

51"20"(042)

Edita: Fundació Caixa Sabadell, 2003
Carrer d'en Font, 25, 08201 - Sabadell
Tel.: 93 725 95 22 A/e: *fundacio@caixasabadell.org*

© Joan Bagaria, Manuel Castellet, Xavier Messeguer, Carles Puente, James Stirling, Sebastià Xambó

Fotocomposició i producció: *Addenda*, s.c.c.l., Pau Claris, 92, Barcelona
Impressió: Winihard Gràfics, sl, Av. del Prat, 7. 08180 Moià
Imprès sobre paper ecològic
ISBN: 84-95166-50-X
Dipòsit legal: B. 5.269-2004

Índex

PRÒLEG 7

DE HILBERT ALS «PROBLEMES DEL MIL·LENNI» 9
Manuel Castellet

UNA PETITA EXCURSIÓ AL PARADÍS DE CANTOR 21
Joan Bagaria

RECREACIONS COMPUTACIONALS DE LA TEORIA DE CODIS CORRECTORS D'ERRORS 41
Sebastià Xambó Descamps

MATHEMATICAL MODELLING IN SPORT 73
J. R. Stirling - M. S. Zakynthinaki

LA TECNOLOGIA FRACTAL: DE LES MATEMÀTIQUES A L'EMPRESA DE TECNOLOGIA 99
Carles Puente i Baliarda

GENÒMICA COMPUTACIONAL, UNA CIÈNCIA INTERDISCIPLINAR 109
Xavier Messeguer

Pròleg

Del formalisme rigorós de Hilbert a l'ús d'eines sofisticades de les matemàtiques en la genòmica, no hi ha només un segle llarg de distància, sinó un canvi de mentalitat i, sobretot, el reconeixement i el convenciment que la recerca en matemàtiques, des de les àrees més bàsiques i abstractes a aquelles més fàcils de comunicar a la societat, pot jugar un paper cada cop més cabdal en el desenvolupament científic i tecnològic.

Si ara fa més de setanta anys, un rector d'universitat a Alemanya es permetia dir en la inauguració d'un curs acadèmic que «la matemàtica és la base de tot el coneixement i el contenidor de tota l'alta cultura», ara hi podem afegir que també és la base de l'alta tecnologia o la base dels grans moviments financers o de les solucions a les complexes realitats socials, per citar tres exemples.

El nostre món és cada cop més interconnectat, i l'increment de la interconnexió en un sistema fa augmentar-ne el grau de complexitat. Estem, doncs, immersos en un sistema complex, que al mateix temps és fràgil i inestable. Només una bona eina per a la gestió dels sistemes complexos (siguin tecnològics, financers o socials) podrà donar resposta als problemes plantejats. La matemàtica és una d'aquestes possibles eines, amb un paper cada cop més decisiu en la gestió dels sistemes complexos. I aquesta és la imatge que hem volgut transmetre en el Tercer Cicle Ferran Sunyer i Balaguer, el contingut del qual recull el volum que ara presentem.

Començant amb el Hilbert de l'any 1900, hem pretès fer un recorregut des dels aspectes més formals als més aplicats, per donar als assistents al cicle una visió de diverses línies cap on es poden dirigir llurs interessos.

Si Joan Bagaria ens parla de la recerca puntera en aspectes formals de la teoria de conjunts i la fonamentació, amb aplicacions a la teoria de grafs i a l'anàlisi funcional, Sebastià Xambó ens ha mostrat com les poderoses eines de la geometria juguen avui un paper decisiu en alguns aspectes de la tecnologia de la informació i la comunicació, mitjançant el desenvolupament de codis correctors d'errors.

La part més aplicada l'han constituït les tres darreres conferències del cicle, sobre la interacció dels sistemes complexos en l'esport, per James Stirling, sobre l'ús de matemàtiques en els processos de seqüenciació del genoma, per Xavier Messeguer, i oferint

l'exposició d'un model empresarial que, a partir d'eines de l'anàlisi i els sistemes dinàmics, produeix antenes fractals comercialitzades.

Una darrera reflexió sobre aquest Tercer Cicle Ferran Sunyer i Balaguer. Un cicle de conferències no és una unió disjunta de xerrades sobre temes dispersos; cal que hi hagi un fil conductor, i aquest ha estat un recorregut per la història del darrer segle, analitzant alguns problemes centrals de les matemàtiques que motiven la realitat actual. Però, a més, si hi ha un «Tercer Cicle», vol dir que n'hi hagué un «Primer» i un «Segon». També en aquest context cal interpretar el contingut, com una mena d'ésser viu que cada dos anys vol donar a conèixer el rol important de les matemàtiques en l'avenç de la nostra societat.

Manuel Castellet
Director del Centre de Recerca Matemàtica

De Hilbert als «Problemes del Mil·lenni»

Manuel Castellet
Director del Centre de Recerca Matemàtica

Fa poc més de 100 anys, el 8 d'agost de 1900, David Hilbert va donar una conferència al Congrés Internacional de Matemàtics de París, en la qual formulava i raonava 23 problemes matemàtics. Aquesta conferència programàtica, que ha influït en el desenvolupament de les matemàtiques durant el segle, és famosa encara avui, més de cent anys després.

Les matemàtiques tenen un triple paper en el món actual: es basen en una lògica estricta i, per tant, estableixen l'estàndard de la veritat objectiva; són, com l'art, una creació del nostre esperit i, per tant, una part de la nostra tradició cultural, i, finalment, es poden utilitzar en aplicacions concretes. Els profans perceben només aquesta darrera funció pràctica; la recerca en matemàtica fonamental, que és a la base de tot, tanmateix resta desconeguda.

Amb la idea de canviar una mica això, el *Clay Mathematics Institute* —una fundació que té com a objectiu la difusió del coneixement matemàtic— ha instaurat recentment set premis d'un milió de dòlars cada un per la resolució dels set «problemes del mil·lenni». Aquests problemes, seleccionats amb molta cura per experts reconeguts, són de les més difícils qüestions obertes de diversos camps de les matemàtiques. N'hi ha que són problemes nous, d'altres fa més de cent anys que esperen la seva resolució.

Encara sobre la conferència de Hilbert

Els «Problemes del Mil·lenni» foren presentats públicament el mes de maig de l'any 2000 a París, com una mena de nova versió de la famosa conferència de Hilbert de 100 anys abans, al segon Congrés Internacional de Matemàtics, celebrat també a París, com he dit. Hi ha, però, una gran diferència: mentre que els problemes plantejats ara afecten només aspectes parcials de les matemàtiques actuals, Hilbert abastava la totalitat de les matemàtiques d'aleshores. També és cert que l'any 1900 hi havia un Hilbert i ara, malgrat l'avenç increïble que han fet les matemàtiques, no tenim una figura com ell, capaç d'entendre la globalitat.

El programa de Hilbert ha influït en el desenvolupament de les matemàtiques durant tota la primera meitat del segle XX i bona part de la segona; per aquesta raó, el record de la seva famosa conferència ha restat molt viu.

Fa 100 anys, David Hilbert, de 38 anys, era professor a la Universitat de Göttingen, aleshores la fortalesa de les matemàtiques. Entre els seus col·legues, Hilbert era considerat el matemàtic alemany més important. A començament d'any va rebre la invitació per impartir la conferència inaugural del congrés de París. Hermann Minkowski, professor a l'ETH de Zürich i amic íntim de Hilbert des dels anys en què ambdós eren estudiants a Königsberg, li demanava de «donar una ullada al futur de les matemàtiques i als seus problemes»; d'un tema com aquest, segons Minkowski, se'n continuaria parlant durant dècades.

La carta de Minkowski no fou mai contestada. Hilbert va acabar el text de la seva conferència a l'últim minut, com si fos més llatí que germànic, i l'envià a Minkowski i a Adolf Hurwitz, antic mestre de Hilbert i també professor a l'ETH. Era, però, massa tard; la conferència de Hilbert no fou la conferència inaugural del congrés i s'inclogué en el tercer dia de sessions. Hilbert mesurà malament la durada de la seva intervenció i es va haver d'acontentar a presentar 10 dels 23 problemes que havia preparat. Convé destacar dos aspectes de la presentació general i de cadascun dels problemes: la relació entre problemes particulars i teories generals i la necessitat de demostracions rigoroses (*strenge Beweise*).

La resolució d'un problema particular i la construcció d'una teoria general han d'anar, segons Hilbert, de bracet. Encara més, apuntava, els esforços per resoldre un problema difícil són necessaris per a l'avenç de les matemàtiques, encara que aquests esforços no portin a la resolució del problema. I esmentava dos exemples paradigmàtics: la hipòtesi de Fermat i el problema dels tres cossos de la mecànica.

Malgrat que la hipòtesi de Fermat no fou demostrada fins als anys 1993 i 1995, per Andrew Wiles, ja s'havien fet intents interessants per a la seva resolució a partir del segle XVIII. Uns intents que no donaven els fruits desitjats, però que, per exemple, portaren a la teoria algebraica de nombres, una nova teoria amb una multitud de conceptes que més tard s'anaven desenvolupant independentment. Els mètodes utilitzats per Wiles foren, però, d'un tipus totalment diferent; i la seva demostració està lligada a consideracions sobre la geometria de corbes.

De manera semblant va passar amb el problema dels tres cossos de la mecànica, que avui encara planteja preguntes obertes. El problema va inspirar el parisenc Henri Poincaré en nous mètodes de la mecànica celest, que s'han convertit actualment en eines importants per a la navegació de satèl·lits.

Hilbert i Poincaré es coneixien i s'admiraven, però les seves concepcions de les matemàtiques eren ben diferents. Per a Poincaré hi tenien un paper essencial la intuïció i les analogies físiques; per a Hilbert, era essencial el marc de la lògica més estricta. Per això, Hilbert va destacar en la seva conferència al congrés de París la necessitat que qualsevol afirmació, qualsevol resolució d'un problema, havia d'ésser demostrada. I si un problema no té solució, com la quadratura del cercle, aleshores cal poder demostrar precisament això, que no té solució. Tanmateix, malgrat que afirmava que qualsevol possible resultat era demostrable o no, no tenia cap prova que les coses fossin així.

Una demostració era, per a Hilbert, una successió finita de raonaments lògics. Ben segur que hi ha altres aspectes de la matematització, com consideracions de plausibilitat, analogies, experiències físiques i intuïció; però res de tot això no pot substituir la demostració. Hilbert demanava la demostració rigorosa no només per a l'aritmètica i la geometria clàssica, sinó també per a totes les altres branques de les matemàtiques, inclosa la matemàtica aplicada. En aquell temps, això no era un posicionament tan clar com ara. La demanda de demostracions rigoroses té només ple sentit en el marc de la construcció axiomàtica de les matemàtiques. Aquesta pretensió de Hilbert posava de manifest la qüestió de si és possible o no bastir les matemàtiques sobre una base que no admeti contradicció.

Als no iniciats els podria sorprendre que, precisament, hom pugui posar en dubte els fonaments de l'edifici matemàtic, la seva consistència. Aquí cal puntualitzar que cap dels conceptes matemàtics que fem servir contínuament al nostre entorn apareixen en el nostre món real. No hi ha ni punts ideals ni línies rectes infinites; són creacions del nostre esperit, com l'espai de la geometria elemental, els espais de dimensions superiors o la sèrie infinita dels nombres enters.

Els objectes matemàtics són, simplement, un conjunt d'elements. Entre ells es defineixen relacions, com la suma, la multiplicació o la proximitat. Aquestes relacions han de satisfer determinats axiomes, que, naturalment, es prenen de l'observació; no són necessàriament fets evidents sinó regles de joc de la teoria; aquestes regles i les conseqüències que se'n deriven és l'únic que hom pot utilitzar per a les demostracions. Això era el que Hilbert entenia per una construcció axiomàtica de les matemàtiques. Aquestes idees van anar agafant cos, especialment sota la influència del grup Bourbaki, i han estat decisives per a l'avenç de les matemàtiques.

Un cop Hilbert va desenvolupar una fonamentació axiomàtica de l'aritmètica i de la geometria clàssica, es va plantejar immediatament la qüestió de la consistència del sistema d'axiomes, és a dir, de la no contradicció. No en tenia cap dubte, però la intuïció no li era suficient. Per això, proposava, en el segon dels seus 23 problemes, la demostració que no és possible, en el sistema axiomàtic de l'aritmètica, mitjançant raonaments lògics, demostrar una afirmació i, al mateix temps, la contrària.

Aquesta demostració que ell demanava no va arribar mai. Al contrari, 30 anys després de la conferència de Hilbert, quan ell encara era matemàticament actiu, el lògic Kurt Gödel va demostrar que no es pot trobar una demostració matemàtica del principi de no contradicció per a l'aritmètica. De fet, el que va demostrar Gödel és que no és possible fer cap demostració de la consistència de l'aritmètica amb mètodes finitaris i, per tant, formalitzables en l'aritmètica mateixa. En general, no es pot demostrar la consistència d'una teoria fent servir mètodes formalitzables dins de la mateixa teoria. Però sí que és possible, per exemple, demostrar la consistència de l'aritmètica en la teoria de conjunts de Zermelo-Fraenkel (ZF); però per demostrar la consistència de ZF cal fer-ho amb una teoria més forta, que aleshores caldrà demostrar que és consistent, i així successivament.

No només el segon, sinó també el primer dels problemes proposats per Hilbert

obtingué una resposta ben diferent de la que ell pensava. Un subconjunt no numerable d'un continu (és a dir, dels punts d'una recta real), es pot posar en correspondència bijectiva amb tot el continu? És la hipòtesi del continu. Hilbert suggeria que qualsevol subconjunt d'un continu era o bé numerable o bé un continu, però l'any 1963 el matemàtic Paul Cohen va demostrar quelcom sorprenent: tant l'afirmació com el seu contrari són compatibles amb els axiomes. De fet, Cohen va demostrar que la negació de la hipòtesi del continu és consistent amb els axiomes; la consistència de la hipòtesi del continu amb els axiomes ja l'havia demostrada Gödel, el 1938. Hom pot prendre, per tant, com a nou axioma tant una afirmació com l'altra.

Tampoc no és cert, tal com pensava Hilbert, que qualsevol problema o bé es pot resoldre o bé en podem demostrar la insolubilitat. Durant el segle XX, els lògics van treballar aquesta qüestió amb mètodes molt subtils.

Per fer un judici històric de la conferència de Hilbert de l'any 1900, és irrellevant que les respostes a les seves preguntes fonamentals siguin ben diferents de les que ell havia imaginat. Allò realment decisiu fou que havia formulat les preguntes d'una manera tan clara i precisa que, després, els investigadors van poder treballar-hi. El que és important és, per exemple, el fet que hom accepta que en el sistema axiomàtic no hi ha contradicció, encara que no ho puguem demostrar d'una manera finitística. Les aportacions inestimables que han generat les matemàtiques des de fa centenars d'anys han creat una atmosfera de credibilitat.

Aquestes qüestions fonamentals i algunes altres que no cito constituïen només una part petita del programa de Hilbert. Paral·lelament, va formular una llarga sèrie de problemes específics. L'espectre és molt ampli i inclou totes les parts del pensament matemàtic. 100 anys després, s'han resolt tots els problemes específics llevat de tres. Un d'aquests és la hipòtesi de Riemann, que actualment està considerat el més important i més difícil dels problemes no resolts i figura, per tant, entre els «Problemes del Mil·lenni». La hipòtesi de Riemann tracta dels zeros d'una funció relacionada amb els nombres primers i està connectada amb gairebé totes les disciplines de les matemàtiques. Malgrat que la corroboren enormes càlculs amb ordinador, no ha pogut ésser demostrada.

Podríem continuar amb una mostra dels problemes específics que Hilbert va plantejar, que van des de les equacions diofàntiques fins als fonaments axiomàtics de la física, però el que ara vull destacar és que el programa de Hilbert anava des de teoremes d'existència fonamentals, passant per qüestions abstractes i per concretes, fins a les aplicacions. Per aquesta amplitud, la seva conferència continua essent considerada com un esdeveniment històric.

La nova situació

Però, 100 anys després, la situació ha canviat. No només ara no tenim un Hilbert capaç d'entendre la globalitat, que no el tenim, sinó que el món de les matemàtiques

compta actualment amb uns 50.000 investigadors que, gràcies als mitjans de comunicació i d'informació, constitueixen un organisme en el qual cada part reacciona ràpidament als estímuls de les altres. Per això ara no és possible un programa com el de Hilbert.

Tanmateix, el Clay Mathematics Institute de Cambridge, Massachusetts, amb un comitè científic format per Alain Connes, Arthur Jaffe, Andrew Wiles i Eduard Witten, va encarregar a set dels més destacats matemàtics del moment el plantejament de set grans problemes oberts. No és un programa com el de Hilbert, que abastava el conjunt de totes les matemàtiques, amb mentalitat europea, que combinava aspectes fonamentals amb resultats molt específics. És, per a mi, una mostra de la realitat matemàtica actual i del domini que exerceix la concepció liberal nord-americana. El Clay Institute ha dotat set premis amb un milió de dòlars cada un per a aquells que resolguin cadascun dels problemes plantejats. Resulta mediàticament molt atractiu, però no crec que faci avançar la recerca com ho feu el programa de Hilbert.

Els mateixos promotors dels premis declaren que els «Problemes del Mil·lenni» no pretenen marcar la direcció de les matemàtiques durant aquest segle que ha començat; només volen centrar l'atenció en un petit conjunt de qüestions matemàtiques pendents des de fa temps, cada una d'elles central, que s'estan resistint a ésser resoltes, com, per exemple, la hipòtesi de Riemann, que ja he esmentat.

Els «Problemes del Mil·lenni»

1. La hipòtesi de Riemann (formulada per Bernhard Riemann l'any 1859)

«Els zeros de la funció zeta de Riemann tenen part real igual a un mig».

És, possiblement, el més famós dels resultats pendents des de fa temps (i s'ha comprovat que és cert per als «primers» mil cinc-cents milions de zeros).

La demostració de la hipòtesi de Riemann donaria informació definitiva a diverses qüestions sobre la freqüència dels nombres primers.

Preparat per Enrico Bombieri.

2. P versus NP (formulat per Stephen Cook l'any 1971)

«És cert que P és igual a NP?»

La pregunta equival a determinar si qualsevol llenguatge acceptat per un algorisme no-determinístic en un temps polinomial és també acceptat per algun algorisme determinístic en temps polinomial (el contrari sempre és cert).

Una resposta afirmativa permetria disposar d'algorismes útils per a molts problemes computacionals pràctics, però al mateix temps destruiria la seguretat de transaccions financeres fetes a través d'Internet. Per tant, afecta essencialment la criptografia.

Preparat per Stephen Cook.

3. La conjectura de Hodge (formulada per William Hodge l'any 1950)

«En una varietat algebraica projectiva no singular sobre els complexos, qualsevol classe de Hodge és una combinació lineal racional de classes de cicles algebraics.»

Se sap que un cicle algebraic és el mateix que un subespai analític tancat.

Si la conjectura és certa, els cicles de Hodge admetran una interpretació geomètrica, i això permetria conèixer com són les peces que s'adjunten a determinats espais per obtenir-ne de nous, tècnica que es va mostrar molt potent durant tot el segle passat.

Preparat per Pierre Deligne.

4. La conjectura de Poincaré (formulada per Henri Poincaré l'any 1904)

«Qualsevol 3-varietat tancada simplement connexa és homeomorfa a la 3-esfera.»

La conjectura inicialment plantejada per Poincaré exigia només que la varietat inicial tingués la mateixa homologia que l'esfera, però ell mateix va demostrar que era falsa, en construir aquesta esfera de Poincaré.

La versió n-dimensional de la conjectura, exigint que la varietat inicial tingui els mateixos grups d'homotopia que l'esfera, és certa per a $n=2$, per a $n\geq 5$ (S. Smale 1960) i per $n=4$ (M. Freedman 1980).

Preparat per John Milnor.

Durant el període de correcció de proves d'aquest article, el matemàtic rus Grisha Perelman ha donat a conèixer una demostració de la conjectura de Poincaré en dimensió tres, utilitzant mètodes sofisticats de geometria diferencial, que està essent analitzada pels experts.

5. La teoria de Yang-Mills

«Demostrar que per a qualsevol grup *gauge* simple compacte, la teoria quàntica de Yang-Mills a l'espai de dimensió 4 existeix i té defecte de massa positiu.»

La demostració d'aquest resultat permetria un coneixement matemàtic millor de les teories quàntiques de camps 4-dimensionals, molt importants per a la física. Permetria, per exemple, donar una justificació matemàticament satisfactòria a la invisibilitat dels quarks.

Preparat per Arthur Jaffe i Eduard Witten.

6. Les equacions de Navier-Stokes

«Existeixen solucions diferenciables físicament raonables per a les equacions de Navier-Stokes en tres dimensions?»

El resultat és conegut en dos dimensions, fins i tot per a les equacions d'Euler i diversos càlculs computacionals numèrics hi avalen una resposta afirmativa.

La importància de les equacions d'Euler i de Navier-Stokes és que descriuen el moviment d'un fluid a l'espai.

Preparat per Charles Fefferman

7. La conjectura de Birch i Swinnerton-Dyer

La conjectura és tècnicament molt difícil de formular, però diu més o menys que les solucions racionals de determinades equacions algebraiques estan íntimament relacionades amb una certa funció zeta, com la de Riemann, de manera que si la funció s'anul·la en el punt 1, aleshores hi ha una infinitat de punts racionals, i si no s'anul·la, només n'hi ha un nombre finit.

Constitueix, en certa manera, una reformulació del desè problema de Hilbert, que l'any 1970 Yu V. Matiyasevich va demostrar que era irresoluble.

Preparat per Andrew Wiles.

Però..., i a casa nostra què?

«A casa nostra» vol dir el conjunt de departaments de matemàtiques (amb els noms que tinguin) integrats a la xarxa d'universitats de l'Institut Joan Lluís Vives o, dit en un altre llenguatge, el conjunt de departaments de matemàtiques de les terres de llengua catalana, sens prejudici de les denominacions tradicionals o estatutàries que aquesta llengua tingui en els diversos territoris.

A casa nostra —siguem realistes— no ens podem fixar com a objectiu per als propers anys demostrar cap dels 7 problemes que acabo de plantejar, com no es podien plantejar els nostres avantpassats resoldre cap dels 23 problemes de Hilbert. Però, encara que em reafirmo en el que acabo de dir, cal aclarir que la situació ara, a casa nostra, referint-nos al context internacional, és ben diferent de la de fa 100 anys; fins i tot és ben diferent de la de fa 30 anys. A inicis dels anys setanta, la probabilitat de trobar al *Mathematical Reviews* o a la *Zentralblatt für Mathematik* un article signat per un matemàtic de casa nostra, o àdhuc espanyol, era gairebé 0. Actualment, en canvi, en totes les àrees abunden els treballs de recerca de matemàtics de les nostres universitats publicats en revistes de prestigi, i en totes les àrees tenim matemàtics científicament destacats.

Això que acabo d'afirmar és una realitat constatable, per exemple, referit només a les universitats de Catalunya, en el *report* sobre la recerca matemàtica, elaborat l'any 1998 per l'Institut d'Estudis Catalans, sota la direcció de Joan Girbau: la relació entre el nombre total d'articles publicats a Catalunya i el nombre d'habitants pel producte interior brut del país és similar al de Noruega, Gran Bretanya o Alemanya, i el 8,2% de tots els articles de matemàtiques són publicats a revistes considerades d'excel·lència, dades que situen Catalunya, pel que fa a la recerca matemàtica i a aquests paràmetres, en els primers llocs dels països científicament més significatius.

Per tant, el gran esforç realitzat pels matemàtics de les darreres generacions no ha estat inútil. I ens n'hem de congratular, però al mateix temps ens ha de servir d'estímul per continuar en aquesta línia iniciada ara fa gairebé 30 anys. No resoldrem, probablement, cap dels 7 «Problemes del Mil·lenni», no guanyarem, probablement, cap Medalla Fields, o en tot cas no serà fàcil aconseguir-ho, però de moment estem tenint

una presència important en la bibliografia matemàtica i una consideració més que notable en el context internacional. L'organització, del Tercer Congrés Europeu de Matemàtiques l'estiu de l'any 2000, a càrrec de la Societat Catalana de Matemàtiques, n'és un bon exemple.

No em toca a mi entrar ara aquí a fer una reflexió sobre el que convé investigar en les diverses àrees, ni sobre la desitjable articulació dels grups concrets d'investigació. Vull limitar-me a fer unes reflexions sobre la consideració dels nostres països com a comunitats científicament considerades internacionalment. És cert que tenim investigadors de casa nostra coneguts internacionalment; també és cert que els nostres grups de recerca comencen a ésser visibles bibliogràficament. Però una cosa és que individualment siguem ja coneguts i reconeguts i una altra de ben diferent és que ho siguem com a comunitat. Sense la primera no podríem aconseguir la segona, però aquest hauria d'ésser un objectiu irrenunciable. Hauríem de voler aconseguir que aquesta comunitat que he anomenat «casa nostra» no sigui coneguda internacionalment només pel Gaudí barceloní, per les platges illenques, per la ceràmica castellonenca, per les taronges valencianes o pel Barça, sinó coneguda, també, com una comunitat científica, com una comunitat matemàtica.

I tenim eines per fer-ho, i podem, si volem, crear noves eines per donar-nos més força i més presència. Em limitaré a esmentar les que conec, però, sens dubte, n'hi ha d'altres d'iniciades o assajades aquí o allà, que hauríem d'intentar compartir tots.

Eina 1. La xarxa d'universitats de l'Institut Joan Lluís Vives. Explotem les seves potencialitats! Quants de nosaltres ens hem acollit a les possibilitats de mobilitat —febles, certament— que ofereix? Podríem, per exemple, proposar un programa estable d'intercanvi d'estudiants de llicenciatura o de doctorat o un petit paquet de beques postdoctorals. Segur que un reduït grup de treball dedicat a explorar aquestes possibilitats podria acabar plantejant propostes serioses amb èxit.

Eina 2. La Societat Catalana de Matemàtiques. Aquí, l'adjectiu no es refereix a catalana de Catalunya, sinó a catalana del conjunt de territoris on s'estén aquesta llengua. És cert que la Societat té socis arreu del territori i que intenta, també, organitzar activitats arreu; per exemple, les «Proves Cangur» a les Illes Balears o a Castelló, la «Tercera Trobada Matemàtica» a València. Però crec que, entre tots, hauríem d'exigir-nos una distribució de les activitats més equitativa i una presència més gran de tots el territoris i a tots els territoris. Per exemple, amb beques per a estudiants o per a doctorants esteses a tot l'àmbit lingüístic, és a dir, a tot l'àmbit d'actuació de la societat, que coincideix amb el de l'Institut Joan Lluís Vives.

Eina 3. Una comunitat es pot donar a conèixer com a tal per mitjà d'una revista científica que aconsegueixi un bon nivell. I no la tenim. Tenim algunes revistes, una amb més antiguitat que les altres, que ens permeten un cert intercanvi editorial, però que són

escassament conegudes internacionalment. Proposta que llenço ara: creació d'una revista nova, amb visió de futur, tant des del punt de vista editorial com de prestigi científic, publicada per la nova empresa editorial de la Societat Europea de Matemàtiques, que es presentés com la continuïtat de totes les revistes que ara publiquem. Aquesta revista hauria d'ésser liderada i editada per la Societat Catalana de Matemàtiques. Només si renunciem a protagonismes estèrils podrem progressar; altrament, ens estancarem en cadascun dels nostres petits regnes de taifes.

Eina 4. El Centre de Recerca Matemàtica. Com que és el que conec millor, en parlaré una mica més. Deixeu-me dir abans, però, que si l'he inclòs com una d'aquestes possibles eines és amb el desig i la voluntat que el CRM tingui més presència en tota la comunitat matemàtica d'aquesta «casa nostra» i amb l'oferiment a tots els matemàtics d'utilitzar i aprofitar les interaccions que pot proporcionar el CRM.

El Centre de Recerca Matemàtica (CRM) fou creat l'any 1984 per l'Institut d'Estudis Catalans, amb l'objectiu principal de posar a disposició dels matemàtics catalans un centre d'investigació que estimulés la millora de la recerca matemàtica a Catalunya, tant en l'aspecte qualitatiu com en el quantitatiu. Per assolir aquest objectiu, el CRM invita destacats científics d'arreu del món a fer-hi estades de recerca, facilita el contacte dels nostres joves investigadors amb ells i amb institucions científiques, duu a terme programes de recerca, organitza congressos, seminaris i altres reunions científiques i difon els resultats de la recerca.

El CRM acull cada any una quarantena d'investigadors que treballen amb els matemàtics catalans, en períodes que van d'un a dotze mesos. Així es potencien els nostres grups de recerca i així es potencia Catalunya com un país científicament —en aquest cas, matemàticament— desenvolupat. L'activitat del CRM es complementa amb l'organització de congressos temàtics i de cursos avançats d'abast internacional. Només aquests darrers tres anys s'han celebrat 15 esdeveniments d'aquest tipus, 14 dels quals han estat finançats per la Comunitat Europea.

Des de l'any passat, amb la voluntat de consolidar el CRM com un dels instituts d'investigació més destacats d'Europa, s'ha reformulat la seva forma jurídica, i actualment el Centre de Recerca Matemàtica és un consorci integrat per l'Institut d'Estudis Catalans i la Generalitat de Catalunya (mitjançant el Departament d'Universitats, Recerca i Societat de la Informació) amb personalitat jurídica pròpia.

De la tradició al futur

Abans de l'any 1000, Gerbert d'Orlhac, un monjo dels Pirineus catalans, que seria ben conegut com a papa Silvestre II, estudià el *Quadrivium* d'una manera profunda i renovà els sistemes de càlcul a l'àmbit europeu dos segles abans de Fibonacci, construint un àbac original.

El mallorquí Ramon Llull, un dels escriptors més importants de la llengua catalana del segle XIII, comprenia ja l'esfericitat de la Terra i escrivia, al final d'aquest segle, l'*Ars combinatoria* i l'*Art de navegar*, un llibre, que no fou superat fins al segle XVI, en el qual es descriu l'astrolabi i s'esmenta l'ús de l'agulla magnètica, entre altres coneixements. Alexander von Humboldt afirmava, ara fa poc més de cent anys, que aquests progressos de la ciència es van transmetre a la resta del món civilitzat des de Catalunya, a través dels altres pobles de la Mediterrània.

Pocs anys després, el jueu mallorquí Abraham Cresques dibuixà, era l'any 1375, el mapamundi anomenat *Atles català*, una representació del món conegut aleshores i que fou fonamental per a tots els navegants i viatgers de l'època. La novel·la històrica d'Alfred Bosch *L'Atles furtiu*, premi Sant Jordi de literatura, reflecteix l'entorn en el qual es desenvolupà l'obra de Cresques.

Un segle més tard, el segon llibre de matemàtiques que s'imprimia a Europa (després de l'aritmètica anònima de Treviso) és la *Summa de l'art d'Aritmètica* de Francesc Santcliment, un llibre imprès en llengua catalana a Barcelona, l'any 1482, la traducció al castellà del qual constitueix la primera impressió d'un llibre de matemàtiques a Espanya. Eumo Editorial publicà, l'any 1998, una edició facsímil de l'obra de Santcliment, a cura d'Antoni Malet.

La decadència d'una cultura està normalment vinculada a la decadència del país que la té com a pròpia. No és estrany, doncs, que no puguem tornar a parlar de matemàtiques a Catalunya ni als Països Catalans fins molt recentment, si exceptuem el cas del valencià Josep Chaix, autor de treballs de càlcul diferencial i integral, que realitzà amb Pierre Méchain, l'any 1793, els càlculs per a la mesura de l'arc meridià entre els Pirineus i Barcelona.

Ja en ple segle XX, fruit dels moviments de renaixença i de la consolidació d'una identitat pròpia, apareixen dos destacats investigadors de relleu internacional: Lluís Santaló, nascut a Girona i emigrat a l'Argentina, pioner de la geometria integral, l'estereologia i les probabilitats geomètriques, que morí ara fa poc més d'un any; i Ferran Sunyer i Balaguer, a qui està dedicat aquest cicle de conferències, possiblement el millor matemàtic a Catalunya durant la dècada dels cinquanta i els seixanta, tetraplègic de naixement, amic d'Hadamard, Mandelbrot i Malliavin, entre altres, en honor del qual l'Institut d'Estudis Catalans, a través de la Fundació Ferran Sunyer i Balaguer, atorga anualment el premi que duu el seu nom a una monografia que recopili i exposi els avenços més recents en una àrea activa de les matemàtiques.

Hem d'aprendre del passat per poder encarar el futur. A començament dels anys vuitanta, les universitats catalanes —la de Barcelona, l'Autònoma i la Politècnica— començaren a descobrir que es podia apostar per una recerca matemàtica de qualitat. Els diversos departaments de matemàtiques han intensificat la formació de doctors i la producció de treballs d'investigació. El Centre de Recerca Matemàtica ha servit de pont de contacte amb una bona part de la recerca matemàtica que es fa al món desenvolupat.

Per primera vegada a Catalunya —i, tal vegada, a Espanya—, s'estan entenent les paraules pronunciades per Konrad Knopp en la sessió inaugural de curs de la Universitat de Tübingen, l'any 1927: «la matemàtica és la base de tot el coneixement i el contenidor de tota l'alta cultura». Certament, el nostre món és cada cop més complex, i la complexitat de qualsevol sistema augmenta amb el grau d'interconnexió. Un món més interconnectat és, doncs, un sistema complex i, al mateix temps, fràgil i inestable. La matemàtica té un paper cada cop més decisiu en la gestió dels sistemes complexos (tecnològics, financers o socials) del segle que hem començat. Per tant, la matemàtica serà cada cop més un instrument de poder, que de vegades s'està subestimant perillosament.

Ara, iniciat ja el segle XXI, no tenim un Hilbert, com fa cent anys, però disposem d'una xarxa d'investigadors i de coneixement que són la clau per al desenvolupament científic i tecnològic. I, després de segles, els matemàtics catalans estem en condicions de començar a tenir-hi un paper important, com ho feren alguns dels nostres avantpassats.

Una petita excursió al paradís de Cantor

Joan Bagaria

Institució Catalana de Recerca i Estudis Avançats (ICREA) i Departament de Lògica, Història i Filosofia de la Ciència, Universitat de Barcelona

1. INTRODUCCIÓ

El 4 de juny de 1925, en una reunió de la Societat Matemàtica de Westphalia en honor de Karl Weierstrass, David Hilbert (1862-1943) va pronunciar la seva famosa conferència «Sobre l'infinit». Després d'assenyalar el paper cabdal de Weierstrass en l'eliminació de les idees vagues derivades de la noció d'infinitèssim i la conseqüent creació d'una base sòlida per al càlcul infinitessimal, insisteix que l'*infinit* encara té un paper fonamental en l'anàlisi matemàtica. En efecte, l'infinit encara apareix en la definició dels nombres reals com a successions infinites convergents de nombres racionals, i en la noció mateixa de conjunt dels nombres reals. Segons Hilbert, per tal de completar el procés de fonamentació de l'anàlisi matemàtica, cal aclarir la noció d'infinit actual. Diu Hilbert:

> *«L'aclariment definitiu de la* naturalesa de l'infinit *ha esdevingut necessari, no només per als interessos especials de les ciències particulars, sinó sobretot per a* l'honor del mateix enteniment humà.»

En l'anàlisi matemàtica l'infinit apareix com una noció límit, això és, l'*infinit potencial*. Però aquest no és l'infinit real, l'infinit que tenim quan considerem, per exemple, la totalitat dels nombres naturals, o la totalitat dels punts de la recta, com un objecte donat i complet. Això és el que anomenem *infinit actual*. La disciplina matemàtica que tracta de l'infinit actual és la teoria de conjunts, que té els seus orígens en la teoria dels nombres transfinits de Georg Cantor (1845-1918). Aquesta teoria és, segons Hilbert,

> *«...la flor més admirable de l'intel·lecte matemàtic i, en general, un dels més alts assoliments de la pura activitat racional humana».*

En aquesta conferència explorarem alguns dels aspectes més fascinants de la teoria de conjunts. Començarem amb els seus orígens en l'obra de Cantor i la seva formulació axiomàtica de Zermelo-Fraenkel, amb l'axioma d'elecció. Veurem com tots els

objectes matemàtics: nombres, funcions, grups, espais topològics, etc. són, de fet, conjunts, i com tota l'activitat matemàtica de demostració de teoremes es pot reduir, formalment, a fer deduccions en lògica de primer ordre a partir dels axiomes de la teoria de conjunts. D'aquesta manera, la teoria de conjunts constitueix el model estàndard de la matemàtica. Això ens permet fer metamatemàtica, és a dir, estudiar matemàticament la matemàtica mateixa. La teoria de conjunts és sobretot una teoria dels conjunts infinits. Veurem doncs fins on pot arribar la successió dels nombres infinits de Cantor, els àlephs, des del primer nombre infinit $\aleph_0$, que mesura la quantitat de nombres naturals, fins als grans cardinals: els cardinals inaccessibles, mesurables, forts, de Woodin, supercompactes, etc. Mostrarem com es poden demostrar teoremes sobre els nombres reals amb jocs infinits, i com l'existència o no de grans cardinals determina què es pot i què no es pot demostrar. També veurem com es poden donar situacions aparentment paradoxals en matemàtiques, per exemple, com es pot trencar una esfera en 5 trossos i, després de tornar-los a enganxar, obtenir dues esferes iguals a l'anterior. Explicarem com es poden afegir objectes a l'univers matemàtic mitjançant el mètode de *forcing* de Cohen i, d'aquesta manera, demostrar que quelcom no es pot demostrar. Com a exemple, explicarem per què la famosa Hipòtesi del Continu de Cantor, formulada el 1878, encara no està resolta i per què encara no sabem quants nombres reals hi ha.

