INDOCRYPT 2010, 13 December 2010

Getting a Few Things Right and Many Things Wrong

Neal Koblitz, Univ. of Washington, koblitz@math.washington.edu

Outline of Talk

Outline of Talk

- Isogenies – common wisdom might (or might not) be wrong
- Survey of my own experiences being wrong most of the time
- History of embarrassing moments in “provable security”
- What is to be done? -- the keyword that unlocks the solution to this conundrum

Reference for the first part of the talk:

"ECC: The Serpentine Course of a Paradigm Shift"
by Ann Hibner Koblitz, N.K., & Alfred Menezes,
to appear in the ECC issue of J. Number Theory.

Reference for the first part of the talk:

"ECC: The Serpentine Course of a Paradigm Shift"
by Ann Hibner Koblitz, N.K., & Alfred Menezes,
to appear in the ECC issue of J. Number Theory.

In the meantime it's available at
http://eprint.iacr.org/2008/390.pdf

Reference for the first part of the talk:

"ECC: The Serpentine Course of a Paradigm Shift"
by Ann Hibner Koblitz, N.K., & Alfred Menezes,
to appear in the ECC issue of J. Number Theory.

In the meantime it's available at
http://eprint.iacr.org/2008/390.pdf

See especially Section 11 concerning the
security implications of isogeny walks.

Reference for the first part of the talk:

"ECC: The Serpentine Course of a Paradigm Shift"
by Ann Hibner Koblitz, N.K., & Alfred Menezes,
to appear in the ECC issue of J. Number Theory.

In the meantime it's available at
http://eprint.iacr.org/2008/390.pdf

See especially Section 11 concerning the
security implications of isogeny walks.

Also see our videoabstract at
http://www.youtube.com
(search for "elliptic curve cryptography")

But I also recommend Section 13, which contains discussions of such topics as

But I also recommend Section 13, which contains discussions of such topics as

- The family life of gorillas

But I also recommend Section 13, which contains discussions of such topics as

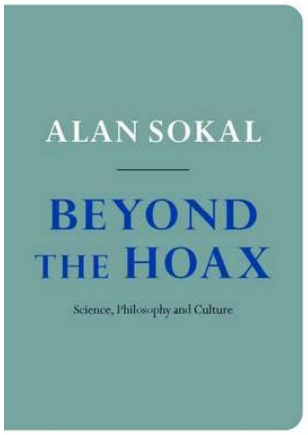
- The family life of gorillas

- Smart houses and robotic butlers

- Controversy over whether or not Arizona's ancient Native American tribes were warlike

- Controversy over whether or not Arizona's ancient Native American tribes were warlike

- The impact of Alan Sokal's "hermeneutics of quantum gravity"



Conventional wisdom

Conventional wisdom

- In cryptography, for greatest security choose parameters as randomly as possible.

Conventional wisdom

- In cryptography, for greatest security choose parameters as randomly as possible.
- In elliptic/hyperelliptic curve cryptography it's safest to choose the defining equation to have random coefficients.

Conventional wisdom

- In cryptography, for greatest security choose parameters as randomly as possible.
- In elliptic/hyperelliptic curve cryptography it's safest to choose the defining equation to have random coefficients.
- It's okay to use special curves for reasons of efficiency if you insist, but some day that choice might come back to bite you.

In 1991, I proposed the use of the non-supersingular $\mathbf{F}_2$ -curves (also called anomalous binary curves)

$y^2 + xy = x^3 + 1$ or $y^2 + xy = x^3 + x^2 + 1$

In 1991, I proposed the use of the non-supersingular $\mathbf{F}_2$ -curves (also called anomalous binary curves)

$$y^2 + xy = x^3 + 1 \quad \text{or} \quad y^2 + xy = x^3 + x^2 + 1$$

because they seemed to have some efficiency advantages over random curves.

In 1991, I proposed the use of the non-supersingular $\mathbf{F}_2$ -curves (also called anomalous binary curves)

$$y^2 + xy = x^3 + 1 \quad \text{or} \quad y^2 + xy = x^3 + x^2 + 1$$

because they seemed to have some efficiency advantages over random curves.

The U.S. National Security Agency (NSA) liked these curves, and at Crypto 1997 J. Solinas gave a talk presenting a thorough and definitive treatment of how to optimize ECC operations on these curves.

At present these curves are one of the three sets of curves recommended by the U.S. National Institute of Standards and Technology (NIST).

At present these curves are one of the three sets of curves recommended by the U.S. National Institute of Standards and Technology (NIST).

(Each set contains 5 curves over different-size fields for different security levels.)

At present these curves are one of the three sets of curves recommended by the U.S. National Institute of Standards and Technology (NIST).

(Each set contains 5 curves over different-size fields for different security levels.)

Some people have been mistrustful of this family of curves, in part because of the "conventional wisdom" given above.

Another reason for mistrust is the syllogism:

Another reason for mistrust is the syllogism:

"NSA wants us to use these curves.

Another reason for mistrust is the syllogism:

"NSA wants us to use these curves.

"We don't trust the NSA.

Another reason for mistrust is the syllogism:

"NSA wants us to use these curves.

"We don't trust the NSA.

Another reason for mistrust is the syllogism:

"NSA wants us to use these curves.

"We don't trust the NSA.

"Therefore we don't trust these curves."

My analysis is: NSA isn’t a monolith.

My analysis is: NSA isn't a monolith.

Just because an NSA mathematician recommends a family of curves, it doesn't necessarily follow that they're no good.

WHITE HATS

BLACK HATS

N.S.A.

Mathematicians & Engineers

high intelligence, high ethics

Spies, Spooks & Bureaucrats

low intelligence, low ethics

WHITE HATS

BLACK HATS

Corporate world

Researchers & Engineers

high intelligence, high ethics

Marketing People & Executives

low intelligence, low ethics

WHITE HATS

BLACK HATS

University world

Professors

Administrators

high intelligence, high ethics

low intelligence, low ethics

But please note:

But please note:

The last slide doesn’t apply to any university administrators who are present today.

