Coppersmith's Theorem

Background, Generalizations, and Applications

Joshua Hill

Department of Mathematics
University of California, Irvine

Number Theory Seminar
2010-Oct-07 and 2010-Oct-21
http://bit.ly/CuSmith

UNIVERSITY *of* CALIFORNIA · IRVINE

Outline

Introduction Outline

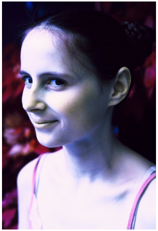
Introduction

- Presentation of the paper “Ideal forms of Coppersmith’s theorem and Guruswami-Sudan list decoding” by Henry Cohn and Nadia Heninger
- Includes significant material from “Using LLL-Reduction for Solving RSA and Factorization Problems: A Survey” by Alexander May

Henry Cohn

- is a principal researcher at Microsoft Research New England, and an affiliate professor at the MIT department of mathematics.
- is interested in discrete geometry, coding theory, cryptography, combinatorics, computational and analytic number theory, and theoretical computer science.
- is not Henri Cohen.

Nadia Heninger



(photo by Jacob Appelbaum)

- is a graduate student in computer science at Princeton University and is currently a visiting graduate student at MIT.
- is one of the authors of the very interesting paper "Lest We Remember: Cold Boot Attacks on Encryption Keys".
- presented this paper at the Crypto 2010 rump session.
- is also not Henri Cohen.

Don Coppersmith

- described this original theorem in his 1997 paper “Small solutions to polynomial equations, and low exponent RSA vulnerabilities”.
- has had a profoundly wide-ranging impact on both the theory and practice of cryptography (e.g., helped design the DES S-boxes, and was one of the designers of the AES submission MARS)

Background Outline

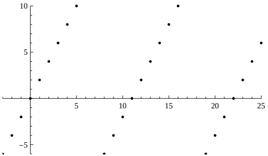
Lattice Definition

- Much of the work involves computations over lattices.
- Starting definition:

Definition

A lattice is an additive discrete subgroup of $\mathbb{R}^n$ that spans $\mathbb{R}^n$.

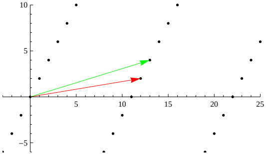
- This can be thought of as a free $\mathbb{Z}$-module.
- Example:



UNIVERSITY of CALIFORNIA · IRVINE

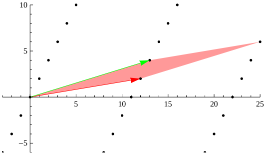
Basis of a Lattice

- If L is an n-dimensional lattice, it has n basis elements.
- There are an infinite number of possible bases
  - They are not all equally good!
  - We prefer a basis where the basis elements are as short as possible (under the ℓ^2 norm).
    - A basis made up of minimum-length vectors is called reduced.
  - We prefer a basis where the basis elements are close to orthogonal (using the standard inner product).
- Example of a (suboptimal) basis:

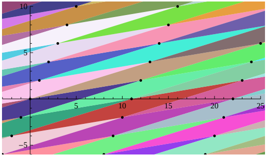


Determinant of a Lattice

- Given a basis to an n-dimensional lattice, we can arrange the basis elements into an $n \times n$ matrix, and take the determinant of that matrix.
- This can be thought of as calculating the (signed) n-volume of the fundamental parallelepiped.
  - The fundamental parallelepiped is a shape defined by the lattice that can be tiled to cover all of $\mathbb{R}^n$.



$\rightarrow$



  - The determinant of a lattice is (up to sign) independent of choice of basis.

Getting a "Good" Basis

- Sadly this is a very hard problem.
- Finding the shortest vector in a lattice (SVP) is hard.
  - Even finding a vector that is "too close" is not RP-time unless $\mathrm{RP} = \mathrm{NP}$
- All bases for a lattice have the same determinant, so if we have the shortest basis possible, its elements are "nearly orthogonal".
- A polynomial time algorithm is desired

Lenstra-Lenstra-Lovász to the Rescue

- In 1982 LLL introduced:
  - A new notion of “reduced” called “LLL-reduced”.
  - A polynomial-time algorithm that produces an LLL-reduced lattice basis from any basis.
- LLL-reduced:

Definition

Let $\{\mathbf{b}_i\}_{i=1}^n$ be a basis for the lattice L, $\{\mathbf{b}_i^*\}_{i=1}^n$ be the corresponding Gram-Schmidt orthogonal basis, and $\mu_{i,j}$ be the component of $\mathbf{b}_i$ along $\mathbf{b}_j^*$. A basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ is LLL-reduced if

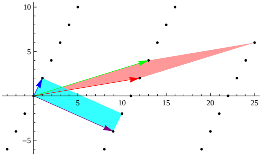
- $\left|\mu_{i,j}\right| \leq \frac{1}{2}$ for $1 \leq j < i \leq n$ and
- $\left\|\mathbf{b}_i^* + \mu_{i,i-1}\mathbf{b}_{i-1}^*\right\|^2 \geq \frac{3}{4}\left\|\mathbf{b}_{i-1}^*\right\|^2$ for $1 < i \leq n$ (The Lovász condition)

The LLL Reduced Basis: Not "Good", but "Good Enough"

- The Lovász condition assures that if two adjacent vectors are swapped prior to the Gram-Schmidt orthogonalization, the norm can't decrease too much.
- It's not really clear how this is related to the Shortest Vector Problem until you examine some consequences. The ones we need are:
  - $\|\mathbf{b}_1\| \leq 2^{\frac{n-1}{4}} \det(L)^{\frac{1}{n}}$
  - For all $\mathbf{x} \in L$ with $\mathbf{x} \neq \mathbf{0}$, $\|\mathbf{b}_1\| \leq 2^{\frac{n-1}{2}} \|\mathbf{x}\|$
- The second consequence is telling us that we "almost" have a solution to the SVP.

The LLL Algorithm: Born to Run

- The LLL algorithm takes a lattice basis, and produces a corresponding LLL-reduced lattice basis.
  - LLL runs in (worst case) $O\left(n^6 \log^3\left(\max_i \|b_i\|\right)\right)$.
  - In practice, it generally does better than this.
- A toy example:

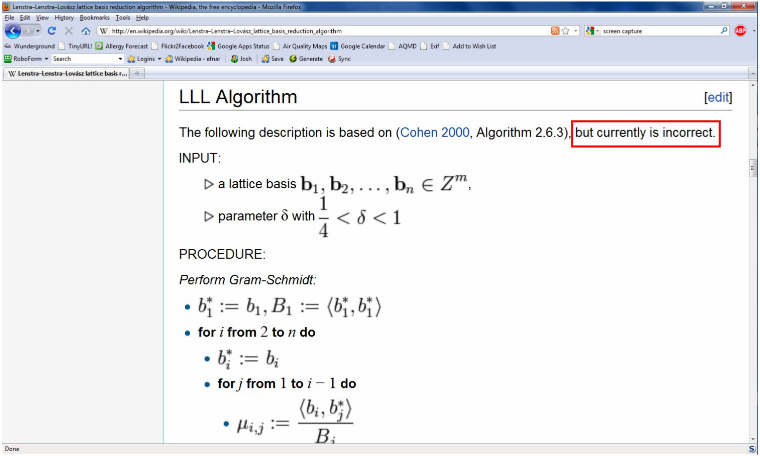


UNIVERSITY of CALIFORNIA · IRVINE

The LLL Algorithm: To Infinity (and beyond!)

- The LLL algorithm has many, many uses.
- Want more information?
  - There's always Wikipedia...

The LLL Algorithm: Wikipedia



LLL Algorithm

[edit]

The following description is based on (Cohen 2000, Algorithm 2.6.3), but currently is incorrect.

INPUT:

▷ a lattice basis $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n \in Z^m$,

▷ parameter δ with $\frac{1}{4} < \delta < 1$

PROCEDURE:

Perform Gram-Schmidt:

- $b_1^* := b_1, B_1 := \langle b_1^*, b_1^* \rangle$
- **for** i **from** 2 **to** n **do**
  - $b_i^* := b_i$
  - **for** j **from** 1 **to** $i-1$ **do**
    - $\mu_{i,j} := \dfrac{\langle b_i, b_j^* \rangle}{B_j}$

The LLL Algorithm: To Infinity (and beyond!)