2. L'ORIGEN DE LA TEORIA DE CONJUNTS. CANTOR

El 1874, Cantor va demostrar que el conjunt dels nombres reals algebraics és numerable, això és, bijectable amb $\mathbb{N}$. Els nombres reals algebraics són les solucions reals de les equacions de la forma

$$a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \ldots + a_1 x + a_0 = 0$$

on els coeficients a_i són enters. Donada una equació com l'anterior, sigui el seu *índex* el nombre natural

$$|a_n| + |a_{n-1}| + |a_{n-2}| + \ldots + |a_1| + |a_0| + n$$

és clar que per a cada k només hi ha un nombre finit d'equacions amb índex k. Per exemple, només hi ha una equació d'índex 2: $x = 0$, que té 0 com a solució. Quatre equacions d'índex 3: $2x = 0$, $x+1 = 0$, $x-1 = 0$ i $x^2 = 0$, que tenen com a solucions 0, –1 i 1, etc. Per tant, podem numerar tots els nombres algebraics d'acord amb l'índex de l'equació de la qual són solucions, i, per a una mateixa equació, en ordre creixent.

Així doncs, els conjunts dels nombres naturals, els nombres enters, els racionals i els algebraics són tots numerables i, per tant, bijectables entre ells.

D'altra banda, i de manera sorprenent, Cantor va demostrar que el conjunt de punts de l'interval [0,1] no es pot numerar. Vegem-ne una demostració senzilla: Suposem que [0,1] i $\mathbb{N}$ fossin bijectables. Cada x de [0,1] té una representació binària, i.e., una funció $f: \mathbb{N} \to \{0,1\}$ tal que

$$x = \sum_{n \in \mathbb{N}} f(n)\, 2^{-(n+1)}$$

Com que cada x de [0,1] té com a màxim dues representacions binàries diferents, tindríem una bijecció $F: \mathbb{N} \to \{0,1\}^{\mathbb{N}}$. Ara sigui $f: \mathbb{N} \to \{0,1\}$ tal que

$$f(n) = 0 \text{ si i només si } F(n)(n) = 1$$

Aleshores, $f \neq F(n)$ per a tot n. Però això és impossible, ja que qualsevol funció $f: \mathbb{N} \to \{0,1\}$ pertany al recorregut de F.

Cantor va demostrar també que $\mathbb{R}$, $\mathbb{R}^n$ i fins i tot $\mathbb{R}^{\mathbb{N}}$ són bijectables entre ells i amb [0,1]. «Ho veig i no m'ho crec!», va escriure Cantor. Tots aquests conjunts són, per tant, no numerables. Així doncs, Cantor, per primera vegada a la història, va descobrir que hi havia almenys dos tipus diferents d'infinit: l'infinit numerable i l'infinit no numerable.

Del 1878 al 1883 Cantor publica una sèrie de treballs que constitueixen l'origen de la teoria de conjunts. Hi introdueix la noció d'ordinal transfinit, o infinit. Quan comptem amb els nombres naturals, el que fem és assignar un nombre a cada objecte que estem comptant, i això ho fem de manera ordenada, respectant l'ordre dels nombres naturals. Quan acabem de comptar, obtenim un nombre, l'últim, que ens diu quants objectes tenim. Però en comptar, també estem ordenant els objectes que comptem, d'acord amb l'ordre dels nombres naturals. Així, els nombres naturals són alhora nombres cardinals, això és, indiquen una quantitat, i nombres ordinals, és a dir, indiquen l'ordre. Però si volem comptar els nombres reals, haurem esgotat tots els nombres naturals abans de comptar-los. Res no impedeix, però, com s'adonà Cantor, d'introduir-hi nombres ordinals infinits i continuar comptant. Per a això, ens cal un nombre ordinal que ens indiqui l'ordre que s'obté quan s'acaben els nombres naturals. D'aquest, Cantor en digué ω. Pertant, els nombres ordinals comencen així, segons Cantor:

$$1, 2, 3, 4, 5, 6, 7, \ldots\ldots, \omega, \omega{+}1, \omega{+}2, \omega{+}3, \ldots\ldots, \omega{\cdot}2, \omega{\cdot}2{+}1, \omega{\cdot}2{+}2, \ldots\ldots,$$
$$\omega{\cdot}3, \omega{\cdot}3{+}1, \omega{\cdot}3{+}2, \ldots\ldots, \omega^2, \omega^2{+}1, \ldots\ldots, \omega^2{+}\omega, \ldots\ldots, \omega^2{+}\omega{\cdot}2, \ldots\ldots, \omega^2{\cdot}2, \ldots\ldots,$$
$$\omega^2{\cdot}2{+}\omega, \ldots\ldots, \omega^3 \ldots\ldots, \omega^\omega, \ldots\ldots, \omega^{\omega^\omega}, \ldots\ldots$$

Disposem ara de més nombres ordinals per continuar comptant, un cop exhaurits els nombres naturals. Cal fer-hi, però, una observació important: en el cas dels nombres infinits, les nocions d'ordinal i de cardinal no coincideixen. Per exemple, ω i $\omega+1$ són ordinals diferents, però els correspon el mateix cardinal. En efecte, ω i $\omega+1$ són numerables, en el sentit que si tenim dos conjunts A i B tals que A té ω elements i B en té $\omega+1$,

aleshores A i B són bijectables amb $\mathbb{N}$. De fet, tots els ordinals infinits que hem descrit són numerables i els correspon, per tant, el mateix cardinal, el cardinal de $\mathbb{N}$, que Cantor representà per $\aleph_0$.

Cantor anomenà *nombres de la segona classe* aquells nombres ordinals infinits que s'obtenen a partir de ω mitjançant els processos d'afegir una unitat i de passar al límit, de tal manera que el conjunt de tots els seus predecessors és un conjunt numerable. Tots els ordinals infinits que hem descrit abans són nombres de la segona classe. Els de la primera classe són els nombres naturals, els ordinals finits. Cantor va descobrir, però, que el conjunt dels nombres de la segona classe no és numerable i que, de fet, el cardinal que li correspon és el següent cardinal infinit, $\aleph_1$. Així tenim els dos primers cardinals infinits: $\aleph_0$, que és el cardinal de $\mathbb{N}$ i coincideix amb l'ordinal ω, i $\aleph_1$, que és el cardinal del conjunt dels nombres de la segona classe.

Aquest procés es pot anar repetint, això és, podem considerar el nombre ordinal corresponent al conjunt d'ordinals de la segona classe, que Cantor representà per ω_1, i podem definir els ordinals de la tercera classe com aquells que s'obtenen mitjançant els processos d'afegir una unitat i de passar al límit, de tal manera que el conjunt dels seus predecessors és de cardinalitat $\aleph_1$, això és, bijectable amb el conjunt de nombres de la segona classe. Al conjunt dels ordinals de la tercera classe correspon el següent cardinal infinit $\aleph_2$, etc. D'aquesta manera s'obtenen els cardinals transfinits:

$$\aleph_0, \aleph_1, \aleph_2, \aleph_3, \ldots\ldots, \aleph_\omega, \aleph_{\omega+1}, \ldots\ldots, \aleph_{\omega_\omega}, \ldots\ldots, \aleph_{\omega_{\omega_\omega}}, \ldots\ldots$$

on per a cada ordinal α tenim un cardinal $\aleph_\alpha$.

L'any 1885 Cantor va sofrir la seva primera crisi mental, deguda sobretot als atacs de Leopold Kronecker (1823-1891) a la nova teoria de conjunts, i als seus intents d'impedir que Cantor publiqués els seus treballs. A aquesta crisi, hi contribuiren també, sens dubte, els intents infructuosos de Cantor de resoldre la *Hipòtesi del Continu*, hipòtesi que havia formulat ja el 1878 i que diu que qualsevol conjunt infinit de nombres reals és o bé numerable o bé bijectable amb el continu $\mathbb{R}$. Com a pas previ, Cantor havia conjecturat i intentat demostrar que qualsevol conjunt, i en particular el conjunt dels nombres reals, és ben ordenat, això és, es pot ordenar totalment, de tal manera que qualsevol subconjunt no buit té un element mínim. Aquest es coneix com el *Principi del Bon Ordre*. Qualsevol conjunt ben ordenat es pot comptar, és a dir, es pot bijectar amb un segment inicial dels nombres ordinals. Si acceptem aquest principi, la Hipòtesi del Continu diu que el conjunt dels ordinals menors a ω_1 és bijectable amb el conjunt dels nombres reals o, en altres paraules, que la cardinalitat de $\mathbb{R}$ és la menor possible, això és, $\aleph_1$. Hilbert va posar la Hipòtesi del Continu en el primer lloc de la seva famosa llista de problemes de l'any 1900.

3. LES PARADOXES

El 1895, Cantor publicà el primer volum del seu treball *Contribucions a la fonamentació de la teoria de conjunts transfinita*, i el 1897, el segon, on desenvolupà la teoria general de conjunts i hi exposà tota la teoria dels ordinals i cardinals transfinits. Però el mateix any 1897 Cesare Burali-Forti (1861-1931) publicà una paradoxa de la teoria de conjunts, que ja coneixia Cantor. Aquesta apareix en considerar el *conjunt OR* de tots els nombres ordinals. A aquest conjunt correspon un nombre ordinal que hauria de ser un membre d'*OR* i alhora més gran que tots els membres d'*OR*. En una carta a Dedekind, del 1899, Cantor solucionà la paradoxa distingint entre el que anomenà multiplicitats consistents, que són els conjunts, i multiplicitats inconsistents, com la classe de tots els ordinals, que duen a inconsistències o paradoxes si es consideren com un conjunt.

Que la noció de conjunt arbitrari és problemàtica es posà de manifest el 1902, amb la famosa paradoxa de Bertrand Russell (1872-1970): sigui A el conjunt de tots aquells conjunts que no són elements d'ells mateixos. Aleshores, si A pertany a A, A no pertany a A, i si A no pertany a A, A pertany a A. La conclusió és que A no pot ser, doncs, un conjunt.

Se segueix que el conjunt de tots els ordinals, el conjunt de tots els conjunts que no es pertanyen a ells mateixos, o el conjunt universal, o conjunt de tots els conjunts, no existeixen!

És clar, doncs, que calia establir de manera precisa quins són els criteris de formació de conjunts arbitraris. Calia una axiomatització de la teoria de conjunts per tal d'evitar les paradoxes o contradiccions derivades de considerar multiplicitats inconsistents com a conjunts. «Ningú no serà capaç d'expulsar-nos del paradís que Cantor va crear per a nosaltres», digué Hilbert en la seva conferència de 1925.

4. LA TEORIA AXIOMÀTICA DE CONJUNTS

La primera axiomatització de la teoria de conjunts fou proposada per Ernst Zermelo (1871-1953), publicada el 1908. Zermelo ja havia demostrat el Principi del Bon Ordre de Cantor, utilitzant el que es coneix com l'*Axioma d'Elecció*, axioma que, de fet, és equivalent al Principi del Bon Ordre.

L'axiomatització de Zermelo fou posteriorment millorada per A. Thoralf Skolem (1867-1963) i A. Abraham H. Fraenkel (1891-1965) i ha esdevingut el que es coneix com a teoria de conjunts de Zermelo-Fraenkel (ZF), amb l'Axioma d'Elecció (Choice), o ZFC.

Els axiomes de ZFC es poden veure com una descripció de l'univers de tots els conjunts purs. Aquest univers forma una jerarquia ben ordenada i cumulativa, indexada pels nombres ordinals, començant amb el 0. Això és, sigui

$V_0 = \emptyset$, el conjunt buit.

$V_{\alpha+1} = P(V_\alpha)$, el conjunt de tots els subconjunts de V_α.

$V_\lambda = \bigcup_{\beta<\lambda} V_\beta$ la unió de tots els V_β, β menor que λ, si λ és un ordinal límit.

L'univers de tots els conjunts és la unió de tots els V_α:

$$V = \bigcup_{\alpha \in OR} V_\alpha$$

Els axiomes de ZFC són els següents:

(1) Extensionalitat. Si dos conjunts tenen els mateixos elements són iguals.
(2) Potència. Donat un conjunt x, hi ha un conjunt $P(x)$ que té com a elements tots els subconjunts de x.
(3) Unió. Donat un conjunt x, hi ha un conjunt $\bigcup x$ que té com a elements tots els elements dels elements de x.
(4) Infinitud. Hi ha un conjunt infinit.
(5) Axioma de substitució. Tota funció definible que té com a domini un conjunt, el seu recorregut també és un conjunt.
(6) Regularitat. Tot conjunt pertany a V.
(7) Axioma d'elecció (AC). El Principi del Bon Ordre, això és, qualsevol conjunt es pot ordenar bé.

Els axiomes de ZFC estan especialment dissenyats per construir V. Per exemple, l'axioma de substitució serveix per poder continuar en els estadis λ límit, això és, per posar junts en un conjunt tots els V_β, β menor que λ, per tal de poder aplicar l'axioma de la unió i així obtenir V_λ. En ZFC, els ordinals s'identifiquen, segons la definició de John Von Neumann (1903-1957), amb el conjunt dels seus predecessors. Així, $0 = \emptyset$, $1 = \{0\}$,..., $n + 1 = n \cup \{n\}$,..., $\omega = \{0, 1, 2, 3, ...\}$, $\alpha + 1 = \alpha \cup \{\alpha\}$, etc. Els nombres cardinals són aquells ordinals que no són bijectables amb cap dels seus elements. Així, $\aleph_0 = \omega$, $\aleph_1 = \omega_1$,...., $\aleph_n = \omega_n$,...., $\aleph_\omega = \bigcup_n \aleph_n$, etc. AC afirma que qualsevol conjunt A es pot ordenar bé i és, per tant, bijectable amb un cardinal.

En la teoria axiomàtica de conjunts, aquests axiomes es formulen de manera precisa utilitzant el llenguatge formal de la lògica de primer ordre amb un únic símbol no lògic, el símbol $\in$, que representa la relació de pertinença. Per exemple, en aquest llenguatge formal, l'axioma d'extensionalitat s'escriu així:

$$\forall x \forall y \left(\forall z \left(z \in x \leftrightarrow z \in y \right) \rightarrow x = y \right)$$

Aquesta formalització és especialment necessària per poder formular correctament l'axioma de substitució, ja que cal precisar què entenem per *definible*. El que

entenem és: «definible per mitjà d'una fórmula del llenguatge formal». Així tenim infinits axiomes de substitució, un per a cada fórmula del llenguatge formal.

5. LA TEORIA DE CONJUNTS I LA FONAMENTACIÓ DE LA MATEMÀTICA

ZFC permet no solament desenvolupar tota la teoria dels nombres transfinits de Cantor sinó que, com a teoria màximament general i abstracta, basada en la noció fonamental de conjunt arbitrari, permet desenvolupar tota la matemàtica clàssica. Això vol dir que tots els objectes matemàtics es poden veure com a conjunts, i que tots els teoremes matemàtics es poden demostrar a partir de ZFC mitjançant les regles lògiques de demostració usuals, que són formalitzables en el llenguatge formal de la teoria de conjunts.

Els nombres naturals són els ordinals finits. Les relacions binàries en un conjunt s'identifiquen amb conjunts de parells ordenats d'elements del conjunt. Per exemple, la relació de divisibilitat entre nombres naturals s'identifica amb el conjunt de tots els parells ordenats $\langle m, n\rangle$ de nombres naturals en els quals m divideix n. El parell ordenat $\langle m, n\rangle$ no és res més que el conjunt $\{\{m\}, \{m, n\}\}$. Els nombres enters es construeixen com a classes d'equivalència de parells ordenats de nombres naturals, i els racionals com classes d'equivalència de parells ordenats d'enters. Els nombres reals no són res més que talls de Dedekind, que són conjunts no buits i afitats de racionals tancats sota predecessors. Les funcions n-àries no són res més que un tipus especial de relacions $n + 1$-àries, on identifiquem la funció f amb el conjunt de n + 1-tuples ordenades $\langle x_1,..., x_n,y\rangle$ tals que $f(\langle x_1,..., x_n\rangle) = y$. Un espai topològic X no és res més que un conjunt amb una topologia, que és un conjunt de subconjunts de X tancat sota unions arbitràries i interseccions finites, etc.

Prenem qualsevol teorema clàssic, per exemple, el teorema de Bolzano, que diu que si $a < b$ són nombres reals, $f: [a,b] \to \mathbb{R}$ és contínua, i $f(a) < 0 < f(b)$, llavors existeix un $c \in (a,b)$ tal que $f(c) = 0$. Acabem d'observar que tots els objectes matemàtics que apareixen en aquest teorema es poden veure com a conjunts, i les propietats bàsiques d'aquests objectes es poden demostrar formalment a partir de ZFC, per exemple, la continuïtat dels nombres reals. Aleshores, la demostració usual del teorema es pot veure com una demostració formal, amb l'aplicació de les regles de la lògica de primer ordre, a partir de ZFC. Per descomptat, escriure la demostració completa fent servir el llenguage formal és alhora difícil i desaconsellable per raons pràctiques, ja que la demostració seria no solament llarguíssima sinó intuïtivament incomprensible. El que és important, però, és veure que, en principi, això és possible.

El fet que tota la matemàtica clàssica es pugui formular i desenvolupar dins de la teoria de conjunts axiomàtica de ZFC fa possible la metamatemàtica, que és l'estudi matemàtic de la matemàtica mateixa. Qüestions com la demostració matemàtica de la possibilitat o impossibilitat de demostrar o no un enunciat matemàtic ara tenen ple sen-

tit, si entenem com a enunciat matemàtic un enunciat del llenguatge formal de la teoria de conjunts, i com a demostració matemàtica una demostració formal a partir de ZFC.

6. ENUNCIATS INDECIDIBLES. ELS TEOREMES D'INCOMPLETESA DE GÖDEL

Sigui φ un enunciat matemàtic, ens podem preguntar si φ és vertader o no. Però en matemàtiques, *vertader* vol dir *demostrable*. La veritat d'un enunciat matemàtic s'estableix amb una demostració a partir de principis bàsics o axiomes. Com podem saber, doncs, si φ és demostrable? Si existeix una demostració de φ, caldrà trobar-la. Però, i si no existeix? Com podem saber que no existeix una demostració de φ? Una manera seria trobant una demostració de la negació de φ. Això ens diria que no hi pot haver cap demostració de φ, si no és que la matemàtica fos inconsistent! Però, i si la negació de φ tampoc no fos demostrable? Aleshores φ seria *indecidible* o *independent* dels axiomes.

Gödel va demostrar el 1931 els seus famosos teoremes d'incompletesa. El primer diu que en qualsevol sistema formal consistent que permeti desenvolupar l'aritmètica bàsica hi ha enunciats indecidibles, enunciats que ni ells ni la seva negació no són demostrables en el sistema. En particular, hi ha enunciats del llenguatge formal de la teoria de conjunts que no són ni demostrables ni refutables a partir dels axiomes de ZFC, suposant, és clar, que ZFC sigui consistent.

És consistent ZFC? El segon teorema d'incompletesa de Gödel, encara més sorprenent que el primer, ens diu que en qualsevol sistema formal consistent que permeti desenvolupar l'aritmètica bàsica, un dels enunciats indecidibles és precisament l'enunciat que afirma la pròpia consistència del sistema. Per tant, si ZFC és consistent, aleshores no podem demostrar la seva consistència a partir dels axiomes de ZFC. L'enunciat que afirma la consistència de ZFC, diguem-ne CON(ZFC), és:

$$0=1 \textit{ no és demostrable en } \text{ZFC}$$

Aquest és un enunciat aritmètic, ja que diu que la successió de símbols $0 = 1$ no és l'últim pas de cap demostració a partir de ZFC, i tant els símbols com els enunciats i les demostracions es poden codificar fàcilment mitjançant successions finites de nombres naturals. Per tant, la conclusió és que, en qualsevol teoria formal axiomàtica que contingui ZFC i sigui consistent, sempre hi haurà enunciats artimètics indecidibles.

ZFC és acceptat actualment com el sistema formal estàndard en el qual es desenvolupa la matemàtica. Això vol dir que un enunciat matemàtic és un teorema, és *vertader* si és demostrable en ZFC. Però, què passa amb els enunciats indecidibles? Si tots els enunciats indecidibles fossin del tipus CON(ZFC) probablement això no ens preocuparia gaire, ja que sembla que no afecten directament els tipus de problemes matemàtics

que hom considera habitualment. Però, per bé o per mal, això no és així. Com veurem més endavant, hi ha molts enunciats matemàtics naturals que són indecidibles en ZFC.

Davant un enunciat matemàtic, tenim tres possibilitats: que sigui demostrable en ZFC, que la seva negació sigui demostrable en ZFC, o que sigui indecidible en ZFC. Però, com es fa per veure que un enunciat matemàtic és indecidible? Això és possible gràcies al fet que ZFC és una teoria formal de primer ordre, i el teorema de completesa de la lògica de primer ordre de Gödel ens diu que per veure que un enunciat φ no és demostrable en ZFC n'hi ha prou de trobar un *model* de ZFC on φ és fals. Un model M de ZFC és un parell ordenat $\langle M, E_M \rangle$ on M és un conjunt no buit, E_M és una relació binària en M, i tots els axiomes de ZFC són vertaders en $\langle M, E_M \rangle$, interpretant els elements de M com a conjunts i E_M com la relació de pertinença. Per tant, si podem trobar models M i N de ZFC en els quals φ és vertader en M i fals en N, voldrà dir que φ és indecidible.

Cal observar que una conseqüència del segon teorema d'incompletesa de Gödel és que no és possible demostrar en ZFC l'existència d'un model de ZFC, ja que l'existència d'un model de ZFC equival a la consistència de ZFC. Per tant, quan parlem de models de ZFC estem suposant que ZFC és consistent. El teorema de completesa diu, precisament, que ZFC és consistent si, i només si, té un model.

Un dels resultats més sorprenents de la matemàtica del segle XX és que la Hipòtesi del Continu de Cantor és indecidible.

7. LA HIPÒTESI DEL CONTINU I

Recordem que la Hipòtesi del Continu diu que qualsevol conjunt infinit de nombres reals o bé és numerable o bé és bijectable amb $\mathbb{R}$. En ZFC, com que tenim el Principi del Bon Ordre com a axioma, la Hipòtesi del Continu és equivalent a afirmar que la cardinalitat de $\mathbb{R}$ és $\aleph_1$.

El 1938, Gödel va construir un model de ZFC en què val la Hipòtesi del Continu. Aquest és l'anomenat *Univers construïble de Gödel* i es representa amb la lletra L. L es construeix com V, però així com per passar de V_α a $V_{\alpha+1}$ preníem *tots* els subconjunts de V_α, per passar de L_α a $L_{\alpha+1}$ només prenem aquells subconjunts de L_α que són *definibles* en L_α. Per construir L no cal fer ús de l'Axioma d'Elecció (AC), però un cop construït es pot comprovar que AC val a L. Això mostra, doncs, que si ZF és consistent, ZFC també ho és, i que, per tant, podem fer servir AC sense por d'introduir cap contradicció a ZF. Que la Hipòtesi del Continu valgui a L és una conseqüència del fet que a L pertanyen només els conjunts estrictament necessaris i, per tant, essent el model de ZFC més petit possible, la cardinalitat de $\mathbb{R}$ hi és també la menor possible, $\aleph_1$.

Com que la Hipòtesi del Continu val a L, i L és un model de ZFC, no es pot demostrar en ZFC la negació de la Hipòtesi del Continu, ja que si això fos possible, la Hipòtesi del Continu seria falsa en tots els models de ZFC.

Després del resultat de Gödel, i davant dels repetits intents infructuosos de demostrar la Hipòtesi del Continu en ZFC, va anar prenent força la idea que potser era indecidible. El que calia era trobar una manera de construir un model de ZFC on la Hipòtesi del Continu fos falsa. Això ho aconseguí 25 anys després Paul J. Cohen (nascut el 1934), amb un nou mètode revolucionari, el *forcing*, per al qual li fou concedida la medalla Fields.

8. EL *FORCING*

La tècnica del *forcing*, inventada per Cohen, és un mètode de construcció de models de ZFC extraordinàriament flexible. Permet construir models amb les propietats més diverses i amb un gran control dels enunciats que acabaran sent vertaders en el model que es construeix, demostrant, d'aquesta manera, la consistència i la indecidibilitat de molts enunciats matemàtics. De manera semblant a com es passa d'un cos K a una extensió algebraica $K[a]$, per exemple, del cos dels racionals $\mathbb{Q}$ al cos $\mathbb{Q}[\sqrt{2}]$, així es passa d'un model M de ZFC a una extensió de *forcing* $M[G]$. Tot i l'analogia, però, les diferències són enormes. El mètode de *forcing* és d'una extraordinària sofisticació i complexitat tècnica, en què intervenen tant aspectes combinatoris o topològics, com lògics i metamatemàtics. Per donar-ne una petita idea, considerem el cas en què volem obtenir, partint d'un model M, una extensió de *forcing* on no valgui la Hipòtesi del Continu. Haurem d'afegir a M, doncs, almenys $\aleph_2$ nombres reals nous. El primer problema que tenim és que, com que M és un model de ZFC, no podem demostrar en ZFC que hi hagi nombres reals que no estiguin ja a M. D'on els traiem, doncs, els reals nous que volem afegir-hi? Per simplificar el problema, suposem de moment que només volem afegir-hi un real nou i, per simplificar-ho encara més, que el que volem afegir-hi és la seva representació binària, una successió infinita de zeros i uns. La idea és treballar amb el conjunt d'aproximacions finites a la successió infinita de zeros i uns que volem afegir-hi. Ordenant les successions finites de manera natural, el que tenim és l'arbre binari complet, i la successió infinita que volem afegir-hi no és res més que una branca infinita d'aquest arbre. La qüestió és trobar una branca que no pertanyi a M. Això ocorrerà si la branca és el que s'anomena genèrica sobre M, que vol dir que interseca tots els conjunts densos i oberts de l'arbre binari, amb la topologia de l'ordre, que pertanyen a M. Però, com podem trobar una branca genèrica sobre M? La resposta és que n'hi haurà si M és numerable! Però com pot ser M numerable si és un model de ZFC, i en ZFC hem demostrat que hi ha conjunts no numerables, com el conjunt dels nombres reals? La resposta ens la dóna el teorema de Löwenheim-Skolem, que diu que per a qualsevol model M de ZFC n'hi ha un altre $N \subseteq M$ que és numerable i que satisfà exactament els mateixos enunciats. Per tant, podem suposar que M és numerable, i prenem una branca b de l'arbre binari que sigui genèrica sobre M. El model resultant $M[b]$ continua sent un model

de ZFC i conté els mateixos ordinals que M. Comprovar tot això no és gens trivial, i per fer-ho cal fer servir la relació de *forcing*, representada per $\Vdash$, establerta entre les aproximacions a la branca genèrica b i les fórmules del llenguatge de la teoria de conjunts, ampliat amb noms per als elements de l'extensió genèrica $M[b]$. El *Teorema de forcing* diu que un enunciat φ val en $M[b]$ si i només si hi ha una aproximació finita $p \subseteq b$ tal que $p \Vdash \varphi$. Però, per fer que la Hipòtesi del Continu falli, cal afegir-hi almenys $\aleph_2$ noves branques genèriques sobre M. Això es fa de manera semblant, però simultàniament amb $\aleph_2$ còpies de l'arbre binari complet. Hi ha una última dificultat, i és que ens hem d'assegurar que en afegir-hi les noves $\aleph_2$ branques, el cardinal $\aleph_2$ de M continuarà sent el $\aleph_2$ de l'extensió genèrica. Podria passar, i de fet passa molt sovint, que alguns cardinals del model base M ja no són cardinals en l'extensió genèrica. Afortunadament, en aquest cas això no passa, a causa de les propietats topològiques de l'arbre binari. Però per comprovar-ho cal altra vegada fer ús de la relació de *forcing* $\Vdash$.

La Hipòtesi del Continu val en el model L de Gödel, i no val en el model construït per Cohen. Per tant, és independent de ZFC.

La tècnica de *forcing* permet afegir a un model M numerable de ZFC gairebé qualsevol tipus d'objecte. A més, el procés es pot iterar i així obtenir extensions genèriques successives. D'aquesta manera es poden construir models de ZFC amb propietats molt diverses, i es pot demostrar la independència de ZFC d'un gran nombre d'enunciats matemàtics. L'allau de resultats d'independència obtinguts amb el *forcing* durants els anys seixanta i setanta van posar de manifest la insuficiència de ZFC per respondre moltes qüestions matemàtiques fonamentals. Calia, doncs, trobar axiomes nous que, afegits a ZFC, permetessin resoldre totes aquestes qüestions.

9. GRANS CARDINALS

Com ja hem vist al començament, l'univers de tots els conjunts V es defineix com la unió de la successió transfinita dels V_α. Ja hem vist que V no pot ser un conjunt, perquè aleshores apareix la paradoxa de Russell. Però si V fos un conjunt, aleshores li correspondria un ordinal, κ, que de fet coincidiria amb el κ-èssim cardinal $\aleph_\kappa$. A més, V_κ seria un model de ZFC. És clar que no podem demostrar en ZFC que hi hagi un κ amb aquestes propietats, ja que, com que V_κ és un model de ZFC, hauríem demostrat en ZFC que ZFC té un model i que, per tant, és consistent, i això ja hem vist que és impossible pel segon teorema d'incompletesa de Gödel. Per què, doncs, no afegim a ZFC l'axioma que diu que hi ha un cardinal κ que és el κ-èssim cardinal i tal que V_κ és un model de ZFC?

Aquest axioma fou proposat el 1930 per Waclaw Sierpiński (1882-1969) i Alfred Tarski (1902-1983), i és el primer dels *axiomes de grans cardinals*. D'un cardinal κ amb les propietats esmentades se'n diu *inaccessible*.

Altres axiomes d'existència de grans cardinals, molt més potents que els inaccessibles, varen anar apareixent successivament al llarg del segle XX. Els més importants, els *cardinals mesurables*, foren descoberts, també el 1930, per Stanislaw M. Ulam (1909-1984), descobridor també, d'altra banda, del mètode d'ignició de la bomba d'hidrògen, com a conseqüència dels seus treballs sobre extensions de la mesura de Lebesgue.

10. CONJUNTS MESURABLES

Recordem que un conjunt de nombres reals és Borel si es pot obtenir a partir dels conjunts oberts, prenent complements i interseccions numerables. Recordem també que un conjunt de nombres reals A té mesura zero, o és nul, si per a qualsevol $\varepsilon > 0$ hi ha una successió d'intervals oberts $\{I_n\}_n$ tal que $A \subseteq \bigcup_n I_n$ i $\sum_n |I_n| < \varepsilon$. A és *Lebesgue mesurable* si és gairebé un conjunt Borel, això és, si difereix d'un conjunt Borel en un conjunt de mesura zero. A cada conjunt mesurable A correspon un nombre $\mu(A) \in [0, \infty]$, la seva mesura, que és invariant per traslació i és σ-additiva, això és, la mesura d'una unió numerable de conjunts mesurables i disjunts dos a dos és la suma de les mesures dels conjunts. A més, la mesura d'un interval és la seva longitud.

En ZFC es pot demostrar l'existència de conjunts de nombres reals que no són Lebesgue mesurables. Per exemple, el conjunt descobert el 1905 per Giuseppe Vitali (1875-1932): sigui A un subconjunt maximal de $[0,1]$ amb la propietat que per a qualsevol $x, y \in A$ diferents, $x - y$ no és racional. L'existència d'A la garanteix l'axioma d'elecció. Per veure que A no és mesurable, per a cada racional p, considerem el conjunt $A_p = \{x+p : x \in A\}$. Tots aquests conjunts són disjunts dos a dos. Sigui B la unió de tots els A_p, on p és un racional de l'interval $[-1,1]$. A no pot tenir mesura zero, ja que aleshores B tindria mesura zero, i això és impossible perquè $[0,1] \subseteq B$. D'altra banda, A tampoc no pot tenir mesura positiva, ja que aleshores B tindria mesura infinita, i això no pot ser perquè $B \subseteq [-1, 2]$.

Trivialment, tots els conjunts Borel són mesurables, i el 1917 Lusin va demostrar que totes les imatges contínues de conjunts Borel, els anomenats *conjunts analítics*, també són mesurables. Com que si un conjunt és mesurable també ho és el seu complement, tots els complements de conjunts analítics, els anomenats conjunts coanalítics, són també mesurables. La pregunta natural és, doncs, si les imatges contínues dels conjunts coanalítics, els anomenats conjunts Σ^1_2, són també mesurables. Però la resposta és indecidible en ZFC! En L hi ha conjunts Σ^1_2 que no són mesurables, i amb *forcing* es poden construir models on tots els conjunts Σ^1_2 són mesurables.

L'existència de conjunts que no són Lebesgue mesurables és deguda al fet que la mesura de Lebesgue és invariant per traslació. De fet, no hi pot haver cap mesura que estengui la mesura de Lebesgue, sigui σ-additiva i invariant per traslació i que mesuri tots els conjunts de nombres reals. La qüestió de si hi pot haver alguna mesura que estengui la mesura de Lebesgue, que sigui σ-additiva i que mesuri tots els subconjunts d'algun con-

junt A, es coneix com el *problema de la mesura*. Ulam va demostrar el 1930 que si existia un conjunt A amb una mesura com aquesta, aleshores o bé la cardinalitat de $\mathbb{R}$ era immensa, o bé existia un cardinal κ anomenat *mesurable*, que és inaccessible i, de fet, molt més que inaccessible, és el κ-èssim inaccessible, etc. L'existència de cardinals mesurables no és, doncs, demostrable en ZFC, fins i tot si hi afegim l'axioma d'existència de cardinals inaccessibles.

11. CARDINALS MÉS GRANS

Després dels resultats d'independència arran de la invenció del *forcing*, es va intentar descobrir si els axiomes de grans cardinals permetien decidir algunes de les qüestions que eren indecidibles en ZFC. Ben aviat es va veure que els axiomes de grans cardinals no permetien dir res de nou sobre la cardinalitat de $\mathbb{R}$ i, per tant, no resolien la Hipòtesi del Continu. Però, sorprenentment, Robert M. Solovay va demostrar el 1969 que si existeix un cardinal mesurable, aleshores tots els conjunts Σ^1_2 de nombres reals són Lebesgue mesurables. El que és sorprenent és que els cardinals mesurables, tan allunyats dels nombres reals en l'univers V, tinguin tanta influència sobre les propietats bàsiques dels nombres reals.

Durant els anys setanta es van proposar nous axiomes de grans cardinals encara molt més forts que l'axioma d'existència de cardinals mesurables. Curiosament, l'existència d'un cardinal mesurable és equivalent a l'existència d'una immersió elemental (diferent de la identitat) de l'univers V en un subunivers M. Això és, una funció

$$j : V \to M$$

que respecta tots els enunciats, i.e., si x és un conjunt i P una propietat definible en el llenguatge formal de la teoria de conjunts, aleshores x té la propietat P en V si i només si $j(x)$ té la propietat P en M. El cardinal mesurable és l'anomenat *punt crític* de la immersió, això és, el primer ordinal κ tal que $j(\kappa) \neq \kappa$. Un resultat famós de Kenneth Kunen de 1971 diu que no hi pot haver cap immersió elemental de V en V. Però exigint que M sigui molt semblant a V s'obtenen axiomes de grans cardinals més i més forts. Per exemple, si demanem que per a cada ordinal α hi hagi una immersió elemental $j : V \to M$ amb punt crític κ i amb $V_\alpha \subseteq M$, κ és el que s'anomena un cardinal *fort*. Si exigim que per a cada α hi hagi una immersió elemental $j : V \to M$ amb punt crític κ i amb totes les successions de longitud α d'elements de V contingudes en M, κ és el que s'anomena un cardinal *supercompacte*, etc. L'existència d'un cardinal supercompacte implica l'existència de molts cardinals forts més petits que ell, i l'existència d'un cardinal fort l'existència de molts cardinals mesurables més petits que ell.

Quin efectes té l'existència de cardinals molt grans, com els supercompactes, sobre qüestions com la mesurabilitat de Lebesgue de conjunts de nombres reals?

12. LA PARADOXA DE BANACH-TARSKI

L'any 1924, Stefan Banach (1892-1945) i Alfred Tarski, basant-se en un teorema de Fèlix Hausdorff (1868-1942) de 1914, van demostrar el teorema següent, conegut com la Paradoxa de Banach-Tarski:

Hi ha una partició de la bola unitat B de $\mathbb{R}^3$, *en 5 peces* $B_0, \ldots, B_4$, *i hi ha isometries* $\rho_0, \ldots, \rho_4 : \mathbb{R}^3 \to \mathbb{R}^3$ *tals que*

$$B=\rho_0[B_0] \cup \rho_1[B_1] \cup \rho_2[B_2] = \rho_3[B_3] \cup \rho_4[B_4]$$

En altres paraules, es pot partir la bola unitat en 5 peces de tal manera que després de tornar-les a enganxar tenim dues boles iguals a la primera (!). De fet, a partir de dos subconjunts afitats A i B amb interior no buit, de l'espai tridimensional $\mathbb{R}^3$, es pot partir A en un nombre finit de peces i reordenar-les mitjançant moviments rígids, de tal manera que formin B. Com és possible que això es pugui fer? No és clarament fals? La raó que es pugui fer és que no totes les peces són Lebesgue mesurables! En la definició de les peces s'utilitza de manera essencial AC. El teorema de Banach-Tarski és potser la conseqüència més antiintuïtiva d'AC. Tot i això, aquest axioma continua tenint un paper fonamental i (gairebé) universalment acceptat per la comunitat matemàtica. Una gran part de la matemàtica actual seria impossible sense l'axioma d'elecció. Les conseqüències antiintuïtives d'AC són degudes al fet que l'axioma afirma l'existència de conjunts que no són explícitament definibles. Per exemple, afirma que existeix un bon ordre de $\mathbb{R}$, però, quin és aquest ordre? Es pot definir d'alguna manera? AC ens permet partir B en peces no mesurables. Però, com són aquestes peces? Les podem definir explícitament? Com veurem, la resposta és que no, si existeixen cardinals supercompactes!

