ON THE RESEARCH CONTRIBUTIONS OF HUGH C. WILLIAMS

Andrew Granville

At this special computational number theory meeting here in the delightful surroundings of Banff in the spring, we are listening to some interesting lectures while taking time to celebrate the career of one of Canada's most distinguished computational number theorists. Hugh Williams, author of almost two hundred articles, an extraordinary book, colleague and advisor to many researchers, iCORE[1] chair in algorithmic number theory and cryptography, recipient of various honours and member of several editorial and advisory boards, is sixty very soon. In this article we give a brief sampling of the many contributions of Hugh C. Williams. (**Photo 1, by the side**)

Hugh was born on July 23^{rd}, 1943 in London, Ontario. His father was a self-taught civil engineer and his mother a legal secretary. (**Photo 2, by the side**) Neither of his parents were particularly interested in mathematics and neither was the young Hugh, but he remembers at ten years of age being kept in after school to work on his mathematics, after unsatisfactory performance in class, which irked him. As a young teen he started to read popular mathematics books, by such authors as E.T. Bell and Constance Reid. But it was his Grade 10 math teacher. Mr. W. Russell, who changed Hugh's perspective on mathematics. Mr. Russell was in reality the football coach drafted in to teach the class. However by persuading Hugh to view mathematics as a big challenge to be overcome, he got Hugh to work hard. This coach introduced Hugh to Euclidean geometry, successfully encouraging him to come to grips with the difficulties. From then on Hugh was hooked on the beauty of mathematics. Hugh's general interest in mathematics was made more specific by Constance Reid's discussion of early primality testing in her book "*From Zero to Infinity*", a gift from his parents, who supported his awakening interest in mathematics. His love of the Pell equation was inspired by Carmichael's "*The Theory of Numbers and Diophantine Analysis*"

After graduating from Burlington Central High School in 1967, Hugh went on to the University of Waterloo where he completed his three degrees, marrying Lynn in the summer of 1967, and obtaining his doctorate in 1969. Hugh's interests in computational number theory were uncommon at the time in Canada, so he was pretty much on his own as a student. Having learnt about the work of Lehmer, Shanks, Brillhart and others from

[1](Alberta) Informatics Circle of Research Excellence

Reid's book, he looked south for his inspiration, and we will see how so many of the topics developed by Hugh's ideas was inspired by their brilliance.

After a brief posting at York working with Ralph Stanton, Hugh spent the next thirty one years at the University of Manitoba, where his first child, Helen, was born in March 1971, followed by Cassandra two-and-a-half years later. Upon retiring he move on to the University of Calgary and the new iCORE chair as well as becoming director of the "Centre for Information Security and Cryptography".

Hugh's first publication was in 1965, as an undergraduate, in joint work with R.A. German and C.R. Zarnke. In this paper, a harbinger of things to come, they wrote down the solution to the Cattle Problem of Archimedes[2]. In fact the solution had been known for about a century in terms of a high power of a certain algebraic integer, but no-one had written down the actual integer in decimal, as it contains over two hundred thousand digits! Following up on these old results the authors of [WGZ65] wrote down all 206, 545 decimal digits, an extraordinary accomplishment when dealing with machines that only contained 32K of memory! The trick in those days was to take information on and off of punch cards, and remarkably this gigantic calculation took just 469 C.P.U. minutes[3]. The paper is a model of tidy description, and discussion of the relevant issues for performing such a giant calculation on the then available machinery. (**Photo 3, by the side**)

At the time of writing Hugh has had 195 papers reviewed on MathSciNet. His level of output has made this biographer's task difficult and it is with reluctance that I have been forced to choose a small subset of his articles, hopefully including the most important and influential, to give the reader some idea of Hugh's career. In this task I have been helped by the recommendations of various colleagues[4]. Hugh's primary interest has been in the Mathematics of Computation: To date he has published over sixty articles in the journal of that title, and has served with great diligence as an editor for 25 years. Hugh has had more than fifty co-authors, many of them students, testament to the positive energy he puts into the younger generation: I have frequently witnessed him making the effort to congratulate a new researcher on their first conference paper and to ask for a preprint. His favorite co-author has been Richard Mollin (30) then Johannes Buchmann (14), C.R. Zarnke (10) and Dan Shanks (7); and he has written often with his postdocs and students: Renate Scheidler (7), Gordon Cormack, Gilbert Fung, Michael Jacobson, Richard Lukes (5), Gunter Dueck, Andreas Stein, A.J. Stephens (4), C.D. Patterson, Eric Seah (3) R. Holte, J.S. Judd (2), Len Baniuk, John Broere, Peter Buhr, Greg Matthew, Brian Schmid and Marsha Tennenhouse (1).

In the rest of this article I will briefly touch on several different subjects that have been visited by Hugh on various occasions, discussing some of the background and glimpsing

[2]Hugh's interest in this problem was inspired by his early reading of Beiler's, *Recreations in the Theory of Numbers: The Queen of Mathematics Entertains*".

[3]In comparison, the same (mathematical) algorithm written in Maple on my Pentium IV, without serious consideration of dealing with such enormous integers, still took 531 seconds to run.

[4]Special thanks go to Renate Scheidler, as well as John Brillhart, Johannes Buchmann, Duncan Buell, Hendrik Lenstra, Carl Pomerance, Gary Walsh and Alf van der Poorten, for their recommendations, particularly one of these who answered: "*They're all good, read all 190*". For the photographs herein I would like to thank Lynn Williams and Richard Mollin

Hugh's contributions.

Lucas' functions.