- The LLL algorithm has many, many uses.
- Want more information?
  - There's always Wikipedia...
  - ...
  - Perhaps not...
- Try some good books:
  - A Course in Computational Algebraic Number Theory by Henri Cohen
  - The LLL Algorithm edited by Phong Q. Nguyen and Brigitte Vallée

Theorem Outline I

University of California · Irvine

A rephrasing of Coppersmith's original theorem:

Theorem (Coppersmith-Howgrave-Graham-May)

Let $f(x)$ be a monic polynomial of degree d with coefficients modulo an integer $N > 1$, and $\beta \in (0, 1]$. One can find all integers such that $|w| \leq N^{\frac{\beta^2}{d}}$ and $\gcd(f(w), N) \geq N^{\beta}$ in time polynomial in $\log N$ and d.

- Note that if we set $\beta = 1$, then we get all sufficiently small solutions where $f(w) \equiv 0 \pmod{N}$.

Outline of Proof I

- k is chosen to help satisfy a bounding lemma.
- If an integer B divides N and also divides $f(w)$, then B^k divides $w^j f(w)^i N^{k-i}$.
- $Q(x) = \sum_{i,j} a_{i,j} x^j f(x)^i N^{k-i} = \sum_i q_i x^i$.
- If we can get a suitable lower bound for B, we are done. Why?
  - If $|Q(w)| < N^{\beta k} \leq B^k$ and $Q(w) \equiv 0 \pmod{B^k}$ then $Q(w) = 0$.
- Find w by factoring $Q(x)$ over the integers (this is polynomial time by Berlekamp-Zassenhaus-van Hoeij)

UNIVERSITY *of* CALIFORNIA · IRVINE

Outline of Proof II

- We bound our desired roots: $|w| < X$ and apply the triangle inequality, giving us $|Q(w)| \leq \sum_i |q_i| X^i < N^{\beta k}$
- This is done by finding a suitably short vector in the the lattice generated by the coefficients of polynomials of the form $(xX)^j f(xX)^i N^{k-i}$.
- The LLL algorithm will produce just such a vector.

Applications of Coppersmith's Theorem

Coppersmith's theorem can be used to:

- Attack stereotyped messages in RSA (sending messages whose difference is less than $N^{\frac{1}{e}}$ can compromise RSA)
- Security proof of RSA-OAEP (constructive security proof).
- Affine Padding
- Polynomially related RSA messages (sending the same message to multiple recipients)
- Factoring $N = pq$ if the high bits of p are known.
- An algorithm that can get the private key for RSA in deterministic polynomial time can be used to factor N in deterministic polynomial time.
- Finding integers with a large smooth factor in a proscribed interval.
- Finding roots of modular multivariate polynomials (heuristic)

Theorem Outline II

University of California · Irvine

Impediments to Generalization

- The notion of size must be established. (here: the absolute value)
- The notion of a vector norm must be generalized. (here: the ℓ_1 norm)
- The lattice that we are working must be established (here: an integer lattice)
- The polynomial time method of extracting a suitably short vector must be established (here: LLL)
- The method of factoring the resulting polynomial must occur in polynomial time. (here: Berlekamp-Zassenhaus-van Hoeij)

Get There From Here: Polynomial Rings

- Comparison: z-degree of the polynomial.
- Vector Norm: the maximal z-degree of the polynomials in the vector (this is a non-Archimedean norm)
- Lattice: a polynomial lattice ($F[z]$ is a ring, so the lattice is a free $F[z]$-module of finite rank.)
- Finding the shortest vector is much easier in this context.
- SVP can be solved in polynomial time
  - This true for all non-Archimedean norms; the SVP reduces to solving a system of linear equations.
- Factoring bi-variate polynomials must occur in polynomial time
  - This is the case for $\mathbb{Q}$, number fields, finite fields (in RP-time)

UNIVERSITY *of* CALIFORNIA · IRVINE

The SVP in a Polynomial Lattice

- A set of basis vectors is column-reduced if the degree of the determinant of the lattice is equal to the sum of the degrees of the basis vectors.
- A set of column-reduced basis vectors always contains a shortest vector for the lattice.
- Column basis reduction of an m-dimensional lattice can be carried out in $m^{\omega+o(1)}D$ field operations, where D is the maximal degree column vector in the lattice and ω is the run time exponent for matrix multiplication. The best known exponent is for Coppersmith-Winograd ($\omega = 2.376$) and the largest reasonable value would be $\omega = 3$ (for naïve matrix multiplication).
- The analogous determinant inequality for our calculated shortest vector, v, is $\deg_z v < \frac{1}{m} \deg_z \det(L)$

University *of* California · Irvine

Proof for Polynomial Rings

This proof is very similar to the integer case!