13. DETERMINACIÓ

Una manera eficient de demostrar que un conjunt de nombres reals té una certa propietat és jugant jocs infinits. A partir d'$A \subseteq [0, 1]$, podem jugar al següent *joc* infinit associat a A: hi ha dos jugadors, Adam i Eva, que tiren de manera alternada $n_i \in \{0,1\}$. Adam tira n_0, llavors Eva tira n_1, seguidament Adam tira n_2, i així successivament. Una partida té aquest aspecte:

Adam	n_0	n_2	n_4	...	n_{2k}	...
Eva	n_1	n_3	...	...	n_{2k+1}	...

Al final de la partida, els jugadors han produït una successió infinita n_0, n_1, n_2,... de zeros i uns. Adam guanya el joc si

$$\sum_{i=0}^{\infty} \frac{n_i}{2^{i+1}} \in A$$

I Eva guanya en cas contrari.

El joc està *determinat* si un dels dos jugadors té una estratègia guanyadora. És a dir, si hi ha una funció f que assigna 0 o 1 a cada successió finita de zeros i uns, de tal manera que si un dels jugadors, diguem Eva, juga d'acord amb aquesta funció tira sempre $f(n_0, n_1, \ldots, n_{2k})$ en el k-èssim torn, aleshores guanya el joc, independentment d'allò que tiri Adam. Diem que el conjunt A està determinat si ho està el joc associat a A.

En ZFC es pot demostrar que hi ha conjunts que no estan determinats. Però Donald Martin va demostrar, el 1975, que qualsevol conjunt Borel està determinat, i si existeix un cardinal mesurable, aleshores qualsevol conjunt analític està determinat.

14. ELS CONJUNTS PROJECTIUS I LA TEORIA DESCRIPTIVA DE CONJUNTS

Els conjunts *projectius* de nombres reals són aquells que s'obtenen a partir dels conjunts Borel amb complements i imatges contínues. Els conjunts analítics, els coanalítics i els Σ^1_2 són, doncs, projectius. Els conjunts projectius formen una jerarquia de complexitat creixent, d'acord amb el nombre de passos que cal fer per obtenir-los a partir dels conjunts Borel. L'estudi de les propietats dels conjunts projectius, com la mesurabilitat de Lebesgue, constitueix el nucli del que s'anomena *teoria descriptiva de conjunts*. Si volem demostrar que els conjunts de nombres reals tenen una certa propietat, podem fer-ho demostrant-ho primer per als conjunts Borel, després per als analítics, després per als coanalítics, després per als Σ^1_2, etc. Fins que es demostri per a tots els projectius. Com que tots els conjunts de nombres reals que apareixen normalment en la pràctica matemàtica són projectius, poder demostrar que tenen certes propietats és de gran interès. A més, tots els resultats sobre conjunts projectius de nombres reals tenen una traducció immediata a conjunts projectius en qualsevol espai polonès (separable i completament metritzable), com $\mathbb{R}^n$, $\mathbb{C}$, espais de Banach separables, etc.

Durant els anys seixanta i setanta es van demostrar un seguit de resultats que mostraven que pràcticament qualsevol tipus de propietat que tenen els conjunts Borel també la tenen tots els conjunts projectius, suposant que tots els conjunts projectius estiguin determinats. Així, la determinació dels conjunts projectius, tot i que no es pot demostrar en ZFC, constitueix un excel·lent candidat com a nou axioma de la teoria de conjunts, ja que resol de manera satisfactòria pràcticament qualsevol qüestió sobre les

propietats dels conjunts projectius. Per exemple, si tot conjunt projectiu està determinat, aleshores tots els conjunts projectius són Lebesgue mesurables, no hi ha cap bon ordre projectiu dels nombres reals, i si són no numerables són bijectables amb $\mathbb{R}$.

L'axioma que diu que tots els conjunts projectius estan determinats es coneix amb el nom de *Determinació Projectiva* (PD). PD implica, doncs, que no hi ha cap contraexemple projectiu a la Hipòtesi del Continu, que qualsevol conjunt projectiu és Lebesgue mesurable, i que, per tant, les peces en què es parteix l'esfera en el teorema de Banach-Tarski no poden ser totes projectives. PD està considerat com un axioma alhora natural i extremadament útil per resoldre qüestions sobre els conjunts projectius. De fet, W. Hugh Woodin ha demostrat que, en un cert sentit, PD respon totes les qüestions sobre els conjunts projectius.

Un dels més grans i sorprenents avenços recents de la teoria de conjunts ha estat la demostració que PD se segueix de l'existència de grans cardinals. Donald Martin i John Steel van demostrar el 1988 que si existeixen infinits cardinals de Woodin, els quals es troben entre els cardinals forts i els supercompactes, aleshores PD val. Més encara, W. Hugh Woodin va demostrar que si existeixen infinits cardinals de Woodin i un cardinal mesurable més gran que tots ells, aleshores no només els conjunts projectius sinó tots els conjunts de reals que són definibles a partir dels nombres reals i els ordinals estan determinats. I també va demostrar que l'existència d'infinits cardinals de Woodin és una hipòtesi necessària per obtenir la consistència de PD. Aquests resultats mostren que els axiomes de grans cardinals, almenys l'existència d'un nombre infinit de cardinals de Woodin, són axiomes alhora naturals i raonables, que decideixen la majoria de les qüestions relatives als conjunts definibles de nombres reals.

Tot i l'èxit enorme dels axiomes de grans cardinals per decidir qüestions sobre els conjunts definibles de nombres reals, i moltes altres qüestions en altres àrees de la matemàtica, hi ha moltes preguntes fonamentals que els axiomes de grans cardinals no poden resoldre. La més important és la Hipòtesi del Continu.

15. AXIOMES DE *FORCING*

A part de la Hipòtesi del Continu i del problema de la mesura, un altre problema central per al desenvolupament de la teoria de conjunts ha estat la *Hipòtesi de Suslin*. Cantor havia demostrat que qualsevol ordre lineal S dens, sense punts extrems, separable i complet és isomorf a la recta real. Mikhail Y. Suslin (1894-1919) va conjecturar que si en lloc de suposar que S és separable suposem que és ccc, això és, que qualsevol col·lecció dijunta d'intervals oberts de S és numerable, aleshores S és també isomorf a $\mathbb{R}$. Ronald Jensen va demostrar el 1968 que la Hipòtesi de Suslin falla a L. D'altra banda, Robert M. Solovay i Stanley Tennenbaum van construir el 1971 amb *forcing* un model on val la Hipòtesi de Suslin, i van demostrar així que és indecidible en ZFC. A partir de la construc-

ció del model de Solovay i Tennenbaum, on val la Hipòtesi de Suslin, Donald A. Martin va aïllar un nou principi conegut com l'*Axioma de Martin* (MA). Aquest principi és equivalent a la següent generalització natural del *Teorema de Categoria de Baire* i diu que:

En qualsevol espai compacte Hausdorff i ccc, la intersecció de $\aleph_1$ *conjunts densos i oberts és densa.*

És fàcil de veure que MA implica la negació de la Hipòtesi del Continu. MA s'ha utilitzat amb gran èxit per resoldre moltes qüestions indecidibles en ZFC. Per exemple, implica la Hipòtesi de Suslin i que qualsevol conjunt Σ^1_2 de nombres reals és Lebesgue mesurable. Tot i això, MA ha estat considerat com un principi útil per demostrar que determinats enunciats eren consistents amb ZFC, però no pas com un axioma.

Com a culminació d'un seguit de generalitzacions cada vegada més fortes de MA introduïdes durant els anys setanta i vuitanta, Matthew Foreman, Menachem Magidor i Saharon Shelah van descobrir el 1988 la versió més forta possible de MA, l'anomenat *Axioma de Martin Màxim* (MM), i van demostrar que era consistent amb ZFC, suposant que ho fos l'existència d'un cardinal supercompacte. MM es formula com MA, però per a espais que tenen una propietat anomenada de *preservació de conjunts estacionaris*. El més sorprenent i interessant és que van demostrar que MM implica que la cardinalitat de $\mathbb{R}$ és $\aleph_2$!

La pregunta és: fins a quin punt MA i MM són axiomes naturals de la teoria de conjunts i no només principis útils per obtenir resultats de consistència amb els axiomes de ZFC? D'entrada, no semblen pas principis evidents o naturals, en el mateix sentit que ho són els axiomes de grans cardinals.

16. ABSOLUTESA GENÈRICA

El mètode de *forcing* permet construir extensions de models numerables. Les extensions que s'obtenen depenen d'un ordre parcial $\mathbb{P}$, que és l'ordre parcial format per les aproximacions a l'objecte nou, o genèric, que és el que volem afegir. Però podem considerar també extensions de *forcing*, o genèriques, de V. És clar que aquestes són extensions *ideals*, ja que tots els conjunts ja són a V. A partir d'un ordre parcial $\mathbb{P}$, podem definir el *model booleà* $V^{\mathbb{P}}$, que està contingut en V, però que representa el model que s'obtindria si hi hagués una extensió genèrica de V obtinguda mitjançant $\mathbb{P}$.

Donat un cardinal κ infinit, sigui $H(\kappa)$ el conjunt de tots els conjunts que són hereditàriament de cardinalitat menor que κ, és a dir, el conjunt de tots els conjunts X tals que X, els elements de X, els elements dels elements de X, etc. tenen cardinalitat menor que κ. Així, $H(\omega)$ és el conjunt de tots els conjunts hereditàriament finits i, per tant, $H(\omega) = V_\omega$. $H(\omega_1)$ és el conjunt de tots els conjunts hereditàriament numerables. Sempre tenim que

$H(\kappa) \subseteq V_\kappa$, però no al revés. Si κ és inaccessible, aleshores $H(\kappa)=V_\kappa$. Els conjunts $H(\kappa)$ també formen una jerarquia cumulativa, de la mateixa manera que els V_κ, i la unió de tots els $H(\kappa)$ és V.

Sempre hi ha un cert grau d'absolutesa entre $V^{\mathbb{P}}$ i V. Per exemple, pel Teorema d'absolutesa de Levy-Shoenfield tenim que, a partir d'un enunciat existencial que parla de conjunts de $H(\omega_1)$, si es pogués fer valdre en una extensió genèrica, aleshores valdria també a V. En altres paraules, si mitjançant *forcing* poguessim fer valdre un enunciat que diu que existeix un conjunt X que té determinades propietats que depenen exclusivament de conjunts en $H(\omega_1)$, aleshores un conjunt X amb aquestes propietats ja existeix.

Recordem que l'axioma d'existència d'un cardinal inaccessible diu que si es pogués continuar amb la construcció de V més enllà de tots els ordinals, encara un pas més, aleshores el cardinal que s'obtindria, que seria un cardinal inaccessible, existeix. Generalitzant de manera semblant el teorema d'absolutesa de Levy-Shoenfield, diríem que qualsevol enunciat existencial que parla de conjunts de $H(\omega_2)$ i que val a $V^{\mathbb{P}}$, val també a V. És a dir, si mitjançant *forcing* poguéssim fer valdre un enunciat que diu que existeix un conjunt X que té determinades propietats que depenen exclusivament de conjunts en $H(\omega_2)$, aleshores un conjunt X amb aquestes propietats ja existeix. Representem aquesta afirmació per

$$H(\omega_2) \prec_\exists V^{\mathbb{P}}$$

Aquesta generalització seria, doncs, un nou axioma natural de la teoria de conjunts, si fos consistent. Malauradament, no ho és, si no és que restringim el tipus d'extensions genèriques.

Ara bé, resulta que MA és equivalent a $H(\omega_2) \prec_\exists V^{\mathbb{P}}$ per a qualsevol $\mathbb{P}$ que sigui *ccc*, això és, per a qualsevol $\mathbb{P}$ en el qual tot conjunt d'elements incompatibles dos a dos és numerable. Així, doncs, MA és, de fet, una axioma natural de la teoria de conjunts. Si formulem l'axioma corresponent a MM, és a dir, $H(\omega_2) \prec_\exists V^{\mathbb{P}}$ per a qualsevol $\mathbb{P}$ que té la propietat de preservació de conjunts estacionaris, aleshores obtenim l'anomenat *Axioma de Martin Màxim Afitat* (BMM).

BMM és consistent amb ZFC, suposant que ho sigui l'existència d'un cardinal supercompacte. A diferència de MM, BMM és un axioma natural de la teoria de conjunts, d'un tipus semblant als axiomes de grans cardinals, però que, a diferència d'aquells, decideix moltes qüestions que els axiomes de grans cardinals no responen. Com a culminació d'una sèrie de resultats de W. Hugh Woodin i David Asperó, Stevo Todorčević va demostrar, fa només uns mesos, que BMM implica que la cardinalitat de $\mathbb{R}$ és $\aleph_2$. Aquest és un resultat remarcable, ja que per primera vegada una extensió natural de ZFC decideix la cardinalitat del continu.

17. LA HIPÒTESI DEL CONTINU II

Per acabar, mencionarem breument, i de manera molt informal, alguns resultats recents de Woodin. El primer resultat diu que si hi ha una classe pròpia de cardinals de Woodin, això és, si per a cada ordinal hi ha un cardinal de Woodin més gran, aleshores la teoria de $H(\omega_1)$ és genèricament absoluta. Això vol dir que no podem canviar amb forcing la teoria dels nombres reals, i implica, per exemple, que tot conjunt de nombres reals definible és Lebesgue mesurable.

El segon resultat diu que hi ha un axioma, anomenat $(*)$, que és una generalització de BMM, que decideix *pràcticament* totes les qüestions sobre $H(\omega_2)$, de tal manera que la teoria de $H(\omega_2)$ és genèricament absoluta. L'axioma $(*)$ és l'equivalent a PD, però per a $H(\omega_2)$. L'axioma $(*)$ implica que la cardinalitat del continu és $\aleph_2$.

Finalment, Woodin va demostrar l'any 2000 que si hi ha una classe pròpia de cardinals de Woodin, aleshores qualsevol axioma que decideixi pràcticament totes les qüestions sobre $H(\omega_2)$, de tal manera que la teoria de $H(\omega_2)$ sigui genèricament absoluta, necessàriament ha d'implicar la negació de la Hipòtesi del Continu.

La noció de *pràcticament totes les qüestions sobre* $H(\omega_2)$, per poder ser formulada de manera precisa, necessita de la introducció d'una nova lògica, l'anomenada *Lògica* Ω de Woodin. El problema més important que queda per resoldre sobre la Hipòtesi del Continu és l'anomenada *Conjectura* Ω. La formulació d'aquesta conjectura és de caràcter molt tècnic. La conclusió és que si la conjectura Ω és verdadera, aleshores qualsevol teoria de conjunts compatible amb l'existència de grans cardinals i que faci la teoria de $H(\omega_2)$ genèricament absoluta, ha d'implicar necessàriament la negació de la Hipòtesi del Continu.

Així, si la conjectura Ω és verdadera, aleshores la Hipòtesi del Continu és, en un sentit molt natural, falsa. Les teories: ZFC més la Hipòtesi del Continu i ZFC més la seva negació no són igualment raonables. Els resultats de Woodin mostren que, en presència de grans cardinals, si acceptem la Hipòtesi del Continu, aleshores no podrem respondre moltes qüestions sobre $H(\omega_2)$.

18. OBSERVACIONS FINALS

En aquesta petita excursió pel paradís de Cantor, hem vist que la teoria de conjunts ha experimentat al llarg del segle XX un desenvolupament extraordinari. El que va començar com una teoria matemàtica dels nombres transfinits s'ha desenvolupat fins que ha constituït una teoria general dels conjunts infinits i una fonamentació de tota la matemàtica. Les tècniques desenvolupades per la teoria de conjunts, com el *forcing*, la combinatòria infinita, la teoria de grans cardinals, etc., l'han convertida en una disciplí-

na d'una gran complexitat i bellesa, amb resultats fascinants que estimulen i alhora desafien la imaginació. 125 anys després de la formulació de Cantor, i 103 anys després que Hilbert la posés en el primer lloc de la seva llista de problemes per al segle XX, la Hipòtesi del Continu continuarà essent, sens dubte, un dels problemes més importants de la matemàtica del segle XXI. La reflexió sobre qüestions matemàtiques fonamentals ha generat la teoria de conjunts tal com la coneixem avui.

El fet mateix que hagi estat possible la unificació de tota la matemàtica clàssica dins d'una única teoria màximament general és una fita remarcable. Però més enllà d'això, la riquesa, originalitat i potència de les idees originades en la teoria de conjunts, les fan aplicables a gairebé totes les àrees de la matemàtica. Cada vegada més, el *forcing*, els grans cardinals, els axiomes de determinació, la teoria descriptiva de conjunts, la combinatòria infinita, etc. troben aplicacions en la resolució de problemes que havien estat oberts durant molts anys en àrees com la Topologia, l'àlgebra, l'Anàlisi Real i Complexa, l'Anàlisi Funcional, la Teoria de la Mesura, etc. Al segle XXI, les idees recentment introduïdes per Shelah, Todorčević, i sobretot Woodin, entre d'altres, no solament contribuiran a la resolució de molts problemes matemàtics, tant els coneguts com els nous que sorgiran, sinó que augmentaran substancialment, com ja ho ha fet la teoria de conjunts al llarg del segle XX, la nostra comprensió de l'univers matemàtic.

Recreacions computacionals de la teoria de codis correctors d'errors

Sebastià Xambó Descamps
Universitat Politècnica de Catalunya

RESUM

En el procés d'enviar informació a través d'una canal de comunicació (mòdems telefònics, telèfons mòbils, televisió digital, sistemes d'emmagatzematge d'informació com ara els CD o DVD, comunicacions per satèl·lit o interplanetàries), cal usar codis correctors d'errors per tal d'evitar-ne la degradació provocada pel *soroll* del canal. Això vol dir que la informació ha de ser codificada apropiadament abans de la transmissió i descodificada en el lloc de destinació. Aquests processos, i molt especialment la descodificació, són remarcablement intensius des del punt de vista matemàtic i computacional. En aquesta conferència s'il·lustrarà aquesta afirmació, amb l'explicació dels aspectes matemàtics, algorísmics i computacionals de tres línies de resultats relacionades:

Fins a quin punt pot ser bo un codi corrector d'errors. Aquí s'explicarà la noció de capacitat d'un canal, les dues fórmules fonamentals de Shannon que expressen aquesta quantitat i el celebrat teorema de codificació de canal, també degut a Shannon. En particular es veurà de quina manera aquest teorema precisa el límit teòric de millora que els codis correctors d'errors poden aconseguir en la seva eliminació dels efectes del soroll.

Com es creen bons codis en la pràctica. Després de formular el problema general, s'explicaran els esquemes més importants que s'han introduït en els darrers cinquanta anys i es considerarà amb més detall el cas dels codis de Reed-Solomon generals.

Quins algorismes s'han inventat per descodificar els bons codis. Com en el punt anterior, s'explicaran les idees principals que s'han anat proposant al llarg dels anys i després es considerarà amb més detall el cas de l'algorisme de Berlekamp-Massey-Sugiyama per als codis de Reed-Solomon.

PRELIMINARS

Acrònims

BCH	Bose-Chaudhuri-Hocquenghem
BMS	Berlekamp-Massey-Sugiyama
CBS	Canal Binari Simètric
CS	Canal Simètric
CD	*Compact Disk*
DVD	*Digital Video Device*
ISBN	*International Standard Book Number*
RS	Reed-Solomon

Aspectes computacionals

Els exemples computacionals s'inclouen en forma de llistats. Són transcripcions al català de les parts més rellevants d'exemples elaborats amb la versió anglesa de WIRIS amb què s'han preparat els exemples continguts en el llibre de l'autor, titulat *Block Error-Correcting Codes: A Computational Primer* (CC en aquest article),[1] i no hauria de ser difícil, a qui s'ho proposi, adaptar-los per tal que es puguin executar en un WIRIS en català.[2]

L'ERA DIGITAL

A la base de l'era digital en la qual vivim, d'aquesta era audiovisual i multimèdia, hi ha un article remarcable: *A Mathematical Theory of Communication.* Aquest treball fou publicat el 1948 per Claude E. Shannon (1916-2001) i una de les seves conseqüències fonamentals, entre d'altres que veurem posteriorment, és que, a tots els efectes pràctics, *qualsevol informació es pot representar digitalment*, és a dir, com una successió finita (però usualment molt llarga) de símbols de qualsevol alfabet finit amb almenys dos símbols. Entre els alfabets ocupa un lloc molt destacat el *binari,* $B=\{0,1\}$. Els seus elements, 0 i 1, s'anomenen *dígits binaris* o *bits*. Amb la suma i el producte binaris B és el cos de dos elements. Notem que $1+b$ és la negació de b (la negació de 0 és 1 i la d'1 és 0).

Per exemple, el cinema (i la televisió) ens té acostumats al fet que una pel·lícula consisteix en una successió de fotogrames presentats amb una certa freqüència, usualment 24 per segon. Un fotograma es pot considerar format, tenint en compte la màxima resolució de la vista, per una matriu de píxels, cadascun dels quals queda determinat per un nivell de gris (o de vermell, verd i blau en el cas de píxels de color). Això és encara més explícit en una pantalla de televisió, ja que està formada físicament per píxels. Finalment, el nivell de gris (o d'un dels tres colors bàsics) s'expressa mitjançant un nombre escrit en binari. En els píxels de color, actualment s'usa un nombre entre 0~00000000 i 255~11111111, de manera que cada píxel de color queda representat per 24 bits (el que fa $2^{24} = 16777216$ colors possibles). Així, doncs, el món visual és digitalitzable, i en principi es pot fer de manera que l'ull humà no pugui distingir la imatge

1. Col·lecció Universitext, Spring-Verlag, 2003. Des de http://www.wiris.com/cc/ se'n pot descarregar una versió digital en format pdf. En aquesta presentació, els exemples que en la versió en paper es donen en forma de llistat (de manera similar a com ho fem en aquest article) han estat substituïts per vincles, que hi donen accés via la interfície de WIRIS. Per a una introducció a la plataforma WIRIS, podeu mirar l'apèndix del llibre esmentat, o el manual en línia (menú Editar, icona ?).

2. Els podeu obtenir si escriviu a *sebastia.xambo@upc.es*.

digital de l'original analògic (Figura 1).[3] Les il·lustracions següents s'han extret del *web* de Kodak.

Figura 1. *Fotografies digitals (la primera és d'una càmera fotogràfica digital).*

En el cas de la veu, i dels sons en general, es pren la pressió de l'ona sonora 44.100 vegades per segon[4] i s'assigna un nombre binari de 16 bits per a cada lectura.[5] La llarga seqüència de bits que en resulta és suficient per reproduir el so fins on l'oïda humana pot distingir. La conclusió és que el món *àudio* també és digitalitzable, i això és a la base de la telefonia digital (incloent-hi la cel·lular) o dels CD per a música. En resum, tot el món audiovisual és digitalitzable, i tecnologies tan actuals com ara la del DVD o la de la televisió digital es basen en aquest fet.

SISTEMES DE COMUNICACIÓ DIGITALS

El problema que han de resoldre els *sistemes de comunicació* és posar a l'abast en un cert punt (*destinació*) la informació generada en un altre punt (*font* o *origen*). La transmissió de la informació es fa a través d'un cert *canal*.

Aquí estem interessats en sistemes de comunicació *digitals*. En aquests sistemes, la informació es representa com a seqüències de símbols d'un determinat alfabet finit.

3. L'Adobe Dynamic Media Group publica una bona introducció al vídeo digital, amb el títol *A Digital Video Primer* (estic agraït a Anna Xambó i Sedó per haver-me proporcionat aquesta referència i dades sobre àudio digital).

4. Atès que l'oïda humana no sent sons de més de 20.000 Hz, l'anomenat teorema del mostratge (Nyquist, Shannon) ens assegura que un mostratge equidistant al doble d'aquesta freqüència és suficient per reproduir totes les característiques del so que l'oïda pot distingir. Per a més informació, incloent-hi notes històriques, vegeu el llibre *Communication Systems in Engineering* (2a edició), de J.G. Proakis i M. Salehi (Prentice-Hall, 2002), pàg. 10 i 46.

5. En el DVD-audio la freqüència és 96000 Hz i s'usen 24 bits en cada lectura.

Aquesta frase, per exemple, està formada per una seqüència de lletres i signes de puntuació. Els números de telèfon o els dels comptes corrents es donen com a seqüències de dígits decimals. Els ISBN dels llibres estan formats per dígits i, al final, possiblement la lletra X. Els codis de barres, malgrat la seva aparença, no són més que seqüències de *bits*. En tot cas, de la secció anterior es desprèn que *tota comunicació es pot realitzar mitjançant un sistema digital* i és per això que, d'aquí en endavant, per sistema de comunicació ja entendrem sistema de comunicació digital.

El model de sistema de comunicació que examinarem en aquest article és el que es representa a la Figura 2. L'element més crític d'aquest sistema és el *canal*, ja que és responsable de transportar la informació de l'origen a la destinació. Pot ser un cable (clàssic o de fibra òptica) o pot ser sense fils (enllaços per ràdio, telefonia mòbil, transmissions interplanetàries). Una de les característiques del canal és l'*ona portadora*, que podem imaginar com una ona electromagnètica d'una freqüència determinada W (d'aquesta freqüència, se'n diu *amplada de banda* del canal). El canal no transporta símbols d'informació, sinó *senyals*. Els senyals formen un conjunt finit de *formes d'ona*, és a dir, d'ones electromagnètiques que tenen una forma i una duració predefinides. Per exemple, per a transmissions binàries podem usar els senyals $\{A\cdot\sin(\omega t), -A\cdot\sin(\omega t)\}$, t en un interval $[0,T]$. Per a un alfabet de quatre elements, podem afegir-hi dos senyals, la fase dels quals difereixi en $\pi/2$ dels dos senyals anteriors: $\{A\cdot\cos(\omega t), -A\cdot\cos(\omega t)\}$.[6]

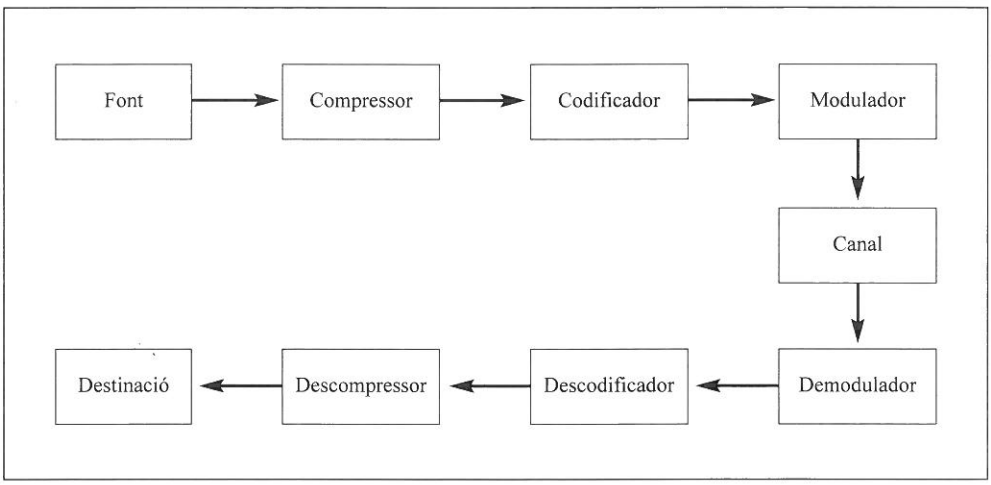

Figura 2. *Model d'un sistema de comunicacions.*

6. Remarcablement, la teoria no depèn de la forma dels senyals usats, sinó només de la seva energia mitjana E_b. Altres senyals emprats en la pràctica inclouen funcions nul·les arreu, llevat en un interval $[0,T]$ en què són constants.

L'element que transforma símbols d'informació en senyals s'anomena *modulador*, i el que fa l'operació inversa a la destinació és el *demodulador*. A l'efecte d'aquest article, podem pensar que el modulador, el canal i el demodulador formen una unitat. La funció d'aquesta unitat és rebre símbols d'informació a l'entrada i retornar símbols d'informació a la sortida. El mecanisme conjunt, però, és útil per explicar l'origen de l'anomenat *soroll* del canal. En efecte, en el seu viatge a través del canal, els senyals experimenten pertorbacions (com ara soroll electromagnètic o soroll tèrmic dels components electrònics) i això fa que hi hagi una certa probabilitat $p(S'|S)$ que el demodulador detecti un senyal S' diferent del senyal S enviat. El resultat és que també tenim una certa probabilitat $p(s'|s)$ de rebre un símbol s' diferent del símbol enviat s. En el cas d'un *canal binari simètric* (CBS), els símbols són bits i

$$p(0|1) = p(1|0) = p,$$

on $0 \leq p \leq 1$. En resulta que

$$p(0|0) = p(1|1) = 1 - p.$$

Això vol dir que si enviem un nombre gran de bits, N, podem esperar (com a mitjana) que n'hi hagi pN d'erronis i $(1-p)N$ de correctes. Anàlogament, un canal per a q símbols es diu simètric (CS) si $p(s'|s)$ té el mateix valor per a totes les parelles $s' \neq s$, i en aquest cas posem p per denotar la probabilitat que un símbol qualsevol sigui alterat en un altre. Vegeu el llistat 1 per a una simulació computacional.

La funció del *compressor*, també anomenat *codificador de font*, és extreure tanta redundància estadística com sigui possible de la informació generada per la font. En el cas del vídeo, per exemple, entre els successius fotogrames sol haver una correlació molt alta, i el mateix sol passar entre els píxels de determinades zones d'un fotograma. Això en la pràctica vol dir que la informació continguda en un vídeo es pot comprimir molt. No essent el tema d'aquest treball, aquí només direm que els algorismes de compressió són molt interessants des del punt de vista matemàtic, i que sense ells el vídeo i la televisió digitals no serien possibles. Naturalment, un dels efectes importants de la compressió és que les demandes sobre el canal de transmissió són molt menors del que haurien de ser si la informació s'hagués d'enviar tal com raja de la font.

Llistat 1. Simulador d'un canal simètric d'alfabet T i probabilitat d'error p

```
CS(X:Llista,T:Llista, p:Real)
comprovar 0≤p & p≤1 & subconjunt?(conjunt(X),T)
:= {cs(x,T,p) amb x en X};

CBS(X,p) := CS(X,{0,1},p);

# Funcions auxiliars
cs(x,T,p):=
 inici
  local r=aleatori()
  si r>p aleshores
   x
  altrament
   element_aleatori(T/{x})
 fi
fi;

element_aleatori(S) := S.aleatori(1,longitud(S));

# Exemple
T={0,1,2}; p=0.2; X=llista_constant(20,0);
CS(X,T,p)
→ {0,2,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0} :: Llista
```

La funció CS(X,T,p) simula un canal simètric d'alfabet T i probabilitat d'error p: si X és la llista dels símbols x que cal tranmetre, retorna la llista dels valors cs(x,T,p). Aquesta funció auxiliar genera un nombre real aleatori r de l'interval [0,1] i retorna el mateix símbol x si $r \leq p$ o un símbol de T diferent de x, escollit aleatòriament, si $p > r$. Notem que si T i U són llistes, T/U dóna la llista obtinguda eliminant de T tots els elements que són de U. Per escollir un element aleatori d'una llista S, es genera un nombre enter aleatori, diguem-ne j, entre 1 i la longitud de S i es retorna el j-èsim element de S. La funció CBS(X,p) simula la transmissió de la llista de bits X per un canal binari simètric amb probabilitat d'error p. Per simplificar, hem dit nombres aleatoris als nombres generats per funcions com ara aleatori() o aleatori(m,n) que en realitat proporcionen nombres pseudoaleatoris.

Simètricament, la funció del *descompressor* és fer l'operació inversa del compressor. Generalment (però no sempre), l'algorisme de compressió és invertible i tant el directe com l'invers solen ser molt eficients.

La part més rellevant per aquest treball és el *codificador* (de canal) i el corresponent *descodificador*. Contràriament al compressor, el codificador introdueix redundància en la informació, amb la idea que el descodificador la pugui usar per corregir els errors produïts pel soroll del canal.

Exemple: el codi de repetició binari de longitud 3

Abans de considerar el cas general, il·lustrem la idea amb un exemple molt simple. Es tracta del codi de repetició d'ordre 3, Rep(3), el qual, tot i la seva simplicitat, ens permet veure que la correcció d'errors és possible. El canal és el binari simètric i la codificació consisteix a repetir cada bit b tres vegades: $b \mapsto bbb$. Per exemple, el resultat de codificar la seqüència 1100101 és la seqüència 111111000000111000111. Veiem que amb aquest esquema de codificació, la informació que s'ha d'enviar pel canal és tres vegades més gran que l'original, i per això diem que la *taxa de transmissió* és 1/3. El

temps de transmissió és, doncs, tres vegades superior, i que ens convingui acceptar aquest sobrecost o no depèn de diversos factors, entre els quals hi ha allò que es pot esperar del descodificador.

En aquest cas, la descodificació es fa extraient de cada grup de tres bits el bit majoritari. És a dir, 000, 100, 010 i 001 es descodifiquen com a 0, mentre que 111, 011, 101 i 110 es descodifiquen com a 1. Si posem p' per indicar la probabilitat que un bit resulti erroni amb aquest esquema, és clar que p' és la probabilitat que el canal introdueixi 2 o 3 errors en un grup de tres bits, és a dir $p'=3p^2\cdot(1-p)+p^3$. El *factor de reducció* d'errors, és a dir, la proporció dels errors que es produeixen usant el codi respecte dels errors sense usar el codi és $p'/p=3p\cdot(1-p)+p^2$. A la Figura 3 hem representat p' i p'/p per $p\in[0,1/2]$. Notem que si p és petit, aleshores $p'\cong 3p^2$ i el factor de reducció d'errors és $p'/p\cong 3p$.

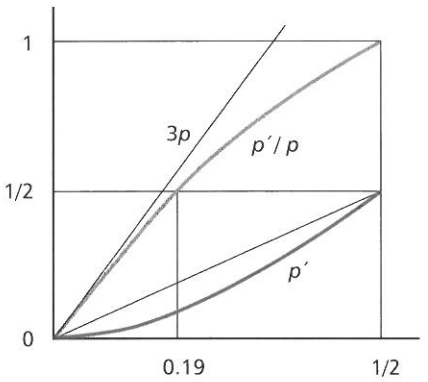

Figura 3. *Gràfiques de p' en funció de p (corba inferior) i de p'/p (corba superior). Aquesta segona funció ens dóna el factor de reducció d'errors del codi de repetició d'ordre 3, ja que representa la proporció d'errors que es produeixen usant el codi respecte dels errors que es produeixen sense usar el codi. Veiem, per exemple, que aquest factor és 1/2 quan $p=0.19$. Per $p=0.001$, $p'/p=0.003$, la qual cosa significa que per mil errors produïts sense codificar, en mitjana se'n produeixen 3 usant el codi.*

Llistat 2. Implementació del codificador i descodificador del codi binari de repetició d'ordre 3

```
f(m:Cadena) :=
inici
  local R=""
  per b en m fer
    R=R|b|b|b
  fi
fi;

g(m:Cadena)
comprovar longitud(m) mod 3 == 0
:=
inici
  local R="", l=longitud(m)/3,
        a, b, c
  per j en 1..l fer
    a=m_{3j-2} ; b=m_{3j-1} ; c=m_{3j}
    si a==b | a==c aleshores
    R=R|a
    altrament_si b==c aleshores
        R=R|b
    altrament
        R=R|"?"
    fi
  fi
fi;
```

La funció de codificació f(m) transforma una cadena m (per exemple de bits) en una cadena R, en la qual cada caràcter b de m s'ha repetit 3 vegades. A l'inici, R es defineix com la cadena nul·la i després, per a cada b de m, s'afegeix a R la cadena bbb. En general, si R és una cadena i b un objecte qualsevol, R|b ens dóna el resultat d'afegir l'objecte al final de R.

La funció de descodificació, g(m), primer comprova que el nombre de caràcters de m sigui un múltiple de 3. Si aquest és el cas, tot seguit trenca m en blocs de 3 caràcters contigus abc, blocs que són examinats successivament. El resultat és R, que a l'inici és la cadena nul·la i que a cada pas es perllonga amb un caràcter x, si x apareix almenys dues vegades en el bloc abc o amb el caràcter d'error ? si els tres caràcters d'abc són distints.

La funció test_rep3(X,p) codifica la cadena X en Y=f(X), tot seguit simula la transmissió de Y per un canal binari simètric de probabilitat p, i el resultat Y´ és descodificat amb g, posant X´=g(Y´). Finalment, es defineix P com la proporció d'errors a Y´, P´ com la proporció d'errors a X, i es retorna {{p,P}, $3\cdot p^2\cdot(1-p)+p^3$,P´}}, la qual cosa permet comparar els valors esperats d'aquestes proporcions (p i $3\cdot p^2\cdot(1-p)+p^3$, respectivament) amb els observats.

Test per al canal binari simètric amb el codi binari de repetició d'ordre 3

```
test_rep3(X,p):=
inici
  local Y=f(X),
        Y´=CBS(Y,p),
        X´=g(Y´),
        P, P´
  P =longitud(posicions_error(Y,Y´)) / longitud(Y) : Flotant
  P´=longitud(posicions_error(X,X´)) / longitud(X) : Flotant
  {{p,P }, {3.p^2.(1-p) + p^3, P´}}
fi;

posicions_error(Y,Z) := {j en 1..longitud(Y) tal_que Y.j!=Z.j};
```

Amb la paraula binària X de longitud N=10000 que només conté 0, una crida de la funció test_rep3(X,p) amb p=0.01 va donar el resultat següent:

```
test_rep3(X,0.01)
→{{0.01, 0.0090667}, {0.000298, 0.0003}}
```

Això mostra que en la «transmissió» del vector codificat s'hi van produir $3N \cdot 0.0090667 = 272$ errors (per terme mitjà n'esperem 300), i que en el vector descodificat hi van quedar 3 errors (en mitjana se n'esperen 2.98).