Hugh was long interested in the subject of primality testing:

"*My interest in this problem first developed on reading Constance Reid's book,* From Zero to Infinity *when I was in my early teens. In Chapter 3 she briefly described Lucas' method for determining the primality of the 39-digit* $2^{127}-1$*, and I was hooked. At the time it seemed absolutely incredible to me that a very large number like this could be established as a prime without the need for an unimaginably large number of trial divisions. I must confess that it still does. The problem of primality testing has continued to enthrall me ever since.*" — Hugh C. Williams (1998) [Wil98].

Lucas' functions are defined as follows: Given integers a, b with $(a,b)=1$ define

$$u_n = \frac{\alpha^n - \beta^n}{\alpha - \beta} \qquad \text{and} \qquad v_n = \alpha^n + \beta^n,$$

where $X^2 - aX - b = (X-\alpha)(X-\beta)$. Note that

$$\begin{aligned} u_0 = 0,\ u_1 = 1, \text{ and } u_{n+2} = au_{n+1} + bu_n \ \text{ for all } n \geq 0; \\ \text{and } \ v_0 = 2,\ v_1 = a, \text{ and } v_{n+2} = av_{n+1} + bv_n \ \text{ for all } n \geq 0; \end{aligned}$$

Well-known examples include the Fibonacci numbers $F_n = u_n(1,1)$, the Mersenne numbers $M_n = u_n(3,-2) = 2^n - 1$, the Fermat numbers $v_{2^k}(3,-2) = 2^{2^k}+1$, and a special sequence $s_k = v_{2^k}(1,1)$. Such sequences were defined and investigated before Lucas, but the reason they are called Lucas, or Lucas-Lehmer sequences, is in honour of the terrific work these authors did using these sequences. For primality testing Lucas proved the remarkable

Theorem. (Lucas) *If* $p \equiv \pm 3 \pmod{10}$ *and* p *divides* u_{p+1}*, but* p *does not divide* u_d *for* $d \leq p$ *with* $d | p+1$*, then* p *is prime.*

This remarkable criterion often allows one to test whether a number p is prime. Lucas was very interested in the primality of Mersenne numbers: His criterion involves the numbers s_k mentioned above which can be defined by $s_1 = 3$ and then $s_{k+1} = s_k^2 - 2$ for each $k \geq 1$. Lucas proved that if $n \equiv 3 \pmod 4$ and M_n divides s_{n-1} then M_n is prime, and used this to show that M_{127} is prime. One might ask how Lucas could do such a calculation with just his bare hands? First note that one only needs to compute the $s_j \pmod{M_n}$. Next note that if we multiply together two positive integers $< M_{127}$ written in binary, then by using $2^{127} \equiv 1 \pmod{M_{127}}$, we can write the multiplication down as a 127-by-127 table of 0's and 1's, and then add these numbers as in binary. This technique is beautifully described in chapter 3 of Hugh's book [Wil98], as well as the connections Lucas found with the geometry of weaving and other topics.

Primality Testing

In 1928 D.H. Lehmer gave the following primality test (based on earlier work of Pocklington and Proth): Suppose that $N-1=FR$ where F is completely factored with $F>R$ and $(F,R)=1$. If, for all q dividing F, there exists a such that $N|a^{N-1}-1$ and $\left(a^{(N-1)/q}-1,N\right)=1$, then N is prime. This test proved useful for creating tables of primes of the form $2^n+1,\ k\cdot 2^n+1,\ k\cdot q^n+1$ where $k<q^n$, and so on. Williams ([MW77, CW80, BCW81]) has used such criterion to determine primality of many such integers.

When it is difficult to factor $N-1$, it may be that $N+1$ is easy to factor. In 1975 Morrison showed that in this situation one can create a primality test: Suppose that $N+1=FR$ where F is completely factored with $F>R$ and $(F,R)=1$. In this case we cannot use the arithmetic of $(\mathbb{Z}/N\mathbb{Z})^*$ to verify the primality of N. Instead, suppose we have an integer $\triangle$ with $\left(\frac{\triangle}{N}\right)=-1$ such that for all q dividing F there exists a,b satisfying $a^2-4b=\triangle$ for which $N|u_{N+1}(a,b)$ and $\left(u_{(N+1)/q}(a,b),N\right)=1$: if so then N is prime. This test proved useful in creating tables of primes of the form $k\cdot 2^n-1,\ k\cdot q^n-1$ with $k<q^n$, 2^n3^mk-1, $(p-1)p^n-1$ and so on. Again Williams ([Wil72, 78, 81b, 87, SW00]) has used such criterion to determine primality of many such integers, and a combination of such ideas was used to show $156\cdot 5^{202}\pm 1$ are twin primes.

One practical difficulty with this plan: We all know how to obtain $a^R \pmod N$ quickly by fast exponentiation. But if this analogous method is to work we are going to need to be able to determine $u_R \pmod N$ quickly. Williams showed how this can be done in no more than double the time it takes to do fast exponentiation. As an example remember that we can compute $a^{1093} \pmod N$ rapidly by computing

$$a, a^2, a^4, a^8, a^{16}, a^{17}, a^{34}, a^{68}, a^{136}, a^{272}, a^{273}, a^{546}, a^{1092}, a^{1093} \pmod N$$

in turn; simply using the two identities $a^{2n}\equiv a^na^n \pmod N$ and $a^{2n+1}\equiv a^{2n}a \pmod N$. Thus this calculation takes 13 rather than 1093 steps! Similarly one can use the identities

$$u_{2n}=u_nv_n, \qquad u_{2n+1}=\frac{1}{2}\left(v_{2n}+au_{2n}\right),$$
$$v_{2n}=v_n^2-2b^n, \qquad v_{2n+1}=\frac{1}{2}\left(av_{2n}+\triangle u_{2n}\right),$$

when wishing to quickly compute $u_{1093} \pmod N$, or indeed any $u_R \pmod N$ or $v_R \pmod N$. Williams and Judd (1976, 77) subsequently generalized primality testing with $N\pm 1$ to other polynomials in N such as: $N^2+1,\ N^2+N+1$ and N^2-N+1.