- k is chosen to help satisfy a bounding lemma.
- If $b(z)$ divides $p(z)$ and also divides $f(w(z))$, then $b(z)^k$ divides $w(z)^j f(w(z))^i p(z)^{k-i}$.
- $Q(x) = \sum_{i,j} a_{i,j}(z) x^j f(x)^i p(z)^{k-i} = \sum_i q_i(z) x^i$.
- If we can get a suitable lower bound for $b(z)$, we are done. Why?
  - If $\deg_z Q(w(z)) < n\beta k \leq k \deg_z b(z)$ and $Q(w(z)) \equiv 0 \pmod{b(z)^k}$ then $Q(w(z)) = 0$.
  - Bound the upper degree, ℓ, of any root that we will get.
  - Construct a polynomial lattice of coefficient vectors of the form $(xz^\ell)^j f(xz^\ell)^i p(z)^{k-i}$, find the shortest vector.
  - This vector can be used to satisfy the desired bound.
- Find $w(z)$ by factoring $Q(x)$

Application: Guruswami-Sudan

- Guruswami-Sudan is an algorithm for list decoding of Reed-Solomon codes
  - Codes generally return the most likely message. In some cases there isn't a single "best message".
  - List decoding instead provides a list of likely messages, one of which is likely correct.
- Each (likely) code word is a root of a constructed polynomial. This theorem extracts these code words.
- The same error rate bounds are attained as in Guruswami-Sudan.
- Runtime is improved.
  - The (original) first stage of Guruswami-Sudan runs in $O(n^{15})$ (worst case)
  - This theorem provides a worst case bound of $O\left(n^{7.752+o(1)}d\right)$.
  - The previously best known method ran in (heuristically conjectured) time $O(n^8)$.

Theorem Outline III

Number Fields I

- A number field, K, is a finite extension of $\mathbb{Q}$
- $K = \mathbb{Q}(\alpha)$, for some α algebraic over $\mathbb{Q}$ (by the PET).
- $m_\alpha(x)$ (the minimal polynomial) is the minimal degree monic polynomial with a root at α.
  - $[K : \mathbb{Q}] = \deg m_\alpha(x) = n$
  - $m_\alpha(x) = (x - \alpha_1) \dots (x - \alpha_n)$ with $\alpha_i \in \mathbb{C}$
- $\mathbb{Q}(\alpha) = \left\{a_0 + a_1\alpha + \dots + a_{n-1}\alpha^{n-1} : a_i \in \mathbb{Q}\right\}$
- Each root corresponds to an embedding of $\mathbb{Q}(\alpha)$ into $\mathbb{C}$
  - σ_i is the map $\alpha \mapsto \alpha_i$, extended $\mathbb{Q}$-linearly.
  - If there are r_1 real roots and r_2 complex (conjugate) root pairs, $n = r_1 + 2r_2$

Size I

- With all these embedding, how do we establish a notion of size?
  - For each embedding of K into $\mathbb{C}$, we have a different "size", namely $|\gamma|_i = |\sigma_i(\gamma)|$.
  - There are $r_1 + r_2$ distinct such "sizes".
  - No one embedding is "the correct one", so we must use them all.

Algebraic Ring of Integers

- If K is analogous to $\mathbb{Q}$, what is analogous to $\mathbb{Z}$?
- $\mathbb{Z}$ has the field of quotients $\mathbb{Q}$.
  - In number fields, there can be many such subrings. Which would we choose?
- We could also look at the algebraic numbers...
  - Roots of monic polynomials with integer coefficients
- Those algebraic numbers which are in K are called the algebraic integers, denoted $\mathcal{O}_K$.
- The algebraic integers form a subring of our number field.
- $\mathcal{O}_{\mathbb{Q}} = \mathbb{Z}$.
- $\mathcal{O}_K$ is a free $\mathbb{Z}$ module of rank n (generators $\omega_1, \ldots, \omega_n$).
  - Finding such a basis is hard (See the algorithms of Zassenhaus or van Hoeij). We assume such an integral basis is known.