Espai de Hamming, codis, codificadors i descodificadors

Si T és un alfabet finit i n un nombre enter positiu, els elements del conjunt T^n són les *paraules* de longitud n formades amb elements de T. Sovint dels elements de T^n també se'n diu *vectors* de longitud n.

La *distància de Hamming*, $dh(x,y)$, entre dues paraules x i y és el nombre d'índexs i en el recorregut $1..n$ que compleixen $x_i \neq y_i$.

Llistat 3. Distància de Hamming

```
dh(x,y)
  comprovar longitud(x)==longitud(y)
  := longitud({i en recorregut (x) tal_que x_i≠y_i})
dh("00110","10101")
→3
```

La distància de Hamming és una distància en el conjunt T^n, és a dir, compleix les propietats següents:

1. $dh(x,y) = dh(y,x)$ per a qualsevol $x,y \in T^n$.
2. $dh(x,y) > 0$ si $x \neq y$ i $dh(x,y) = 0$ si $x = y$.
3. $dh(x,z) \leq dh(x,y) + dh(y,z)$ per a qualssevol $x,y,z \in T^n$.

Un *T-codi de longitud n* és un subconjunt C de T^n. La *distància mínima* de C, $d = d_C$, és el mínim dels valors $dh(x,x')$ per $x,x' \in C$, $x \neq x'$.

La seqüència de símbols a la sortida del compressor se sol agrupar en blocs de longitud k, als quals anomenarem *missatges*. Si els símbols formen l'alfabet T, els missatges formen el conjunt $M = T^k$. Un *codificador* de tipus $[n,k]$ és una aplicació injectiva $f: M \to T^n$. La imatge d'aquesta aplicació és un subconjunt C *de* T^n al qual anomenarem *codi* de f. Si $d = d_C$ direm que C és de tipus $[n, k, d]$, i sovint ho indicarem $C \sim [n, k, d]$. Per exemple, en el cas del codi binari de repetició d'ordre 3 tenim $k=1$, $M=\{0,1\}$, $n=3$ i $C=\{000,111\}$, de manera que el tipus de C és [3, 1, 3].

Un *descodificador* per al codificador f és una aplicació $g : T^n \to C \cup \{?\}$ tal que $g(x)=x$ per a tot $x \in C$. Direm que els vectors del subconjunt $D = g^{-1}(C)$ són *descodifica-*

bles per g, mentre que dels vectors $E=g^{-1}(?)$ diem que donen lloc a un *error del descodificador g*. Direm que *g corregeix s errors* (s un nombre enter no negatiu) si tot vector y per al qual existeix un $x\in C$ tal que $dh(x,y)\leq s$ és descodificable i $g(y)=x$. El màxim s tal que g corregeix s errors s'anomena capacitat correctora de g i usualment es denota t. Per exemple, en el cas del codi binari de repetició d'ordre 3, tot vector de $\{0,1\}^3$ és descodificable i $t=1$.

Un codificador f de tipus $[n,k]$ i un descodificador g per a f de capacitat correctora t es poden usar en un sistema de comunicacions de la manera següent:

1. La informació a la sortida del compressor es divideix en blocs u de longitud k.
2. Per a cada bloc d'informació u es calcula $x=f(u)$ i es transmet pel canal.
3. Si y és el vector rebut, calculem $x'=g(y)$.
4. Si $x'=?$, retornem un missatge d'error; altrament retornem x', que és un vector de C.

És clar, directament de les definicions, que si $dh(x,y)\leq t$, és a dir, si el nombre d'errors produïts en el canal no supera t, llavors $x'=x$, la qual cosa vol dir que podem corregir fins a t errors.

Notem que si el nombre d'errors és $>t$, llavors pot passar que y sigui descodificable i que $x'\neq x$, cas en el qual diem que és un *error indetectable.*

Exemple (Codi de Hamming [7,4,3]). Considerem la matriu binària $G=I_4|R^T$, on

$$R=\begin{pmatrix}1&0&1&1\\1&1&0&1\\1&1&1&0\end{pmatrix},$$

i la funció de codificació $f\colon T^4\to T^7$ tal que $u\mapsto uG=u|uR^T$. Sigui $H=R|I_3$ i notem que les columnes de H són els 7 vectors binaris de longitud 3 no nuls. Es comprova immediatament que el codi C de f és $\{x\in B^7|xH^T=0\}$. El descodificador g es defineix mitjançant l'algorisme següent:

1. Donat $y\in B^7$, posem $s=yH^T\in B^3$. Si $s=0$, $y\in C$ i retornem y. Suposem, doncs, que $s\neq 0$.
2. El vector s coincideix amb una columna de H. Si aquesta columna és la j-èsima, retornem el resultat de negar el j-èsim bit de y, y_j, operació que es pot expressar com $y_j\mapsto 1+y_j$.

Vegem que la capacitat correctora de g és 1. En efecte, si $x\in C$ i $y\in B^7$ és tal que $hd(x,y)\leq 1$, aleshores $y=x+e$, on $e\in B^7$ (el *vector d'error*) té com a molt un 1. Si e és nul, llavors $y=x$, $s=0$ i $g(y)=x$ (pel pas 1 de l'algorisme). Si e té un 1 en la posició j, aleshores $s=yH^T=xH^T+eH^T=eH^T=h_j$, on h_j és la columna j-èsima de H.

Notem que això prova que tot element de B^7 és o bé un element del codi o bé està a distància 1 d'un element del codi, i que això implica que la distància mínima de C és 3, de manera que és de tipus [7,4,3]. Veiem, doncs, que aquest codi té la mateixa capacitat correctora que el codi de repetició [3,1,3], però que en canvi la seva taxa de transmissió, 4/7, és força millor que 1/3. Per a una simulació computacional d'aquest codi, vegeu l'exemple 7 (secció 1.1) de WIRIS/*cc*.

Exemple (Un codi [6,2] que corregeix 2 errors). Veurem ara un exemple d'un codi que té taxa de transmissió 1/3, com el codi binari [3,1,3], però que corregeix 2 errors en lloc d'un. L'alfabet d'aquest codi, però, no és B, sinó $T=Z_7$, que és el conjunt $\{0,1,2,3,4,5,6\}$ amb la suma i el producte mòdul 7.

Considerem la matriu

$$G=\begin{pmatrix}1 & 1 & 1 & 1 & 1 & 1\\ 1 & 2 & 3 & 4 & 5 & 6\end{pmatrix}$$

i el codificador $f\colon T^2\to T^6$ tal que $u\mapsto uG$. Com que els elements x del codi C de f tenen la forma

$$x=(\lambda+\mu,\lambda+2\mu,\lambda+3\mu,\lambda+4\mu,\lambda+5\mu,\lambda+6\mu),$$

veiem que un $x\neq 0$ pot tenir com a molt una coordenada no nul·la, d'on resulta que la distància mínima de C és 5. Això ens permet introduir un descodificador $g\colon T^6\to C\cup C^*$, on C^* denota el conjunt de subconjunts de C, que corregeix 2 errors: donat un element $y\in T^6$, si existeix un $x\in C$ tal que $dh(x,y)\leq 2$, aleshores x és únic i posem $g(y)=x$; altrament $g(y)$ és la llista dels elements de C que assoleixen el mínim de $dh(x,y)$ per $x\in C$ (vegeu el Llistat 4 per a una presentació computacional).

Aquest descodificador és una variació de l'anomenat *descodificador de mínima distància*. Si la distància mínima d'un T-codi C de tipus $[n,k]$ és $d=2t+1$ o $d=2t+2$, de manera que $t=\lfloor (d-1)/2\rfloor$, aleshores podem definir un descodificador $g\colon T^n\to C\cup\{?\}$ que corregeix t errors de la manera següent: donat un element $y\in T^n$, si existeix un $x\in C$ tal que $dh(x,y)\leq t$, aleshores x és únic i posem $g(y)=x$; altrament posem $g(y)=?$. Atès que el cardinal del conjunt de missatges T^k és q^k, on q és el cardinal de T, el tractament computacional d'aquest descodificador no pot fer-se, en general, amb les idees del Llistat 4, ja que per descodificar un vector y s'han de trobar de l'ordre de q^k distàncies de Hamming entre vectors de longitud n, la qual cosa és inviable a partir de valors de k no gaire grans.

Llistat 4. Codi [6, 2] sobre Z_7, que corregeix 2 errors

```
K=ℤ7;
T={element(j,K) amb j en 0..6};
G=⌈1 1 1 1 1 1⌉ : Matriu(K);
  ⌊1 2 3 4 5 6⌋
H=nucli(G);   # compleix que C = {y ∈ T^6 | y H=0}
C = {λ·G1+μ·G2 amb (λ,μ) en (T,T)};
g(y):=
inici
  local M=6, m
  si zero?(y·H) aleshores mostrar("Vector del codi"); retornar y fi
  per x en C fer
    m=dh(x,y)
    si m≤2 aleshores mostrar("Vector d'error:", y-x); retornar x fi
    si m<M aleshores M=m fi
  fi
  {x en C tal_que dh(x,y)==M}
fi;

prova(s):=
inici
  x=combinació_lineal(G,K)
  e=vector_error(s,K)
  {x, e, g(x+e)}
fi;

prova(2)
→{[3,5,0,2,4,6], [0,0,4,2,0,0], [3,5,0,2,4,6]}
```

La funció prova(s) primer construeix una combinació lineal aleatòria x de les files de G a coeficients de K, després genera un vector aleatori e de pes s a coeficients de K i finalment retorna la llista {x,e,g(x+e)}. Notem que la suma x+e simula que en la transmissió de x s'han produït s errors aleatoris, tant pel que fa a la posició com pel que fa al seu valor. Si $s \leq 2$, g(x+e) coincideix amb x (vegeu el text). Altrament, és la llista d'elements del codi a distància mínim de x+e.

CAPACITAT D'UN CANAL I EL TEOREMA DE CODIFICACIÓ AMB SOROLL DE SHANNON

L'exemple del codi de repetició d'ordre 3 ens mostra que existeixen esquemes de codificació-descodificació que permeten reduir substancialment, al preu d'augmentar l'ús o les prestacions del canal de transmissió, el nombre d'errors que s'obtindrien sense aquest esquema.

De fet, és fàcil trobar esquemes que milloren el codi de repetició amb un petit augment del cost computacional. Si només ens haguéssim de preocupar de la disminució d'errors, no és difícil veure que els codis binaris de repetició d'ordre senar $n=2t+1$ serien suficients, ja que es pot veure que $p'\to 0$ quan $t\to\infty$. El cost computacional del descodificador és lineal en n, però és clar que la taxa de transmissió és $1/n$, que tendeix a 0 quan $t\to\infty$.

D'altra banda, els exemples del final de la secció anterior mostren que podem corregir un error amb una taxa de transmissió de 4/7 o que, amb un alfabet de 7 símbols, podem corregir dos errors amb una taxa de transmissió d'1/3. Sembla, doncs, que hi ha espai per a les millores i en principi voldríem poder seguir en aquesta direcció, ja que idealment voldríem una taxa de transmissió tan propera a 1 com es pugui i amb una capacitat correctora tan gran com sigui possible.

És clar, així, que ens hem de preguntar:

¿Quins límits es poden assolir, amb esquemes de codificació-descodificació convenients, pel que fa a:

1. *la disminució d'errors,*
2. *la taxa de transmissió, i*
3. *el cost computacional?*

Durant molt de temps es va creure que la condició de tenir una proporció d'errors arbitràriament petita només es podria aconseguir amb taxes de transmissió que tendissin a zero. Però com exposem tot seguit, Shannon no només va provar, en l'article que hem citat al principi, que aquesta suposició era errònia, sinó que va respondre les preguntes 1 i 2 d'una manera òptima.

Per fer-ho va introduir el concepte de capacitat d'un canal. Intuïtivament, és la fracció c de la informació enviada que es pot extreure a l'arribada. Aquesta idea admet una formulació precisa en termes de la matriu $p(s'\mid s)$ de probabilitats de transició d'un símbol enviat s a un símbol rebut s'.[7] Per a canals simètrics és una funció de p. Més concretament, per a canals binaris simètrics la capacitat la dóna la fórmula

$$c(p)=1+p\log_2 p+(1-p)\log_2(1-p)$$

A la Figura 4 hem representat la gràfica d'aquesta funció. També n'hem donat alguns valors que apareixen freqüentment a la pràctica, i que proven que en moltes ocasions $c(p)\cong 1$. En particular tenim la relació $c(0)=1$, que és intuïtivament clara perquè si $p=0$ no hi ha pèrdua d'informació. També tenim $c(1/2)=0$, que també és clara, ja que, per $p=1/2$, de la seqüència de bits a la sortida del canal no podem preveure millor els

7. Per a detalls, vegeu, per exemple, el llibre *Codes and Cryptography*, de D. Welsh (Oxford University Press, 1988).

bits enviats que si ho féssim apostant per cara en el llançament d'una moneda no esbiaixada. La simetria de la gràfica respecte de la recta $p=1/2$ també és clara, ja que si $p>1/2$ aleshores podem obtenir un canal amb la mateixa capacitat (i probabilitat d'error per bit $1-p$) sense més que canviar el paper de 0 i 1.

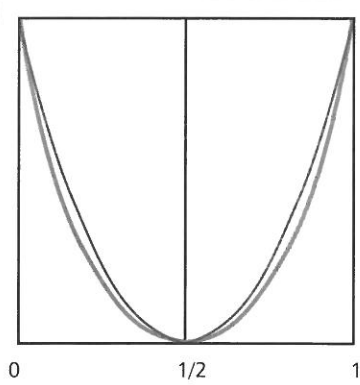

Figura 4. *Gràfica de la capacitat c(p) d'un canal binari simètric amb probabilitat p d'error en un bit. És simètrica respecte de la recta p=1/2. És lleugerament inferior a la gràfica de la paràbola* $4\cdot(p-1/2)^2$ *que interpola els punts* (0,1), (1/2,0) i (1,1).

Valors de $c(p)$ **per** $p=10^{-k}$, $k=1..6$

p	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
$c(p)$	0.531	0.919	0.989	0.9995	0.9998	0.9999

Teorema de codificació de canal (Shannon, 1948). Donat un canal binari simètric de capacitat c, un nombre real $\varepsilon>0$ i un nombre real τ tal que $0<\tau<c$, existeix un codificador $f: B^k \to B^n$ i un descodificador $g: B^n \to C\cup\{?\}$, on C és el codi de f, com ara que $k/n \geq \tau$ i de manera que $p' \leq \varepsilon$, on p' és la probabilitat d'error per bit a la sortida del descodificador.[8]

El teorema de Shannon ens diu que podem tenir probabilitat d'error per bit a la sortida del descodificador arbitràriament petita amb una taxa de transmissió tan propera a la capacitat com vulguem. A més, Shannon també va demostrar que per a taxes de transmissió iguals o superiors a la capacitat, no és possible fer l'error arbitràriament petit. La capacitat és així una mesura del límit que pot assolir la taxa de transmissió si insistim en el fet que la proporció d'errors pugui ser tan petita com vulguem.

Malauradament, la teoria de Shannon és existencial i no dóna cap indicació sobre com es poden obtenir d'una manera constructiva esquemes que gaudeixin de les propietats de l'enunciat i per als quals *el cost computacional sigui acceptable*. La longitud n dels esquemes proposats per Shannon es fa més i més gran conforme reduïm ε o augmentem τ. Això fa que $k=\tau n$ també creixi i, atès que el descodificador considerat és el de mínima distància, la complexitat computacional és exponencial.

8. Per a una demostració, podeu veure el llibre de Wells citat a la nota 7.

Així doncs, haurem de reprendre el camí constructiu iniciat amb els exemples que hem donat anteriorment i mirar d'acostar-nos a l'ideal de Shannon, tot mantenint el cost computacional sota límits manejables. Ho farem en seccions posteriors, en les quals analitzarem algunes de les idees més rellevants que s'han anat proposant per resoldre aquestes qüestions, des de Shannon i Hamming fins al moment present.[9]

La capacitat i el límit de Shannon en la terminologia de telecomunicacions

Fins ara hem usat el llenguatge matemàtic estàndard per exposar l'abast del teorema de Shannon. Però resulta que en els textos d'enginyeria de telecomunicació, com ara el de Proakis-Salehi citat a la nota 4, s'usa una terminologia força diferent que sovint és la causa que sigui costós per als matemàtics esbrinar-ne el significat. En aquesta subsecció considerem quina és la traducció entre els dos llenguatges, pel que fa a la capacitat d'un canal.

La definició que hem donat de la capacitat c no té en compte el ritme en què el canal transmet informació, sinó només quina proporció de la informació que entra en el canal està disponible a la sortida. En la pràctica, però, la capacitat s'expressa en bits/segon, valor que aquí anomenem el *cabal* del canal. El cabal depèn no només de la probabilitat p d'error d'un bit, sinó també de l'amplada de banda W. A més, la probabilitat p no hi apareix d'una manera explícita, sinó implícitament, en el quocient S/N de la potència mitjana del senyal, S, per la potència mitjana del soroll, N. Més concretament, la fórmula de Shannon per al cabal d'un canal binari simètric és [10]

$$c = W \cdot \log_2 \left(1 + \frac{S}{N}\right) \text{ bits/segon.}$$

Posant R per designar la taxa de transmissió del canal en bits/segon, la relació entre S/N i p és[11]

$$p = Q\left(\sqrt{\frac{2E_b}{N_0}}\right), \quad Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-x^2/2}\, dx,$$

on $E_b = S/R$ i $N_0 = N/W$ (Figura 5).[12]

9. Per aprofundir en aquesta matèria vegeu, per exemple, *The Theory of Information and Coding: A Mathematical Framework for Communication*, de R.J. McEliece (Addison-Wesley, 1977).

10. Proakis-Salehi, pàg. 586.

11. *Ibíd.*, pàg. 406.

12. En relació a E_b, vegeu la nota 6.

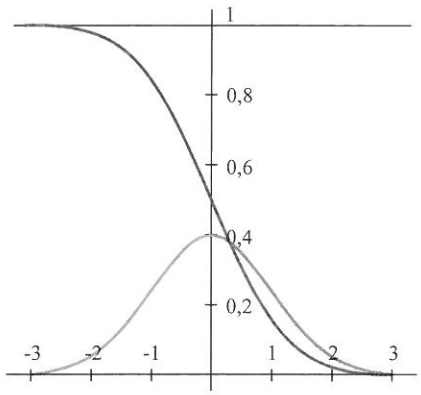

Figura 5. *Gràfica de la funció* $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ (campana de Gauss) *i de la funció* $Q(x)=\frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-x^2/2}dx$. *Per a un* CBS, *el valor de* $Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$ *ens dóna la probabilitat p d'obtenir un bit erroni a la sortida del canal. La funció Q(x) ha estat aproximada amb la fórmula* $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\cdot t\cdot f(t)$, *on* $t=1/(1+kx)\,i$, *per a certes constants reals k,* a_i (vegeu el Llistat 5), $f(t)=a_0+a_1t+a_2t^2+a_3t^3+a_4t^4$.

La quantitat E_b/N_0 no té dimensions i s'anomena *raó senyal-soroll*. Aquí la denotarem ρ.[13] En la pràctica de les telecomunicacions, és costum donar ρ en dB.[14] També és habitual escollir una escala logarítmica per a l'eix d'ordenades p (vegeu la Figura 6).

La gràfica resultant és especialment útil per visualitzar els beneficis de la codificació pel que fa a la disminució d'errors, com ara amb el codi de repetició [3,1,3]. En efecte, per a cada ρ tenim valors p i p' de la probabilitat que un bit sigui erroni sense codificar i amb codificació, respectivament. Per exemple, en el cas del codi de repetició [3,1,3], sabem que $p'=3p^2(1-p)+p^3\cong 3p^2$ i la corba que resulta és la que hem indicat amb R a la Figura 6. Anàlogament, la corba H correspon al codi de Hamming [7,4,3], amb $p'\cong 21p^2$.

13. La notació típica dels llibres de telecomunicació és SNR (*Signal to Noise Ratio*).

14. Si X és un nombre adimensional, $X \sim x$ dB, on $x=10\cdot\log(X)$. Així, multiplicar (dividir) per 10 es tradueix en sumar (restar) 10 dB. Viceversa, la quantitat que correspon a x dB és $X=10^{x/10}$.

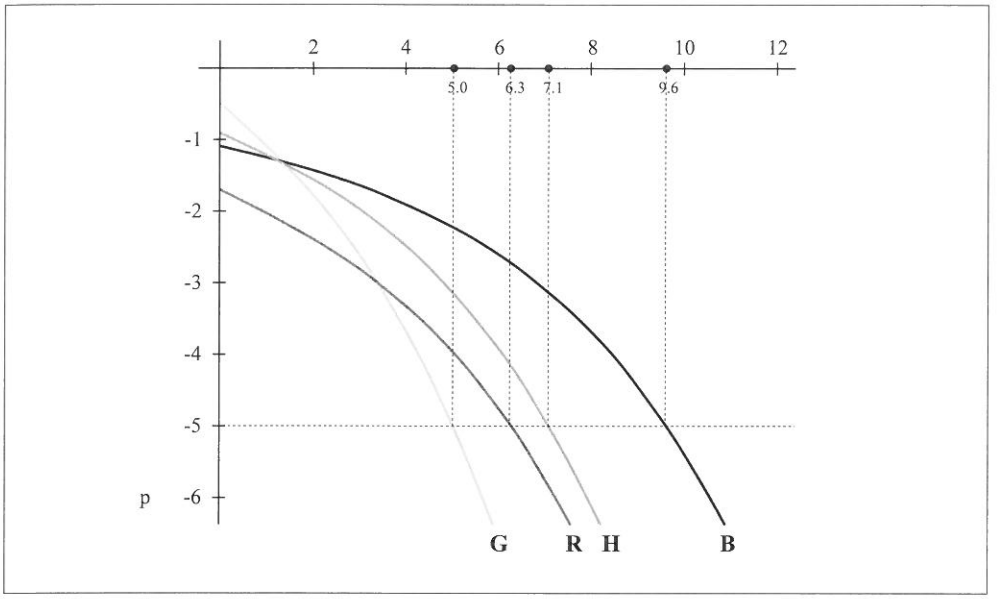

Figura 6. *La corba* B *és la gràfica de la funció* $p=\beta(\rho)$, *on* $\beta(\rho)=Q(\sqrt{2\rho})$, $\rho=E_b/N_0$ (eix d'abscisses, en dB). *L'eix de p* (eix d'ordenades) *s'ha graduat a escala logarítmica. La corba* R *és el resultat de representar el punt* (ρ, p') *corresponent al codi binari de repetició* [3,1,3], *és a dir, amb* $p'\simeq 3p^2$. *La corba* H *és el resultat de representar el punt* (ρ, p') *corresponent al codi de Hamming* [7,4,3], $p'\simeq C(7,3)=21p^2$. *Finalment, i d'una manera similar, la corba* G *correspon al* codi de Golay binari [23,12,7], *amb* $p'\simeq C(23,4)p^4=8855p^4$.

La posició relativa de les corbes H *i* R *no permet concloure que el codi de Hamming és pitjor que el de repetició, ja que en aquestes gràfiques només es té en compte la probabilitat d'error dels codis, i no la seva taxa de transmissió: el codi de Hamming té una taxa d'errors pitjor que la del de repetició, però la seva taxa de transmissió és superior. El codi de Golay, en canvi, té una probabilitat d'error baixa i una taxa superior a* $1/2$, *la qual cosa indica que és millor que* R *i* H.

A la figura també hem inclòs les projeccions sobre l'eix d'abscisses dels punts d'intersecció de la recta $p = 10^{-5}$ *amb les quatre corbes. Per als càlculs corresponents a aquestes gràfiques, vegeu el* Llistat 5.

Llistat 5. La funció Q(x)

```
Q(x : ℝ) :=
inici
  local k=0.2316419, t=1/(1+ k.x), f
  si x<0 aleshores
    retornar 1-Q(-x)
  fi
  f=polinomi([0.31981530, –0.356563782, 1.781477937, –1.821255978, 1.330274429], t)
  e^(-x^2/2)/√(2π) ·t·f
fi
```

Gràfiques de la Figura 5

C=corba(Q,-3..3); N=corba ($\frac{e^{-x^2/2}}{\sqrt{2\pi}}$,-3..3);

Gràfiques B, R, H i G de la Figura 6

dBi(x) := $10^{x/10}$
ber(f) := log(evaluate(f, Q($\sqrt{2\cdot dBi(x)}$)))
B=corba(ber (p),x, 0..12);
R=corba(ber ($3\cdot p^2$), x, 0..12);
H=corba(ber ($21\cdot p^2$), x, 0..12);
G=corba(ber ($8855\cdot p^4$), x, 0..12);

Punts d'intersecció de B, R, H i G amb p=10^{-5}

δ= 0.00001;
BER(f) := evaluate(f, Q($\sqrt{2\cdot dBi(x)}$))
ρ(f) := numerical_solve(BER(f) – δ, x);

$\rho(p)$, $\rho(3\cdot p^2)$, $\rho(21\cdot p^2)$, $\rho(8855\cdot p^4)$

→9.5886, 6.2588, 7.0901, 5.0334

El cabal $c\to\infty$ quan $S\to\infty$, mantenint W i N constants, però l'augment és logarítmic. Això és raonable, perquè a més energia del senyal, més fina pot ser la subdivisió en nivells distingibles a la sortida i, per tant, més gran el nombre de bits transmesos per segon. Aquest possibilitat és l'única disponible en el cas de línies d'amplada de banda limitada.

D'altra banda, si fem $W\to\infty$ mantenint S constant (que en la pràctica vol dir que es disposa d'una amplada de banda il·limitada i d'una potència limitada), aleshores $c\to\log_2(e)\,S/N_0$. Vegeu la Figura 7.

Per tal que la codificació sigui acceptable per corregir errors, és necessari que $R=k/n$ compleixi

$$R<W\cdot\log_2\left(1+\frac{S}{N_0W}\right).$$

Tenint en compte que $S=E_bR$, obtenim, posant $\eta=R/W$,

$$\eta<\log_2\left(1+\eta\,\frac{E_b}{N_0}\right),$$

que equival a

$$\frac{E_b}{N_0} > \frac{2^\eta - 1}{\eta}.$$

Aquesta desigualtat és una condició fonamental que ha de complir el sistema de comunicació per tal que la correcció d'errors sigui possible en un grau satisfactori. En direm el *límit de Shannon.* L'hem representat a la Figura 8 de dues maneres diferents. La quantitat η, que té unitats de bit/seg/Hz, és un paràmetre essencial del sistema de comunicació i s'anomena *eficiència espectral.* Per exemple, en el canal binari simètric sense codificació $\eta = 1$ (en cada cicle de l'ona portadora viatja un bit). Si en el mateix canal usem una codificació $[n,k]$, l'eficiència espectral és k/n, la taxa de transmissió del codi. Atès que

$$\lim_{\eta \to 0} \frac{2^\eta - 1}{\eta} = \ln(2),$$

veiem que $\ln(2)=0.693 \sim -1.6$ dB és el mínim valor de ρ per al qual la correcció d'errors és possible.

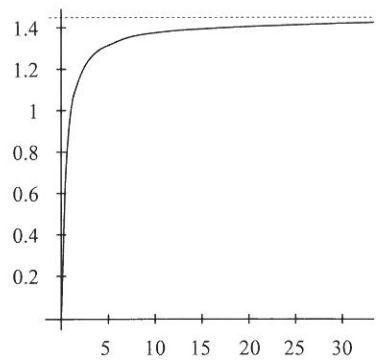

Figura 7. *Cabal d'un canal (per a una potència del senyal S donada) en funció de l'amplada de banda. Com a unitat sobre els dos eixos hem pres S/N_0.*

Notem que el límit superior és, en aquestes unitats, tal com hem vist anteriorment, $\log_2(e) = 1.44$. Per tant, l'augment de la capacitat que resulta d'augmentar l'amplada de banda comença a ser negligible a partir d'un múltiple relativament petit de S/N_0.

La Figura 8 ens proporciona un marc per comparar diversos sistemes de codificació en un mateix canal binari simètric. Fixem una probabilitat d'error δ com a referència (en els textos es pren convencionalment $\delta=10^{-5}$, i aquí farem el mateix). Donat un codi, sigui $f=f(p)$ el polinomi que ens dóna la probabilitat d'error p' del codi i considerem el punt $P=(\rho',\eta)$ definit de manera que η és l'eficiència espectral del sistema codificat amb aquest codi i ρ' el valor definit per les relacions $Q(\sqrt{2\rho'})=p$, $f(p)=\delta$, $p>0$. Direm que P és el *punt de Shannon* del codi, i de la diferència entre ρ' i el valor del límit de Shannon per a η en direm el *dèficit* del codi. Notem que $\rho'=\rho\,(f)$, amb notacions del Llistat 5. Donada una eficiència espectral, un codi serà millor com més proper del límit de Shannon estigui el seu punt de Shannon, és a dir, com més petit sigui el seu dèficit. Anàlogament, amb un valor de ρ', com més gran sigui η millor serà el codi. A la Figura 9 n'hem inclòs alguns exemples.

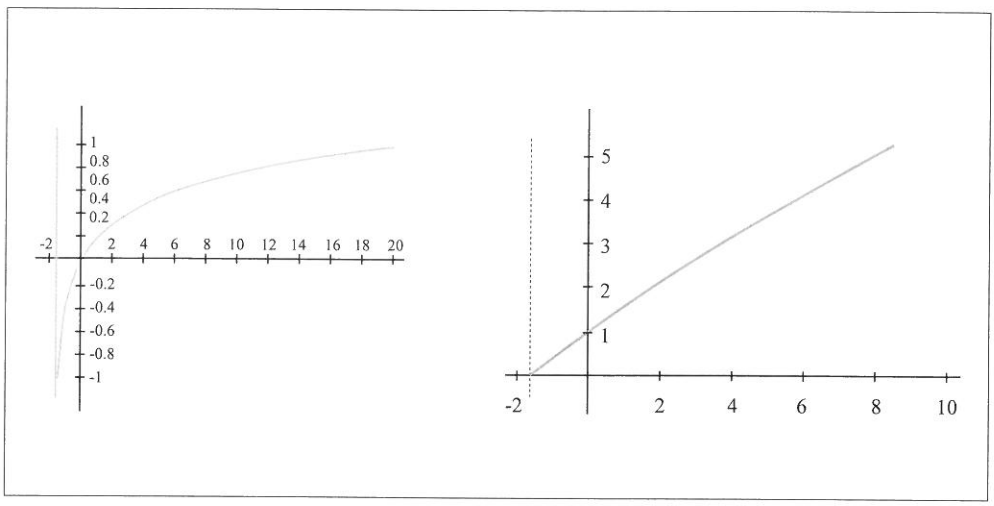

Figura 8. *Límit de Shannon. A l'eix d'abscisses representem* $\rho=E_b/N_0$, *en* dB, *i a l'eix d'ordenades* η. *A la figura de l'esquerra,* η *s'ha disposat a escala logarítmica i a la figura de la dreta, a escala normal (*per a la generació dels gràfics, vegeu el Llistat 6*).*

Llistat 6. Cabal i límit de Shannon

```
cabal(w) := w·log₂(1+ 1/w );
corba(cabal,0.01..22)
límit_shannon(r) := dB( (2^r−1)/r );
corba({límit_shannon(r), r}, r, 0..5)
```

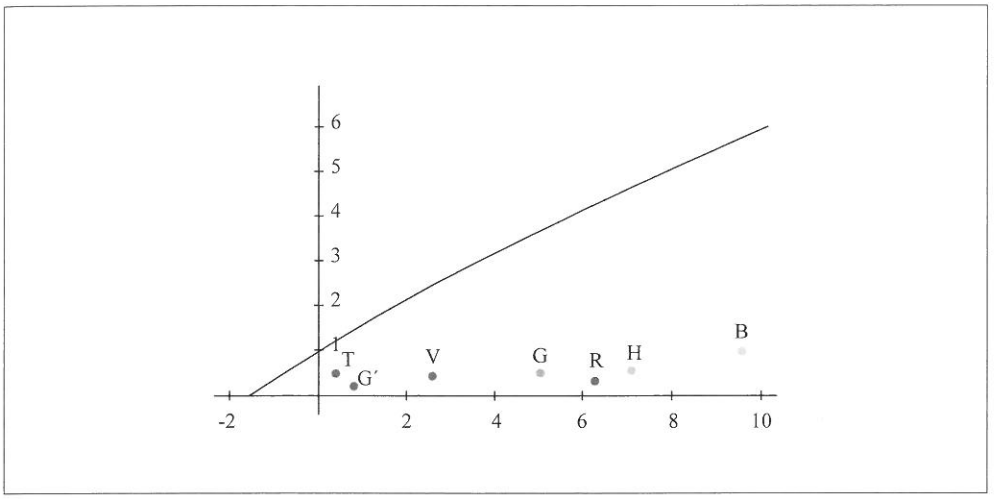

Figura 9. *Límit de Shannon i 7 punts de Shannon: el punt* B, *amb dèficit 9.6* dB, *correspon al* CBS *sense codificar;* H, *amb dèficit* 7.8 dB, *al codi de Hamming* [7,4,3]*;* R, *amb dèficit* 7.3 dB, *al codi de repetició* [3,1,3]*;* G, *amb dèficit* 5.8 dB, *al codi de Golay binari* [23,12,7] (per al càlcul del dèficit d'aquests codis, vegeu el Llistat 7)*; també hi hem inclòs els punts* V (Voyager, 1986); G' (Galileo, 1989) *i* T (un *turbo-codi* actual).

Llistat 7. Punts de Shannon i dèficits

```
dèficit(f,r) :=ρ (f) – límit_shannon(r)
```

CONSTRUCCIÓ DE BONS CODIS

En aquesta secció primer il·lustrem les tècniques algebraiques de construcció de codis en el cas dels codis de Reed-Solomon, subratllant-ne el tractament computacional. Després donem algunes indicacions generals sobre altres famílies de codis que es construeixen d'una manera similar.

Preliminars

Cossos finits. Posarem $\mathbb{Z}_n$ per denotar l'anell de classes de nombres enters mòdul n. Una representació convenient d'aquest anell és el conjunt $\{0,1,...,n-1\}$ amb les operacions de suma i producte ordinàries d'enters, però reduïdes mòdul n. Si p és un nombre primer, aleshores $\mathbb{Z}_p$ és un cos.

Donat un nombre finit q, existeix un cos de cardinal q, si i només si q és potència

d'un nombre primer, és a dir, si i només si $q=p^r$ amb p primer i r enter positiu. A més, si q té aquesta forma, aleshores hi ha un únic cos, llevat d'isomorfismes, el cardinal del qual és q. Aquest cos serà denotat $\mathbb{F}_q$.[15] Notem que si $q=p$ és primer, aleshores $\mathbb{F}_p=\mathbb{Z}_p$. En general, si $f\in\mathbb{F}_p[T]$ és un polinomi irreductible de grau r, el quocient

$$F=\mathbb{F}_p[T]/(f)$$

és un cos de p^r elements. A més, si t és la classe de T mòdul f, els elements de F tenen la forma $a_0+a_1t+\cdots+a_{r-1}t^{r-1}$, $a_i\in\mathbb{F}_p$.

Finalment es compleix que $\mathbb{F}_{p^s}$ és isomorf a un subcòs de $\mathbb{F}_{p^r}$ si i només si s és un divisor de r, i en aquest cas $\mathbb{F}_{p^r}$ té un únic subcòs de cardinal p^s.

La construcció amb WIRIS de l'anell $\mathbb{Z}_n$ es fa amb el símbol $\mathbb{Z}$ de la paleta *Símbols* i l'enter n introduït amb la icona de subíndex de la paleta *Operacions*. Per construir el cos $F=\mathbb{F}_{p^r}$ ho podem fer directament amb la funció *cos_finit*($\mathbb{Z}_p$, r, t). Per exemple, *cos_finit*(*B, 8,* α) construeix el cos F_{256} i els seus elements tenen la forma

$$a_0+a_1\alpha+\cdots+a_7\alpha^7,\ a_i\in B=\mathbb{Z}_2$$

Si coneixem un polinomi mònic irreductible $f=f(t)$ de grau r sobre $\mathbb{Z}_p$, l'expressió $\mathbb{F}_p[t]/(f)$ és una expressió vàlida de WIRIS que construeix el quocient $\mathbb{F}_p[T]/(f)$ i assigna a t la classe de T mòdul f. Per exemple, si $B=\mathbb{Z}_2$ llavors el polinomi $1+\alpha+\alpha^3+\alpha^5+\alpha^8\in B[\alpha]$ és irreductible, i així l'expressió $B[\alpha]/(1+\alpha+\alpha^3+\alpha^5+\alpha^8)$ construeix $\mathbb{F}_{256}$ amb un element distingit α que compleix

$$\alpha^8=1+\alpha+\alpha^3+\alpha^5.$$

Algorisme d'Euclides-Sugiyama. L'entrada d'aquest algorisme està formada per dos polinomis r_0, $r_1\in F[z]$, F un cos, i un nombre enter positiu t tal que $\deg(\delta)<t\le\deg(r_1)$, on $\delta=\mathrm{mcd}(r_0,r_1)$. La sortida està formada per una llista de dos polinomis $\{\sigma,\omega\}$ tals que $\omega\equiv\sigma r_1 \bmod r_0$ i amb $\deg(\sigma)<\deg(r_0)-t$, $\deg(\omega)<t$. La descripció de l'algorisme:

1. Siguin r_i, $i=0,...,j$, els polinomis calculats a partir de r_0 i r_1 amb l'algorisme d'Euclides, però sent j el primer índex, que compleix $\deg(r_j)<t$. Així doncs, per $i=2,...,j$, tenim $r_i=r_{i-2}-q_ir_{i-1}$, on q_i denota el quocient de la divisió entera de r_{i-2} per r_{i-1}.
2. Definim v_0, v_1,..., v_j de manera que $v_0=0$, $v_1=1$ i $v_i=v_{i-2}-q_iv_{i-1}$.
3. Retornem $\{\sigma,\omega\}$, on $\sigma=v_j$ i $\omega=r_j$.