But please note:

The last slide doesn't apply to any university administrators who are present today.

Actually, the empirical basis for that slide comes from the U.S., not from India.

Moreover, in the random-vs-special debate about curve selection, Menezes and I found reason to question the conventional wisdom that random is always more secure.

Moreover, in the random-vs-special debate about curve selection, Menezes and I found reason to question the conventional wisdom that random is always more secure.

There are various scenarios in which someone who chooses ECC with a special curve might end up better off than someone else who chooses a random curve.

Moreover, in the random-vs-special debate about curve selection, Menezes and I found reason to question the conventional wisdom that random is always more secure.

There are various scenarios in which someone who chooses ECC with a special curve might end up better off than someone else who chooses a random curve.

Some such scenarios are suggested by recent work on isogenies.

Moreover, in the random-vs-special debate about curve selection, Menezes and I found reason to question the conventional wisdom that random is always more secure.

There are various scenarios in which someone who chooses ECC with a special curve might end up better off than someone else who chooses a random curve.

Some such scenarios are suggested by recent work on isogenies. (For more details see Section 11 of the “serpentine course” paper.)

Isogenies

Isogenies

By an “isogeny” we mean an algebraic map between two (usually non-isomorphic) elliptic curves over $\mathbf{F}_q$;

Isogenies

By an “isogeny” we mean an algebraic map between two (usually non-isomorphic) elliptic curves over $\mathbf{F}_q$; by Tate’s theorem the two isogenous curves must have the same number of $\mathbf{F}_q$-points,

Isogenies

By an “isogeny” we mean an algebraic map between two (usually non-isomorphic) elliptic curves over $\mathbf{F}_q$; by Tate’s theorem the two isogenous curves must have the same number of $\mathbf{F}_q$-points, and, conversely, curves with the same number of points are isogenous.

Isogenies

By an “isogeny” we mean an algebraic map between two (usually non-isomorphic) elliptic curves over $\mathbf{F}_q$; by Tate’s theorem the two isogenous curves must have the same number of $\mathbf{F}_q$-points, and, conversely, curves with the same number of points are isogenous.

(An isogeny between isomorphic curves is called an “endomorphism.”)

The “isogeny class” of a curve E consists of all the curves over $\mathbf{F}_q$ with the same number of points.

The “isogeny class” of a curve E consists of all the curves over $\mathbf{F}_q$ with the same number of points.

By constructing isogenies from E to other curves in its isogeny class,

The “isogeny class” of a curve E consists of all the curves over $\mathbf{F}_q$ with the same number of points.

By constructing isogenies from E to other curves in its isogeny class, we can transport its discrete log problem to equivalent discrete log problems on the other curves.

The elliptic curves isogenous to a given E can be further partitioned into “endomorphism classes.”

The elliptic curves isogenous to a given E can be further partitioned into “endomorphism classes.” The endomorphism class of E consists of all curves in its isogeny class with the same endomorphism ring.

The elliptic curves isogenous to a given E can be further partitioned into "endomorphism classes." The endomorphism class of E consists of all curves in its isogeny class with the same endomorphism ring.

The endomorphism ring is an order in the ring of integers of the CM-field of E.

The elliptic curves isogenous to a given E can be further partitioned into "endomorphism classes." The endomorphism class of E consists of all curves in its isogeny class with the same endomorphism ring.

The endomorphism ring is an order in the ring of integers of the CM-field of E.

Within the isogeny class of E, how widely different in size can two endomorphism rings be?

That depends on the square part of the discriminant of E.

That depends on the square part of the discriminant of E.

The index of the endomorphism ring in the ring of integers of the CM-field is a divisor of the square part of the discriminant.

That depends on the square part of the discriminant of E.

The index of the endomorphism ring in the ring of integers of the CM-field is a divisor of the square part of the discriminant.

A random curve is almost certain to have a discriminant with very small square part.

For example, the high-security NIST random binary curve B-571 happens to have squarefree discriminant.

For example, the high-security NIST random binary curve B-571 happens to have squarefree discriminant.

By a result of Jao, Miller, and Venkatesan, one can use sequences of isogenies ("isogeny walks") to efficiently travel randomly and uniformly throughout the isogeny class of B-571,

For example, the high-security NIST random binary curve B-571 happens to have squarefree discriminant.

By a result of Jao, Miller, and Venkatesan, one can use sequences of isogenies ("isogeny walks") to efficiently travel randomly and uniformly throughout the isogeny class of B-571, which consists of approximately 2^{285} curves.

In contrast, NIST's anomalous binary curve for the same security level, denoted K-571, is very different.

In contrast, NIST's anomalous binary curve for the same security level, denoted K-571, is very different.

It has discriminant $-7(P_{22}P_{263})^2$ divisible by the square of a 22-bit prime times a 263-bit prime.

In contrast, NIST's anomalous binary curve for the same security level, denoted K-571, is very different.

It has discriminant $-7(P_{22}P_{263})^2$ divisible by the square of a 22-bit prime times a 263-bit prime.

Most of the isogenous curves have very small endomorphism ring with index divisible by the 263-bit prime,

In contrast, NIST's anomalous binary curve for the same security level, denoted K-571, is very different.

It has discriminant $-7(P_{22}P_{263})^2$ divisible by the square of a 22-bit prime times a 263-bit prime.

Most of the isogenous curves have very small endomorphism ring with index divisible by the 263-bit prime, and no one knows how to construct such a large degree isogeny from K-571 to such a curve.

At the very most, one can map K-571 to the curves whose endomorphism ring has index divisible at most by the 22-bit prime.