Size II

- Another notion of size is the norm of an element: if $\gamma \in K$ then $N(\gamma) = \prod_{i=1}^{n} \sigma_i(\gamma)$.
- In $\mathcal{O}_K$, this is especially nice: $\gamma \in \mathcal{O}_K$, $\gamma \neq 0$ then $N(\gamma) = |\mathcal{O}_K/\gamma\mathcal{O}_K|$.
- This last notion suggests the general meaning for ideals of $\mathcal{O}_K$: if I is a non-zero ideal of $\mathcal{O}_K$, then $N(I) = |\mathcal{O}_K/I|$.
- This norm is multiplicative.
- We can't ignore the absolute values. $\mathcal{O}_K$ may contain infinite units (elements of norm 1).

Lattice Analog

- We'll examine finitely generated $\mathcal{O}_K$-submodules of K^r, which we'll call Λ.
- This may not have a basis, but it will have a pseudo-basis:
  - $v_1, \ldots, v_s \in \Lambda$ and ideals $I_1, \ldots, I_s \subset \mathcal{O}_K$ so that $\Lambda = I_1 v_1 + \ldots + I_s v_s$.
- We can apply an analog of LLL (due to Fieker and Stehlé), by embedding as a $\mathbb{Z}$-lattice.
- We apply only the first portion of this algorithm, which finds a set of short module elements.

The Embedding

- First, the notion of an embedding of $\mathcal{O}_K$ into $\mathbb{R}^{r_1} \oplus \mathbb{C}^{2r_2}$

$$\sigma(\omega) = \left(\sigma_j(\omega_i)\right)_{i,j} = \begin{pmatrix} \sigma_1(\omega_1) & \sigma_2(\omega_1) & \cdots & \sigma_n(\omega_1) \\ \sigma_1(\omega_2) & \sigma_2(\omega_2) & \cdots & \sigma_n(\omega_2) \\ \vdots & & & \vdots \\ \sigma_1(\omega_n) & \sigma_2(\omega_n) & \cdots & \sigma_n(\omega_n) \end{pmatrix}$$

- Every element of $\mathcal{O}_K$ is a $\mathbb{Z}$-linear combination of these rows.
- A principal ideal (γ) embeds as $\sigma(\omega)\left(\delta_{i,j}\sigma_i(\gamma)\right)_{i,j}$
- An ideal B generated by the integral basis $b_1, \ldots, b_n$ is embedded as

$$\sigma(b) = \left(\sigma_j(b_i)\right)_{i,j}$$

- This embeds into $\mathbb{R}^n$.

UNIVERSITY of CALIFORNIA · IRVINE

A Number Field Analog to Coppersmith's Theorem

Theorem (Cohn-Heninger)

Let K be a number field of degree n with ring of integers $\mathcal{O}_K$, $f(x) \in \mathcal{O}_K[x]$ a monic polynomial of degree d, and $I \subsetneq \mathcal{O}_K$ an ideal of $\mathcal{O}_K$. For $\beta \in (0, 1]$ and $\lambda_1, \ldots, \lambda_n > 0$ we can find all $w \in \mathcal{O}_K$ with $|w|_i < \lambda_i$ such that $N\left((f(w)\mathcal{O}_K + I)\right) > N(I)^{\beta}$ provided that $\prod_i \lambda_i < N(I)^{\beta^2/d}$ in time polynomial in d, and exponential in n^2. Further, if we can bound $\prod_i \lambda_i < (2 + o(1))^{-n^2/2} N(I)^{\beta^2/d}$ then we can find all such w.

UNIVERSITY *of* CALIFORNIA · IRVINE

Get There from Here: Number Fields

- Comparison: Norm.
- Vector Norm: A ℓ_1 norm in our embedded space.
- Lattice: finitely generated $\mathcal{O}_K$-submodules of K^r.
- LLL in our embedding (first part of Fieker-Stehlé)
- Polynomials over number fields can be factored in polynomial time (Lenstra)

Proof for Number Fields

- k is chosen to help satisfy a bounding lemma.
- Generate $Q(x)$ using terms of the form $x^j f(x)^i I^{k-i}$
- We wish to bound our possible roots:
  - Bounding is with respect to all of the $r_1 + r_2$ distinct absolute values.
  - These bounds are the λ_i
- Find a suitable short vector using LLL.
- The LLL produced-short vector (mapped back) is such a bound.

