Steampunk canonical forms

Bruce Reznick
University of Illinois at Urbana-Champaign

ALGECOM5
IUPUI, Indianapolis, October 29, 2011

1. Introduction

Webpage for this talk:
www.math.uiuc.edu/∼reznick/algecom5.html

1. Introduction

Webpage for this talk:
www.math.uiuc.edu/∼reznick/algecom5.html

Much of this material will appear in the paper *Steampunk canonical forms*, which I hope to submit by the end of the year.

1. Introduction

Webpage for this talk:
www.math.uiuc.edu/∼reznick/algecom5.html

Much of this material will appear in the paper *Steampunk canonical forms*, which I hope to submit by the end of the year.

The approach to canonical forms in this talk owes a lot to a wonderful 1993 paper of Richard Ehrenborg and Gian-Carlo Rota. Alexander-Hirschowitz hovers over all modern work on the rank of forms.

1. Introduction

Webpage for this talk:
www.math.uiuc.edu/∼reznick/algecom5.html

Much of this material will appear in the paper *Steampunk canonical forms*, which I hope to submit by the end of the year.

The approach to canonical forms in this talk owes a lot to a wonderful 1993 paper of Richard Ehrenborg and Gian-Carlo Rota. Alexander-Hirschowitz hovers over all modern work on the rank of forms.

Many thanks to the organizers for the invitation to give this talk to ALGECOM5.

2. Overview

What is steampunk? It is a style based on combining 19th century Victorian culture with a few elements of modern life, such as computers.

2. Overview

What is steampunk? It is a style based on combining 19th century Victorian culture with a few elements of modern life, such as computers.

What are canonical forms? Before I give a definition, here's the Main Example for this talk.

2. Overview

What is steampunk? It is a style based on combining 19th century Victorian culture with a few elements of modern life, such as computers.

What are canonical forms? Before I give a definition, here's the Main Example for this talk.

Any binary quadratic form over $\mathbb{C}$ can be written as

$$p(x, y) = ax^2 + 2bxy + cy^2.$$

There also exist $t_j \in \mathbb{C}$ so that

$$p(x, y) = (t_1 x + t_2 y)^2 + (t_3 y)^2,$$

at least unless $a = 0$ and $b \neq 0$.

2. Overview

What is steampunk? It is a style based on combining 19th century Victorian culture with a few elements of modern life, such as computers.

What are canonical forms? Before I give a definition, here's the Main Example for this talk.

Any binary quadratic form over $\mathbb{C}$ can be written as

$$p(x,y) = ax^2 + 2bxy + cy^2.$$

There also exist $t_j \in \mathbb{C}$ so that

$$p(x,y) = (t_1 x + t_2 y)^2 + (t_3 y)^2,$$

at least unless $a = 0$ and $b \neq 0$.

That is, a "general" binary quadratic form can be written with the square completed.

2. Overview

What is steampunk? It is a style based on combining 19th century Victorian culture with a few elements of modern life, such as computers.

What are canonical forms? Before I give a definition, here's the Main Example for this talk.

Any binary quadratic form over $\mathbb{C}$ can be written as

$$p(x, y) = ax^2 + 2bxy + cy^2.$$

There also exist $t_j \in \mathbb{C}$ so that

$$p(x, y) = (t_1 x + t_2 y)^2 + (t_3 y)^2,$$

at least unless $a = 0$ and $b \neq 0$.

That is, a "general" binary quadratic form can be written with the square completed.

You don't need a proof of this, but I'll give you several anyway.

2. Overview

What is the as-yet-nonexistent subject of **steampunk canonical forms**? First approximation: 19th century algebra plus the concept of vector spaces plus the belief that there is still some juice left in the algebraic geometry of binary forms.

2. Overview

What is the as-yet-nonexistent subject of **steampunk canonical forms**? First approximation: 19th century algebra plus the concept of vector spaces plus the belief that there is still some juice left in the algebraic geometry of binary forms.

Note that there were three parameters in the two representations of binary quadratic forms on the last slide. This is necessary, but not sufficient. Obviously, $\sum_{i=1}^{3}(t_i x)^2$ isn't a canonical form. More subtly, there are three parameters in

$$(t_1 x + t_2 y)^2 + (it_1 x + t_3 y)^2,$$

but the coefficient of x^2 is 0, so the sum above must have a factor of y, which a general binary quadratic form does not.

2. Overview

What is the as-yet-nonexistent subject of **steampunk canonical forms**? First approximation: 19th century algebra plus the concept of vector spaces plus the belief that there is still some juice left in the algebraic geometry of binary forms.

Note that there were three parameters in the two representations of binary quadratic forms on the last slide. This is necessary, but not sufficient. Obviously, $\sum_{i=1}^{3}(t_i x)^2$ isn't a canonical form. More subtly, there are three parameters in

$$(t_1 x + t_2 y)^2 + (it_1 x + t_3 y)^2,$$

but the coefficient of x^2 is 0, so the sum above must have a factor of y, which a general binary quadratic form does not.

This example may look silly, but it isn't.

3. Well-known old canonical forms

There are two sets of well-known 19th century canonical forms; both could be construed as generalizations of the Main Example.

Every quadratic form $p \in H_2(\mathbb{C}^n)$ is a sum of n squares, but since the naive number of coefficients, $n \times n$, is $> N(n,2) = \frac{n(n+1)}{2}$, a sum of n squares is not, *per se*, a canonical form. However, the standard "upper triangular" representation is a canonical form.