Factoring

In the 1930s D.N. and D.H. Lehmer came up with an idea for factoring a given integer n which would work well provided one of the prime factors p of n had the property that $p-1$ only has small prime factors. They did not see this as practical in the calculating environment of the time and so did not pursue or publish it. The test went as follows: Note that if R is the least common multiple of the integers $\le B$ and if $p-1$ divides

R then $a^R \equiv 1 \pmod{p}$ for any integer a which is coprime to p; in particular p divides $\gcd(a^R - 1, n)$ and if we are lucky these are equal. To make this into a "test" we write $R = r_0 r_1 \dots r_m$ and let $a_0 = a$ with $a_{i+1} \equiv a_i^{r_i} \pmod{n}$ for $i = 0, 1, \dots, m-1$; we hope that $1 < \gcd(n, a_i - 1) < n$ for some i, giving us a factorization of n (into two parts). The Lehmers' method was rediscovered by Pollard (1974), who also added a second step if $p-1$ has just one big prime factor, making the method very practical for the computing power of that time.

In 1976 Conway and Guy [Guy76] suggested that there should be an analagous $p+1$ method of factoring using u_n and v_n. Williams [Wil82b] showed how to make this practical using Lucas sequences. He showed, somewhat surprisingly, that one can assume (without loss of generality) that $b = 1$ (in the definition of the Lucas sequences). To make this a "p+1 method" one evidently needs that $\left(\frac{\triangle}{p}\right) = -1$ which cannot be guaranteed (indeed since we don't know p without factoring n this might seem to be a serious obstacle). However Williams noted that there is a 50/50 chance that $(\triangle/p) = -1$ for each random choice of a, and thus suggested randomly choosing a. We now hope that $1 < \gcd(n, u_{r_i}) < n$ for some i or, equivalently, $1 < \gcd(n, v_{r_i} - 2) < n$. Williams generalized Pollard's second step too. He ran trials with great success on Cunningham project numbers, and Fibonacci numbers, discovering quite a few new factors.

Should one publish tables ?

Many of Williams' works have ended with tables of data, though often he restricts himself to a very accurate and detailed description of the algorithm and a precise analysis of the computed data. What to publish in an article on algorthmic and computational matters is a question of some taste (and perhaps too rarely recognized as such). Williams' articles are artful in their composition. If one wants to truly understand the matter at hand, Williams always provides sufficient detail, but nothing superfluous. In fact the question of what to include has long been an issue amongst the leaders of our field, and indeed Gauss himself responded as follows when asked why he did not publish a complete table of the data he obtained when calculating all quadratic forms of each given discriminant up to several thousand:

"*I think it quite superfluous to preserve ... it, and much more so to print it, because*

• *Anyone, after a little practice, can easily, without much expenditure of time, compute for himself a Table of any particular discriminant, if he should happen to want it ...*

• *Because the work has a certain charm of its own, so that it is a real pleasure to spend a quarter of an hour in dong it for oneself; ... and the more so because*

• *It is very seldom that there is any occasion to do it.*" — K.F. Gauss (1841).

Our subject added his own thoughts on the matter some years later:

"*It is ... the author's view ... that this kind of activity is actually a great deal of fun.*" — Hugh C. Williams (1998).

The M.I.T. cryptosystem

The most famous of all number theory cryptosystems, now known as R.S.A., was

originally known by the title of this section. Early on, many suggestions were made as to how R.S.A. might be breakable by various attacks, suggesting some care would be needed in implementing it. In "*Some remarks concerning the M.I.T. public-key cryptosystem*", Schmid and Williams [WS79] showed how to implement M.I.T. (R.S.A.) so as to avoid any such attacks. This was one of the first papers of this type, a foretaste of things to come!

The paper includes, as an aside, a nice way to find large, provably prime, primes; indeed this seems to be the first time this, by now, very standard procedure appeared. The idea is to iterate the following: Suppose we are given a prime q. What we want to do is find a "small" prime $p \equiv 1 \pmod q$, in other words of the form $p = 1 + aq$. We expect to be able to do so with a not much bigger than $\log q$, and in any case with $a < A := [2 \log^2 q]$. To do this the authors suggest to first sieve the integers $1 + aq$ with $a \le A$ using the small primes. Next to determine which of the remaining integers are also base 3 pseudoprimes. When one finds such a number $p = 1 + aq$ one is more-or-less certain that it is prime, and one can easily verify this by the method of Brillhart-Lehmer-Selfridge since prime q is a large factor of $p - 1$. Note that by iterating this algorithm $\log q$ times we quickly obtain a prime which is $> q^2$.

If one has a fast factoring algorithm then R.S.A. is evidently insecure. On the other hand it is not clear whether one can find a fast factoring algorithm if one has a fast algorithm to break any R.S.A. implementation. However it does seem to be tantalizingly close to telling us exactly that, and the question then arises to find a public-key cryptographic scheme whose security (when properly implemented) is equivalent to the difficulty of factoring. In 1980 Williams (and independently Rabin) came up with such a cryptosystem. Again Williams used Lucas sequences to do this: R.S.A. revolves around the order of $a \pmod n$, and Williams looked again (as in the "$p+1$ factoring algorithm" mentioned above) at the orders of Lucas sequences $\pmod n$ (which is, essentially, the same as looking at the order of $a + b\sqrt{d} \pmod n$, though the use of Lucas sequences allows us to only deal with integers). Williams used the idea that $a + b\sqrt{d}$ has order $p - (d/p)$ in the appropriate multiplicative group $\pmod p$, which typically equals $p + 1$ or $p - 1$, adding significant flexibility. In a certain sense this allowed him to take squareroots $\pmod{pq}$ if one can break the cryptosystem which allows him to factor pq (and thus one sees that, at heart, this involves much the same set of ideas as Rabin's well-known algorithm).