Applications (Number Fields)

- Solve some instances of the bounded-distance-decoding problem in number fields.
- Generating smooth numbers over number fields (generalizing Boneh's approach)

Theorem Outline IV

Function Fields I

- A function field is a finite extension of the field $\mathbb{F}_q(x)$.
- χ is an algebraic curve over $\mathbb{F}_q$ which is smooth, projective, and irreducible over the algebraic closure of $\mathbb{F}_q$.
- $\chi(\mathbb{F}_q)$ is the set of points of χ, with coordinates in $\mathbb{F}_q$.
- K is the field of rational functions on χ defined over $\mathbb{F}_q$.
- S is a non-empty subset of $\chi(\mathbb{F}_q)$, and $\mathcal{O}_S$ is the subring of K with poles confined to S.

Function Fields II

- Every point in $\chi(\mathbb{F}_q)$ corresponds to a valuation, which produces an absolute value $|f|_p = q^{-v_p(f)}$.
- The norm of $f \in \mathcal{O}_S$ is $N(f) = \prod_{p \in S} |f|_p$.
- The Riemann-Roch space is $\mathcal{L}(D) = \{0\} \cup \{f \in K^* : (f) + D \succeq 0\}$
  - If the coefficient of $p \in D$ is k, then f can have a pole of order at most k at the point p.
  - This is a finite dimensional $\mathbb{F}_q$-vector space.
- Running time bounds rely on the ability to efficiently compute bases of the Riemann-Roch spaces for divisors of χ.
  - This works for smooth plane curves
  - This is reasonable for applications (Encoding problem for algebraic-geometric codes requires a basis for a Riemann-Roch space)

A Function Field Analog to Coppersmith's Theorem

Theorem (Cohn-Heninger)

Let χ be a smooth, projective, absolutely irreducible algebraic curve over $\mathbb{F}_q$, and let K be its function field over $\mathbb{F}_q$. Let D be a divisor on χ whose support is contained in the $\mathbb{F}_q$-rational points $\chi(\mathbb{F}_q)$, let S be a subset of $\chi(\mathbb{F}_q)$ that properly contains the support for D, let $\mathcal{O}_S$ denote the subring of K consisting of functions with poles only in S, and let $\mathcal{L}(D)$ be the Riemann-Roch space. Let $f(x) \in \mathcal{O}_S[x]$ be a monic polynomial of degree d, and let I be a proper ideal in $\mathcal{O}_S$. Then we can find all $w \in \mathcal{L}(D)$ such that $N(\gcd(f(w)\mathcal{O}_S, I)) \geq N(I)^\beta$, provided that $q^{\deg(D)} < N(I)^{\beta^2/d}$. These can be found in probabilistic polynomial time.

Application (Function Fields)

When S contains a single point, this is equivalent to Guruswami-Sudan list decoding for any algebraic-geometric code.

Summary of Talk I

- Learned a bit about Lattices
- Learned about LLL
  - The meaning of an LLL-reduced lattice basis.
  - Why LLL is useful
  - The runtime of the LLL algorithm
- Learned about Coppersmith's Theorem
  - An outline of the proof
  - Some Applications
- Learned a generalization of Coppersmith's Theorem to polynomial rings
  - An outline of the proof
  - Some Applications

Summary of Talk II

- Learned some background on Number Fields
- Introduced a number-field analog to Coppersmith's Theorem and discussed applications
- Summarized a function-field analog to Coppersmith's Theorem and discussed an application

Fin

- Questions?
- Comments? This is my first seminar presentation. Please provide any input on:
  - the level of the presentation
  - logistics and typesetting
- Presentation materials and slides are here:
  `http://bit.ly/CuSmith`
- Thanks!

Colophon

- The principal font is Evert Bloemsma's 2004 humanist san-serif font Legato. This font is designed to be exquisitely readable, and is a significant departure from the highly geometric forms that dominate most san-serif fonts. Legato was Evert Bloemsma's final font prior to his untimely death at the age of 46.
- Equations are typeset using the MathTime Professional II (MTPro2) fonts, a font package released in 2006 by the great mathematical expositor Michael Spivak.
- The serif text font in this presentation is Jean-François Porchez's wonderful 2002 Sabon Next typeface. Sabon Next is a redesign of Jan Tschichold's 1967 Sabon, which is in turn based on Claude Garamond's 16th century typefaces.
- The URLs are typeset in Luc(as) de Groot's 2005 Consolas, a monospace font with excellent readability.
- Diagrams were produced in Mathematica.

UNIVERSITY of CALIFORNIA · IRVINE