15. Aquests cossos es coneixen com a *cossos de Galois*, en honor al seu descobridor, Evarist Galois, i sovint es denoten $GF(q)$.

Llistat 8. Algorisme d'Euclides-Sugiyama

```
sugiyama(r0,r1,t) :=
inici
  local r, v0=0,v1=1,v, q
  mentre t ≤ degree(r1) fer
    q=quocient(r0,r1); r=residu(r0,r1); v=v0 - q·v1
    r0=r1; r1=r
    v0=v1; v1=v
  fi
  {v1,r1}
fi

sugiyama(z^6, z^3+z+1,2)
→{z^4- z^3-z^2+z+3, 4z+3}
```

Codis de Reed-Solomon

Sigui F un cos finit i q el seu cardinal. Siguin $\alpha_1,...,\alpha_n \in F$ elements no nuls distints de F (per tant ha de ser $q>n$) i k un nombre enter positiu tal que $k \leq n$. El codi de Reed-Solomon de dimensió k associat a $\alpha_1,...,\alpha_n$, que denotarem $\mathrm{RS}(\alpha_1,...,\alpha_n; k)$, és la imatge per l'aplicació lineal $f: F^k \to F^n$ tal que $u \mapsto uG$, on $G=V_k(\alpha_1,...,\alpha_n)$ és la matriu de Vandermonde de k files formada amb els elements $\alpha_1,...,\alpha_n$. El codi $C=\mathrm{RS}(\alpha_1,...,\alpha_n; k)$ és un F-codi de tipus $[n, k]$. En el cas que $\alpha_1,...,\alpha_n$ siguin tots els elements no nuls de F, escriurem simplement $\mathrm{RS}(F,k)$.

Recordem que $V_k(\alpha_1,...,\alpha_n)$ està formada pels elements $\alpha_i^{j-1} (1 \leq i \leq n,\ 1 \leq j \leq k)$ on i és l'índex de columna i j l'índex de fila. Explícitament,

$$V_k(\alpha_1,...,\alpha_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \vdots & \vdots & & \vdots \\ \alpha_1^{k-1} & \alpha_1^{k-1} & \cdots & \alpha_n^{k-1} \end{pmatrix}$$

La matriu $V_n(\alpha_1,...,\alpha_n)$ és quadrada d'ordre n i el seu determinant, que denotarem $D(\alpha_1,...,\alpha_n)$, ve donat per la fórmula

$$D(\alpha_1,...,\alpha_n) = \prod_{1 \leq i < j \leq n} (\alpha_j - \alpha_i)$$

En particular resulta que $V_n(\alpha_1,...,\alpha_n)$ és una matriu invertible.

Llistat 9. Construcció de la matriu de Vandermonde

```
vandermonde(a:Vector, r:Enter):=
  inici
  local n=longitud(a),
        h=vector_constant(n,1),
        H=h, # una seqüència de vectors amb primer terme h
        f=h # el vector corrent de H
  per i en 1..(r-1) fer
    f=prod(f, a); H=(H,f)
  fi
  [H]
fi

prod(a:Vector|Llista, b:Vector|Llista) := [a_i· b_i amb i en recorregut(a)]
```

Exemple

a =[element(j,$\mathbb{Z}_{17}$) amb j en 1..15..2]

→

[1,3,5,7,9,11,13,15]

vandermonde(a,4)

→

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 11 \\ 1 & 3 & 5 & 7 & 9 & 11 & 13 & 15 \\ 1 & 9 & 8 & 15 & 13 & 2 & 16 & 4 \\ 1 & 10 & 6 & 3 & 15 & 5 & 4 & 9 \end{pmatrix}$$

Teorema. *La distància mínima de* RS($\alpha_1,...,\alpha_n$; k) és $n-k+1$.

Prova: Sigui $u\in F^k$ i $x=uG$. La component i-èsima de x és

$$x_i=u_1+u_2\alpha_i+\cdots+u_k\alpha_i^{k-1}=\hat{u}(\alpha_i),$$

on $\hat{u}=u_1+u_2T+\cdots+u_kT^{k-1}$. Com que un polinomi no nul de grau $\leq k-1$ té com a molt $k-1$ arrels, veiem que x_i pot ser 0, si $u\neq 0$, per a $k-1$ de les α_i com a molt. Això significa que x_i és no nul per a, almenys, $n-k+1$ valors de i. Per tant, qualsevol element no nul de C té un pes no inferior a $n-k+1$ (el *pes* d'un vector és el nombre de components d'aquest vector que són no nuls). En resulta que la distància de Hamming entre dos elements distints qualsevol x, x' de C és almenys $n-k+1$, ja que $dh(x,x')$ és el pes de $x-x'$ i $x-x'$ és un element de C. Així doncs, $d_C\geq n-k+1$. Com que la desigualtat contrària és certa per a qualsevol codi (desigualtat de Singleton; cf. CC, secció 1.1), aquesta desigualtat és una igualtat (aquest fet s'expressa dient que els codis RS són de *màxima distància*). □

Observació. Suposem que volem un codi RS que corregeixi almenys t errors (amb el descodificador de mínima distància) i una taxa de transmissió $\geq r>0$. Quin és el mínim q que podem usar? Vegem-ho. Sabem que ha de ser $q>n$ i volem que $k/n\geq r$ i $d\geq 2t+1$ (d la distància mínima). Com que $d=n-k+1$, la darrera condició equival a dir que

$$2t+1\leq n-k+1,$$

i com que $k\geq nr$, serà $2t+1\leq n-nr+1=n(1-r)$, és a dir, $n\geq 2t/(1-r)$. Així doncs, la resposta és que ha de ser

$$q\geq 1+\frac{2t}{1-r}.$$

Exemple. Si volem un codi RS amb taxa de transmissió almenys $3/4$ i que corregeixi 10 errors, s'haurà de complir $q\geq 81$. Si prenem $q=81$, el codi RS(F,60), $F=\mathbb{F}_{81}$, satisfà les condicions, ja que $n=80$, $k/n=60/80=3/4$ i $d=n-k+1=80-60+1=21=2\cdot 10+1$. També podem usar qualsevol cos finit F amb $q>81$ i el codi $C=\mathrm{RS}(\alpha_1,...,\alpha_n; 60)$ format amb $n=80$ elements no nuls distints $\alpha_1,...,\alpha_n\in F$. Això pot tenir avantatges computacionals, com ara amb $\mathbb{Z}_{83}$ en lloc de $\mathbb{F}_{81}$, ja que les operacions de $\mathbb{F}_{81}$ són una mica més costoses que les de $\mathbb{Z}_{83}$.

Exemple. El codi RS(F, 233), on $F=\mathbb{F}_{256}$, té tipus [255,233,23] i s'usa en sistemes de transmissió interplanetària, com ara el del Voyager (1986) i Galileu (1989). En aquest cas els símbols de F s'interpreten, respecte d'una base apropiada de F sobre $B=\mathbb{Z}_2$, com a vectors de 8 bits. La capacitat correctora d'aquest codi és d'11 símbols.

Exemple. En el disc compacte per a música (CD d'àudio) s'empren dos $\mathbb{F}_{256}$-codis C_1 i C_2 de la forma $\mathrm{RS}(\alpha_1,...,\alpha_n; k)$: C_1 amb $n=28$ i $k=24$ i C_2 amb $n=32$ i $k=28$. Així, doncs, $C_1\sim[28,24,5]$ i $C_2\sim[32,28,5]$, de manera que cadascun d'aquests codis corregeix 2 errors. En essència, la informació, segmentada en blocs de 24 símbols de 8 bits, es codifica primer mitjançant C_1 i després mitjançant C_2. El resultat de la primera codificació està segmentat en blocs de 28 símbols de 8 bits, just el que es necessita per a l'entrada del codificador de C_2. La sortida del codificador de C_2 està segmentada en blocs de 32 símbols de 8 bits. Naturalment, en el procés de descodificació se segueix el camí invers: primer es descodifica C_2 i després C_1. Aquest esquema és un cas particular de la *concatenació en sèrie* de dos codis, la qual és usada en una o altra forma en molts sistemes de comunicació.

Matriu de control

Considerem la matriu $H=\tilde{V}_r(\alpha_1,...,\alpha_n)$, on $r=n-k$,

$$\tilde{V}_r(\alpha_1,...,\alpha_n)=\begin{pmatrix} h_1 & h_2 & \cdots & h_n \\ h_1\alpha_1 & h_2\alpha_2 & \cdots & h_n\alpha_n \\ \vdots & \vdots & & \vdots \\ h_1\alpha_1^{r-1} & h_2\alpha_2^{r-1} & \cdots & h_n\alpha_n^{r-1} \end{pmatrix}$$

$$h_i=(-1)^{i-1}D(\alpha_1,...,\alpha_{i-1},\alpha_{i+1},...,\alpha_n)\,/\,D(\alpha_1,...,\alpha_n)=\frac{1}{\prod_{j\neq 1}(\alpha_j-\alpha_i)}$$

A partir d'aquesta definició, no és difícil veure que $GH^T=0$ i d'això resulta que $C=\{x\in F_n \mid xH^T=0\}$ (vegeu CC, Secció 1.2). Direm que H és una *matriu de control* de C.

Generalitzacions

La matriu de control H té la forma

$$\tilde{V}_r(\alpha_1,...,\alpha_n)=V_r(\alpha_1,...,\alpha_n)\cdot \text{diagonal}(h_1,...,h_n),$$

amb les h_i donades en funció de $\alpha_1,...,\alpha_n$ per la fórmula del paràgraf anterior. Resulta, però, que és convenient considerar matrius amb aquesta forma, però amb les $h_i\neq 0$ escollides de manera independent de $\alpha_1,...,\alpha_n$. Les matrius $H=A_r(h_1,...,h_n;\alpha_1,...,\alpha_n)$ que s'obtenen d'aquesta manera s'anomenen *matrius alternants* i els codis

$$A_F(\boldsymbol{h},\boldsymbol{\alpha},r)=\{x\in F^n \mid xH^T=0\}$$

s'anomenen *codis alternants*. Notem que aquesta definició té sentit fins i tot quan els components $\boldsymbol{h}$ i $\boldsymbol{\alpha}$ són d'un cos F' que conté F. Amb aquesta generalitat, però, només es poden establir cotes per a la distància mínima i la dimensió:

Teorema. *Siguin k i d la dimensió i la distància mínima de $A_F(\boldsymbol{h},\boldsymbol{\alpha},r)$. Aleshores es compleixen les desigualtats següents:*

$$n-r\geq k\geq n-mr,\ d\geq r+1.$$

Prova: Vegeu, per exemple, CC, secció 4.1. □

La construcció de codis alternants ens dóna accés, com a casos particulars, a la construcció de dues famílies importants de codis: els anomenats codis BCH (de Bose-Chaudhuri-Hocquenghem) i els de Goppa clàssics.

Codis BCH. Suposem que tenim un cos F' que conté F, i sigui $\alpha \in F'$. Sigui n l'ordre de α, és a dir, el primer enter positiu tal que $\alpha^n=1$. Escollim dos nombres enters δ i l i formem vectors $\boldsymbol{\alpha}$ i $\boldsymbol{h}$ amb els termes de la progressions geomètriques de n termes de raons α i α^l, respectivament. Llavors el codi $A_F = (\boldsymbol{h}, \boldsymbol{\alpha}, \delta-1)$ resulta ser el codi BCH sobre F d'arrel α, *distància de disseny* δ i *desplaçament* l (en símbols, $BCH_F(\alpha,\delta,l)$). Els codis amb $l=1$ es diu que són codis BCH *en sentit estricte* i es posa $BCH_F(\alpha,\delta)$ en lloc de $BCH_F(\alpha,\delta,1)$. Com a conseqüència del teorema, tenim que $d \geq \delta$ (aquesta desigualtat s'anomena *cota BCH*) i $n-(\delta-1) \geq k \geq n-m(\delta-1)$. En el cas binari ($F=B$), la darrera desigualtat es pot millorar: si $\delta=2t+1$, es compleix $k \geq n-mt$.

Codis de Goppa clàssics. Suposem que tenim un cos F' que conté F, un polinomi $h \in F'[T]$ i elements distints $\alpha_1,\ldots,\alpha_n \in F'$ tals que $h(\alpha_i) \neq 0$ per a qualsevol i. Aleshores el codi de Goppa associat a h i $\boldsymbol{\alpha} = (\alpha_1,\ldots,\alpha_n)$, $\Gamma(\boldsymbol{h}, \boldsymbol{\alpha})$, es pot definir com el codi $A_F(\boldsymbol{h}, \boldsymbol{\alpha}, r)$, on r és el grau de h i $\boldsymbol{h} = (h_1(\alpha_1)^{-1},\ldots,h_n(\alpha_n)^{-1})$. Com a conseqüència del teorema, tenim que $d \geq r+1$ i $n-r \geq k \geq n-mr$. En el cas binari hi ha una millora de la primera desigualtat: si h no té arrels múltiples, aleshores $d \geq 2r+1$.

Llistat 10. Aspectes computacionals

```
matriu_alternant(h:Vector, a:Vector, r:Enter) := vandermonde(a,r)·matriu_diagonal(h)
matriu_alternant(a:Vector, r:Enter) := matriu_alternant(a,a,r)
codi_alternant(h:Vector, a:Vector, r:Enter)
  := {h_=h, a_=a, r_=r, b_=invertir_components(a), H_=matriu_alternant(h,a,r) }
codi_alternant(h:Vector, a:Vector, r:Enter, K:Cos) := codi_alternant(h,a,r) |{K_=K}
codi_alternant(a:Vector, r:Enter):= alternant_code(a,a,r)

RS(a:Vector, k:Enter) :=
inici
  local
    h, n=longitud(a)

  h=[∏_{j≠1} (a_j − a_i) amb i en 1..n]

  h=invertir_components(h)
  codi_alternant(h,a,n-k)|{G_=vandermonde(a,k)}|{K_= cos(a_1)}
fi
```

```
RS(F:Cos, k:Enter) :=
  RS([element(j,F) amb j en 1..(cardinal(F)-1)], k)

BCH(a:Element(Cos), d:Enter, l:Enter)
:=
inici
  local n=ordre(a)

  codi_alternant(sèrie_geomètrica(a^l,n), sèrie_geomètrica(a,n),d-1)
fi
BCH(a:Element(Cos), d:Enter, l:Enter, K:Cos) :=
  BCH(a,d,l)|{K_=K}
BCH(a:Element(Cos), d:Enter) := BCH(a,d,1)
BCH(a:Element(Cos), d:Enter, K:Cos) := BCH(a,d,1,K)

goppa(g:Polinomi, a:Vector, K:Cos) :=
  goppa(g,a)|{K_=K}
goppa(g:Polinomi, a:Vector) :=
codi_alternant(invertir_components(eval(g,a)),a, grau(g))

C=RS(Zn(11),4)
→
```

$$\{H\_=\begin{pmatrix} 1 & 2 & 4 & 8 & 5 & 10 & 9 & 7 & 3 & 6 \\ 1 & 4 & 5 & 9 & 3 & 1 & 4 & 5 & 9 & 3 \\ 1 & 8 & 9 & 6 & 4 & 10 & 3 & 2 & 5 & 7 \\ 1 & 5 & 3 & 4 & 9 & 1 & 5 & 3 & 4 & 9 \end{pmatrix}, \text{K\_=Z11},$$

```
 a_=[1,2,4,8,5,10,9,7,3,6], b_=[1,6,3,7,9,10,5,8,4,2], h_=[1,2,4,8,5,10,9,7,3,6], r_=4}
```

Algorisme de Berlekamp-Massey-Sugiyama

Sigui C=RS($\alpha_1,...,\alpha_n$; k). Siguin G i H les seves matrius generadora i de control, respectivament, introduïdes a la subsecció anterior. Suposem que $x \in C$ és el *vector enviat* per un canal simètric d'alfabet F, i sigui y el *vector rebut*. Llavors el vector $e=y-x$ és el *vector d'error*. El problema de la descodificació és reconstruir x a partir de y, suposant que el nombre s d'errors (és a dir, el pes de e) no supera la capacitat correctora t de C. L'algoritme de Berlekamp-Massey-Sugiyama (BMS) resol aquest problema d'una manera eficient i es pot descriure de la manera següent:

1. Calculem la *síndrome* de y, que per definició és el vector $S=yH^T \in F^r$.
2. Si S=0, retornem y (com que H és una matriu de control de C, $S=0$ equival a $y \in C$, i en aquest cas, sota la hipòtesi $s \leq t$, ha de ser $x=y$, ja que altrament

$e=y-x\in C$ i el nombre d'errors seria $\geq d\geq 2t+1$). Podem, doncs, suposar que $S\neq 0$. Posarem

$$S=(S_0,S_1,...,S_{r-1}) \text{ i } S(z)=S_0+S_1z+...+S_{r-1}z^{r-1}.$$

3. Sigui $\{\sigma,\omega\}=\text{sugiyama}(z^r, S(z))$.
4. Calculem la llista J dels índexs $j\in\{1,...,n\}$ que compleixen $\sigma(\alpha_j^{-1})=0$, i posem s per denotar el cardinal de J. Si $s<\deg(\sigma)$, retornar ? (el símbol d'error).
5. Per $j\in J$, calculem

$$e_j=\frac{\alpha_j\cdot\omega(\alpha_j^{-1})}{h_j\cdot\sigma'(\alpha_j^{-1})}$$

i formem el vector x' que resulta de canviar y_j per y_j+e_j en el vector y.

6. Si $x'H^T=0$, retornem x'. Altrament retornem ?

Teorema. *L'algorisme* BMS *corregeix t errors.*

Prova: Vegeu la secció 4.3 de CC. □

Remarca. L'algorisme BMS és també vàlid per a codis alternants, llevat que el control d'errors és una mica més delicat (en diversos punts cal comprovar que el valor de certes expressions, que en principi estan definides a F', és de F). En particular, doncs, l'algorisme BMS també és aplicable als codis BCH i de Goppa clàssics.

Llistat 11. Descodificació

```
BMS(y:Vector, C:Codi) :=
inici
local H, h, r, a, b, K,
s,      # síndrome
z,      # variable del polinomi síndrome
S,      # polinomi síndrome
l,      # polinomi localitzador d'errors
e,      # polinomi avaluador d'errors
M,      # llista de posicions d'error
m,      # un element de M
E,      # llista d'errors corresponent a M
x,      # un element de E
y0      # per retenir y

H=H_(C); r=r_(C); h=h_(C); a=a_(C); b=b_(C); K=K_(C)

s=y·H^T
si zero?(s) aleshores
```

```
  dir("BMS >Vector del codi")
  retornar(y)
fi
S=polinomi(s,z)
{l,e}=sugiyama(z^r, S, terra(r/2))
M=positions_arrels(l,b)
si longitud(M)<grau(l) aleshores
  dir("BMS >Vector no descodificable: el nombre d'arrels és incorrecte")
  retornar(0)
fi

E=null
y0=y
per m en M fer
  x = a_m·aval (e,b_m) / h_m·aval (l',b_m)

  E=(E,-x)
  y_m = y_m + x
fi
E={E}
si zero?(y·H^T) aleshores
    dir("BMS >Descodificació correcta, amb taula d'errors"|{M,E})
    retornar(y)
  altrament
    dir("BMS >Error de descodificació: x´ no té síndrome nul·la")
    retornar(0)
  fi
fi
```

Exemple RS[26,16,11]: corregeix fins a 5 errors. La funció prova_descodificador(C,s) genera un element aleatori del codi, x, un error aleatori de pes s, e, i aplica el descodificador BMS a $y=x+e$. Si no hi ha error del descodificador, retorna una matriu amb files x, e, y. Si hi ha un error del descodificador, només retorna una matriu amb files x, e. Finalment, si es produeix un error indetectable, llavors dóna una matriu amb files x, e, y, x'. En aquest exemple, els elements de les matrius no es donen complets, sinó només el seu logaritme discret en base t (aquests logaritmes s'han calculat prèviament a la taula E; el logaritme discret de 0 es representa amb el caràcter "_").

```
K=Z3[t]/(t^3-t+1); E = taula_exponents(t);
C=RS(K,10) ;
```

```
abreujar(prova_descodificador(C,3),E)
12 25 3 23 20 15 9 19 8 5 2 17 2 13 6 2 9 10 0 15 4 13 7 11 2 22
_ _ _ _ _ _ _ _ _ 14_ _ _ _ 7 _ _ _ _ _ 6 _ _ _ _ _
12 25 3 23 20 15 9 19 8 8 2 17 2 13 15 2 9 10 0 15 25 13 7 11 2 22

abreujar(prova_descodificador(C,5),E)
19 1 15 8 7 2 2 25 12 5 _ 15 _ 10 6 4 6 1 _ 9 6 13 20 22 19 23
_ 21_ _ _ _ 14_ _ 15 _ _ 4 23 _ _ _ _ _ _ _ _ _ _ _ _ _
19 6 15 8 7 2 4 25 12 11 _ 15 4 _ 6 4 6 1 _ 9 6 13 20 22 19 23

abreujar(prova_descodificador(C,6),E)
4 18 8 25 23 8 10 12 0 24 9 11 19 4 11 5 8 1 25 8 24 10 7 22 13 18
_ 23 _ _ _ _ 9 9 _ 0 _ 20 _ _ _ _ 3 _ _ _ _ _ _ _ _ _ _
```

En la primera i segona prova s'han simulat 3 i 5 errors, respectivament, i el descodificador els ha sabut corregir. En el tercer exemple s'han simulat 6 errors, i s'ha produït un error del descodificador (notem que el nombre d'errors és superior a la capacitat correctora).

Mathematical modelling in sport

J. R. Stirling
Centre Recerca Matematica

M. S. Zakynthinaki
Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya

ABSTRACT

We give examples of how mathematical modelling is used to model processes in sport. We start by giving a general overview of areas currently employing mathematical tools to analyze the processes involved. We then go onto look in more detail at the use of mathematics to model the physiological response to exercise. It should also be noted that this area is not only fundamental to the science of technique and training methodology in sport but are also fundamental areas of medicine. As a result improvements in the mathematical analysis of such problems will have far reaching benefits.

1. INTRODUCTION

Mathematics has many uses in sport from the statistical analysis of games of football to the design of sports equipment. Here we will concentrate mainly on the use of mathematics to model and analyze the physiological response of the body to exercise. We will start however with a short review of the use of mathematics in biomechanics.

2. MATHEMATICAL MODELLING AND ANALYSIS IN BIOMECHANICS

Biomechanics is the science that examines forces acting upon and within a biological structure and the effects produced by such forces. A introduction to the biomechanical response to exercise can be found in Nigg and Herzog [47], see also Cavanagh [20] for a introduction to the biomechanics of distance running. Such types of analysis are used in many areas of sport to analyze technique so as to improve performance and also to investigate the causes of injuries and look for methods for how to prevent them. Another very interesting area of biomechanics involves looking at the motion of and forces applied to the foot during walking and running, see for example De Clerque et al

[25], De Wit et al [24], Robbins et al [53, 54, 55] and Hennig et al [30]. Such studies investigate the biomechanical functions of the foot and can be used to identify unnatural motions likely to cause injuries in an athlete. Many of these studies also compare the motion of the foot when running barefoot and shod and hence can be used to help improve the design of running shoes and see wether the shoe is preventing or causing injuries.

Other areas of biomechanics involve the analysis of the time series of forces exerted on or by the body such as ground reaction forces. Stirling and Zakynthinakis [72] used tools from dynamical systems to model the ground reaction forces at the feet of a person standing on a force platform and hence show how an athlete regains stability following being pushed out of balance. Zakynthinaki and Stirling [72], and Gajewski [28] used Fourier analysis to analyze tremor during muscular failure. In this work Fourier analysis was used to analyze the frequency spectrum of the time series of the forces produced when carrying an exercise out until muscular failure. This work has practical uses in the prediction of imminent failure and hence maximal training loads. It can also be used to detect premature failure when testing for maximum load an athlete can cope with.

Animation tools are also of much use in biomechanics especially for the visualization and correction of technique. For example such tools can be used to look for and hence correct nonalignment of relevant parts of the body to the direction of motion. Improvements in this area can lead to a more efficient technique and hence a better performance and also a reduction in the number of injuries an athlete suffers. One such tool is that developed by Ng-Thow-Hing, V., Faloutsos [46].

3. MODELS OF THE PHYSIOLOGICAL RESPONSE TO EXERCISE

We present a review of some of the models currently used to describe the physiological response to exercise (see Billat et al [11] for a review of some other models not presented here and also Stirling et al [60] for models using dynamical systems). We shall focus on the modelling of the heart rate $HR(t)$, volume of oxygen uptake (breathed in) per unit time $\dot{V}O_2(t)$, the lactate concentration levels in the blood, $BLa(t)$ and velocity, v as these are the most commonly used variables to define exercise intensity and fitness levels. For more details on exercise physiology see Astrand et al [1], Martin and Coe [41], Wasserman et al [65] and Wilmore and Costill [70].

3.1 Heart rate kinematics

According to Martin and Coe [41] "If there is one single physiological variable that identifies the total stress load under which an athlete is performing, it is the heart

rate". Heart rate monitors are now commonly used by athletes to monitor the training intensity (see Burke [19], Martin and Coe [41], Daniels [22] and Mille-Hamard et al [43]), the bodies response to training sessions with increased fitness levels and the over all health status of the athlete (ie. overtraining and excessive levels of tiredness) found for example via the resting heart rate levels. These monitors are capable of recording beat to beat intervals for the heart and allow one to down load this information for analysis as a time series to a computer, see figures 1 and 2.

From figure 1 we can see that assuming the distance covered in all of the 3 minute repetitions was the same and other conditions (see 3.1.3) were equal then there are two obvious conclusions. Firstly that as the training session progresses then it becomes more demanding for the body, this is reflected in the increase in both the peak heart rate attained during a repetition and the minimum heart rate achieved during the 3 minute recovery. Secondly as the second time series has a lower heart rate through out the whole training session than the first then the body was more able to cope with the exercise and hence the athlete is getting fitter. We can also see when we compare figures 1 and 2 that the timeseries for the heart rate reflects the particular training session, hence the differences in the two figures.

3.1.1 On the relationship between HR and $\dot{V}O_2$

Functionally we have two hearts, with the right heart delivering blood to the lungs and the left everywhere else. The two primary operating variables of each heart are the

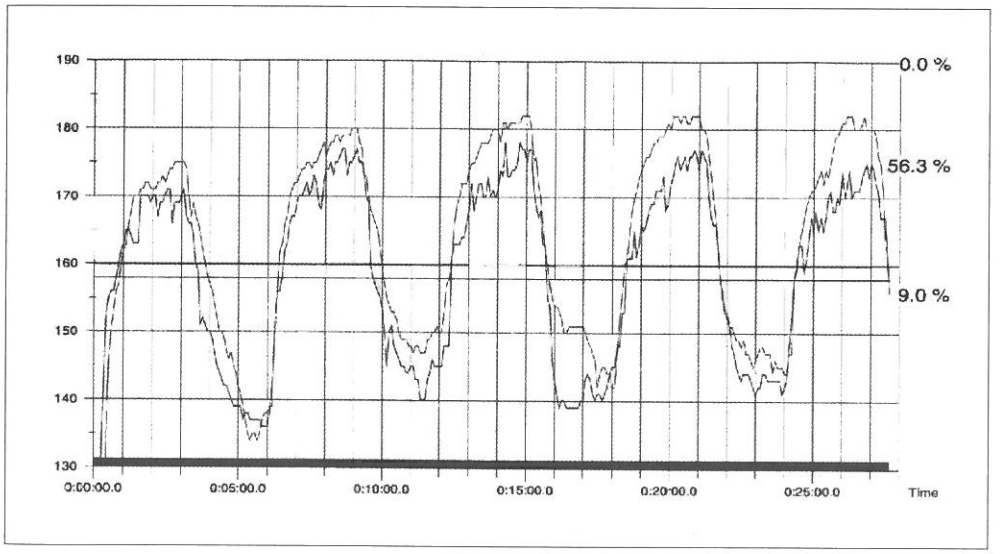

Figure 1. Heart rate data of the author, (5 second average) from two separate training session of 5 repetitions of 3 minutes at $\upsilon\dot{V}O_{2max}$ with 3 minutes recovery at 50% $\upsilon\dot{V}O_{2max}$ between repetitions.

heart rate, HR and the volume pumped per beat, the stroke volume, Q_s. The cardiac output is the product of heart rate and stroke volume

$$CO = HR \cdot Q_s \qquad \text{(ml or L of blood)} \tag{1}$$

The cardiac output must equal venus return as the heart can only deliver to the arterial side of the body blood which returns to it from the venus side. At rest venus return is controlled by 4 principle factors: the tone or caliber of the venus vessels, the position of the body in space, the total blood volume and the depth of breathing. Upon exercising the milking action of the skeletal muscles help push blood through the veins towards the heart.

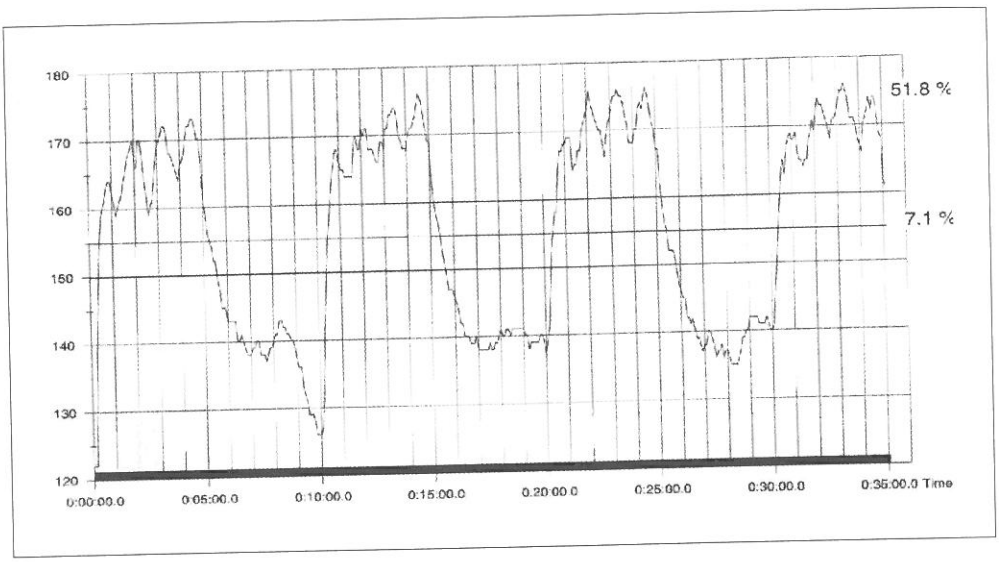

Figure 2. Heart rate data of the author, (5 second average) from a training session of 4 sets of 4 repetitions of 45s at maximum speed with 30 seconds recovery jogging between repetitions and 5 minutes recovery jogging between sets.

When blood is sampled from the coronary arteries and from the coronary venus sinus and analyzed for its O_2 content, we find a very large amount of O_2 has been removed, (termed the arteriovenous O_2, difference $a\bar{V}O_2^d$, ie. the difference between the mixed venous blood O_2 concentration and the arterial blood O_2 concentration).

It is found that the maximum heart rate HR_{max} is either unchanged or slightly reduced by endurance training, and the maximum value of the arteriovenous O_2 difference $a\bar{V}O_2^d$ peaks at about $16ml/dl$. It is observed that in the trained heart as a result of cardiac chamber enlargement, the stroke volume at any given work load is larger and the

heart rate is lower, hence increasing the available perfusion time, resulting in an increase in the bodies aerobic power.

$\dot{V}O_2$ is equal to the product of cardiac output CO and the O_2 extracted from the blood, $a\bar{V}O_2^d$.

$$\dot{V}O_2 = HR \ . \ Q_s \ . \ a\bar{V}O_2^d \tag{2}$$

This relationship between the HR and $\dot{V}O_2$ is relatively simple, though not trivial as neither Q_s nor $a\bar{V}O_2^d$ are constant. Both Q_s and $a\bar{V}O_2^d$ are functions of exercise intensity, fitness and time. It is believed that the cardiac output, CO and heart rate HR normally increase linearly with $\dot{V}O_2$ during increasing work rate exercise. Notice this linear relationship can only be true if we reach HR_{max} at the same time as we reach $\dot{V}O_{2max}$, which isn't necessarily the case. Swain et al [62, 63] found the $\%HR^{res}$ is approximately equal to the $\%\dot{V}O_2^{res}$ on average for a large population of runners, (where the reserve for the heart rate $\%HR^{res} = \frac{HR(t) - HR_{min}}{HR_{max} - HR_{min}} \cdot 100\%$ and $\%\dot{V}O_2^{res} = \frac{\dot{V}O_2(t) - \dot{V}O_{2min}}{\dot{V}O_{2max} - \dot{V}O_{2min}} \cdot 100\%$, with HR_{min} and HR_{max} being the minimum and maximum heart rates respectively and $\dot{V}O_{2min}$ $\dot{V}O_{2max}$ being the minimum and maximum $\dot{V}O_2$ respectively). This relationship however needs to be understood on an individual basis as the deviation from the $\%HR^{res}$ vs $\%\dot{V}O_2^{res}$ curve is large.

3.1.2 A model for HR(t)

It has been found that the heart rate HR, like the $\dot{V}O_2$, is an exponential like function of both the speed (or intensity) V, and the time t. The relationship is such that the HR at which we plateau out at (i.e. $\dot{HR} = 0$) given sufficient time is determined by the intensity of the exercise and an additional slow component and drift. In Engelen et al [26] (see also Linnarsson [39]) they model both the on and off kinetics for the $\dot{V}O_2$ and HR response to exercise. Here we look at the on/off response for the HR. The "on" response corresponds to how the heart rate reacts to an increase in the intensity of the exercise to a new higher level intensity. HR_{on} is modelled as

$$\begin{aligned} HR_{on} &= HR_b + A_0\left(1 - e^{-\frac{t}{\tau_0}}\right) \quad phase1 \\ &+ A_1\left(1 - e^{-\frac{(t - TD_1)}{\tau_1}}\right) \quad phase2 \\ &+ A_2\left(1 - e^{-\frac{(t - TD_2)}{\tau_2}}\right) \quad slow\ component \end{aligned} \tag{3}$$

The off kinetics correspond to how the heart rate reacts to a decrease in the intensity of the exercise to a new lower level of intensity. HR_{off} is modelled as

$$HR_{off} = HR_{End_{ex}} - 0 \quad phase1$$
$$- A_1(1 - e^{-\frac{(t-TD)}{\tau_1}}) \quad phase2$$
$$- A_2(1 - e^{-\frac{(t-TD)}{\tau_2}}). \quad slow\ component \tag{4}$$

The parameters are very similar to those used used to model the on and off $\dot{V}O_2$ kinetics (see section 3.2.5 for a description). The magnitude of the parameters reflects the intensity of the exercise. Note there is no rapidly decreasing phase 1 term unlike in the off kinematics for the $\dot{V}O_2$. However like in the off kinematics for the $\dot{V}O_2$ the time delay is common, for the same reasons as described there (see section 3.2.5).

3.1.3 Other factors effecting the heart rate

Heart rate can be effected by many other variables, other than the intensity of the exercise. Some of these variables can be controlled and some cannot hence to understand the heart rate dynamics one needs to understand how the following variables may effect the time series.

Temperature - High temperature and humidity causes an elevation in the heart rate. Heart rate shows the lowest value at an outside temperature of about 20 degrees centigrade. High surrounding temperate and air humidity place greater demands on the body during physical exercise. For equal work loads an increase in the surrounding temperature and or humidity will result in an increased heart rate. This is because the demands on the bodies heat regulation system are increased, requiring increased production of perspiration, and an increased blood flow in the capillaries of the skin. To do this we require an increase in the fluid loss. This decreases the circulating blood volume and diminishes the blood supply to the heart, which the body will compensate for by increasing the heart rate. This shows why during extended exercise efforts it is essential to intake fluids and cool the body sufficiently so as to prevent dehydration and hence enable peak performances. At equal sub maximal efforts it is often assumed that for each degree centigrade rise in temperature we have an approximate 10 to 15 beats per minute increase in heart rate. Obviously at maximal efforts when the heart rate is at HR_{max} we cannot increase the heart rate. It should also be noted that in hot and humid conditions the time an athlete can remain exercising at a particular intensity will also decrease.

Age - with age the HR_{max} gradually decreases.

Over-training, insufficient recovery and illness - Depending on the type of over-training the morning pulse may be higher or extremely low. As a result of over-training

a lag in the increase of the heart rate *HR* with high intensity exercise may be observed. Another result of over-training is the fact that the maximum heart rate HR_{max} may no longer be obtainable. Over-training therefore results in a completely different heart rate response both during exercise and at rest.

Nutrition and hydration- Adequate nutrition and hydration can improve performance in endurance events resulting in a lower *HR* at a given intensity.

Altitude - HR_{min} decreases upon the first few hours after an athlete arrives at high altitude. This is however followed by an increase, which is approximately 10% at 2000m and 45% at 4500m when compared with sea level values for HR_{min}.

Depending on the altitude the HR_{min} will return to sea level values or lower after some days acclimatization. In fact this is used as a good sign of acclimatization.