At the very most, one can map K-571 to the curves whose endomorphism ring has index divisible at most by the 22-bit prime.

There are approximately 2^{22} – that is, 40 lakhs – such curves.

Now let’s make a hypothetical assumption.

Now let's make a hypothetical assumption.

Let's suppose that an algorithm were found that solves the elliptic curve discrete log problem (ECDLP) in time T_1

Now let's make a hypothetical assumption.

Let's suppose that an algorithm were found that solves the elliptic curve discrete log problem (ECDLP) in time T_1 in a proportion ε (a very small but not negligible proportion) of all elliptic curves over $\mathbf{F}_q$,

Now let's make a hypothetical assumption.

Let's suppose that an algorithm were found that solves the elliptic curve discrete log problem (ECDLP) in time T_1 in a proportion ε (a very small but not negligible proportion) of all elliptic curves over $\mathbf{F}_q$, where the property of being a "weak" curve is independent of isogeny and endomorphism class.

Now let's make a hypothetical assumption.

Let's suppose that an algorithm were found that solves the elliptic curve discrete log problem (ECDLP) in time T_1 in a proportion ε (a very small but not negligible proportion) of all elliptic curves over $\mathbf{F}_q$, where the property of being a "weak" curve is independent of isogeny and endomorphism class.

Let's also suppose that the "weak" property can be spotted quickly once we get to the curve.

NOTE: There are some fields where Weil descent provides such an algorithm.

NOTE: There are some fields where Weil descent provides such an algorithm. However, those fields are composite degree extensions of F_2, not the degree 571 extension.

NOTE: There are some fields where Weil descent provides such an algorithm. However, those fields are composite degree extensions of F_2, not the degree 571 extension. For this reason our assumption is speculative.

NOTE: There are some fields where Weil descent provides such an algorithm. However, those fields are composite degree extensions of F_2, not the degree 571 extension. For this reason our assumption is speculative.

Then a discrete log on E can be found in time roughly $T_1 + T_2/\varepsilon$, where T_2 denotes the time for constructing an isogeny that moves the ECDLP to another curve –

NOTE: There are some fields where Weil descent provides such an algorithm. However, those fields are composite degree extensions of F_2, not the degree 571 extension. For this reason our assumption is speculative.

Then a discrete log on E can be found in time roughly $T_1 + T_2/\varepsilon$, where T_2 denotes the time for constructing an isogeny that moves the ECDLP to another curve – assuming, of course, that the class of curves to which isogenies from E can be constructed in time less than T_2 contains more than $1/\varepsilon$ curves.

Thus, if ε is much less than 1/(40 lakhs), we find that K-571 is safe from the attack, but B-571 is not.

Thus, if ε is much less than 1/(40 lakhs), we find that K-571 is safe from the attack, but B-571 is not.

In other words, it is the possibility of random isogeny walks through an endomorphism class that under certain circumstances might make a random curve less secure than a special curve.

What conclusions do we want to draw?

What conclusions do we want to draw?

Not that we should prefer special curves over random ones.

What conclusions do we want to draw?

Not that we should prefer special curves over random ones.

All we can say is that we don’t really know.

What conclusions do we want to draw?

Not that we should prefer special curves over random ones.

All we can say is that we don’t really know.

It’s a judgment call.

What conclusions do we want to draw?

Not that we should prefer special curves over random ones.

All we can say is that we don't really know.

It's a judgment call.

Despite the reluctance of many cryptographic researchers to admit it, in fact cryptography is as much an art as a science.

By the way, a similar hypothetical scenario can be given over a prime field.

By the way, a similar hypothetical scenario can be given over a prime field. For example, suppose we choose a random prime B and a random even number A such that

By the way, a similar hypothetical scenario can be given over a prime field. For example, suppose we choose a random prime B and a random even number A such that

(i) $p = A^2 + B^2$ is prime;

By the way, a similar hypothetical scenario can be given over a prime field. For example, suppose we choose a random prime B and a random even number A such that

(i) $p = A^2 + B^2$ is prime;

(ii) either $n = (p+1)/2 - A$
or else $n = (p+1)/2 + A$
is prime.

By the way, a similar hypothetical scenario can be given over a prime field. For example, suppose we choose a random prime B and a random even number A such that

(i) $p = A^2 + B^2$ is prime;

(ii) either $n = (p+1)/2 - A$
or else $n = (p+1)/2 + A$
is prime.

Then the elliptic curve E over $\mathbf{F}_p$ defined (for suitable a in $\mathbf{F}_p$) by $y^2 = x^3 - ax$ is isolated, in the sense that it's infeasible to construct isogenies to other curves.

Possibly this means that this curve might be more secure than a random curve over that field.

Possibly this means that this curve might be more secure than a random curve over that field.

However, the only NIST-recommended curves over a prime field are random ones.

Possibly this means that this curve might be more secure than a random curve over that field.

However, the only NIST-recommended curves over a prime field are random ones.

Maybe they're right to do that…

Possibly this means that this curve might be more secure than a random curve over that field.

However, the only NIST-recommended curves over a prime field are random ones.

Maybe they're right to do that…

Or maybe not.

PART II:

PART II: AN ABBREVIATED HISTORY OF MY MISJUDGMENTS AND ERRONEOUS BELIEFS DURING THE PAST 25 YEARS

PART II: AN ABBREVIATED HISTORY OF MY MISJUDGMENTS AND ERRONEOUS BELIEFS DURING THE PAST 25 YEARS

First major one:

PART II: AN ABBREVIATED HISTORY OF MY MISJUDGMENTS AND ERRONEOUS BELIEFS DURING THE PAST 25 YEARS

First major one:

In the late 1980's it seemed (to me at least) that any elliptic curve group would be secure as long as its order is prime or almost prime.

With that condition, I thought, all curves were created equal, and were endowed with an intractable ECDLP.