What is secure?

Other than for the sheer intellectual pleasure, why would one care that a cryptosystem is as secure as the difficulty of factoring an integer? Surely one only wants something to be so complicated that some enemy could not untangle the complications? This question is beautifully addressed by Williams and Buchmann:

"*In modern cryptography the trick is to find a hard problem that can be used as the basis of the security of the scheme. That is, your opponent should, in order to break the system, be forced to solve a problem that is widely thought to be computationally difficult. As . . . a . . . rigorous proof of the difficulty of any of these problems seems not to be forthcoming . . . we can only certify the difficulty . . . by a dubious measure:*

An admissible problem is one that has resisted solution over many years of concerted attack by very knowledgeable and skilled *[sic] practitioners.*"

— J. Buchmann and H. C. Williams (1990)

An early sieving problem – factoring

Fermat, in a letter to Frenicle in 1643, explained a factoring algorithm he created, by working with the example $n = 2027651281$. He wished to write n as the difference of two squares, hopefully with the larger square close to $\sqrt{n}$. Fermat begins by defining $r := [\sqrt{n}] = 45029$ and noting that $n - r^2 = 40440$. With this information to hand he then notes that

$(r+1)^2 - n = \left(r^2 - n\right) + (2r+1) = -40440 + 90059 = 49619.$

But $49619 \equiv 19 \mod 100$ and so is not a square. Next

$(r+2)^2 - n = \left((r+1)^2 - n\right) + (2r+3) = 49619 + 90061 = 139680.$

But $13960 \equiv 80 \mod 100$ and so is not a square. Fermat then writes that one should continue like this until one finds that $(r+12)^2 - n = 1040400 = 1020^2$ so that $n = (r+12)^2 - 1020^2 = 46061 \times 44021$. Note that at each stage one simply adds 90059, then 90061, then 90063, and so on. Thus Fermat remarks that "*instead of* 11 *additions the regular method of factoring would require division by all primes in* $[7, 44021]$". This is perhaps the first recorded comparison of the complexity of different algorithms!

In general we wish to find x such that $x^2 - n$ is a square. Thus we must have that $x^2 - n \equiv \square \mod E$ for all integers E. Like Fermat we can "narrow down" the possibilities for x by determining what residue classes it can belong to $\mod E_1, E_2, E_3, \ldots$ This is an instance of

The Generalized Sieving Problem

Given

- Bounds $A < B$
- Moduli $1 < m_1 < m_2 < \cdots < m_k$ with $(m_i, m_j) = 1$
- Sets $R_i = \{r_{i,j} : \ 0 \le r_{i,j} < m_i\}$ of acceptable residues,

determine all $x,\ A \le x < B$, for which $(x \mod m_i) \in R_i$ for all i.

This is a general formulation of a question at the heart of many important problems. Indeed computationally it has been useful for the following problems :

- Factoring
- Primality testing
- Representation of an integer by a quadratic form
- Finding " pseudosquares"
- Find integer solutions of a polynomial (in several variables)
- Quadratic polynomials generating many primes
- Determining large (small) values of $L(1, \chi)$ and thus $h\left(\sqrt{-d}\right)$

- $\sqrt{d}$ with long period (in its continued fraction).

In any particular generalized sieving problem there may be further requirements on x, such as

"Is x a square ?"; or

"Is x a prime ?", etc.

The process of checking through values of x with these additional restrictions is called *filtering*.

An important consideration is that Patterson (1991) showed in his thesis that the generalized sieving problem considered as a "decision problem", is NP-complete (in other words, it would take a miracle to have an efficient general method!).

Construction of automatic sieving devices

One of Williams' passions is the construction of devices, very very fast devices, that can work on the generalized sieving problem. This subject has a rich history, the early part of which is beautifully described in Williams book "*Édouard Lucas and primality testing*", and the latter part of which is mostly due to Williams and his collaborators at Manitoba. Let me quote Williams own description [LPW95] of the construction of such machines:

"*To see how a machine that performs this sieving operation can be fabricated, we note that each modulus m_i can be represented by* a loop *or* ring, *which is divided into m_i positions. Those positions that represent acceptable residues are tagged in some manner and the machine examines one positions from each ring at a fixed location called* a window *or* tap. *The rings are advanced in unison and a solution is detected when the window is filled exclusively with tags. A* trial counter *records the number of shifts performed by the machine. Thus the machine can be set up to start searching from some point N by setting the i^{th} ring such that $N \pmod{m_i}$ is in the window when the mechanism is turned on; after s shifts the value being tested is $N + s$. A device of this type is called* a number sieve *or, frequently, just* a sieve."

We will list the important such machines, either important for the new ideas they brought to the subject, or for their engineering. There were several machines built or designed early on:

Machine	Year	Description
F.W. Lawrence	1896	Cardboard or Wooden Rings Brass Studs for electrical contact
M. Kraitchik	Feb. 1912	"Model" in wood with working rings and gears
P. Carissan	Mar. 1912	Strips of paper, not rings

The above machines were all based on beautiful ideas that were to influence future construction but these machines themselves produced no significant results!