Medication - Various medications can effect the heart rate (ie. beta blockers have been shown to decrease both the HR_{min} and the HR_{max}).

Infectious disease - A fever can lead to an increase in the HR_{min} of 10 to 15 beats for every degree temperature rise. Also during sub-maximal efforts the *HR* shows a different pattern. HR_{min} is also increased during recovery from infectious deceases.

Mental activity - Mental stress can effect *HR*.

3.2 $\dot{V}O_2$ kinematics

$\dot{V}O_2$ can be defined as the volume of oxygen we breath in (uptake) per unit time. $\dot{V}O_{2max}$ is the maximal limit of the ability to increase $\dot{V}O_2$, further increase in intensity do not yield a larger $\dot{V}O_2$. This is referred to as the aerobic capacity, maximal oxygen uptake or $\dot{V}O_{2max}$. Below the lactate threshold, LT (see section see section 3.3.1 for a definition and discussion of the lactate threshold) the rate of increase in $\dot{V}O_2$ uptake and demand is an approximately linear function of exercise intensity, whilst above the lactate threshold the function is nonlinear.

The $\dot{V}O(t)$ can be measured as a time series using breath by breath analysis equipment. Such recording are often taken in a lab whilst the athlete runs on a treadmill at various speeds however portable devices are also now available. It is also common to measure the volume of other gases other than the O_2 in each intake our outtake of breath. In this way it is possible to measure the chemical content of the gases we inspire and expire.

Figure 3 shows the on transient kinetics of $\dot{V}O(t)$ for four different intensities. It can be seen that moderate intensity shows an exponential like rise and plateau. The same can be seen in the response to a heavy exercise intensity, though this time there is a delay in reaching the plateau or steady state $\dot{V}O_2$. For very heavy intensity exercise the plateau is the $\dot{V}O_{2max}$, there is also a delay in reaching this plateau. For the severe intensity exercise the $\dot{V}O(t)$ rises very steeply until it is limited by the $\dot{V}O_{2max}$, which then becomes the plateau for remaining time the exercise can be carried out at.

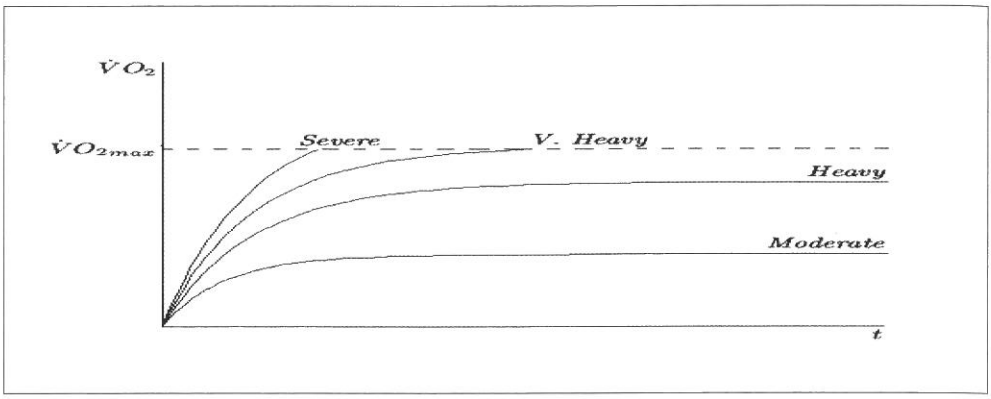

Figure 3. A Sketch of the on transient kinetics of $\dot{V}O_2(t)$ time series response to four different exercise intensities. Where moderate intensity exercise is below the *LT*, heavy is below the critical velocity v_{crit} (see section 3.4), very heavy is such that the oxygen demand is slightly greater than or equal to the $\dot{V}O_{2max}$ and severe is such that the oxygen demand is far greater than the $\dot{V}O_{2max}$.

Note if the exercise is to severe then the athlete may have to stop due to fatigue before they reach the $\dot{V}O_{2max}$. Figure 4 shows the off transient kinetics for the $\dot{V}O_2(t)$. In other words it shows how the $\dot{V}O_2(t)$ reduces as the body recovers to a new lower oxygen demand following exercise of different intensities.

3.2.1 Anaerobic capacity or oxygen deficit

The oxygen deficit is defined as the integral with respect to time of the difference between the oxygen demand for that particular exercise intensity and the oxygen uptake during the whole exercise bout (for a review see Saltin [56, 57], and Bangsbo [3, 4, 5]). It is used to quantify the anaerobic energy contribution (see section 3.3 for definition of the anaerobic energy contribution) to the work performed (see Saltin [56, 57], Bangsbo [3, 4, 5] and Medbo [42]). In short term non intense exercise below the lactate threshold, there is no problem, the problem arises during intense exhaustive exercise. If the work exhausts the subject in a short duration it is likely that the maximum oxygen deficit is reached. Energy costs or oxygen demands must be accurately known to calculate the oxygen deficit. This is not difficult at sub maximal work loads below the lactate threshold where the steady state oxygen uptake (ie the plateau, see in the $\dot{V}O_2$ vs t curve) represents the oxygen demand. At exhaustive exercise, however the nature of the true cost is more complicated. The problem being that we cannot extrapolate from the sub maximal linear relationship as this would underestimate the oxygen deficit. We also cannot assume that the mechanical efficiency the subject remains constant with increasing intensity.

It can be shown that the higher the intensity of the exercise the lower the mechanical efficiency of the athlete (see Saltin [56, 57], and Bangsbo [3]). It can be shown that

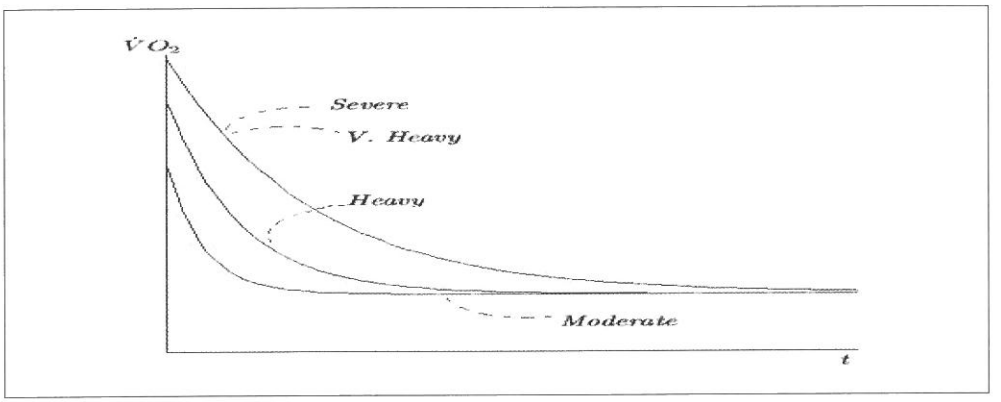

Figure 4. A Sketch of the off transient kinetics of $\dot{V}O_2(t)$ time series response to recovery (ie. to a lesser oxygen demand) from four different exercise intensities.

exercise involving more muscle mass results in a higher oxygen deficit. This can be seen in the fact that whole body exercise such as swimming or rowing have a maximal oxygen deficit more than 50% larger.

3.2.2 Three phase for pulmonary $\dot{V}O_2$

Hill and Lupton [31], recognized that the pulmonary $\dot{V}O_2$ rises approximately exponentially following the onset of moderate exercise. With more modern tools (breath by breath analysis) and "excessive" averaging, three phases can be observed in the rise of $\dot{V}O_2$ (see Whipp [68]) . Phase 1 occurs in the first 15 to 25 seconds. The rise in $\dot{V}O_2$ is believed to be due primary to increased cardiac output and secondarily to changes in mixed venous O_2 content and lung gas stores. Phase 2 follows this, and in this phase there is an exponential rise to a steady state level of $\dot{V}O_2$ for moderate exercise below the lactate threshold, *LT*. Phase 3 represents the new steady state level of $\dot{V}O_2$, when it is achieved (ie. below *LT*). For exercise intensities above that corresponding to the lactate threshold however the $\dot{V}O_2$ kinetics is more complicated than the mono exponential model described above for moderate (below *LT*) intensities of exercise. Instead we have an additional slow component.

Hill and Stevens [33] conclude that "at the onset of exercise in the severe intensity domain, $\dot{V}O_2$ is initially driven toward the O_2 demand, and then is limited by the achievable $\dot{V}O_2\dot{V}O_{2peak}$" Where $\dot{V}O_{2peak} = \dot{V}O_{2max}$ unless we have premature fatigue (see also the curve labelled severe in figure 3).

3.2.3 $\dot{V}O_2$ *slow component*

Below the *LT*, $\dot{V}O_2$ reaches a steady state after 3 minutes of constant work rate exercise (Davies, di Pampero, and Cerretelli [23], Linnarsson [39], Whipp and

Wassermann [67] because the rate of ATP (see section 3.3 for more on ATP) utilization and aerobic regeneration of ATP becomes equal. The $\dot{V}O_2$ response at higher rates of work does not follow the steady state response (ie. a plateau). Instead at power outputs above the lactate threshold the oxygen consumption continues to increase either to the $\dot{V}O_{2max}$ or to a delayed steady state $\dot{V}O_2$, i.e. $\dot{V}O_{2ss}$. This slow drift upwards in $\dot{V}O_2$ during constant work rate exercise has been called the slow component of oxygen uptake kinetics. It has been speculated that this slow component is probably associated with ventilation, shifting of metabolic substrate from fat to carbohydrate, and or and increase in body temperature. While others believe the most likely mechanism however is an alteration in muscle fiber recruitment patterns, with the recruitment of more fast muscle fibers which are less efficient (see Gaesser and Poole [27] for a review).

3.2.4 $\dot{V}O_2$ drift

This is a similar but likely unrelated phenomena to that of the $\dot{V}O_2$ slow component. The phenomena is defined as a slow increase in $\dot{V}O_2$ during prolonged, sub maximal, constant power output exercise. Unlike the slow component $\dot{V}O_2$ drift is observed at power outputs well below lactate threshold and the magnitude of the increase in $\dot{V}O_2$ is much less. It is believed, to be due to an increase in ventilation, an increase in levels of circulating catecholamines and thermo regulatory constraints, see Wilmore and Costill [70].

3.2.5 Characterization of $\dot{V}O_2$ kinematics as a function of exercise intensity, Barstow [6]

During moderate exercise $\dot{V}O_2$ reaches a new steady state within 180 seconds in normal subjects, with little or non rise in blood lactate. The steady state $\dot{V}O_2$ increases linearly with work-rate. The time constant in phase 2 (after the first 15-20 seconds) is constant across work intensities and appears to reflect muscle oxygen utilization kinetics. However when heavier exercise is performed, which elevates blood lactate through the exercise, the $\dot{V}O_2$ response becomes more complex. The predominant phase 2 response continues to rise exponentially with about the same time constant as for moderate exercise, and the amplitude continues to be linearly related to work rate. However an additional slowly developing rise in $\dot{V}O_2$ is also usually observed, beginning 100-200 seconds into exercise. This additional $\dot{V}O_2$ delays attainment of a steady state, increasing the overall O_2 "cost" of the exercise and is statistically associated with the rate and magnitude of increase in blood lactate. Interestingly in children (and in some elite athletes see Billat [14]), neither the slow component nor the blood lactate rise as much during heavy exercise.

$\dot{V}O_2$ has been modelled in Barstow [6] using the following triple exponential form (see also figure 5).

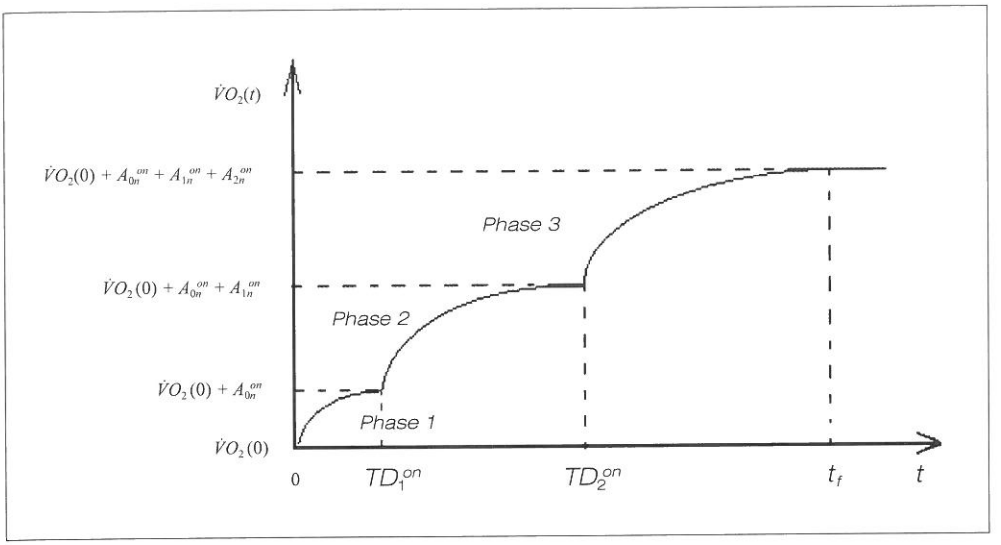

Figure 5. A sketch showing the 3 phases used to model oxygen uptake kinetics.

$$\dot{V}O_2 = \dot{V}O_2(0) + A_0^{on}\left(1 - e^{-\frac{(t)}{\tau_0^{on}}}\right) \quad phase\ 1$$
$$+ A_1^{on}\left(1 - e^{-\frac{(t - TD_1^{on})}{\tau_1^{on}}}\right) \quad phase\ 2$$
$$+ A_2^{on}\left(1 - e^{-\frac{(t - TD_2^{on})}{\tau_2^{on}}}\right) \quad slow\ component \tag{5}$$

where $\dot{V}O_2(0)$ is the initial value of the $\dot{V}O_2$ at time $t = 0$. A_0^{on}, A_1^{on} and A_2^{on} are the asymptotic values, τ_0^{on}, τ_1^{on} and τ_2^{on} are the time constants, and TD_1^{on} and TD_2^{on} are time delays for each exponential term. We define $A_{0n}^{on} = A_0^{on}\left(1 - e^{\frac{-TD_1^{on}}{\tau_0^{on}}}\right)$, $A_{1n}^{on} = A_1^{on}\left(1 - e^{\frac{-(TD_2^{on} - TD_1^{on})}{\tau_1^{on}}}\right)$ and $A_{2n}^{on} = A_2^{on}\left(1 - e^{\frac{-(t_f - TD_2^{on})}{\tau_2^{on}}}\right)$. The 3 phase model is such that after $t = TD_1^{on}$ we terminate phase 1 and put $\dot{V}O_2(TD_1^{on}) = \dot{V}O_2(0) + A_{0n}^{on}$, the second phase then starts and is added to $\dot{V}O_2(TD_1^{on})$. Phase 2 is terminated after $t = TD_2^{on}$, *giving* $\dot{V}O_2(TD_2^{on}) = \dot{V}O_2(0) + A_{0n}^{on} + A_{1n}^{on}$, the third phase then starts and is added to $\dot{V}O_2(TD_2^{on})$. Phase 3 is terminated after $t = t_f$ giving $\dot{V}O_2(t_f) = \dot{V}O_2(0) + A_{0n}^{on} + A_{1n}^{on} + A_{2n}^{on}$. Note that below the *LT* there is no slow component or phase 3 and hence the model will have only 2 phases. Note also when phase 1 is active phase 2 and the slow component are inactive due to the time delays TD_1^{on} and TD_2^{on} respectively. Phase 2 becomes active after $t = TD_1^{on}$, which approximately 10 seconds, followed by the slow component after $t = TD_2^{on}$, which approximately 180 seconds.

It should also be noted that many models have $A_0^{on} = 0$ as this initial cardiodynamic first phase is does not directly represent muscle O_2 utilization (Krogh and Lindhard [37]), instead they start the modelling after TD_1^{on}. Also $TD_1^{on} = TD_2^{on}$ as its not necessarily proven that the slow component should really start after a different time delay than phase 2. In Engelen et al [26] and Ozyener et al [49] they model the off exercise kinematics as well, in the following form

$$\dot{V}O_2 = EE\dot{V}O_2 - A_0\left(1 - e^{-\frac{(t)}{\tau_0}}\right) - A_1\left(1 - e^{-\frac{(t-TD)}{\tau_1}}\right) - A_2\left(1 - e^{-\frac{(t-TD)}{\tau_2}}\right) \tag{6}$$

where the parameters are as before apart from $EE\dot{V}O_2$ which is the end exercise level of $\dot{V}O_2$ and the time delay $TD_1 = TD_2 = TD$ as the phase 2 dynamics and the slow component begin at different times during the intense exercise, but in the off phase they exist together at the time the off phase begins. Phase 1 is believed to be present in both the on and off responses to exercise.

3.2.6 Oxygen deficit solutions

In Bearden and Moffatt [7] they define the oxygen deficit in two ways, the traditional way O_{2trad}^{defi} and their modified way $O_{2b\&m}^{defi}$, from equations 5.

$$O_{2trad}^{defi} = (\dot{V}O_2(0) + A_{0n}^{on} + A_{1n}^{on} + A_{2n}^{on}).(t_f - t_0) - \int_{t_0=0}^{t_f} \dot{V}O_2 dt \tag{7}$$

$$O_{2b\&m}^{defi} = (\dot{V}O_2(0) + A_{0n}^{on} + A_{1n}^{on} + A_{2n}^{on}).(t_f - t_0) - (A_{2n}^{on}.TD_2^{on}) - \int_{t_0=0}^{t_f} \dot{V}O_2 dt \tag{8}$$

$$\int_{t_0=0}^{t_f} \dot{V}O_2 dt = \dot{V}O_2(0).t_f + A_{0n}^{on}(t_f - TD_1^{on}) + A_{1n}^{on}(t_f - TD_2^{on}) +$$

$$+ \int_{t_0=0}^{TD_1^{on}} A_0^{on}\left(1 - e^{-\frac{(t)}{\tau_0^{on}}}\right)dt + \int_{TD_1^{on}}^{TD_2^{on}} A_1^{on}\left(1 - e^{-\frac{(t-TD_1^{on})}{\tau_1^{on}}}\right)dt + \int_{TD_2^{on}}^{t_f} A_2^{on}\left(1 - e^{-\frac{(t-TD_2^{on})}{\tau_2^{on}}}\right)dt.$$

$$O_{2trad}^{defi} = -A_0^{on}.TD_1^{on}\, e^{-\frac{TD_1^{on}}{\tau_0^{on}}} + A_1^{on}\left(TD_1^{on} - TD_2^{on}.e^{-\frac{TD_2^{on}-TD_1^{on}}{\tau_1^{on}}}\right) + A_2^{on}\left(TD_2^{on} - t_f.e^{-\frac{t_f-TD_2^{on}}{\tau_2^{on}}}\right) -$$

$$- A_0^{on}.\tau_0^{on}\left(e^{-\frac{TD_1^{on}}{\tau_2^{on}}} - 1\right) - A_1^{on}.\tau_1^{on}\left(e^{-\frac{TD_2^{on}-TD_1^{on}}{\tau_1^{on}}} - 1\right) - A_2^{on}.\tau_2^{on}\left(e^{-\frac{t_f-TD_2^{on}}{\tau_2^{on}}} - 1\right) \tag{9}$$

$$O_{2b\&m}^{defi} = -A_0^{on}.TD_1^{on}\, e^{-\frac{TD_1^{on}}{\tau_0^{on}}} + A_1^{on}\left(TD_1^{on} - TD_2^{on}.e^{-\frac{TD_2^{on}-TD_1^{on}}{\tau_1^{on}}}\right) + A_2^{on}(TD_2^{on} - t_f)\, e^{-\frac{t_f-TD_2^{on}}{\tau_2^{on}}} -$$

$$-A_0^{on}.\tau_0^{on}\left(e^{-\frac{TD_1^{on}}{\tau_0^{on}}}-1\right)-A_1^{on}.\tau_1^{on}\left(e^{-\frac{TD_2^{on}-TD_1^{on}}{\tau_1^{on}}}-1\right)-A_2^{on}.\tau_2^{on}\left(e^{-\frac{t_r-TD_2^{on}}{\tau_2^{on}}}-1\right) \quad (10)$$

If we look at solutions where we assume we have neither the first or third phase then we have $A_0=A_2=TD_1=TD_2=0$ which gives $O_2^{defi}=A_1.\tau_1\left(1-e^{-\frac{t}{\tau_1}}\right)$ which when taken in the limit that t approaches infinity gives $O_2^{defi}=A_1.\tau_1$. Such solutions where there is no slow component work for intensities below the lactate threshold and for very intense exercise where the oxygen demand is in excess of the $\dot{V}O_{2max}$, see Whipp [68].

3.2.7 Solutions for the time spent at $\dot{V}O_{2max}$

Billat *et al* [12] used the constant velocity mono exponential version to calculate the time spent at $\dot{V}O_{2max}$ (i.e. $A_0 = A_2 = 0$)solving equation 5 for t gives

$$t=-\tau ln\left[1-\frac{\dot{V}O_2(t)-\dot{V}O_{2b}(0)}{A_1}\right] \quad (11)$$

which is the time taken to arrive at $\dot{V}O_2(t)$. For $\dot{V}O_2(t)=\dot{V}O_{2max}-\delta$ the is term t is called the time to arrive within δ of $\dot{V}O_{2max}$, $TA_{\dot{V}O_{2max}-\delta}$. We now use this and the total time spent exercising, *tlim* to calculate the total time above $\dot{V}O_{2max}-\delta$, which we term the $tlim_{\dot{V}O_{2max}-\delta}$

$$tlim_{\dot{V}O_{2max}-\delta}=tlim-TA_{\dot{V}O_{2max}-\delta}. \quad (12)$$

In this manner we can compare the time spent above $\dot{V}O_{2max}-\delta$ when exercising at other exercise intensities, if it is believed they can be modelled sufficiently well with the mono exponential function. For the full three phase work, we need to account for the time delay, *TD*2 before the start of the slow component. We look at the solution for an exercise duration greater than *TD*2

$$\dot{V}O_2(t)=A_1+A_2\left(1-e^{-\frac{(t-TD_2)}{\tau_2}}\right). \quad (13)$$

$$t=TD_2-\tau_2 ln\left[1-\frac{\dot{V}O_2(t)-A_1}{A_2}\right] \quad (14)$$

Where as before we let $\dot{V}O_2(t)=\dot{V}O_{2max}-\delta$ hence t becomes $TA_{\dot{V}O_{2max}-\delta}$. The time spent above $\dot{V}O_{2max}-\delta$ can then be found as before by using equation 12. This shows the limitations in the current model (i.e. equation 5) as we have to be able to find TD_2 before we can predict the $tlim_{\dot{V}O_{2max}-\delta}$ and to do this requires a time series for that intensity. It would be better if we could predict such quantities for other velocities other than those we have data for, the same goes for all the other parameters. This point has been addressed in Stirling et al [60] along with other issues regarding the smoothness of the model.

3.3 Blood lactate BLa(t) kinematics

To understand what blood lactate is we need first to understand the chemical equations for aerobic and anaerobic metabolism in the production of energy. The energy released from complete cellular breakdown of fuels is stored in the form of adenosine triphosphate (ATP). Below we present the aerobic breakdown (i.e. in the presence of oxygen) of a carbohydrate such as glucose is given by

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O + 36ATP. \tag{15}$$

This is a highly efficient process which results in enormous amounts of energy for work (ie. *36ATP*). Where as the anaerobic breakdown (i.e. in the absence of oxygen) of glucose

$$C_6H_{12}O_6 \rightarrow 2ATP + 2 \text{ Lactic acid} \leftrightarrow 2H^+ + 2LAC^- \tag{16}$$

which is far less efficient (ie. only *2ATP* as apposed to 36 is produced). Under physiological conditions the lactic acid produced dissociates almost immediately and completely to lactate ions (Lac^-) and hydrogen ions (H^+). This eventually inhibits the enzymatic breakdown sequence as the enzymes involved in fuel breakdown operate operate within a narrow range of acidity. The accumulation of H^+ ions causes much discomfort leading to an eventual reduction in the intensity of the exercise. Anaerobic metabolism is the main source of energy for very intense short duration exercises, ie. a 400 meter race, races over 1500 meters are have approximately a 50 : 50 balance between the two forms of metabolism and as the duration of the race increases then so does the percentage of the aerobic contribution. Due to the difference in the energy production in aerobic and anaerobic metabolism the more the energy demand can be met by aerobic metabolism the more efficient the process. This is one of the effects of increased fitness.

The concentration of blood lactate during incremental exercise reflects the difference between the release of lactate from the muscle (muscle efflux) and the uptake of blood lactate by muscle and other tissue. This means that the concentration of blood lactate doesn't necessarily reflect muscle lactate production. As a result it should be noted that the accumulation of blood lactate per se does not play an underlying role in the alteration of metabolic processes during exercise, (for a review see Brooks [17, 18]). However it is still believed that blood lactate concentration is a useful tool for predicting endurance performance and intensity of exercise training programs. One very common training session for example is 1 or 2 repetitions of 15 to 20 minutes at a velocity corresponding to the lactate threshold *LT* with a warm up and warm down and 5 minutes recovery easy running between the repetitions. The blood lactate is commonly measured at discrete intervals in the time series in exercise physiology labs along with the heart rate, $\dot{V}O_2$ and velocity while doing physiological testing.

3.3.1 BLa terminology: Lactate threshold and maximum lactate steady state

There are two obvious features of the blood lactate kinematics which are of interest from both a practical and mathematical point of view. One is the lactate threshold *LT* and the other is the individual anaerobic threshold, *IAT*. The lactate threshold, *LT* is not really a threshold in mathematical sense of the word as there is no abrupt change, however in exercise physiology it is defined as the point at which blood lactate begins to accumulate, above resting levels during exercise of increasing intensity. This can be seen as the point at which the gradient of the *BLa vs V* curve begins to rapidly increase its gradient, see figure 6. Another way of defining it is the highest $\dot{V}O_2$ or *HR* that can be attained during incremental exercise before an elevation in blood lactate is observed. This is also called the lactate break point, the onset of plasma lactate accumulation, the anaerobic threshold and the aerobic threshold. The individual anaerobic threshold, *IAT* is defined as the highest $\dot{V}O_2$ or *HR* that can be maintained over time (15 to 20 minutes) without a continual increase in the blood lactate, see figure 7. The term maximal lactate steady state, *MLSS* is also used. The definition of this term is far more rigorous from a mathematical point of view than that of the *LT*, however both terms are commonly used. With training the lactate threshold occurs at a higher speed allowing the runner to run faster without accumulating lactate in the blood. This results from a greater ability to clear lactate from the muscles and also less lactate begin produced for the same work rate. The ability to exercise at a high intensity without accumulating lactate is beneficial to the athlete because lactate formulation contributes to fatigue.

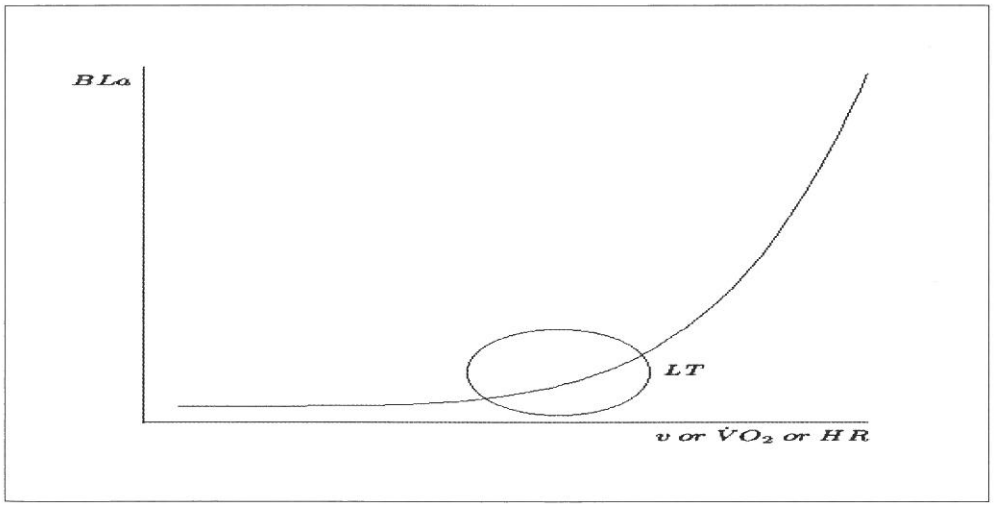

Figure 6. A sketch of the *BLa vs* v or *HR* or $\dot{V}O_2$, highlighting the lactate threshold, *LT*. Either of v, $\dot{V}O_2$ or *HR* are commonly used as the x axis variable.

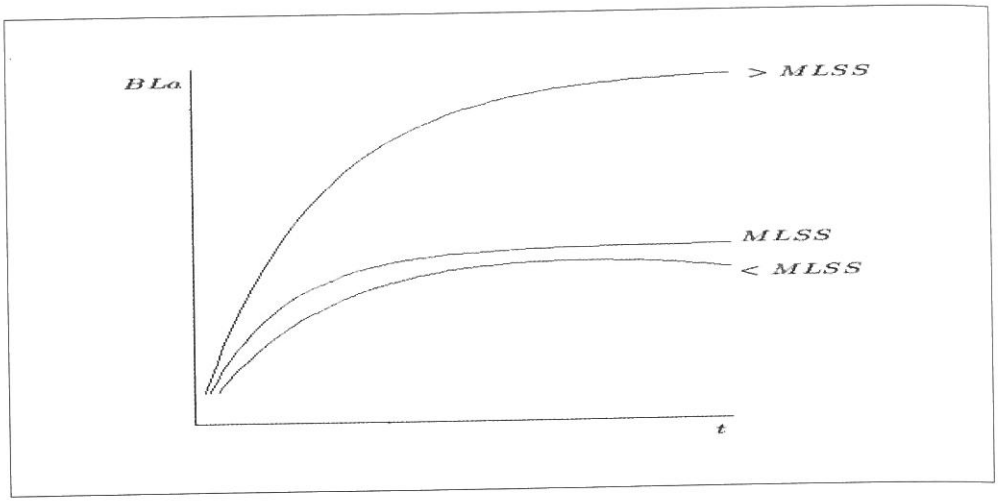

Figura 7. A sketch of the *BLa vs t* highlighting the maximum lactate steady state, *mlss*.

3.3.2 The relationship between the ventilatory threshold and lactate threshold

The so called thresholds for both the ventilation and lactate accumulation have been shown to occur together with a high correlation for metabolically normal people. (examples and references are given in Martin and Coe [41]). There is confusion regarding the terminology as there are two thresholds, and many names to describe them. The first threshold is observed during mild work (accompanied by breathing changes and a small rise in blood lactate concentration) the second threshold however occurs during more intense exercise (accompanied by breathing changes and and a steady accumulation of lactate).

The second threshold is the one often sensed by runners as a change in the breathing intensity such that the conversation stops. At the intensity of work which results in rapidly accumulating lactate levels elevated ventilatory removal of CO_2 can no longer maintain blood acidity within reasonable limits. As blood lactate levels rapidly rise the Ph of the blood begins to fall and the rising H^+ concentration provides a powerful ventilatory stimulus.

Gaesser and Poole [27] stated that "although the blood lactate profile is well correlated with the time course and magnitude of the $\dot{V}O_2$ slow component, the relationship is coincidental rather than causal. The most likely mechanism accounting for the slow component is muscle fibre recruitment". There also exists work however which dispute the coincidental relationship between a blood lactate threshold and the ventilatory threshold. Examples and references of work in which it was found that the anaerobic and lactate threshold do not coincide are show in the book of Weltman [66]. Weltman concludes that "recent studies have examined the anaerobic, ventilatory and lactate

thresholds... The results indicate that the three thresholds are determined by different underlying mechanisms and reflect different phenomena."

3.3.3 BLa vs t

It can be shown that for low values of exercise intensity there is a minimal increase, if any, in blood lactate concentration with time. There are even some cases where the concentration decreases during low-intensity exercise as blood lactate is used as a substrate. For higher intensity exercise a velocity or power output is reached above which the blood lactate concentration is a nonlinear increasing function of time.

3.3.4 BLa vs velocity

It can be shown that blood lactate is an exponential function of intensity or velocity V. If incremental exercise is carried out it can be seen that for low intensities there is no increase from the base or resting levels of blood lactate concentration. This is so up until the lactate threshold, after which for increasing intensities of exercise we get an exponentially increasing value of the blood lactate concentration.

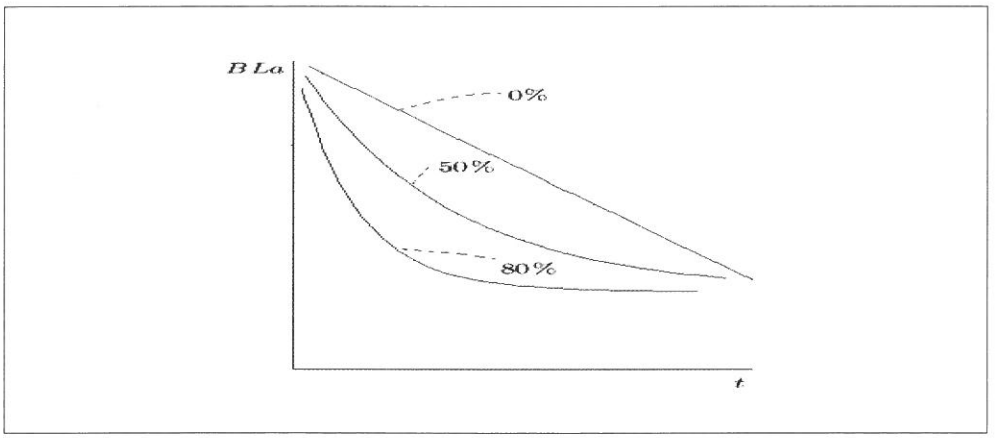

Figura 8. A sketch of the blood lactate time series for recovery at 3 different exercise intensities (given as a $\dot{V}O_{2max}$) from and elevated blood lactate level following hard exercise

In all anaerobic training it is of value to have active rest, ie. low intensity exercise not exceeding a work intensity demanding above 60% of $\dot{V}O_{2max}$. lactate is then transferred quicker from the fatigued muscles because the perfusion of the muscle is maintained at a higher level than at rest. A prerequisite is that the rest exercise intensity is low enough not to cause further lactate formation (ie. must be below the lactate threshold). Note that even faster returns to pre-exercise blood lactate response could be noticed at relative work loads up to 80% of maximal oxygen uptake, $\dot{V}O_{2max}$ in endurance trained

individuals, this is probably below their lactate threshold. Note the difference in the shape of the curves for different recovery intensities, (see figure 8), with 0% being approximately linear and 60% to 80% being exponential, with 80% having the steepest drop in *BLa*, see Saltin [57].

3.3.5 BLa and heart rate or $\dot{V}O_2$ relationship

The relationship between the blood lactate concentration and the $\dot{V}O_2$ has been subject to debate (see Myers etal [45]). Essentially the arguments are related to wether a continuous function can describe the blood lactate kinematics for tests of increasing intensity. Beaver et al [8] used a regression of the log of the lactate vs the log of the $\dot{V}O_2$ to try to highlight the lactate threshold. The lack of a threshold however was reported in Yeh et al [71]. It was found in Hughson et al [35] that the difference in fit between the two models is minor and it was believed the continuous models was a more appropriate model. This continuous model also described in Myers et al [45] is an exponential function in the following form,

$$BLa = a + b.e^{c.\dot{V}O_2} \tag{17}$$

where *a, b, c* are parameters. This shows that the blood lactate can be modelled as an exponential function of $\dot{V}O_2$ (and hence *HR*), for ramp exercise tests. Its should be noted that this function appears similar though not identical to the dependency of the blood lactate on the exercise intensity of velocity, *V*. In fact may figures interchange the velocity, with the $\dot{V}O_2$ as the *x* axis measure of intensity.

Note that Billat [10] provides a method of determining the maximum lactate steady state, based on an interpolation between a value above and one below this state, see Billat [10] for further details.

3.4 Speed: υ_{mara}, υ_{mlss}, υ_{crit} and the minimum velocity to achieve $\upsilon\dot{V}O_2^{max}$ in a incremental step test

Speed is another variable which is recorded when carrying out physiological test. The speed at which certain physiological phenomena such as the lactate threshold *LT*, maximum lactate steady state or $\dot{V}O_{2max}$ occur is often calculated in such tests and then used to set training sessions and record fitness levels. Billat and Koralsztein [9], and Billat [14] provide a detailed review of the subject, including its history, uses, definitions and protocols for determining it.

Daniels et al. [21] introduced the term velocity at $\dot{V}O_2^{max}$, or $\upsilon\dot{V}O_{2max}$ which was found as the first velocity to achieve $\dot{V}O_2^{max}$ in an incremental step test with 3 minute increments. The velocity was found to be close to that in a 3000 m race (with an approximate duration of 9 minutes) for elite female runners. They reported it to be a very useful term which combined $\dot{V}O_2^{max}$ and economy into a single variable, which could be used to identify differences between various runners.

There is a range of velocities for which there is a sustained increase in blood lactate and a decrease in arterial Ph with time. $\dot{V}O_2$ increases in a mono exponential way and stabilizes at approximately 80% for at least 1.5 hours of steady exercise (in top level marathon runners). After this time period we can get an increase in $\dot{V}O_2$ due to the so called $\dot{V}O_2$ drift which is due to thermoregulatory constraints. Such effects happen at a velocity, υ_{mara} which is the best velocity at which to race a marathon it is approximately 80% of $\upsilon\,\dot{V}O_2^{max}$. If we increase the velocity to υ_{mlss} which is equal to approximately 85% of $\upsilon\,\dot{V}O_2^{max}$ the maximal lactate steady state occurs, when the lactate levels are stable as the production of lactate is equal to the consumption. The time limit at this velocity is approximately an hour.