With that condition, I thought, all curves were created equal, and were endowed with an intractable ECDLP.

So for pedagogical reasons why not use the simplest possible curves?

With that condition, I thought, all curves were created equal, and were endowed with an intractable ECDLP.

So for pedagogical reasons why not use the simplest possible curves? And this is what I often did -- in my introductory book published in 1987 and in my articles and talks in the 1980's.

It's an elementary exercise to show that the curve

$y^2 = x^3 - x$ over $\mathbf{F}_p$ with $4|(p+1)$

It's an elementary exercise to show that the curve

$y^2 = x^3 - x$ over $\mathbf{F}_p$ with $4|(p+1)$

or

$y^2 + y = x^3$ over $\mathbf{F}_p$ with $3|(p+1)$

It's an elementary exercise to show that the curve

$y^2 = x^3 - x$ over $\mathbf{F}_p$ with $4|(p+1)$

or

$y^2 + y = x^3$ over $\mathbf{F}_p$ with $3|(p+1)$

has group order $p+1$.

It's an elementary exercise to show that the curve

$y^2 = x^3 - x$ over $\mathbf{F}_p$ with $4|(p+1)$

or

$y^2 + y = x^3$ over $\mathbf{F}_p$ with $3|(p+1)$

has group order $p+1$.

Just choose p so that $(p+1)/4$ or $(p+1)/6$ is prime, and ECC is secure...

It's an elementary exercise to show that the curve

$y^2 = x^3 - x$ over $\mathbf{F}_p$ with $4|(p+1)$

or

$y^2 + y = x^3$ over $\mathbf{F}_p$ with $3|(p+1)$

has group order $p+1$.

Just choose p so that $(p+1)/4$ or $(p+1)/6$ is prime, and ECC is secure...

Or so I thought.

These curves also have some nice efficiency advantages for computing point multiples, especially over extension fields of $\mathbf{F}_2$ and $\mathbf{F}_3$.

These curves also have some nice efficiency advantages for computing point multiples, especially over extension fields of $\mathbf{F}_2$ and $\mathbf{F}_3$.

Then in 1991 Menezes-Okamoto-Vanstone showed that the Weil pairing gives a reduction of the ECDLP to the DLP on the multiplicative group of an extension of the field of definition.

These curves also have some nice efficiency advantages for computing point multiples, especially over extension fields of $\mathbf{F}_2$ and $\mathbf{F}_3$.

Then in 1991 Menezes-Okamoto-Vanstone showed that the Weil pairing gives a reduction of the ECDLP to the DLP on the multiplicative group of an extension of the field of definition.

And for supersingular curves, such as the two written above, the extension degree is very small.

These curves also have some nice efficiency advantages for computing point multiples, especially over extension fields of $\mathbf{F}_2$ and $\mathbf{F}_3$.

Then in 1991 Menezes-Okamoto-Vanstone showed that the Weil pairing gives a reduction of the ECDLP to the DLP on the multiplicative group of an extension of the field of definition.

And for supersingular curves, such as the two written above, the extension degree is very small. Usually it's 2, as in the above cases.

This killed supersingular curves for ECC

This killed supersingular curves for ECC and made me feel foolish for having used them so often as illustrative examples.

This killed supersingular curves for ECC and made me feel foolish for having used them so often as illustrative examples.

NOTE: Much later – Resurrection.

This killed supersingular curves for ECC and made me feel foolish for having used them so often as illustrative examples.

NOTE: Much later – Resurrection.

What surprised me as much as the MOV attack that killed supersingular curves in 1991

This killed supersingular curves for ECC and made me feel foolish for having used them so often as illustrative examples.

NOTE: Much later – Resurrection.

What surprised me as much as the MOV attack that killed supersingular curves in 1991 was that 10 years later they made a roaring comeback from the grave –

This killed supersingular curves for ECC and made me feel foolish for having used them so often as illustrative examples.

NOTE: Much later – Resurrection.

What surprised me as much as the MOV attack that killed supersingular curves in 1991 was that 10 years later they made a roaring comeback from the grave – when pairing-based crypto took the research community by storm.

Another of my misjudgments during that time period:

Another of my misjudgments during that time period:

In 1989, when I first proposed Hyperelliptic Curve Cryptography, if you'd asked me I would have explained what I saw as the main potential advantage of HCC over ECC

Another of my misjudgments during that time period:

In 1989, when I first proposed Hyperelliptic Curve Cryptography, if you'd asked me I would have explained what I saw as the main potential advantage of HCC over ECC by speculating that most likely the higher the genus, the more security you'd get.

In other words, an attack that might work in low genus would be less likely to work in high genus.

In other words, an attack that might work in low genus would be less likely to work in high genus.

That's what I thought…

In other words, an attack that might work in low genus would be less likely to work in high genus.

That's what I thought…

But that turned out to be exactly the opposite of what happened.

I was taken completely by surprise by the Adleman—DeMarrais—Huang algorithm (1994), which gave a subexponential time solution of the HCDLP in large genus.

I was taken completely by surprise by the Adleman—DeMarrais—Huang algorithm (1994), which gave a subexponential time solution of the HCDLP in large genus.

My favorite illustrative example, which was a genus-191 curve over $\mathbf{F}_2$, immediately became totally insecure.

I was taken completely by surprise by the Adleman—DeMarrais—Huang algorithm (1994), which gave a subexponential time solution of the HCDLP in large genus.

My favorite illustrative example, which was a genus-191 curve over $\mathbf{F}_2$, immediately became totally insecure.

After subsequent work by Gaudry, Diem, and others, it now seems that anything bigger than genus 2 is less secure than genus 1 or 2.

The only HCC that is fully competitive with ECC is genus-2 HCC.

The only HCC that is fully competitive with ECC is genus-2 HCC.

I couldn’t have been more wrong in my intuition about the greater security of high genus!