Machine	Year	Rings	Trials/Sec
P.& E. Carissan	1919	14	35-40
Bicycle Chain	1927	19	50
Optical Gears	1932	30	5,000
16mm Movie Film	1936	18	50
DLS-127	1966	31	1,000,000
DLS-157	1969	37	1,000,000
Shift Register	1975	42	20,000,000
UMSU	1983	32	133,000,000
OASiS	1989	16	215,000,000
MSSU	1991	30	192,000,000
Bronson, Buell	1992	22	1,024,000,000

All of these machines worked and produced worthwhile results. The main part of the history is dominated by two figures: D.H. Lehmer and his wonderfully ingenious creations, from the bicycle chain sieve of 1927 through to the shift register sieve of 1975; and H.C. Williams and his teams at Manitoba and now at Calgary. MSSU, the "Manitoba scalable sieve unit", has produced many results of note.

We will now discuss various problems in which Williams' machines have substantially pushed the boundaries.

Pseudosquares.

Let $L_p \equiv 1 \mod 8$ be the least positive non-square for which $(L_p/q) = 1$ for all odd primes $q \le p$.

• If odd $n < L_p$ then n is a square if and only if $(n/q) = 1$ for all odd primes $q \le p$. Thus if L_p grows fast we can create a fast test to determine whether a number is a square.

• (Kraitchik)-Selfridge-Weinberger [Wil78]: If $n < L_p, n \equiv 1 \mod 8$ and $q^{(n-1)/2} \equiv -1$ or $1 \pmod n$ for all primes $q \le p$ (with -1 obtained at least once), then n is a prime or prime power.

If L_p grows fast enough then this gives us a polynomial time primality test. Assuming the Riemann Hypothesis for Dirichlet L-functions we have $L_p > e^{\sqrt{p/2}}$; in fact, we expect L_p grows like $2^{\pi(p)} \approx e^{cp/\log p}$. The sieve results on this problem are as follows:

Kraitchik (1924) determined L_p for $p \le 47$, obtaining $L_{47} = 9,257,329$ for example. Lehmer and various collaborators (1928-1973) determined L_p for $p \le 61$ (bicycle chain sieve), $p \le 79$ (SWAC), $p \le 127$ (DLS-127), $p \le 151$ (DLS-157). Williams and his collaborators (1988-) determined L_p for $p \le 191$ (UMSU), $p \le 223$ (OASIS), $p \le 229$ (OASIS), $p \le 271$ (MSSU), and recently have gone beyond 300. Note that $L_{271} = 10198100582036287689 \approx 10^{19}$

The pseudosquare problem can be run in parallel. Thirty MSSU machines allowed 1.774×10^{12} integers to be sifted each second in 1996!

Large $L(1,(d/.))$ values.

Of course if $\left(\frac{d}{q}\right) = 1$ for lots of small primes q then $L(1,(d/.)) = \prod_{q \text{ prime}} \left(1 - \frac{1}{q}\left(\frac{d}{q}\right)\right)^{-1}$

is "large". If $d<0$ this leads to large $h(-d)$. Littlewood (1928) showed that the Riemann Hypothesis for Dirichlet L-functions implies that $L(1,(d/.))<\{2+o(1)\}e^{\gamma}\log\log|d|$. In 1973 Shanks became intrigued in "testing" this bound by constructing d where $L(1,(d/.))$ is large, and found that the ratio never gets as big as 2. Lukes, Patterson, Williams (1996) did vast calculations of $L(1,(d/.))/e^{\gamma}\log\log|d|$, and found that "large values" were typically $\leq$ 1.5. Recent theoretical evidence suggests that $L(1,(d/.))<e^{\gamma}(\log\log|d|+\log\log\log|d|+C)$ for some constant C, which is now being tested computationally.

Prime Producing Polynomials.

Hardy-Littlewood (1923) conjectured that, for $d=1-4A$,

$$\#\{n\leq N: n^2+n+A \text{ is prime}\}\sim C(d)N/\log N$$

where

$$C(d):=\prod_{p\geq 3}\left(1-\frac{(d/p)}{p-1}\right).$$

Therefore to have $C(d)$ large, we want $(d/p)=-1$ for all small primes p: another generalized sieving problem! Shanks noted that the given product for $C(d)$ does not converge rapidly, and so he rewrote it as

$$C(d)=\frac{\zeta(4)}{L\left(1,\left(\frac{d}{\cdot}\right)\right)L\left(2,\left(\frac{d}{\cdot}\right)\right)}\frac{1}{2}\prod_{p\mid d}\left(1-\frac{1}{p^4}\right)\prod_{q\geq 3}\left(1-\frac{2}{q(q-1)^2}\right),$$

which is much more amenable to rapid calculation (so long as one can compute $L(1,(d/.))$ rapidly, which is another speciality of both Shanks and Williams).

The most famous example of a quadratic polynomial that produces many prime values is the Euler polynomial n^2+n+41 which is prime for $0\leq n\leq 39$. (Heegner (1952), Baker and Stark (late 1960s) showed that $A=41$ is the largest A for which n^2+n+A is prime for all $0\leq n\leq A-2$.). The prime $k-$tuplets conjecture suggests that for any N there exists an integer A for which n^2+n+A is prime for $0\leq n\leq N$. However no example has yet been found with $N=40$, so the Euler polynomial is still the champion! In the nineties, Ruby found the polynomial $36n^2-810n+2753$ which is prime for $0\leq n\leq 42$.