Hill and Lupton [31] said that when using Hill as the subject, "the rate of oxygen intake due to exercise increases as speed increases, reaching a maximum for the speeds beyond about 256 m/min. At this particular speed, for which no further increase in O_2 intake can occur, the heart, lungs, circulation, and the diffusion of oxygen to the active muscle-fibres have attained there maximum activity. At higher speeds the requirement of the body for oxygen is far higher but cannot be satisfied, and the oxygen debt continually increases." They then go on to say, "Considering the case of running, there is clearly a critical speed for each individual at which there is a genuine dynamic equilibrium, breakdown being balanced by restoration, above which, however, the maximum oxygen intake is inadequate, lactic acid accumulating, a continuously increased oxygen debt being occurred, fatigue and exhaustion setting in."

This critical velocity, the minimum velocity to elicit maximum oxygen uptake, or $\dot{V}O_2^{max}$, was defined by Moritani et al. [44] as close to the lactate threshold velocity. Volkov et al. [64] used this minimum velocity, or critical speed to measure the maximal aerobic capacity, the total oxygen consumption at $\dot{V}O_2^{max}$ when the subject is asked to run at the critical speed until exhaustion. Billat [14] defines the critical velocity to be the highest velocity for which the $\dot{V}O_2$ can reach a, delayed, sub maximal steady state, this is true also for the blood lactate Poole et al [50, 51]. Like wise Hill and Ferguson [32] demonstrated that "the υ_{crit} is the threshold intensity above which exercise of sufficient duration will lead to the attainment of $\dot{V}O_2^{max}$". Hill et al. [34] use this in the following definition of the severe intensity domain "the severe exercise intensity domain may be defined as that range of work rates over which $\dot{V}O_{2max}$ can be elicited during constant-load exercise... This upper boundary is the highest work rate for which exercise duration is prolonged sufficiently (in this study, 136 ± 17*s*) to allow $\dot{V}O_2$ to rise to its maximal value. The lower boundary for severe exercise is just above υ_{crit}, which is the highest work rate that is sustainable for a prolonged duration and that will not elicit $\dot{V}O_{2max}$". υ_{crit} is at approximately 90% of $\upsilon\dot{V}O_{2max}$ for well trained distance runners. At this velocity Billat [14], claims there is a slow component but it doesn't cause the $\dot{V}O_2$ to reach $\dot{V}O_2^{max}$, instead the $\dot{V}O_2$ stabilizes at a sub maximal steady state value, (i.e. for an athlete running at 90% of $\upsilon\dot{V}O_{2max}$ the $\dot{V}O_2$ could stabilize at 95% $\dot{V}O_{2max}$). As we are above the MLSS the concentration of blood lactate increases as we accumulate more lactate due to

there being more produced than consumed. The time limit at this velocity is reduced to less than 30 minutes due to rapid glycogen depletion.

Above the critical velocity (i.e. during high intensity exercise) neither the blood lactate or $\dot{V}O_2$ are stabilized and hence the $\dot{V}O_2$ continues to rise till it reaches the $\dot{V}O_{2max}$ unless fatigue sets in before $\dot{V}O_{2max}$ can be achieved. As shown in Billat and Koralsztein [9] there are a range of velocities which will achieve $\dot{V}O_{2max}$ and there are many different ways to estimate $\upsilon\dot{V}O_{2max}$. What can also be seen is that they don't get the minimum velocity to achieve $\dot{V}O_2^{max}$ as it is shown that at 90% of the $\upsilon\dot{V}O_{2max}$ some subjects reach $\dot{V}O_2^{max}$ in a run to exhaustion. The reason for this is in the way the $\upsilon\dot{V}O_{2max}$ is determined by step test where the velocity is incremented every few minutes where as the time to exhaustion at 90% of the $\upsilon\dot{V}O_{2max}$ is of the order of 10 times greater than the step increment.

What is also of much interest is the time limit t_{lim} at $\upsilon\dot{V}O_{2max}$, (this was calculated using the three phase model in section 3.2.7) or the time to exhaustion when running at $\upsilon\dot{V}O_{2max}$. According to Billat and Koralsztein [9] the average t_{lim} is 6 minutes ±25%. This t_{lim} is highly variable for groups of individuals with the same $\upsilon\dot{V}O_{2max}$, but for an individual the t_{lim} is reproducible (for a particular test protocol). It was shown that the t_{lim} at $\dot{V}O_2^{max}$ was correlated with the velocity at the lactate threshold (and maximal lactate steady state) expressed as a percentage of the $\upsilon\dot{V}O_{2max}$. This would explain the inter individual differences of the t_{lim} and suggest that the role of anaerobic contribution should be taken into account. Billat and Koralsztein [9] also showed that the t_{lim} at $\dot{V}O_2^{max}$ was correlated negatively with the $\upsilon\dot{V}O_{2max}$. Such that the runners who attained the highest $\upsilon\dot{V}O_{2max}$ spent the least time t_{lim} at $\dot{V}O_2^{max}$ before exhaustion.

4. CONCLUSIONS

The work presented here introduced some areas of biomechanics but focused mainly on models of the physiological response to exercise. Though the models presented here are not sophisticated from the point of view of modern mathematics, they are what is currently used. It can be seen that there are many open problems in the field of exercise physiology (and also biomechanics) and as a result the application of more modern and sophisticated techniques of analysis and modelling could provide very interesting and far reaching results. This is particularly the case with the use of tools from time series analysis (see Kants and Schrieber [36]) and non linear dynamics (see Guckenheimer and Holmes [29]) as most of the physiological variables have a time series which is a nonlinear functions of time and exercise intensity.

The use of more modern methods of analysis would be of much benefit for understanding the correlations and interactions between variables. These tools are also of much value in improving the identification of certain features of the time series of

physiological variables, such as the lactate threshold, LT and identifying better their relationship with fitness and performance levels.

The use of modern mathematical methods for analysis and modelling of the processes we have described could not only have a large impact in the development and understanding of training methodology and the testing of athletes involved in sport but also in medicine. Exercise physiology and biomechanics are fundamental areas of medicine involving not only the general health of the population but also such things as recover and recuperation following illnesses and injuries etc. The tools described here are also used in the diagnosis of many different disorders causing exercise intolerance, see Wasserman et al [65].

AKNOWLEDGMENTS

The authors would like to acknowledge their collaborator Professor Bengt Saltin.

REFERENCES

1. ASTRAND PO., RODAHL K., DAHL HA., STROMME SB., 2003 "Textbook of Work Physiology: Physiological of Bases of Exercise", *Human Kinetics*.
2. ASTRAND, P.-O. and SALTIN, B., 1961, "Oxygen uptake during the first minutes of heavy muscular exercise", *Journal of Applied Physiology*, 16, 6, 971-976.
3. BANGSBO, J., 1996, "Physiological factors associated with efficiency in high intensity exercise", *Sports Medicine*, 22, 5, 299-305.
4. BANGSBO, J., 1996, "Oxygen deficit: a measure of the anaerobic energy production during intense exercise?," *Canadian Journal of Applied physiology*, 21, 5, 350-363.
5. BANGSBO, J., 1998, "Quantification of anaerobic energy production during intense exercise", *Medicine and Science in Sports and Exercise*, 30,1, 47-52.
6. BARSTOW, T.J., 1994, "Characterization of $\dot{V}O_2$ kinetics during heavy exercise," Medicine and Science in Sport and Exercise, 26, 11, 1327-1334.
7. BEARDEN SE., and MOFFATT RJ., 2000, "$\dot{V}O_2$ kinetics and the O_2 deficit in heavy exercise", *Journal of applied physiology* 88, 1407-1412.
8. BEAVER, WL., WASSERMANN, K., WHIPP, BJ., 1985, "Improved detection of the lactate threshold during exercise using a log-log transformation", *Journal of applied physiology* 59, 1936-1940.
9. BILLAT, V.L., KORALSZTEIN, J., 1996, "Significance of the Velocity at $\dot{V}O_{2max}$ and the time to exhaustion at this velocity", *Sports Med.* 22 (2) 90-108.

10. BILLAT, V.L., 1996, "Use of blood lactate measurements for the prediction of exercise performance and control of training, recomendations for long distance running", *Sports Med.* 22 (3) 157-175.
11. BILLAT, V.L., KORALSZTEIN, J., MORTON HR., 1999, "Time in Human endurance models, from emperical models to physiological models", *Sports Med.* 27 (6) 359-379.
12. BILLAT, V.L., BOCQUET, V., SLAWINSKI, J., LAFFITTE, L., DEMARLE, A., CHASSAING, and KORALSZTEIN, J., 2000, "Intermittent runs at the velocity associated with maximal oxygen uptake enables subjects to remain at maximal oxygen uptake for a longer time than intense but sub maximal runs", *European Journal of Applied Physiology* 81, 188-196.
13. BILLAT, V.L., BOCQUET, V., SLAWINSKI, J., LAFFITTE, L., DEMARLE, A., CHASSAING, P. and KORALSZTEIN, J. P., 2000, "Effect of prior intimittent runs at $\upsilon\dot{V}O_{2max}$ on Oxygen kinetics during an all-out severe run in humans" *J. Sports Med., phys. fitness* 40, 185-194.
14. BILLAT, V.L., 2000, "$\dot{V}O_2$ slow component and performance in endurance sports", *British Journal of sports medicine*, 34 (2), 83-85.
15. BILLAT, V.L., 2001, "Interval training for performance: a scientific and emperical practice, part I, Aerobic interval training", *Sports Med.* 31 (1) 13-31.
16. BILLAT, V.L., 2001, "Interval training for performance: a scientific and emperical practice part II, Anaerobic interval training", *Sports Med.* 31 (2) 75-90.
17. BROOKS GA., 1985, "Anaerobic threshold: a review of the concept and directions for future research", *Med. Sci. in sports and exc.* 17, 1, 22-31.
18. BROOKS GA., 2002, "Lactate shuttles in nature", *Biochemical soc. trans.* 30, 2, 258-264.
19. BURKE, E.R. (Ed), 1998, "Precision heart rate training", *Champaign IL, Human Kinetics*.
20. CAVANAGH PR., (Ed), 1990, "Biomechanics of distance running", *Champaign IL, Human Kinetics*.
21. DANIELS, J., SCARDINA, N., HAYES, J., et al., 1984, "Elite and subelite female middle and long distance runners", in: Landers DM., editor. *Sport and Elite Performers*, Vol. 3, Proceedings of the 1984 Olympic Scientific Congress: 1984 Jul 19-23, Oregon. *Champaign (IL), Human Kinetics*, 57-72.
22. DANIELS, J., 1998, "Daniels' running formula", *Champaign IL, Human Kinetics*.
23. DAVIES, CT., DI PRAMPERO, PE., CERRETELLI, P., 1972, "Kinetics of the cardiac output and respiratory gas exchnage during exercise and recovery" *Journal of applied physiology* 32, 618-625.
24. DE WIT B., DE CLERCQ D., AERTS P., 2000, "Biomechanical analysis of the stance phase during barefoot and shod running", *Journal of Biomechanics*, 33, 269-278.
25. DE CLERCQ D., AERTS P., KUNNEN M., 1994, "The mechanical characteristics of the human heel pad during foot strike in running: a in vivo cineradiographic study", *Journal of Biomechanics*, 27, 1213-1222.

26. ENGELEN, M., PORSZASZ, J., RILEY, M., WASSERMAN, K., MAEHARA, K., and BARSTOW TJ., 1996, "Effects of hypoxic hypoxia on O_2 uptake and heart rate kinetics during severe exercise", *J. Appl. Physiol.* 81(6), 2500-2508.
27. GAESSER and POOLE, 1996, "The slow component of oxygen uptake in humans", *Exercise and sport science reviews*, 24, 35-70.
28. GAJEWSKI, J., WOJCZUK, J., 1989. "Changes in tremor following high intensity exercise", *Biology of Sport*, 6 (3), 225-232.
29. GUCKENHEIMER J. and HOLMES P., 1983, *Nonlinear oscillations, dynamical systems and bifurcations of vector fields*, Springer-Verlag.
30. HENNIG, EM., VALIANT GA., LIU Q., 1996, "Biomechanical variables and the perception of cushioning for running in various types of footware", *Journal of Biomechanics*, 12, 143-150.
31. HILL, A.V., LUPTON, L., 1923, "Muscular exercise, lactic acid and the supply and utilization of oxygen", *Q. J. Med.*, 16 135-171.
32. HILL, D.W., FERGUSON, C.S., 1999, "A physiological description of critical velocity", *Eur. J. Physiol. Occup. Physiol.*, 79 (3), 290-293.
33. HILL, D.W., STEVENS E.C., 2001, "The $\dot{V}O_2$ responce to the onset of severe intensity exercise", *Can. J. Appl. Physiol.* 26 (4), 350-355.
34. HILL, D.W., POOLE D.C., SMITH J.C., 2002, "The relationship between power and tiem to achive $\dot{V}O_{2max}$", 34 (4), 709-714.
35. HUGHSON, RL., WEISIGER, KH., Swanson GD., 1987, "Blood lactate concentration increases as a continuous function in progressive exercise", *Journal of applied physiology*, 62 1975-1981.
36. KANTZ H. and SCHREIBER T., 1999, *Nonlinear Time Series Analysis*, Cambridge University Press.
37. KROGH, A., LINDHARD, J., 1913, "The regulation of respiration and circulation during the initial stages of muscular work" *Journal of Physiology*, 47, 112-136.
38. LAMARRA, I., WHIPP, BJ., WARD, SA., and WASSERMAN, K., 1987, "Effect of interbreath fluctuations on characterizing exercise gas exchangekinetics" *J. Appl. Physiol.*, 62, 2003-2012.
39. LINNARSSON, D., 1974, "Dynamics of pulmonary gas exchange and heart rate changes at start and end of exercise", *Acta Physiol. Scand. Suppl.* 415, 1-68.
40. MARGARIA, R., CERRETELLI, P., HGHENSO, P., SASSI, G., 1963, "Energy cost of running", *Journal of applied Physiology*, 18, 2, 367-370.
41. MARTIN, D.E. and COE, P.N. , 1997, "Better training for distance runners", *Champaign IL: Human Kinetics*.
42. MEDBO, J.I., MOHN, A.C., TABATA, I., 1988, "Anaerobic capacity determined by maximal accumulated O_2 deficit", *J. Appl. Physiol.*, 64, 50-60.
43. MILLE-HAMARD, L., LAFFITE, L., DEMARLE, A., KORALSZTEIN, J. P. and BILLAT, V.L., 2000, "Oxygen-pulser steady state allows estimation of $\dot{V}O_2$, by heart rate monotoring in a sever submaximal run", *Biology of sport,* 17, 3, 389-399.

44. MORITANI, T., NAGATA, A., DE VRIES, H.A., et al., 1981, "Critical power as a measure of physical working capacityt and anaerobic threshold", *Ergonomics*, 24, 339-350.
45. MYERS, J., WALSH, D., BUCHANAN N., MCAULEY P., BOWES E., FROELICHER V., 1994, "Increase in blood lactate during ramp exercise: comparison of continuous and threshold models", *Med. Sci. Sports Exerc.*, 26, 11, 1413-1419.
46. NG-THOW-HING, V., FALOUTSOS, P., 2000, *Dynamic Animation and Control Environment (Dance), Software, version 1.0*, University of Toronto.
47. NIGG BM., HERZOG W., 1999, *Biomechanics of the musculo-skeletal system*, second edition, John Wiley & Sons.
48. NOAKES, T., 1991, "Lore of running. Discover the science and spirit of running", *Champaign IL, Human Kinetics*.
49. OZYENER, F., ROSSITER, HB., WARD, SA., and WHIPP, BJ., 2001, "Influence of exercise intensity on the on -and off- transient kinetics of pulmonary oxygen uptake in humans", *Journal of Physiology*, 53.3, 891-902.
50. POOLE, D.C., WARD, S.A., GARDNER G.W., WHIPP, B.J., 1988, "Metabolic and respiratory profile of the upper limit for prolonged exercise in man", *Ergonomics*, 31, 1265-1279.
51. POOLE, D.C., WARD, S.A., WHIPP, B.J., 1990, "The effects of training on the metabolic and respiratory profile of high intensity cycle ergometer exercise", *Eur. J., Appl., Physiol.* 59, 421-429.
52. DI PRAMPERO, P.E., 1986, "The energy cost of human locomotion on land and in water", *Int. J. Sports Med.*, 7, 55-72.
53. ROBBINS SE., HANNA AM., 1987, "Running-related injury prevention through barefoot adaptations", *Medicine and Science in Sports and Exercise*, 19, 148-156.
54. ROBBINS SE., GOUW GJ., HANNA AM., 1991, "Running-related injury prevention through innate impact-moderating behaviour", *Medicine and Science in Sports and Exercise*, 21, 130-139.
55. ROBBINS SE., GOUW GJ., 1991, "Athletic footware: unsafe due to perceptual illusions", *Medicine and Science in Sports and Exercise*, 23, 217-224.
56. SALTIN, B., ?, "Anaerobic Capacity: Past, present and Prospective", In Taylor, A., and Gollmack, P.D., (Eds), *Biochemistry of exercise VII, Human Kinetics: Champaign IL,* 387-412.
57. SALTIN, B., 1987, "The physiological and biochemical basis of aerobic and anaerobic capacities in man; effect of training and range of adaption", In Maehlum, S., Nilsson, S., Renstrom, P., eds. *An update in sports medicine: Proceedings from the second Scandinavian conference in sports medicine*. 16-59.
58. SALTIN, B., LARSEN, H., TERRADOS, W., BANGSBO, J., BAK, T., KIM, C.K., SVEDENHAG, J. and ROLF, C.J., 1995, "Aerobic exercise capacity at sea level and at altitude in Kenyan boys, junior and senior runners compared with Scandinavian runners," *Scandinavian Journal of Medicine in Sport*, 5, 209-221.

59. SALTIN, B., KIM, C.K., TERRADOS, N., LARSEN, H., SVENDENHAG, J. and ROLF, C.J., 1995, "Morphology, enzyme activities and buffer capacity in leg muscles of Kenyan and Scandinavian runners," *Scandinavian Journal of Medicine in Sport*, 5, 222-230.
60. STIRLING, JR., ZAKYNTHINAKI MS., SALTIN B., 2003, *A model of oxygen up take kinetics: Including a method for caculating oxygen demand, deficit and debt*, CRM prepint.
61. STIRLING, JR., ZAKYNTHINAKI MS., 2003, *Stability in sport,* CRM prepint.
62. SWAIN, D.P. and LEUTHOLTZ B.C. , 1997, "Heart rate reserve is equivalent to $\%\dot{V}O_{2max}$ reserve, not to $\%\dot{V}O_{2max}$", *Medicine and Science in Sport and Exercise*, 29, 410-414.
63. SWAIN, D.P., LEUTHOLTZ, B.C. KING, M.E., HAAS, L.A. and BRAUCH, J.D., 1998, "Relationship between % heart rate reserve and $\%\dot{V}O_2$ reserve in treadmill exercise", *Medicine and Science in Sport and Exercise*, 30, 2, 318-321.
64. VOLKOV, N.L., SHIRKOVETS, E.A., BORILKEVICH, V.E., 1975, "Assessment of aerobic and anerobic capacity of athelets in treadmill running tests", *Eur. J. Appl. Physiol,* 34, 121-130.
65. WASSERMAN K., HANSEN JE., SUE DY., CASABURI R., WHIPP BJ., 1999, *Principles of exercise testing and interpretation, including pathophysiology and clinical applications, third edition*, Lippincott Williams and Wilkins.
66. WELTMAN A., 1995, *The Blood Lactate Response to Exercise (Current Issues in Exercise Science)* Human Kinetics.
67. WHIPP BJ., WASSERMAN, K., 1972, "Oxygen uptake kinetics for various intensities of constant load work", *Journal of applied physiology*, 33, 351-356.
68. WHIPP, B.J., 1994, "The slow component of O_2 uptake kinetics during heavy exercise", *Medicine and Science in Sport and Exercise*, 26, 11, 1319-1326.
69. WHIPP, B.J., ROSSITER, H.B. and WARD, S.A., 2002, "Exertional oxygen uptake kinetics: a stamen or stamina?", *Biochemical Society transactions*, 237-247.
70. WILMORE, J.H. and COSTILL, D.L. , 1999, "Physiology in Sport and Exercise", *Champagne, IL: Human Kinetics*.
71. YEH MP., GARDENER RM., ADAMS TD., YANOWITZ, FG., CRAPO RO., 1983, "Anaerobic threshold, problems of determination and validation," *Journal of applied physiology* 55, 1178-1186.
72. ZAKYNTHINAKI MS., STIRLING JR., 2003, "Appearance of high frequency signals combined with fatigue-induced increases in tremor as a means for the anticipation and detection of muscular failure", submitted, *Journal of Biomechanics*.

La tecnologia fractal: de les matemàtiques a l'empresa de tecnologia

Carles Puente i Baliarda
Director de Tecnologia, FRACTUS, S.A.

INTRODUCCIÓ

Els serveis de telecomunicació han crescut de forma exponencial durant els últims cinquanta anys. Molts d'aquests serveis depenen de les ones de ràdio per transportar un dels béns més apreciats d'avui dia: la informació. D'aquesta manera, la radio, la televisió i els sistemes de comunicació més recents han convertit les antenes en un aparell omnipresent a la nostra vida diària, que pot competir en difusió amb els refrigeradors, els forns i d'altres aparells quotidians.

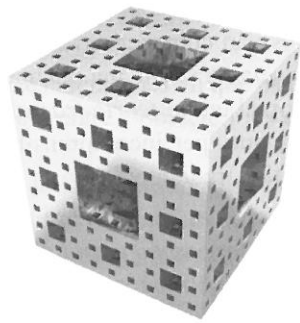

Figura 1. *Les propietats geomètriques dels fractals permeten dissenyar antenes per a aplicacions diverses.*

Les antenes són una de les parts més voluminoses en un sistema de telecomunicació. La longitud de l'ona electromagnètica radiada, i per tant la seva freqüència, imposa una forta restricció en la mida de l'antena. El comportament de l'antena és sovint molt sensible a la seva forma geomètrica i a la seva mida en relació amb la longitud d'ona. Això té dues conseqüències importants. La primera és que un cop l'antena s'ha dissenyat per a operar a una longitud d'ona particular, rarament és útil per a d'altres freqüències (altres longituds d'ona). Com que els sistemes de telecomunicació operen a bandes de freqüència separades, per tal de no interferir-se els uns amb els altres, s'ha de comptar

habitualment amb una antena individual per a cada sistema. La segona és que un cop assignada la banda freqüencial d'un servei en particular, i per tant la longitud d'ona de treball, l'antena no es pot fer arbitràriament petita, a causa de la dependència del comportament de l'antena amb la longitud d'ona.

La tecnologia fractal dóna solucions a aquestes limitacions de les antenes convencionals. Aquesta tecnologia permet dissenyar antenes multibanda o multifreqüència, és a dir, antenes capaces de ser compartides per múltiples serveis de comunicacions que operin simultàniament, i dissenyar antenes miniatura, és a dir, antenes que són capaces de radiar eficientment tot i la seva mida reduïda en relació amb la longitud d'ona.

TECNOLOGIA FRACTAL

Un objecte fractal és un objecte matemàtic abstracte, que, no obstant això, serveix per descriure geometries existents en la naturalesa. Així doncs, diverses formes naturals, com ara les muntanyes, les falgueres, la distribució de galàxies a l'univers, els bronquis del sistema pulmonar, les ramificacions dels arbres, els núvols, els llamps, o les costes continentals, per citar alguns exemples, es caracteritzen molt millor des del punt de vista de la geometria fractal que des del punt de vista de la geometria tradicional euclidiana. Les formes fractals es van començar a estudiar el segle XIX, però no va ser fins al segle XX, a la dècada dels setanta i de la mà del matemàtic Benoit B. Mandelbrot, quan es va utilitzar per primer cop el terme *teoria fractal* com a alternativa per descriure algunes formes geomètriques, que, per incomprensibles o rebuscades, havien estat deixades de banda per la matemàtica tradicional.

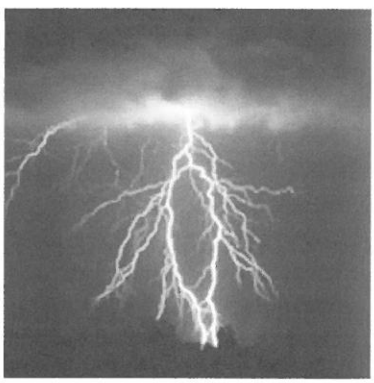

Figura 2. *Exemples de formes fractals que es poden trobar a la naturalesa.*

Les formes fractals es caracteritzen per dues propietats importants: l'autosimilitud i la dimensió fractal.

Autosimilitud: Una estructura és autosimilar si es pot descompondre arbitràriament en petites peces, i cadascuna d'elles és una petita rèplica de l'estructura total. És important destacar que les petites peces puguin, en realitat, obtenir-se de l'estructura total per transformació de versemblança, és a dir, cada petita peça de l'estructura total pot ésser vista com una còpia a diferents escales. Tot i que aquesta propietat d'autosimilitud s'observa en moltes geometries fractals, i és la més fàcil de detectar a simple vista en un objecte d'aquest tipus, no és necessari que un fractal sigui estrictament autosimilar. Les formes fractals són figures geomètriques que es caracteritzen, sobretot, per tenir una dimensió no sencera, o almenys una dimensió més gran que les geometries convencionals.

Dimensió fractal: La dimensió fractal és una extensió de la definició de dimensió convencional utilitzada per a les formes euclidianes. Una possible forma de calcular-ho és: donada una estructura autosimilar, existeix una relació entre el factor de reducció s i el nombre de rèpliques N en què l'estructura es descomposa. Aleshores tenim:

$$D=\frac{\log N}{\log \frac{1}{s}}$$

on:

—D és dimensió d'autosimilitud, encara que sovint es designa per D_s per diferenciar-la d'altres versions de la dimensió fractal.

—N és el nombre de còpies idèntiques en què es pot dividir l'estructura.

—s és el factor de reducció entre l'objecte i les còpies de primer nivell que descomponen l'estructura.

Com es pot comprovar a la figura 3, aplicant la fórmula anterior a dues formes geomètriques euclidianes, un quadrat i un cub, es pot extreure el nombre de còpies idèntiques en les quals es pot dividir la forma geomètrica particular. Un quadrat es pot dividir en N=4 còpies idèntiques i un cub es pot dividir en N=8 còpies idèntiques. Com s'observa, aquesta definició de dimensió ens dóna 2 en cas de l'estructura plana, i 3 en cas de l'estructura volumètrica.

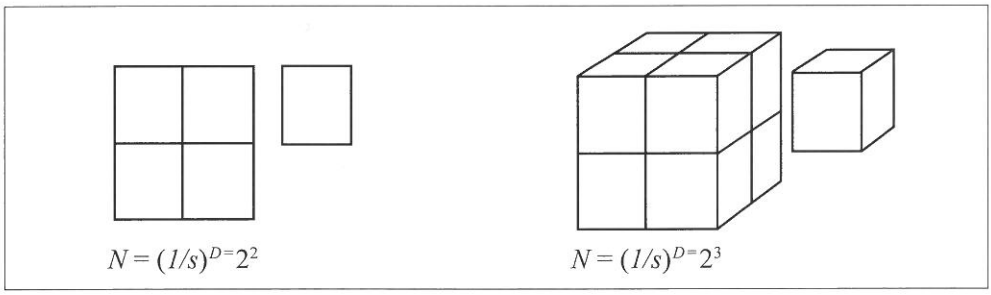

Figura 3. *Exemple d'aplicació de la fórmula de la dimensió D a dos objectes euclidians convencionals.*

En canvi, quan volem fer el mateix càlcul però ara aplicat a dues formes geomètriques fractals (vegeu figura 4), *the Sierpinski Gasket* (la qual pot ser dividida en tres còpies idèntiques amb un factor d'escala de 2) i *the Sierpinski Carpet* (la qual pot ser dividida en 8 còpies idèntiques amb un factor d'escala de 3), s'obtenen dimensions no senceres.

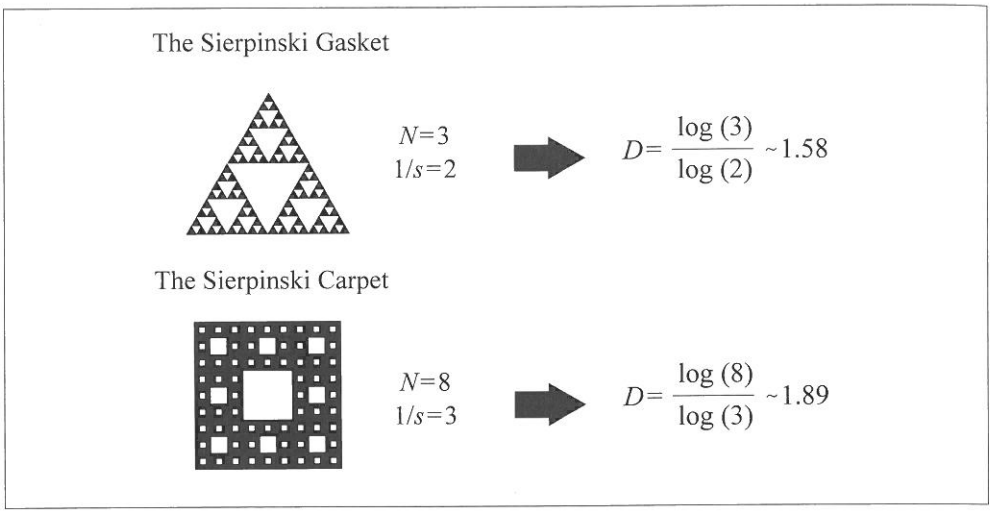

Figura 4. *Exemple d'aplicació de la fórmula de la dimensió D a formes fractals.*

Alguns exemples de formes o corbes fractals són: la corba de Koch, que data de principis del segle XX, la creu gòtica, el conjunt de Cantor, el triangle de Sierpinski o la corba de Hilbert.

Es pot observar a la figura 5 en què consisteix una forma o corba fractal. Amb una dimensió fractal de D=1.26, la corba de Koch creix per 4/3 a cada iteració fractal fins a una longitud infinita. Una corba infinita en un espai finit amb una dimensió més gran que la unitat. Això és possible pel que significa la geometria fractal i confirma que la corba de Koch, tot i ser un objecte curvilini, tendeix a «omplir» l'àrea del pla de forma més eficientment que una corba d'1D convencional.

La raó per la qual el disseny fractal d'antenes apareix com un camí atractiu per al seu disseny doble. D'una banda perquè una antena autosimilar (que conté moltes copies d'ella mateixa en escales diferents) permet operar de manera similar en diferents longituds d'ona. És a dir, que l'antena manté les característiques de radiació similars a través de les diferents bandes, com és l'exemple del monopol de Sierpinski (vegeu figures 6 i 7). D'altra banda, perquè les propietats de farciment espacial d'algunes formes basades en fractals (la dimensió fractal) permet a les antenes petites amb formes fractals aconseguir un millor aprofitament del petit espai circumdant i d'aquesta manera poden radiar més eficientment, com és el cas del monopol de Hilbert.

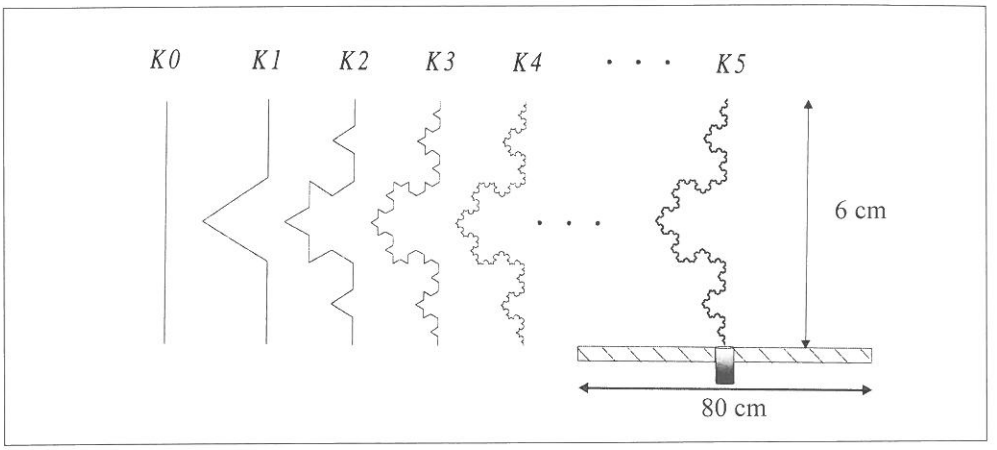

Figura 5. *Construcció de les cinc primeres iteracions de la corba fractal de Koch. La corba creix exponencialment a cada nova iteració.*

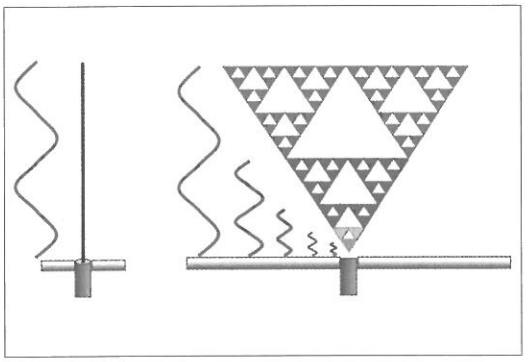

Figura 6. *Monopol tradicional i monopol de Sierpinski. Es pot observar que mentre que el monopol tradicional té una única freqüència de ressonància (o de treball), el de Sierpinski té cinc freqüències de ressonància, és a dir, opera de manera similar en cinc freqüències diferents.*

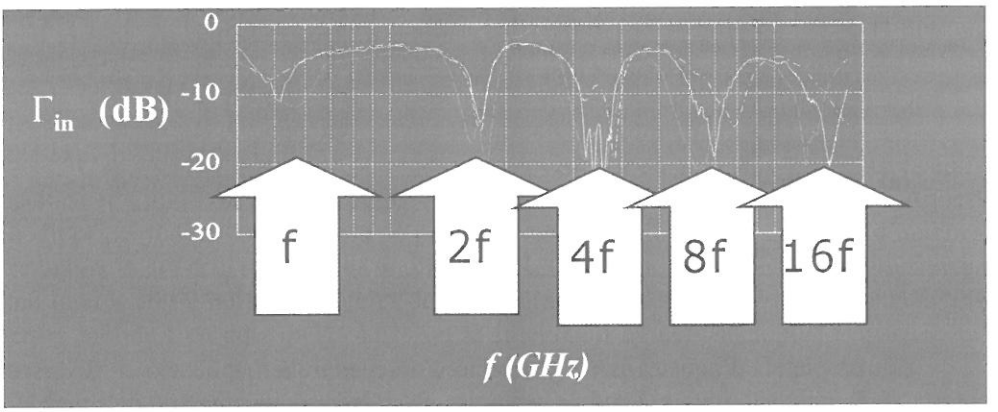

Figura 7. *Comportament multibanda del monopol de Sierpinski, tantes bandes com iteracions fractals.*

Com es pot observar a la figura anterior (fig. 7), el monopol de Sierpinski té un comportament multibanda, ja que es comporta de manera similar per a diferents freqüències, tantes freqüències com iteracions fractals. En aquest cas, la figura representa el comportament multibanda del monopol de Sierpinski de cinc iteracions fractals, fet que es pot veure a la figura 6.

FRACTUS, UNA EMPRESA DE TECNOLOGIA

Fractus és l'empresa capdavantera arreu del món en l'aplicació de la tecnologia fractal al disseny d'antenes multibanda i miniatura. Aprofitant les propietats característiques de la geometria fractal, l'empresa desenvolupa i comercialitza antenes de mida reduïda, que, a més són capaces de combinar en un únic dispositiu diversos sistemes de comunicació que operen simultàniament. L'equip d'investigadors de Fractus va ser el primer a demostrar les propietats multibanda de les antenes inspirades en formes fractals, així com el seu excel·lent comportament en entorns de dimensions reduïdes. Gràcies a les particulars propietats dels fractals, aquestes antenes poden operar simultàniament en múltiples bandes amb paràmetres radioelèctrics semblants; a més, aquestes mateixes propietats permeten reduir el pes i la mida de les antenes, minimitzant l'impacte mediambiental que provoquen avui les estructures actuals.

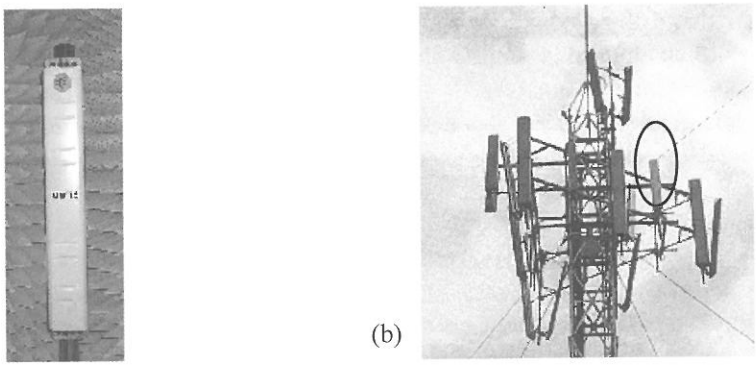

Figura 8. *(a) Antena Fractus de BTS; (b) Antena Fractus de BTS en el cel de la ciutat de Barcelona. L'antena presenta una mida reduïda, comparada amb les antenes que empren la tecnologia tradicional.*

Les propietats d'aquesta nova generació d'antenes troben aplicació en diversos entorns de la vida quotidiana. Un dels exemples clars és la racionalització dels desplegaments de les xarxes d'estacions base (BTS) per a telefonia mòbil. La tecnologia frac-

tal multibanda permet integrar en una única antena els sistemes de telefonia actuals (GSM, DCS), conjuntament amb la futura tercera generació de sistemes de comunicació mòbil (és la també anomenada tecnologia UMTS). D'aquesta manera es facilita el desplegament i l'actualització de les xarxes amb els serveis més avançats, a l'hora que es racionalitza i es minimitza l'impacte mediambiental de les instal·lacions existents. Gràcies als productes desenvolupats amb aquesta tecnologia, Fractus ha estat el principal proveïdor de la xarxa de comunicacions europea més gran que ha desplegat la primera fase del sistema UMTS.