Next embarrassing episode:

Next embarrassing episode:

In the early 1990's, Mike Fellows and I became captivated by the notion that, despite the fiasco with knapsacks, good cryptosystems could in fact be constructed from NP-hard combinatorial problems.

Next embarrassing episode:

In the early 1990's, Mike Fellows and I became captivated by the notion that, despite the fiasco with knapsacks, good cryptosystems could in fact be constructed from NP-hard combinatorial problems.

We even wrote a paper with the exuberant title "Combinatorial Cryptosystems Galore!"

There was only one actual example that we spent some time developing, and it had a sorry history.

There was only one actual example that we spent some time developing, and it had a sorry history.

As I recount in my book *Random Curves*:

There was only one actual example that we spent some time developing, and it had a sorry history.

As I recount in my book *Random Curves*:

“Mike Fellows and I… constructed a system based on… *ideal membership…* that involved polynomials, and we challenged people to try to crack it.

There was only one actual example that we spent some time developing, and it had a sorry history.

As I recount in my book *Random Curves*:

“Mike Fellows and I… constructed a system based on… *ideal membership...* that involved polynomials, and we challenged people to try to crack it.

“The most attractive feature of our cryptosystem was the name that Mike thought up for it: *Polly Cracker.*

There was only one actual example that we spent some time developing, and it had a sorry history.

As I recount in my book *Random Curves*:

"Mike Fellows and I… constructed a system based on… *ideal membership...* that involved polynomials, and we challenged people to try to crack it.

"The most attractive feature of our cryptosystem was the name that Mike thought up for it: *Polly Cracker.*

"It was very inefficient, and before long some papers were published that indeed cracked the code."

Back to ECC:

Back to ECC:

During the first 15 years of ECC my feeling was that it didn’t matter what field you worked over.

Back to ECC:

During the first 15 years of ECC my feeling was that it didn't matter what field you worked over. You had to avoid generic algorithms by working in groups of large prime order, and after MOV you had to avoid supersingular curves.

Back to ECC:

During the first 15 years of ECC my feeling was that it didn't matter what field you worked over. You had to avoid generic algorithms by working in groups of large prime order, and after MOV you had to avoid supersingular curves.

But otherwise you could use whatever field you most enjoy working with, and security is unaffected by that choice.

Back to ECC:

During the first 15 years of ECC my feeling was that it didn't matter what field you worked over. You had to avoid generic algorithms by working in groups of large prime order, and after MOV you had to avoid supersingular curves.

But otherwise you could use whatever field you most enjoy working with, and security is unaffected by that choice.

That's what I thought.

But I was wrong about this.

But I was wrong about this.

Late 1990's: Frey proposes Weil descent to attack the DLP on curves over composite degree extension fields.

But I was wrong about this.

Late 1990's: Frey proposes Weil descent to attack the DLP on curves over composite degree extension fields.

His idea was to transport the ECDLP to the DLP on a high-genus curve over a subextension,

But I was wrong about this.

Late 1990's: Frey proposes Weil descent
to attack the DLP on curves over composite
degree extension fields.

His idea was to transport the ECDLP
to the DLP on a high-genus curve over
a subextension, where it could be attacked
by the faster high-genus algorithms.

Soon Gaudry, Hess, Smart, Galbraith, Menezes, Teske, and others found weak curves over certain binary fields of composite extension degree.

Soon Gaudry, Hess, Smart, Galbraith, Menezes, Teske, and others found weak curves over certain binary fields of composite extension degree.

Fortunately, other people (such as Scott Vanstone) had had better instincts than I had, and all commercial implementations and all ECC standards used prime fields or prime-degree extensions of $\mathbf{F}_2$.

Another example of how bad I am at anticipating future developments:

Another example of how bad I am at anticipating future developments:

In early 1998 I published *Algebraic Aspects of Cryptography.*

Another example of how bad I am at anticipating future developments:

In early 1998 I published *Algebraic Aspects of Cryptography.* In a section titled “Cultural Background” I discussed the Birch and Swinnerton-Dyer Conjecture,

Another example of how bad I am at anticipating future developments:

In early 1998 I published *Algebraic Aspects of Cryptography.* In a section titled “Cultural Background” I discussed the Birch and Swinnerton-Dyer Conjecture, after which I essentially apologized to my readers for taking up their valuable time with something that, while mathematically important, has no relevance for cryptography.

A mere 8 months later I was eating those words, after I received an email from J. Silverman outlining a striking new approach to the ECDLP.

A mere 8 months later I was eating those words, after I received an email from J. Silverman outlining a striking new approach to the ECDLP.

It was a variant – somewhat backwards – version of index calculus, and for that reason Silverman called it “xedni calculus.”

A mere 8 months later I was eating those words, after I received an email from J. Silverman outlining a striking new approach to the ECDLP.

It was a variant – somewhat backwards – version of index calculus, and for that reason Silverman called it “xedni calculus.”

What was most alarming for ECC people was that Silverman used the heuristics of the BSD Conjecture (and an analytic rank formula of Mestre) to boost the likelihood of a successful attack on the ECDLP.

The heuristics of this conjecture were very hard to analyze from a practical computational standpoint.

The heuristics of this conjecture were very hard to analyze from a practical computational standpoint.

So the exact same mathematics that 8 months earlier I had been dismissing as irrelevant to cryptography turned out to be at the center of our fears about xedni.

The heuristics of this conjecture were very hard to analyze from a practical computational standpoint.

So the exact same mathematics that 8 months earlier I had been dismissing as irrelevant to cryptography turned out to be at the center of our fears about xedni.

After a lot of initial worry about xedni (fueled by our concern that RSA would use xedni as a weapon in their public relations battle with ECC, which was still going strong in 1998), I found that we could use the height function to show that xedni wouldn't work.