Although the Euler polynomial is still the monic polynomial with the largest run of primes at the start, it is not the monic polynmial with the highest $C(.)$ value, and so we would expect, by the Hardy-Littlewood conjecture, that it does not have the highest "density" of primes. Indeed this is confirmed by the following examples:

$$C(-163)=3.3197732\ldots \qquad \#\{n\leq 10^6: n^2+n+41 \text{ prime}\}=261080$$
$$C(-111763)=3.631998\ldots \qquad \#\{n\leq 10^6: n^2+n+27941 \text{ prime}\}=286128$$

D.H. Lehmer, E. Lehmer, and D. Shanks (1970), gave extensive tables of d with $C(d)$ larger than any previous, and these tables have been extended by Fung and Williams (1990) who found A_1 such that $\#\{n\leq 10^6:\ n^2+n+A_1 \text{ prime}\}=361841$, and A_2 such that $C(1-4A_2)=5.0976398\ldots$; and by Jacobson and Williams (2003) who found A_3 such that $C(1-4A_3)=5.65726388\ldots$

Periodic continued fractions with long periods.

Williams (1981) suggested that the period of $\sqrt{d}$ is always $< c\sqrt{d}\log\log d$; and that d can be found with the period this large by taking $d = q$ or $2q$, with q prime, $\equiv -1 \mod 4$, and $(d/p) = -1$ for many small primes p.

Patterson and Williams (1985), and Stephens and Williams (1988) constructed lots of examples from this generalized sieving problem construction.

Special quadratic discriminants.

Mollin, Williams and I [GMW00] proved that if $-d < 0$ is a fundamental discriminant such that $\left(\frac{-d}{p}\right) \neq -1$ for all $p \leq \sqrt{d/2}$, then $d = 5, 8, 12, 13, 17, \ldots, 2044, 2244$ or 3705. The proof went as follows:

• For $d \geq 10^{18}$ one proved and used an explicit version of the Polya-Vinogradov inequality.

• For $d < 10^{18}$, this is a generalized sieving problem and was achieved on the MSSU with five months CPU time.

Rabinowicz-Mollin-Williams criteria

At the 1912 International Congress of Mathematicians Rabinowicz announced and proved the following elegant result: If $-d \equiv 1 \mod 4$ is squarefree then $x^2 + x + \frac{d+1}{4}$ is prime for $0 \leq x \leq \frac{d+1}{4} - 2$ if and only if $h(-d) = 1$.

The Heegner-Baker-Stark result, mentioned above, states that this happens only for $-d = -3, -7, -11, -19, -43, -67, -163$ (this last example giving the Euler polynomial $x^2 + x + 41$).

When $d > 0$ with $d \equiv 1 \pmod 4$ one considers the following criterion:

$$(\ddagger) \qquad -x^2 + x + \frac{d-1}{4} \quad \text{is prime for } 1 < x < \sqrt{\frac{d-1}{2}}\ .$$

One can show that $(\ddagger)$ implies $h(d) = 1$.

Mollin-Williams (1989). *Assume the Riemann Hypothesis for Dirichlet L-functions. Then* $(\ddagger)$ *holds if and only if* $d = 5, 13, 17, 21, 29, 37, 53, 77, 101, 173, 197, 293, 437$ *or* 677. *Also* $(\ddagger)$ *implies* $h(d) = 1$.

Dirichlet's class number formula tells us that if $-d < -6$ then $h(-d) = \frac{\sqrt{d}}{\pi} L\left(1, \left(\frac{-d}{\cdot}\right)\right)$; and if $d > 0$ then $h(d)\log\epsilon_d = \sqrt{d}L\left(1, \left(\frac{d}{\cdot}\right)\right)$. In general we expect that $\left|L\left(1, \left(\frac{d}{\cdot}\right)\right)\right| = (\log\log|d|)^{O(1)}$ so, essentially, $h(-d) \approx \sqrt{d}$, and thus we expect that there are finitely many d with $h(-d) = 1$ (which is nonetheless *very* hard to prove!). However $h(d)\log\epsilon_d \approx \sqrt{d}$ so feasibly $h(d)$ is often small and $R = \log\epsilon_d$ is often large: Calculations have suggested that this is usually the case, and conjectures of Gauss and of Cohen and Lenstra give more precise insight. Williams and his collaborators have verified these conjectures in many calculations. With Jacobson, they have even been able to calculate $h(d)$ and ϵ_d with d as large as 10^{100}.

In the result stated above we always have $d > 0$ with $h(d) = 1$ and ϵ_d small:

If $d = m^2 + 4$ then $\epsilon_d = \frac{m-\sqrt{d}}{2}$. In this case one easily finds that $h(d) = 1$ for $d = 5, 13, 29, 53, 173, 293$, and Yokoi conjectured that these were all.

If $d = 4m^2 + 1$ then $\epsilon_d = 2m - \sqrt{d}$. In this case one easily finds that $h(d) = 1$ for $d = 5, 17, 37, 101, 197, 677$, and Chowla conjectured that these were all.

If $d = m^2 - 4$ then $\epsilon_d = \frac{m-\sqrt{d}}{2}$. In this case one easily finds that $h(d) = 1$ for $d = 5, 21, 77, 437$.

These cases account for all the d satisfying $(\ddagger)$ according to the above result of Mollin and Williams, which was proved by using explicit lower bounds on $L\left(1, \left(\frac{\pm d}{\cdot}\right)\right)$: It is known, assuming the Riemann Hypothesis for Dirichlet L-functions that $L\left(1, \left(\frac{\pm d}{\cdot}\right)\right) \geq 1/(40,000 \log^2 d)$ for $d \geq 1001$. A result of Tatuzawa (1951), following up on famous ideas of Siegel, gives the very useful unconditional bound $L\left(1, \left(\frac{\pm d}{\cdot}\right)\right) \geq .655\eta/d^\eta$ for all $d \geq e^{1/\eta} + 75,000$ with *at most one exception* (actually we don't believe that there are any exceptions but we cannot prove that). Thus, with enough computation Mollin and Williams used this to show $(\ddagger)$ implies $h(d) = 1$ with "at most one exception".