Un altre entorn d'aplicació d'aquesta tecnologia és el sector de l'automoció. La ràdio ha deixat d'ésser l'únic sistema de comunicacions que incorporen els cotxes, principalment gràcies a la progressiva introducció en els vehicles de la telefonia mòbil i dels sistemes de radionavegació via satèl·lit (GPS). En aquest cas, la tecnologia fractal permet integrar en un sol component del vehicle els tres serveis esmentats, alhora que elimina qualsevol altre tipus d'antena de la superfície exterior del cotxe. Per introduir-se en el difícil sector de l'automoció, Fractus va establir una aliança estratègica amb una multinacional catalana del sector d'automoció FICOSA INTERNATIONAL, fruit de la qual ha estat la creació de l'empresa conjunta A^3 (Advanced Automotive Antennas, S.L.).

Figura 9. *Implantació de la tecnologia fractal per fer antenes en el sector de l'automoció. El mirall retrovisor és un dels llocs on poden ser col·locades les antenes, per la seva petita mida.*

L'aparició de les antenes fractals permet visionar un gran nombre de noves aplicacions en l'àmbit de les tecnologies de la informació i de les comunicacions. Aquest inèdit principi de disseny suposa un notable avanç en el disseny de dos tipus d'antenes tradicionalment dificultoses: les multibanda i les miniatura. Ambdós tipus són d'especial interès en l'entorn de sistemes de telecomunicació actuals, en especial en els sistemes mòbils i en aquells sistemes on la integració de serveis suposi l'aparició de noves aplicacions, com ara els sistemes d'identificació per radio freqüència (RFID) i els de miniaturització de components i circuits integrats en la indústria dels semiconductors.

Recentment, l'empresa ha estat pionera també en la introducció de la tecnologia en les antenes dels aparells de telefonia mòbil, on la necessitat de miniaturització de l'antena (obligada per la constant reducció de la mida dels aparells) i de la integració de múltiples serveis és una tendència clarament manifesta.

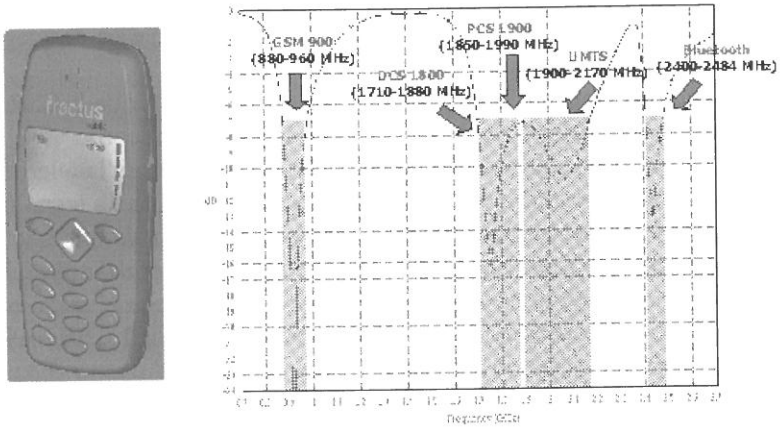

Figura 10. *Implantació de la tecnologia fractal per fer antenes per a telèfons mòbil. Es pot observar clarament com la tecnologia fractal permet tenir en una sola antena un comportament multibanda.*

Actualment la plantilla de Fractus està integrada per trenta-cinc persones, l'àmplia majoria enginyers de telecomunicació o enginyers industrials. L'empresa té tres doctors en enginyeria de telecomunicació i dos enginyers doctorands que estan fent la seva tesi doctoral a l'empresa. El contrapès al perfil necessàriament tecnològic d'aproximadament la meitat dels treballadors de l'empresa el dóna un sòlid equip amb una àmplia formació i experiència en l'àrea de la gestió i la direcció d'empreses. Segons la seva vocació de projecció internacional, l'equip humà de Fractus té un marcat caire multicultural. Així doncs, tot i que la base principal està constituïda per gent del nostre país, l'empresa té treballadors de França, Suïssa, Alemanya, els Estats Units, Costa Rica, el Regne Unit i Austràlia. La seu de l'empresa és a Sant Cugat del Vallès.

Un dels aspectes clau en la creació de l'empresa va ser l'obtenció del Gran Premi IST del desembre de 1998, atorgat per la Comissió Europea i el Consell Europeu per a la Ciència Aplicada i l'Enginyeria (EuroCASE). El premi ha estat una magnífica targeta de presentació per als diversos clients i *partners* tecnològics d'arreu del món. Això, juntament amb la vocació i l'experiència internacional del nucli de l'empresa, ha estat un factor decisiu per trencar barreres i donar des de l'inici un caire internacional al projecte, projecte que, d'altra banda, no hauria estat possible engegar-lo en un àmbit només local.

Figura 11. *Fractus és una empresa de tecnologia amb seu a Sant Cugat del Vallès.*

Des dels seus orígens, Fractus ha mantingut l'esperit d'innovació tecnològica, cercant en cada cas els diversos entorns de la vida quotidiana que millor es podien beneficiar dels avantatges de la tecnologia. D'aquesta manera, sectors tan diversos com la telefonia mòbil o l'automoció s'estan beneficiant ja avui dia d'avenços científics tan abstractes però tan propers com poden ser les matemàtiques fractals.

Genòmica computacional, una ciència interdisciplinar

Xavier Messeguer
Universitat Politècnica de Catalunya

El desenvolupament de la ciència no segueix un ritme constant sinó que més aviat el podríem imaginar com una successió d'esglaons: de cop i volta, un descobriment fonamental accelera el ritme de moltes disciplines científiques durant un curt espai de temps, fins que s'incorpora al corpus científic i el ritme torna al seu curs anterior. Sembla que la possibilitat de conèixer i manipular els genomes representa aquest descobriment fonamental que està accelerant els desenvolupaments en altres disciplines, com la física, la química, la informàtica o l'estadística.

Aquesta conferència pretén mostrar que és molt necessari el desenvolupament d'aquestes altres disciplines per a la genòmica, però primer de tot caldrà establir quin és el seu objecte d'estudi. Per aquesta raó, en primer lloc s'introduiran els conceptes biològics i informàtics fonamentals: genoma, gen, proteïna, cost... Seguidament, s'introduiran tres casos d'estudi, per mostrar com aquesta interdisciplinarietat ha permès resoldre'ls o, si més no, progressar en la seva comprensió. Concretament, parlarem de com se seqüencien els genomes, de com es pot predir l'activació dels gens i de com es poden comparar genomes.

Internet és una eina essencial avui dia en qualsevol àrea de la ciència. El nostre portal *http://www.lsi.upc.es/~alggen* n'és un exemple, perquè hi apareixen les aplicacions informàtiques que hem desenvolupat tant en l'àmbit de la recerca com en el de la docència. Relacionades amb aquesta conferència es poden consultar les aplicacions PROMO, sobre la predicció i activació dels gens, i MALGEN, sobre la comparació de genomes.

1. INTRODUCCIÓ

En aquesta secció introduirem els conceptes necessaris per poder entendre la conferència.

1.1 Genòmica

El genoma és el llibre d'instruccions que el cos humà necessita per viure i reproduir-se. Es troba dins el nucli de totes les cèl·lules del cos.

Com es llegeix el genoma? Qualsevol text segueix unes pautes de lectura, per exemple, aquest text que el lector està llegint està construït sobre l'alfabet *a,b,c...* i segueix unes pautes que permeten la seva comprensió (paraules separades per espais, frases separades per signes de puntuació...). De la mateixa manera, el genoma està constituït sobre un alfabet i té unes pautes de lectura que tot seguit explicarem.

Les lletres, les paraules i les frases

Les lletres: els cromosomes estan formats per dues cadenes de DNA entrellaçades i que cadascuna té quatre estructures bàsiques, anomenades bases, que es representen amb les lletres: *A*, *C*, *G*, *T*.

Llavors, una cadena de DNA és una tirallonga de lletres, com

...AACGTTGCTCGATTCGTATCCGTAGCTGACGCTGCATGCA...

Com que la composició de la segona cadena ve determinada per la primera, intercanviant les *A* per les *T* i les *C* per les *G*, només cal enregistrar el contingut d'una cadena, que llegirem d'esquerra a dreta.

Per donar una idea de la grandària d'un cromosoma, podríem dir que els cromosomes humans tenen la llargada de tota una enciclopèdia catalana, ara bé, de mitjana, ja que hi ha una gran variabilitat en les seves longituds. I seguint amb el paral·lelisme, un cromosoma és una enciclopèdia que no té ni fotografies, ni esquemes, ni mapes, ni signes de puntuació ni espais en blanc. El lector pot agafar un text i, amb l'ordinador, treure-li tots els signes de puntuació; observarà que la comprensió del text és més difícil perquè la divisió en paraules pot no ser única, és a dir, hi apareix l'ambigüitat. Si, a més, es canvia totes les vocals per només una, la dificultat augmenta... i això que continuem tenint moltes consonants. Doncs bé, ara es pot entendre la dificultat de llegir un text de només quatre lletres i sense cap signe de puntuació, com el genoma.

Les frases: podem equiparar les frases d'un text amb certes estructures que es troben al genoma, com es veu a la figura 1.

Aquestes estructures estan formades per dues parts: la «zona promotora» i el «gen». La zona promotora determina sota quines circumstàncies s'activa el gen, és a dir, el quan i el perquè, i el gen conté la composició de la proteïna que ha de ser creada (o sintetitzada). Efectivament, com mostra la figura 1, les frases poden estar molt allunyades les unes de les altres.

Les paraules: la figura 1 representa les paraules amb quadrets, i mostra que les paraules també estan separades entre elles. Les paraules de la zona promotora acostu-

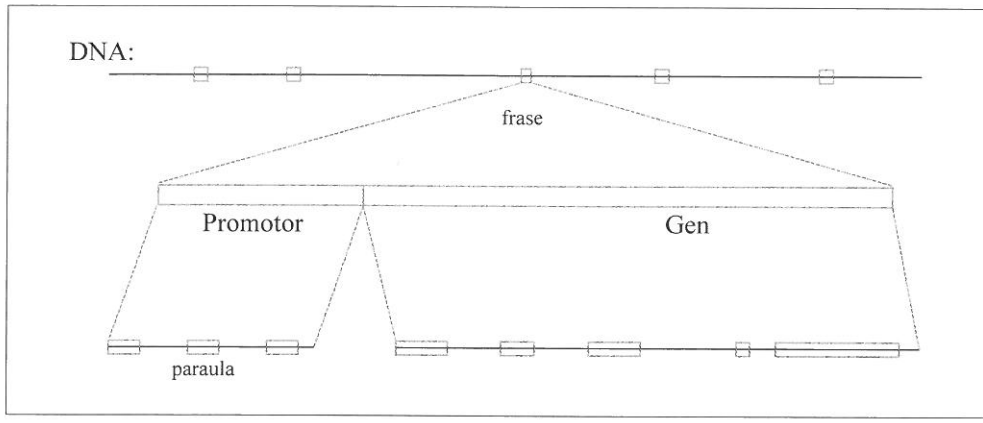

Figura 1. *Les paraules i les frases del genoma.*

men a ser molt curtes (entre 5 i 20 caràcters) i les del gen poden arribar a ser molt llargues (milers de bases).

L'activació del gen i la síntesi de les proteïnes

El procés que es desenvolupa quan el promotor s'ha activat té les tres fases que mostra la figura 2:

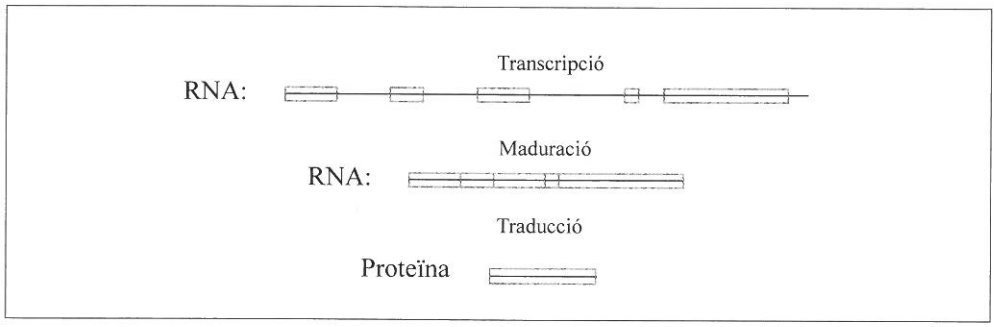

Figura 2. *Síntesi de la proteïna.*

Transcripció: es copia tot el gen en una altra cadena, anomenada RNA missatger.
Maduració: s'ajunten les paraules.
Traducció: es tradueix aquesta informació, per tal de sintetitzar la proteïna que executarà la funció per a la qual ha estat creada. Concretament, passem de l'alfabet de les bases de 4 lletres a l'alfabet dels aminoàcids de 20 lletres, i la traducció es fa de la

manera següent: cada tres bases consecutives determinen un únic aminoàcid. A mesura que la proteïna es va creant, es va plegant sobre ella mateixa i crea una estructura tridimensional.

1.2 Cost d'un algorisme

En una primera aproximació, la noció de cost temporal d'un algorisme va lligada als segons, minuts o hores que triga a executar-se. Però amb això no n'hi ha prou, perquè no ens diu com creixerà el temps si augmentem els casos que es tracten. És a dir, si per fer un tractament determinat sobre unes cadenes triguem 1 mseg, com podem tenir una idea del que trigaríem sobre un nombre de cadenes mil vegades més gran o sobre mil cadenes com la primera? Cal remarcar que aquesta pregunta té sentit, ja que un increment d'aquesta magnitud ha estat el creixement real de les necessitats de computació dels biòlegs.

Per donar resposta a aquesta pregunta, associem una funció matemàtica al cost de l'algorisme. Fonamentalment, parlem de tres menes de costos: lineal, quadràtic i exponencial. Passem a explicar-los:

— **Cost lineal:** si per tractar *m* cadenes necessiten un temps *t*, llavors per tractar *10 m* cadenes necessitarem *10 t*, i per tractar *1000 m* cadenes necessitarem *1000 t*. És a dir, d'un mil·lisegon passem a 10 mil·lisegons i a un segon, tal com ho indica el diagrama següent:

$$\begin{aligned} m &\longrightarrow & t &= 1 \text{ mseg} \\ 10\ m &\longrightarrow & 10\ t &= 10 \text{ mseg} \\ 1000\ m &\longrightarrow & 1000\ t &= 1 \text{ seg} \end{aligned}$$

— **Cost quadràtic:** si per tractar *m* cadenes es triga un temps *t*, llavors per tractar *10 m* cadenes es trigarà *100 t*... tal com ho indica el diagrama següent:

$$\begin{aligned} m &\longrightarrow & t &= 1 \text{ mseg} \\ 10\ m &\longrightarrow & 100\ t &= 100 \text{ mseg} \\ 1000\ m &\longrightarrow & 1000000\ t &= 16 \text{ min} \end{aligned}$$

— **Exponencial:** si per tractar *m* cadenes es triga un temps *t*, llavors per tractar *10 m* cadenes es trigarà 2^{10} *t*... tal com ho indica el diagrama següent:

$$\begin{aligned} m &\longrightarrow & t &= 1 \text{ mseg} \\ 10\ m &\longrightarrow & 2^{10}\ t &= 1 \text{ seg} \\ 1000\ m &\longrightarrow & 2^{1000}\ t = 10^{30}\ t &= 10^{18} \text{ anys} \end{aligned}$$

Així, es pot observar que un programa que triga 1 mseg, si multipliquem per mil les dades passa a trigar 1 seg, si el seu cost és cost lineal; 16 min, si és quadràtic i 1.000000.000000.000000 anys si té cost exponencial. Amb aquest exemple es pot veure que els programes amb cost exponencial passen a ser intractables molt ràpidament, per molta capacitat de computació que es tingui.

2. SEQÜENCIACIÓ DE LES CADENES DE DNA

Entenem per *seqüenciar una cadena* l'establiment de la successió de bases que la formen. Els mètodes informàtics que podem aplicar estan determinats per les manipulacions que són possibles al laboratori:

— Hibridació: determina quins mots d'una longitud fixa apareixen a la cadena.
— Trets: permet copiar la cadena i trencar-la de forma aleatòria.

Passem a explicar-les.

2.1 Hibridació

Amb aquest mètode, sabem quins mots d'una longitud fixa apareixen a la cadena i el que cal és enllaçar-los d'alguna forma. Per exemple, suposem que la cadena que volem conèixer és

XXXXXXXXXXXXX

i només sabem que conté els mots

AAC	CAA	GAT	TGC
ACG	CGG	GCC	TTG
ATT		GGA	

Com podem reconstruir-la?

La idea es mirar de superposar tots els mots de tal manera que les dues últimes bases d'un coincideixen amb les dues primeres del següent:

AAC
ACG
CGG
...

Aquesta mateixa idea es pot expressar amb el graf sufix-prefix que es mostra a la figura 3:

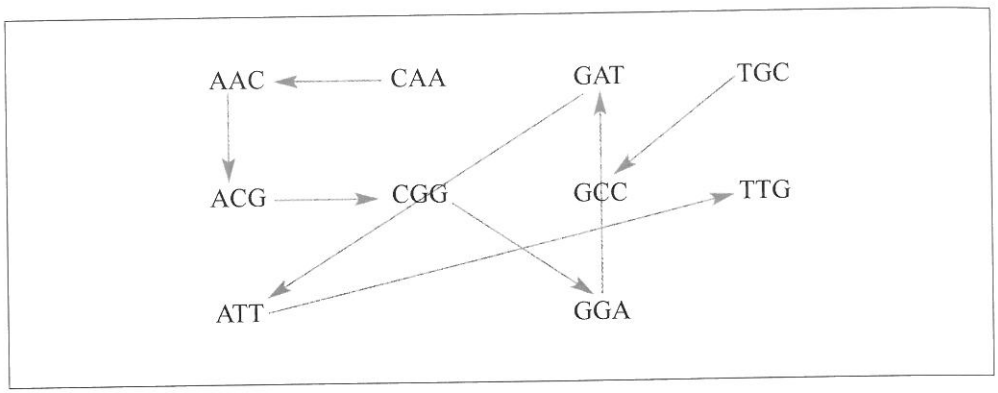

Figura 3. *Graf sufix-prefix, superposant dues bases.*

Aquest graf ens mostra l'existència d'un camí que permetrà conèixer la cadena que busquem

CAACGGATTGCC

Anem a estudiar el seu cost: tenim dos processos:

— Comparar tots els mots de dos en dos per trobar-hi les coincidències. Aquest procés té cost quadràtic.
— Buscar el camí en el graf. Aquest procés té cost lineal en el nostre exemple.

I ara anem als problemes que ens podem trobar amb aquest mètode. A l'exemple que hem proposat ha estat senzill recuperar la cadena original, però què hauria passat amb la cadena

CAACGGACGGATTGCC

que conté un mot repetit? Doncs que apareix un cicle en el graf i a l'hora de recuperar la cadena original no sabríem si es tracta de l'anterior o de qualsevol de les següents cadenes

CAACGGACGGACGGATTGCC, CAACGGACGGACGGACGGATTGCC...

És a dir, no sabem quants cops hem de recórrer el cicle. Per aquesta raó aquest mètode d'hibridació és bo si es poden evitar les repeticions, però, quina és la probabilitat que un mot es repeteixi? Aquesta qüestió comporta la definició d'un model estadístic que sigui prou proper a la realitat però que, alhora, permeti obtenir una resposta aproximada a la qüestió que hem plantejat.

Suposem que una cadena de longitud N té $N–L+1$ mots de longitud L i que cada mot ha estat generat aleatòriament, amb una distribució equiprobable de les bases. Llavors la probabilitat que entre $N–L+1$ hi hagi dos mots iguals es pot aproximar pel nombre de possibilitats que tinc d'agafar dos mots (nombre combinatori) i multiplicar-lo per la probabilitat que dos mots siguin iguals 4^{-L}. Si volem que la probabilitat de trobar dos mots iguals sigui de l'1% per $L=8$, això vol dir que $N–L+1$ ha de ser bastant petit (~30), la qual cosa indica que aquest mètode és convenient per a cadenes curtes. El valor de 8 representa el que s'utilitza al laboratori.

2.2 Trets (shot gun)

Amb aquest mètode, podem copiar la cadena tants cops com vulguem i, posteriorment, partir-la aleatòriament en trossos de diferent llargada i sense saber-ne l'ordre. Com podem aprofitar-ho? Per exemple, tenim la cadena:

xxxxxxxxxxxxxxxxxxxxxxxxxx

fem les còpies

xxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxx

i les partim en trossos a l'atzar

xxxx|xxxxxxxxx|xxxxxxx|xxxxxx
xxxxxxx|xxxxx|xxxxx|xxxxxx|xxx
xxxx|xxxxx|xxxxxxx|xxxx|xxxxxx

Tots els trossos es troben barrejats, però, si trobem parelles de mots en què el sufix de la primera coincideix amb el prefix de la segona, com en el diagrama següent:

xxxx|**xxxxxxxxx**|xxxxxxx1xxxxxx
xxxxxxx|xxxxx|xxxxx|xxxxxx|xxx

podem esperar la reconstrucció de la cadena. Suposem que en partir les tres còpies n'he obtingut els trossos:

accgt, aggt, acgatac, accttta, tttaac, gataca, accgtacc, ggt, acaggt, taacga, accg, tacctt

llavors podem construir el graf sufix-prefix de la figura 4 i recorre'l per obtenir la cadena que buscavem:

accgtacctttaacgatacaggt.

Es pot observar que els trossos que estan inclosos en un altre no apareixen al graf.

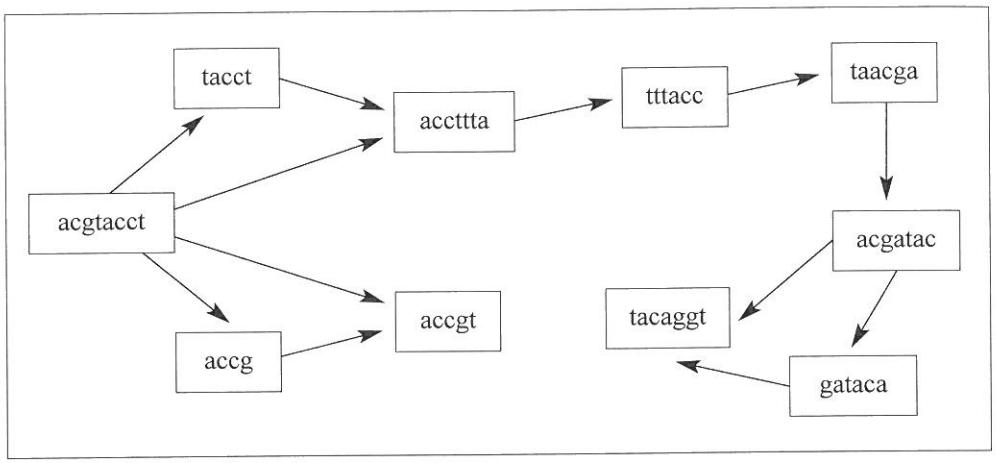

Figura 4. *Graf prefix-sufix.*

Ara estudiem el cost d'aquest mètode per trets: comparar tots els trossos dos a dos té un cost quadràtic, però buscar el camí al graf té un cost exponencial en un cas real; per aquesta raó, aquest mètode per trets només es pot fer servir en grafs no gaire grans i amb molta capacitat de computació.

Finalment, tractarem una de les moltes qüestions que queden pendents: quantes còpies és aconsellable fer de la primera cadena? És clar que com més còpies en fem, més probabilitat tindrem de trobar la cadena que busquem, però, alhora, les dimensions del graf augmenten i el cost exponencial el pot tornar intractable, així que cal buscar quin pot ser el nombre mínim de còpies per obtenir un resultat prou bo.

En aquest cas, el model teòric que es proposa és el següent: suposem que tenim M trossos de longitud d d'una cadena de longitud N escollits aleatòriament, llavors la probabilitat que una base de la cadena sigui recoberta per almenys un dels trossos té la distribu-

ció binomial $B(M,d/N)$. Si M es fa gran i d/N petita (és a dir, considerem cadenes molt llargues) però el seu producte es manté constant, llavors la distribució binomial tendeix a la distribució de Poisson $P(Md/L)$. Llavors la probabilitat que almenys un tros cobreixi cada punt és $1-e^{(NMd/L)}$, la qual cosa indica que si volem un recobriment del 99%, Md/L ha de valer 4.6. Observem que Md/L representa el nombre de vegades que recobrim la cadena original; per tant, ens dóna el nombre de còpies que cal fer de la cadena original.

3. PREDICCIÓ DE L'ACTIVACIÓ DELS GENS

Recordem de la figura 1 que la zona promotora té paraules que són curtes i estan separades. Perquè el procés de transcripció es dugui a terme s'ha d'activar una maquinària concreta, que es crea en la zona promotora i que activarà el mecanisme de transcripció. Cada peça d'aquesta maquinària és una proteïna que busca la paraula del promotor que li correspon i s'hi enganxa. Així, en el nostre exemple de la figura 1, tres proteïnes s'enganxaran a les tres paraules, altres proteïnes s'enganxaran a les tres primeres, i quan tot l'aglomerat de proteïnes estigui acoblat, s'activarà una darrera proteïna, que començarà a transcriure el gen.

El nostre interès és el següent: podem predir les bases que formen cada paraula per mètodes computacionals? Per poder-ho fer, hem de conèixer la paraula i no és un fet senzill. El problema és que en el laboratori es troben cadenes petites de DNA que se sap que contenen una paraula concreta, però per a la mateixa paraula els experimentalistes troben diferents composicions. Imaginem que diferents laboratoris han seqüenciat una zona al voltant de la paraula concreta d'una zona promotora i n'han obtingut les cadenes següents:

atgagtcaga
gtttatctctgagtcacatcagca
gctgagtcatgatgagtcatgctg
gtgactaa
gcagtcacgcttgg

Amb mètodes informàtics, cal buscar què tenen en comú aquestes cadenes i el que es fa és buscar el mot més llarg que es troba en la majoria d'elles (els * indiquen coincidències)

```
        atgagtcaga
gtttatctctgagtcacatcagca
      gctgagtcatgatgagtcatgctg
        gcagtcacgcttgg
        *****
```

Com podem observar, el mot *agtca* es troba en quatre de les cinc cadenes; per tant, suposem que és el nucli de la paraula que busquem, i que la cadena que hem exclòs era errònia. Ara bé, en la realitat pot ser que una proteïna s'hi enganxi encara que alguna base sigui diferent, llavors cal fixar-se en el fet que en la posició anterior a la primera columna de *a* hi ha tres *g* i una *c* i a la columna anterior, tres *t* i una *g*. Si en la realitat pot ser que s'enganxi en qualsevol d'aquestes paraules hem de mirar de guardar tota aquesta informació; per aquesta raó, es defineix el que s'anomena «matriu de pesos» que dóna la freqüència de cada base per a cada columna significativa. En aquest cas, obtindríem la matriu

$$\begin{array}{cc} \text{a:} \\ \text{b:} \\ \text{c:} \\ \text{d:} \end{array} \begin{pmatrix} 0 & 0 & 4 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & 0 & 4 & 0 \\ 1 & 3 & 0 & 4 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 4 & 0 & 0 \end{pmatrix}$$

on la primera fila indica el nombre de *a* en cada columna i les files següents el nombre de *c*, *g* i *t*. Observem que després de l'última *a*, apareix una columna amb *c*, *g* i *t* tan repartides que no afegeix informació útil.

Suposem que amb aquesta matriu hem predit paraules subratllades per la matriu anterior a la cadena

atcggcttgagtcacccaacgtcagtcaaaaccttccgtaaagcagtca

Observem que la primera paraula subratllada és la que coincideix millor amb la matriu, la segona ja no hi coincideix tan bé, ja que la segona lletra és una *c* en lloc d'una *g*, i la tercera és la que té pitjor coincidència. Aquest exemple mostra que cal assignar a les prediccions un valor de semblança i que hem d'acceptar com a prediccions prou bones aquelles que no s'allunyen gaire de la millor. En aquest exemple, podríem acceptar fins quatre paraules, totes aquelles en les quals es conserva el nucli i varien les dues primeres lletres.

En aquest procés, queden diverses qüestions obertes, comentarem dues:

— Fins a quin punt és significativa la matriu de pesos?

És a dir, si generem aleatòriament les cadenes inicials amb la mateixa distribució de probabilitat que les experimentals i les alineem, quin és el nombre esperat de posicions que es conserven? Experimentalment, s'ha trobat que cal esperar 4 posicions conservades en l'exemple que us hem proposat, per la qual cosa tenir-ne 5, com en el nostre cas, és estadísticament significatiu.

— Quantes prediccions cal esperar sobre una cadena de 5.000 bases?

En una primera aproximació, i acceptant que la matriu de pesos del nostre exem-

ple accepta quatre possibles paraules, sabem que de 47 paraules de longitud 7 només n'hi haurà quatre d'acceptades, i això ve a ser una de cada 4.100. Per tant, no és significatiu que prediem una paraula sobre una cadena de 5.000 bases.

4. COMPARACIÓ DE GENOMES

Primer cal explicar què entenem per comparar dues cadenes. Si una és molt més curta, s'entén com el fet de buscar aquells llocs en què la cadena petita apareix de manera semblant a la gran

```
CAATCACTCTGACGACTATCGTAGCGCGGCTACATTTCGCCAACTAATG
CA–TCTG                                  CATCT--G
**  ****                                 ***  *
```

Però si són de longitud semblant, significa intentar establir una relació un-a-un entre les bases de cada cadena, i d'entre totes les possibles relacions, cal trobar la millor.

```
CATCACTACTGACGACTATCGTAGCGCGGCTATACATCTACGCC-AA-
C-TGACTGC--ACGACTATCGT-TT-GCGGCTACACA-CTACGCACAAC
*  *****  *  **********     ******  ***  ******  **
```

Per poder comparar alineaments, cal puntuar-los, i per aquesta raó s'assigna un valor a cada columna. El valor de l'alineament és la suma dels valors de les columnes. Concretament, una columna es valora de la manera següent: si hi ha coincidència, un valor positiu; si hi ha substitució, un valor més negatiu, i si hi ha inserció/deleció (representat per un guió) un valor més negatiu.

L'algorisme que busca el millor alineament té cost quadràtic, la qual cosa vol dir que podem alinear seqüències fins a unes 10.000 bases en un temps raonable (pocs segons). Però si el que volem fer és comparar genomes de milions de bases, és a dir cadenes mil vegades més grans, el cost quadràtic significa que el temps de computació es multiplica per un milió la qual cosa el fa intractable amb els ordinadors actuals.

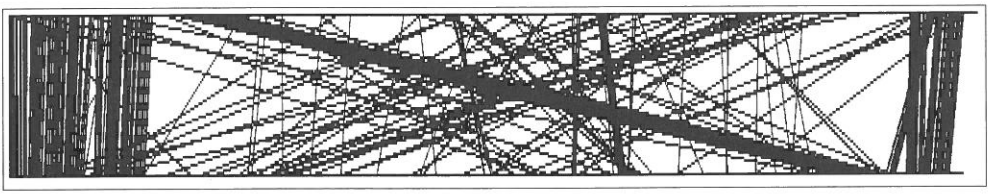

Figura 5. *Comparació de les dues bacteries* Pyrococcus abyssi de *1.790.334 bases i* Pyrococcus horikoshii *de 1.763.341 bases.*

En el cas de genomes semblants, hi ha un algorisme que permet donar una idea de la semblança entre els genomes en temps lineal. (Vegeu Figura 5).

Les línies horitzontals representen els genomes, i les línies inclinades indiquen que hi ha trossos exactament iguals de més de 20 bases consecutives. Per tant, la imatge creada suggereix que les parts inicials i finals dels dos genomes són molt semblants.

BIBLIOGRAFIA COMENTADA

El llibre clàssic d'aquesta disciplina és:
WATERMAN, M.S.: *Introduction to Computational Biology*, Chapman & Hall, 1995, London.

Ara bé, una introducció als aspectes algorísmics més entenedora i agradable es pot trobar al llibre:
SETUBAL, J. i J. MEIDANIS: *Introduction to Computational Molecular Biology*, PWS Publishing Company.

I, més formalment, des del punt de vista algorísmic es pot consultar el llibre:
PEVZNER, P.A.: *Computational Molecular Biology*, MIT press, London.

Des del punt de vista estadístic, es pot consultar:
EWWENS, W.J i G.R. GRANT: *Statistical Methods in Bioinformatics*, Springer.

I, per acabar, un darrer llibre molt complert en tots els aspectes:
MOUNT, D.W: *Bioinformatics*, Cold Spring Harbor Laboratory Press, Nova York.

Col·lecció «Aula de Ciència i Cultura»

1. Casanovas, Pompeu (ed.): *Les dimensions polítiques de la moral contemporània*, Cicle Aranguren (1).
2. Bonada, Lluís (ed.): *Cinc notes sobre Josep Pla, Homenatge a Josep Pla,* Sabadell, 1997.
3. Casanovas, Pompeu (ed.): *Globalització econòmica i experiència politicomoral*, Cicle Aranguren (2), Sabadell, 1998.
4. Solà, Joan (ed.): *L'obra de Joan Coromines*, Cicle d'Estudi i Homenatge, Sabadell, 1999.
5. Solervicens, Josep (ed.): *Miquel Batllori, historiador humanista*, Sabadell, 1998.
6. Benaul, Josep M. (ed.): *El 1898, 100 anys després.* Cicle en homenatge al Prof. Tuñón de Lara, Sabadell, 1998.
7. Brunet i Mauri, Joan (ed.): *Els mitjans de comunicació i la societat de la informació*. Cicle Joan Costa i Deu, Sabadell, 1998.
8. Creus, Maia (ed.): *El disseny contemporani. De les idees i el projecte a la pràctica professional*, Cicle Disseny (1), Sabadell, 1999.
9. Casanovas, Pompeu (ed.): *Societat civil i transformació de l'Estat al llindar del segle XXI*, Cicle Aranguren (3), Sabadell, 1999.
10. Aguadé, Jaume (ed.): *Les bases matemàtiques de la civilització tecnològica.* Cicle Ferran Sunyer i Balaguer (1), Sabadell, 1999.
11. Casanovas, Pompeu (ed.): *Filosofia del segle XX a Catalunya: mirada retrospectiva*, Cicle Aranguren (4), Sabadell, 2001.
12. Jou, David i Josep Enric Llebot (eds.): *Física i vida.* Cicle Francesc Salvà i Campillo (1), Sabadell, 2001.
13. Casacuberta, Carles (ed.): *Models matemàtics en la ciència i la societat,* Cicle Ferran Sunyer i Balaguer (2), Sabadell, 2002.
14. Bosch, Antoni (dir.): *El futur del cristianisme. El pensament teològic als inicis del segle XXI.* Cicle Ramon Llull, Sabadell, 2001.
15. Payrató, Lluís (ed.): *Les llengües a Catalunya.* Cicle Joan Coromines (En premsa).

16. CREUS, Maia (ed.): *El disseny avui. De l'objecte al seu context,* Cicle Disseny (2), Sabadell, 2002.
17. JOU, David i Josep Enric LLEBOT (eds.): *Física de la quotidianitat.* Cicle Francesc Salvà i Campillo (2), Sabadell, 2002.
18. LLEBOT, Josep Enric (ed.): *A cavall del 2000,* Sabadell, 2003
19. CASANOVAS, Pompeu i Marta POBLET (eds.): *Gestió i resolució de conflictes al segle XXI: ciutadania i violència,* Cicle Aranguren (5), Sabadell, 2003.
20. CASTELLET, Manuel: *Matemàtiques del segle XXI: dels fonaments a la tecnologia,* Cicle Ferran Sunyer i Balaguer (3), Sabadell, 2004.

Del formalisme rigorós de Hilbert a l'ús d'eines sofisticades de les matemàtiques en la genòmica no hi ha només un segle llarg de distància, sinó un canvi de mentalitat i, sobretot, el reconeixement i el convenciment que la recerca en matemàtiques, des de les àrees més bàsiques i abstractes a aquelles més fàcils de comunicar a la societat, pot jugar un paper cada cop més cabdal en el desenvolupament científic i tecnològic.

Fa més de setanta anys, un rector d'universitat a Alemanya deia que «la matemàtica és la base de tot el coneixement i el contenidor de tota l'alta cultura». Ara hi podem afegir que també és la base de l'alta tecnologia, dels grans moviments financers, de les solucions a les complexes realitats socials...

El nostre món es cada cop més interconnectat, i aquesta interconnexió augmenta el grau de complexitat. Estem, doncs, immersos en un sistema complex que, alhora, és fràgil i inestable. Només una bona eina per a la gestió d'aquests sistemes (siguin tecnològics, financers o socials) pot donar resposta als problemes plantejats. La matemàtica és una d'aquestes possibles eines, amb un paper cada cop més decisiu en la gestió dels sistemes complexos. I aquesta és la imatge que hem volgut transmetre en el Tercer Cicle Ferran Sunyer i Balaguer, el contingut del qual recull aquest volum.

CENTRE DE RECERCA MATEMÀTICA
CRM
INSTITUT D'ESTUDIS CATALANS
MCMVII
FUNDACIÓ Caixa Sabadell
ISBN 84-95166-50-X
9 788495 166500