I was so thrilled about this success in defending ECC that I gave a talk at ECC 2000 titled

"Miracles of the Height Function: A Golden Shield Protecting ECC"

At around the same time a paper by Silverman and Suzuki made a detailed examination of index calculus and explained why it wouldn’t work.

At around the same time a paper by Silverman and Suzuki made a detailed examination of index calculus and explained why it wouldn't work.

Essentially, the Silverman-Suzuki paper elaborated on the argument that Vic Miller made in his original ECC paper in 1985.

At around the same time a paper by Silverman and Suzuki made a detailed examination of index calculus and explained why it wouldn't work.

Essentially, the Silverman-Suzuki paper elaborated on the argument that Vic Miller made in his original ECC paper in 1985.

At ECC 2007, Silverman made a similar analysis for all 4 ways one could try index or xedni with liftings to global fields.

But alas! Index calculus has reared its evil head during the last few years.

But alas! Index calculus has reared its evil head during the last few years.

For example, Gaudry and Diem found subexponential index calculus algorithms for the ECDLP on elliptic curves defined over the degree-m extension of $\mathbf{F}_q$ as m and q grow suitably.

But alas! Index calculus has reared its evil head during the last few years.

For example, Gaudry and Diem found subexponential index calculus algorithms for the ECDLP on elliptic curves defined over the degree-m extension of $\mathbf{F}_q$ as m and q grow suitably.

I've learned the hard way that it's foolish to go around talking about "Golden Shields" that protect security of ECC protocols or anything else.

But alas! Index calculus has reared its evil head during the last few years.

For example, Gaudry and Diem found subexponential index calculus algorithms for the ECDLP on elliptic curves defined over the degree-m extension of $\mathbf{F}_q$ as m and q grow suitably.

I've learned the hard way that it's foolish to go around talking about "Golden Shields" that protect security of ECC protocols or anything else. And I don't believe in miracles anymore.

In each case of misjudgment, I thought
I had a good mathematical reason to feel
confident about what the future would bring.

In each case of misjudgment, I thought
I had a good mathematical reason to feel
confident about what the future would bring.

But history is capricious.

In each case of misjudgment, I thought
I had a good mathematical reason to feel
confident about what the future would bring.

But history is capricious.

It likes to play jokes on us.

In each case of misjudgment, I thought I had a good mathematical reason to feel confident about what the future would bring.

But history is capricious.

It likes to play jokes on us.

I learned that, much as we might wish to convey an impression of self-confidence and mathematical certainty to the outside world, to our colleagues, and to ourselves, such self-confidence is rarely justified.

PART III:

PART III: SOME SCANDALS IN THE HISTORY OF "PROVABLE SECURITY"

PART III: SOME SCANDALS IN THE HISTORY OF "PROVABLE SECURITY"

For more details, please see the series of "Another Look" papers by Menezes and me on the eprint.iacr.org website.

Mihir Bellare and Phil Rogaway are leading cryptographic researchers, and the originators of the notion of “practice-oriented provable security.”

Mihir Bellare and Phil Rogaway are leading cryptographic researchers, and the originators of the notion of “practice-oriented provable security.”

In 1994 Bellare and Rogaway “proved security” for a version of RSA that they called “Optimal Asymmetric Encryption Padding” (OAEP).

Mihir Bellare and Phil Rogaway are leading cryptographic researchers, and the originators of the notion of “practice-oriented provable security.”

In 1994 Bellare and Rogaway “proved security” for a version of RSA that they called “Optimal Asymmetric Encryption Padding” (OAEP).

Soon after, MasterCard and Visa included OAEP in their SET electronic payment standard.

In 2001 – seven years later – Victor Shoup examined the Bellare-Rogaway "proof of security" and showed that it was fallacious.

In 2001 – seven years later – Victor Shoup examined the Bellare-Rogaway "proof of security" and showed that it was fallacious.

This type of thing causes a credibility problem.

A case that I think is even more scandalous occurred in 2005.

A case that I think is even more scandalous occurred in 2005.

In February of that year, Hugo Krawczyk – one of IBM's top cryptographers – submitted a paper to Crypto 2005 in which he claimed to have found flaws in the Menezes—Qu—Vanstone (MQV) key agreement protocol.

A case that I think is even more scandalous occurred in 2005.

In February of that year, Hugo Krawczyk – one of IBM's top cryptographers – submitted a paper to Crypto 2005 in which he claimed to have found flaws in the Menezes—Qu—Vanstone (MQV) key agreement protocol.

He especially criticized MQV for lacking a "proof of security."

Krawczyk replaced MQV with a modified version he called HMQV that he claimed was both more efficient and **provably secure.**

Krawczyk replaced MQV with a modified version he called HMQV that he claimed was both more efficient and **provably secure.**

If his claims had been valid, this would have been a major embarrassment not only to Menezes and his coauthors, but also to the NSA, which had licensed MQV (and 25 other patented protocols) from Certicom and whose experts had studied it closely.

Neither Krawczyk nor the Crypto Program Committee sent a preprint to any of the MQV authors before the Program Committee accepted it.

Neither Krawczyk nor the Crypto Program Committee sent a preprint to any of the MQV authors before the Program Committee accepted it.

When Menezes was finally sent a copy – after the Program Committee had already accepted it – he saw that, first of all, Krawczyk's objections to MQV were without foundation.

Neither Krawczyk nor the Crypto Program Committee sent a preprint to any of the MQV authors before the Program Committee accepted it.

When Menezes was finally sent a copy – after the Program Committee had already accepted it – he saw that, first of all, Krawczyk's objections to MQV were without foundation.

Moreover, he discovered that the paper's main argument ("proof") was fallacious.

Krawczyk had claimed that in his HMQV he could increase efficiency by discarding a “public key validation” step that had been put into MQV to prevent known attacks.