Rather surprisingly there have been recent significant developments in this area without deep analytic or computational tools:

Biro (2003). *The Chowla and Yokoi conjectures are true.*

The techniques described above have allowed Mollin and Williams to prove many results about $h(d)$ being small when ϵ_d is small, and to provide intriguing connections to prime producing polynomials. As one final example:

Louboutin, Mollin, Williams (1993). *If $d = m^2 + r$ where $r|4m$ and if the square of every ideal in $\mathbb{Q}\left(\sqrt{d}\right)$ is principal then d is from a given list of 227 possibilities. If GRH is true this list is complete. If not we are missing at most one value of d.*

Mollin and Williams have an analogous 1992 result for *all* d with $\epsilon_d < 2d$.

Arithmetic in $\mathbb{Q}(\sqrt{d})$ where d is a fundamental discriminant

We will quickly set up some notation (copied from [BW90]). The ring of integers is $[1, \omega]$ where

$$\omega = \begin{cases} \sqrt{d} & \text{if } d \equiv 2, 3 \pmod 4; \\ (1 + \sqrt{d})/2 & \text{if } d \equiv 1 \pmod 4. \end{cases}$$

A *primitive* ideal takes the form $[a, b + \omega]$ where a divides $N(b + \omega)$. A *reduced* ideal is primitive, and either $|\alpha| \geq a$ or $|\overline{\alpha}| \geq a$ for all $\alpha \in I$.

Assume $d > 0$, $d \equiv 1 \pmod 4$. Let $p_0 = 2b + 1$ and $q_0 = 2a$. For each $i > 0$ let $a_i = \lfloor (p_i + \sqrt{d})/q_i \rfloor$ and then $p_{i+1} = a_i q_i - p_i$ and $q_{i+1} = (d - p_{i+1}^2)/q_i$. Let $I_{j+1} = [q_j/2, (p_j + \sqrt{d})/2]$.

The ideals $I_0 = [a, b + \omega], I_1, I_2, \ldots$ give the successive ideals of *Gauss's reduction algorithm*. Note that this sequence is eventually periodic, say with $I_{p+j} = I_j$ for all $j \geq 1$. (There is a quadratic form that "corresponds" to I_j, namely $q_j x^2 + 2p_j xy - q_{j-1} y^2$).

In fact

$$I_j = (\Psi_j)\, I_{j+1} \quad \text{where} \quad \Psi_j := \frac{p_j + \sqrt{d}}{q_j}.$$

Let

$$\theta_n = \Psi_1 \Psi_2 \dots \Psi_{n-1} = (u_{n-2} + v_{n-2}\sqrt{d})/2a$$

where

$$u_n/v_n = [a_0, a_1, \dots, a_n];$$

then θ_{p+1} is the fundamental unit of $\mathbb{Q}(\sqrt{d})$.

The infrastructure. For $1 \le i < j \le p+1$ define the distance

$$\delta(I_i, I_j) = \log(\theta_j/\theta_i) = \sum_{l=i}^{j-1} \log \psi_l$$

Note that $\delta(I_1, I_{p+1}) = \log \theta_{p+1} = \log \epsilon_d$, which is the regulator, R.

If $I_0 = (1) = [1, \omega]$ and $1 \le i, j \le p$ then we can " multiply" I_i by I_j with the composition algorithm, and then reduce (in polynomial time) to obtain some ideal I_k for which

$$\delta(I_k, I_1) \approx \delta(I_i, I_1) + \delta(I_j, I_1)$$

with error bounded in absolute value by $\log d$. This leads to a fast algorithm for moving around the ideals in what may be a very large $\left(\approx \sqrt{d}\right)$ cycle of reduced forms.

Picture 4: I would like a picture here of the cycle of forms.

In 1972 Shanks introduced this new idea, for real quadratic fields. His write-up lacked details, which were provided in full by Lenstra (1980), who gave a thorough analysis. The infrastructure concept (and its use algorithmically) was generalized to all number fields with unit rank 1 by Buchmann and Williams (1988), something which has been used by several authors recently.

Shanks also introduced the "baby steps-giant steps" algorithm. In essence the idea is as follows: Given a cyclic group G of order n generated by an element g, and another element h of G, we wish to determine r such that $g^r = h$. One idea is to simply search through the last $g^1, g^2, g^3, \dots$ until we get h; one would expect this to take around n steps. Shanks' idea is to write down the two lists of numbers $h, hg^{-1}, hg^{-2}, hg^{-3}, \dots, hg^{-(m-1)}$, and $1, g^m, g^{2m}, \dots g^{\ell m}$ where $\ell = [n/m]$, and then compare the two lists finding an element in common; then $hg^{-i} = g^{jm}$ so we can take $r = i + jm$. If $m \approx \sqrt{n}$ then this algorithm will take around $n^{1/2}$ steps, a massive saving. Suitably modified (since the cycle of reduced forms is not a group), this algorithm can be used on the cycle of reduced ideals, using the infrastructure to construct the necessary lists of ideals.

Exploiting "baby steps-giant steps" and "infrastructure".