Krawczyk had claimed that in his HMQV he could increase efficiency by discarding a “public key validation” step that had been put into MQV to prevent known attacks.

It was his security “proof” that had given him the confidence to do this.

Krawczyk had claimed that in his HMQV he could increase efficiency by discarding a “public key validation” step that had been put into MQV to prevent known attacks.

It was his security “proof” that had given him the confidence to do this.

Menezes quickly found not only that the proof was fallacious, but that certain of the HMQV protocols succumb to the same attacks as MQV would have if those checks had not been put in it.

Top experts such as Krawczyk – and the members of the Crypto 2005 Program Committee – would not have made this blunder if they had not been mesmerized by the so-called proof.

Top experts such as Krawczyk – and the members of the Crypto 2005 Program Committee – would not have made this blunder if they had not been mesmerized by the so-called proof.

So-called “proofs of security” often do more harm than good because they lull people into a false sense of security and cause them to take leave of common sense.

Regrettably, much cryptographic writing exudes a brash certainty about the work.

Regrettably, much cryptographic writing exudes a brash certainty about the work.

Abstracts and introductions to papers often read as if they were written by marketing people or as part of a patent application, full of hype with little connection to reality.

Regrettably, much cryptographic writing exudes a brash certainty about the work.

Abstracts and introductions to papers often read as if they were written by marketing people or as part of a patent application, full of hype with little connection to reality.

In particular, they highlight their reductionist security argument (that reduces a supposedly intractable problem to a successful attack of a specified sort)

Regrettably, much cryptographic writing exudes a brash certainty about the work.

Abstracts and introductions to papers often read as if they were written by marketing people or as part of a patent application, full of hype with little connection to reality.

In particular, they highlight their reductionist security argument (that reduces a supposedly intractable problem to a successful attack of a specified sort) using terminology designed to convince the reader that their protocols have been “proved” to be secure.

Here’s an example:

Here's an example: A paper by Boldyreva—Gentry—O'Neill—Yum available from eprint.iacr.org/2007/438 gave a pairing-based construction of sequential aggregate signatures

Here's an example: A paper by Boldyreva—Gentry—O'Neill—Yum available from eprint.iacr.org/2007/438 gave a pairing-based construction of sequential aggregate signatures (in which several people in sequence put together a single compact signature).

Here's an example: A paper by Boldyreva—Gentry—O'Neill—Yum available from eprint.iacr.org/2007/438 gave a pairing-based construction of sequential aggregate signatures (in which several people in sequence put together a single compact signature). They claimed it

Here's an example: A paper by Boldyreva—Gentry—O'Neill—Yum available from eprint.iacr.org/2007/438 gave a pairing-based construction of sequential aggregate signatures (in which several people in sequence put together a single compact signature). They claimed it

> "…permits savings on bandwidth and storage… substantially improves computational efficiency and scalability over any existing scheme with suitable functionality…

“In contrast to the only prior scheme to provide this functionality, ours offers improved security…

“In contrast to the only prior scheme to provide this functionality, ours offers improved security… We provide formal security definitions and support the proposed scheme with security proofs…”

> “In contrast to the only prior scheme to provide this functionality, ours offers improved security… We provide formal security definitions and support the proposed scheme with security proofs…”

The amusing thing about this example is that about a year later Huang, Lee, and Yung showed that a crucial security proof in this paper was fallacious.

> “In contrast to the only prior scheme to provide this functionality, ours offers improved security… We provide formal security definitions and support the proposed scheme with security proofs…”

The amusing thing about this example is that about a year later Huang, Lee, and Yung showed that a crucial security proof in this paper was fallacious.

They also broke the corresponding protocol.

> "In contrast to the only prior scheme to provide this functionality, ours offers improved security… We provide formal security definitions and support the proposed scheme with security proofs…"

The amusing thing about this example is that about a year later Huang, Lee, and Yung showed that a crucial security proof in this paper was fallacious.

They also broke the corresponding protocol. (The flaw was catastrophic – details are given in my March 2010 AMS Notices article with Menezes.)

Leading writers on cryptography have often made the claim that reduction arguments constitute “proofs of security” that can be offered to the public as a guarantee.

Leading writers on cryptography have often made the claim that reduction arguments constitute “proofs of security” that can be offered to the public as a guarantee.

From the preface to the book by Katz and Lindell:

Leading writers on cryptography have often made the claim that reduction arguments constitute “proofs of security” that can be offered to the public as a guarantee.

From the preface to the book by Katz and Lindell:

> “…cryptographic constructions can be proven secure with respect to a clearly-stated definition of security and relative to a well-defined cryptographic assumption.

Leading writers on cryptography have often made the claim that reduction arguments constitute “proofs of security” that can be offered to the public as a guarantee.

From the preface to the book by Katz and Lindell:

> “…cryptographic constructions can be proven secure with respect to a clearly-stated definition of security and relative to a well-defined cryptographic assumption.
>
> “This is the essence of modern cryptography, and what has transformed cryptography from an art to a science.

Leading writers on cryptography have often made the claim that reduction arguments constitute “proofs of security” that can be offered to the public as a guarantee.

From the preface to the book by Katz and Lindell:

> “…cryptographic constructions can be proven secure with respect to a clearly-stated definition of security and relative to a well-defined cryptographic assumption.
>
> “This is the essence of modern cryptography, and what has transformed cryptography from an art to a science. The importance of this idea cannot be over-emphasized.”

Meanwhile, anyone who's dismayed by the large number of fallacious proofs in the provable security literature is supposed to be consoled by the prospect that advances in "theorem-proving" software will soon make it possible to prove the security of our protocols automatically

Meanwhile, anyone who's dismayed by the large number of fallacious proofs in the provable security literature is supposed to be consoled by the prospect that advances in "theorem-proving" software will soon make it possible to prove the security of our protocols automatically, with no longer any possibility of flaws in the proofs.

Meanwhile, anyone who's dismayed by the large number of fallacious proofs in the provable security literature is supposed to be consoled by the prospect that advances in "theorem-proving" software will soon make it possible to prove the security of our protocols automatically, with no longer any possibility of flaws in the proofs.

Human mistakes and failings will supposedly disappear from the process of establishing guarantees of security.

I discuss this naïve and dubious notion in
“Another look at automated theorem-proving,”
http://eprint.iacr.org/2007/401.pdf

I discuss this naïve and dubious notion in "Another look at automated theorem-proving," http://eprint.iacr.org/2007/401.pdf

And anyone who's bewildered by the exotic nature of some of the cryptographic assumptions that underlie security proofs for many of the pairing-based protocols is supposed to be reassured by Boyen's exuberant explanation at the Pairings-2008 conference:

“The newcomer to this particular branch of cryptography will… be astonished by the sheer number, and sometimes creativity, of these assumptions…

“The newcomer to this particular branch of cryptography will… be astonished by the sheer number, and sometimes creativity, of these assumptions…

“…in comparison to the admittedly quite simpler algebraic structures of twentieth-century public-key cryptography…

"The newcomer to this particular branch of cryptography will… be astonished by the sheer number, and sometimes creativity, of these assumptions…

"…in comparison to the admittedly quite simpler algebraic structures of twentieth-century public-key cryptography… the new 'bilinear' groups offer a much richer palette of cryptographically useful trapdoors than their 'unidimensional' counterparts…"

In our March 2010 article in the AMS Notices, Menezes and I explain why we do not share Boyen's enthusiasm for the bold and exotic assumptions that populate the landscape in pairing-based cryptography.

In our March 2010 article in the AMS Notices, Menezes and I explain why we do not share Boyen's enthusiasm for the bold and exotic assumptions that populate the landscape in pairing-based cryptography.

Indeed, some of the assumptions used by leading researchers turned out to be significantly weaker than they had thought.

In our March 2010 article in the AMS Notices, Menezes and I explain why we do not share Boyen's enthusiasm for the bold and exotic assumptions that populate the landscape in pairing-based cryptography.

Indeed, some of the assumptions used by leading researchers turned out to be significantly weaker than they had thought.

And the jury is still out on most of the other assumptions, since hardly any of them have been investigated thoroughly.

Thus, on the one hand, we see the trend of bold and boastful writing by cryptographic researchers.

Thus, on the one hand, we see the trend of bold and boastful writing by cryptographic researchers.

On the other hand, we see a long history of misjudgments, faulty assumptions, uncertainty, and catastrophic blunders in cryptography that continues to the present day.

Thus, on the one hand, we see the trend of bold and boastful writing by cryptographic researchers.

On the other hand, we see a long history of misjudgments, faulty assumptions, uncertainty, and catastrophic blunders in cryptography that continues to the present day.

How can we reconcile the disciplinary culture of our field with reality?

To paraphrase Lenin,

To paraphrase Lenin,

What is to be done?

To paraphrase Lenin,

What is to be done?

Should we retreat into despair and cynicism?

To paraphrase Lenin,

What is to be done?

Should we retreat into despair and cynicism?

That is not the message of my talk.

To paraphrase Lenin,

What is to be done?

Should we retreat into despair and cynicism?

That is not the message of my talk.

There is a better answer to this conundrum.

There is a keyword that unlocks the puzzle!

There is a keyword that unlocks the puzzle!

This keyword is revealed if we look for guidance in the wisdom of ancient India.

There is a keyword that unlocks the puzzle!

This keyword is revealed if we look for guidance in the wisdom of ancient India.

In Chapter 13, Verses 8-12 of the Bhagavad Gita

There is a keyword that unlocks the puzzle!

This keyword is revealed if we look for guidance in the wisdom of ancient India.

In Chapter 13, Verses 8-12 of the Bhagavad Gita, the all-knowing one, Sri Krishna, describes the qualities that are necessary for knowledge.

Here is the list, first in Sanskrit,
then in transliteration, then in English.

Here is the list, first in Sanskrit,
then in transliteration, then in English.

Note what the very first quality is.

अमानित्वमदम्भित्वमहिंसा क्षान्तिरार्जवम् ।
आचार्योपासनं शौचं स्थैर्यमात्मविनिग्रहः ॥ ७ ॥

Amānitvam-adambhitvam-ahimsā kṣāntir-ārjavam
Ācaryopāsanam śaucam sthairyam-ātma-vinigrahaḥ

7. Humility; unpretentiousness; non-violence; forbearance; uprighteness; service of the teacher; purity; steadiness; self-control;

It’s clear that if Krishna were brought in as a consultant on how to improve our understanding of cryptography,

It’s clear that if Krishna were brought in as a consultant on how to improve our understanding of cryptography,

and if He were to glance at some of the boastful introductions to articles on eprint.iacr.org,

It’s clear that if Krishna were brought in as a consultant on how to improve our understanding of cryptography,

and if He were to glance at some of the boastful introductions to articles on eprint.iacr.org, at Krawczyk’s HMQV paper,

It's clear that if Krishna were brought in as a consultant on how to improve our understanding of cryptography,

and if He were to glance at some of the boastful introductions to articles on eprint.iacr.org, at Krawczyk's HMQV paper, at Boyen's *Pairings-2008* paper,

It's clear that if Krishna were brought in as a consultant on how to improve our understanding of cryptography,

and if He were to glance at some of the boastful introductions to articles on eprint.iacr.org, at Krawczyk's HMQV paper, at Boyen's *Pairings-2008* paper, and at the preface to the book of Katz and Lindell,

…He would conclude that what the cryptographic research community badly needs is a healthy dose of

…He would conclude that what the cryptographic research community badly needs is a healthy dose of

Amaanitvam.