One of Williams' primary interests has been in exploiting the infrastructure to obtain fast algorithms for many problems that use the algebra of real quadratic fields. **Picture 5: "Dan Shanks tells Hugh to "pick up the infrastructure baton and run with it".**

Here are a few applications:

Stephens and Williams (1989). *$O\left(D^{1/4+\epsilon}\right)$ algorithm for Eisenstein's 1844 problem : Given fundamental $D \equiv 5 \pmod 8$ determine whether there exists a solution in integers x, y to $x^2 - Dy^2 = 4$ with $(x, y) = 1$.*

Buchmann and Williams (1988). *Given an ideal I of $Q(\sqrt{d})$ determine whether it is principal. Algorithm in time $O\left(\log N(I) + \sqrt{R}\, d^{o(1)}\right)$.*

Buchmann and Williams (1990). *Came up with a beautiful "one-way function" based on determining ideals in the cycle of reduced ideals in the principal class. Can be used for logins for example (instead of a hash function).*

Indeed they showed that if you can invert this one-way function, you can factor d!

Buchmann, Thiel and Williams (1995). *Gave a short representation of elements of a real quadratic order in terms of infrastructure.*

This shows that the question of whether a given ideal I is principal is in complexity class NP.

The Ankeny-Artin-Chowla conjecture. *If prime $p \equiv 1 \mod 4$ and $\left(u + v\sqrt{p}\right)/2$ is the fundamental unit of $\mathbb{Q}\left(\sqrt{p}\right)$, then p does not divide v.*

Calculations on this have a long history:

For all $p \leq$	By	In
2000	Ankeny, Artin, Chowla	1952
10^5	Goldberg	1954
6.25×10^6	Beach, Zarnke and Williams	1971
10^8	Soleng	1986
10^9	Stephens, Williams.	1988 (in fact all $D \leq 10^9$)
10^{11}	te Riele, Van de Poorten, Williams	108971

Ankeny, Artin, and Chowla came up with a delightful formula to determine whether v is divisible by p:

$$2hv/u \equiv \frac{1}{p}\left(\prod_{\substack{1\le r\le p-1\\ \left(\frac{r}{p}\right)=1}} r \;-\; \prod_{\substack{1\le n\le p-1\\ \left(\frac{n}{p}\right)=-1}} n\right) \mod p,$$

though this is impractical in a modern computing environment.

Mollin and Walsh (1986) noted that Erdös' conjecture, that there are no three consecutive powerful numbers, is relevant: For if $x-1 = a^3b^2$, x and $x+1 = A^3B^2$ are all powerful then $x^2 - Dy^2 = 1$ where $D = aA$ divides $y = aAbB$. They observed that there are composite D for which D divides y in the fundamental unit $x + \sqrt{D}y$ (of the order $[1, \sqrt{D}]$), such as $D = 2 \times 23$ and $2 \times 5 \times 43$. Stephens and Williams found all of the other such $D \leq 10^7$, namely 23×79, $2 \times 3 \times 7 \times 19 \times 73$, $3 \times 69997, 41 \times 79 \times 541, 2 \times 1562159, 3 \times 5 \times 273281$.

The algorithm used by Williams and his collaborators uses infrastructure in the cycle of reduced forms in the principal class, and the *baby step-giant steps algorithm*, to determine a multiple of $v \mod p$ in time $D^{1/4+o(1)}$.

Key Exchange Protocols.

Diffie and Hellman's (1976) famous protocol runs as follows:
In $\mathbb{F}_q$ generated by g:

Alf selects x at random and transmits g^x to Renate;

Renate selects y at random and transmits g^y to Alf;

Both Alf and Renate can determine $g^{xy} = (g^x)^y = (g^y)^x$, the secret key, but not eavesdropping Gary, who, even with Canada's best at his disposal, has still not figured out how to determine z from g and g^z.

We still do not know how secure this really is — the discrete log problem has resisted attack to date and so might satisfy the criteria given by Buchmann and Williams (discussed above). However it certainly has not withstood the same scrutiny as more intrinsic mathematical problems.

In 1988, in the first issue of the *Journal of Cryptology*, McCurley gave an analagous key exchange protocol in $\mathbb{Z}/n\mathbb{Z}$ which is as hard to break as factoring n. In the same issue Buchmann and Williams gave an analagous key exchange protocol in the class group of an imaginary quadratic field, which is more complicated than Diffie-Hellman but probably more secure. With Düllman they subsequently implemented this, giving also a new, faster reduction algorithm.

Buchmann, Scheidler and Williams (1994, also *J. Cryptology*) gave an analagous key exchange protocol in the infrastructure of the cycle of reduced ideals in the principal class. This is not a group, but it is close when using the $\delta(.,.)$ metric (described above). The paper contains a beautiful analysis and discussion of all relevant practical and theoretical issues.

The next chapter

We have reviewed some of the key accomplishments in the career of Hugh Williams, so far. There are several main themes in his papers that we have discussed revolving around: Lucas sequences, building sieving machines, and the use of the infrastructure of $\mathbb{Q}(\sqrt{d})$.

On the other hand, I have not done justice to his work on computing class numbers and regulators of quadratic fields, a constant theme throughout his career, instead focusing on applications. Also I have not discussed his important contributions to the use of Morrison

and Brillhart's *continued fraction factoring algorithm*, and in particular his insightful 1987 attempt, with Wunderlich, to parallellize this algorithm. And there are more...

All these works develop concepts that have become increasingly important in our subject in part due to Hugh's influence and they are being used more widely today than ever before, by his students, by his collaborators and by others. There were no greater influences on Hugh, intellectually, than Dick Lehmer and Dan Shanks, and through him we see their ideas proliferating to the modern day.

Hugh has recently moved on, to a new job in a new city, with a different focus, and very recently as a new grandfather. We hope his forthcoming years will be as productive, influential and interesting as the ones to date. **Picture 6: Hugh with Alexa Jayne**

Département de Mathématiques et statistique, Université de Montréal, CP 6128 succ. Centre-Ville, Montréal QC H3C 3J7, Canada

E-mail address: andrew@dms.umontreal.ca