3. Well-known old canonical forms

There are two sets of well-known 19th century canonical forms; both could be construed as generalizations of the Main Example.

Every quadratic form $p \in H_2(\mathbb{C}^n)$ is a sum of n squares, but since the naive number of coefficients, $n \times n$, is $> N(n,2) = \frac{n(n+1)}{2}$, a sum of n squares is not, *per se*, a canonical form. However, the standard "upper triangular" representation is a canonical form. Let $\{t_{ij} : 1 \le i \le j \le n\}$ be the $N(n,2)$ parameters. Then a general quadratic form in n variables can be written as

$$F(\{t_{ij}\})(x) := \sum_{i=1}^{n} (t_{ii}x_i + \cdots + t_{in}x_n)^2$$

3. Well-known old canonical forms

There are two sets of well-known 19th century canonical forms; both could be construed as generalizations of the Main Example.

Every quadratic form $p \in H_2(\mathbb{C}^n)$ is a sum of n squares, but since the naive number of coefficients, $n \times n$, is $> N(n,2) = \frac{n(n+1)}{2}$, a sum of n squares is not, *per se*, a canonical form. However, the standard "upper triangular" representation is a canonical form. Let $\{t_{ij} : 1 \le i \le j \le n\}$ be the $N(n,2)$ parameters. Then a general quadratic form in n variables can be written as

$$F(\{t_{ij}\})(x) := \sum_{i=1}^{n} (t_{ii}x_i + \cdots + t_{in}x_n)^2$$

The usual proof of this is constructive: just iteratively complete the square.

3. Well-known old canonical forms

In 1851, J. J. Sylvester proved a once-famous theorem about canonical forms for binary forms, via a useful algorithm.

Theorem (Sylvester's Canonical Forms)

(i) A general quadratic form of degree $d = 2k - 1$ can be written as

$$\sum_{j=1}^{k} (\alpha_j x + \beta_j y)^{2k-1}.$$

(ii) For any non-zero linear form $\ell(x, y) = \alpha x + \beta y$, a general binary form of degree $d = 2k$ can be written as

$$\lambda \cdot \ell^{2k}(x, y) + \sum_{j=1}^{k} (\alpha_j x + \beta_j y)^{2k}.$$

for some $\lambda \in \mathbb{C}$.

4. Some new steampunk canonical forms

Here are a few examples of what I'll be talking about and proving (the proofs are quite simple):

Theorem

A general binary sextic form is the sum of the cube of a quadratic form and the square of a cubic form.

4. Some new steampunk canonical forms

Here are a few examples of what I'll be talking about and proving (the proofs are quite simple):

Theorem

A general binary sextic form is the sum of the cube of a quadratic form and the square of a cubic form.

4. Some new steampunk canonical forms

Here are a few examples of what I'll be talking about and proving (the proofs are quite simple):

Theorem

A general binary sextic form is the sum of the cube of a quadratic form and the square of a cubic form.

Theorem

A general binary form of degree $d = 2k$ can be written as

$$(\alpha_0\ x^2 + \beta_0\ x\ y + \gamma_0\ y^2)^k + \sum_{j=1}^{k-1}(\alpha_j x + \beta_j y)^{2k}.$$

4. Some new steampunk canonical forms

Theorem (Old Slinky)

A general cubic form $p(x_1, \dots, x_n)$ has a unique representation in the form

$$p(x_1, \dots, x_n) = \sum_{1 \le i \le j \le n} (\alpha_{\{i,j\},i} x_i + \cdots + \alpha_{\{i,j\},j} x_j)^3.$$

Even though this is actually a canonical form, it is not a representation of p as a minimal **number** of cubes.

4. Some new steampunk canonical forms

Theorem (Old Slinky)

A general cubic form $p(x_1, \dots, x_n)$ *has a unique representation in the form*

$$p(x_1, \dots, x_n) = \sum_{1 \le i \le j \le n} (\alpha_{\{i,j\},i} x_i + \cdots + \alpha_{\{i,j\},j} x_j)^3.$$

Even though this is actually a canonical form, it is not a representation of p as a minimal **number** of cubes.

Theorem (Old Slinky)

A general cubic form $p(x_1, \dots, x_n)$ has a unique representation in the form

$$p(x_1, \dots, x_n) = \sum_{1 \le i \le j \le n} (\alpha_{\{i,j\},i} x_i + \cdots + \alpha_{\{i,j\},j} x_j)^3.$$

Even though this is actually a canonical form, it is not a representation of p as a minimal **number** of cubes.

Somewhat better is a wonderful and little-known canonical form was discovered by Boris Reichstein in 1987.

4. Some new steampunk canonical forms

Theorem (Reichstein's Theorem)

A general cubic $p(x_1, \dots, x_n)$ *can be written as*

$$\sum_{k=1}^{n} (\alpha_{k1} x_1 + \cdots + \alpha_{kn} x_n)^3 + q(x_1, \dots, x_{n-2}).$$

4. Some new steampunk canonical forms

Theorem (Reichstein's Theorem)

A general cubic $p(x_1, \dots, x_n)$ *can be written as*

$$\sum_{k=1}^{n} (\alpha_{k1} x_1 + \cdots + \alpha_{kn} x_n)^3 + q(x_1, \dots, x_{n-2}).$$

4. Some new steampunk canonical forms

Theorem (Reichstein's Theorem)

A general cubic $p(x_1,\dots,x_n)$ can be written as

$$\sum_{k=1}^{n}(\alpha_{k1}x_1+\cdots+\alpha_{kn}x_n)^3+q(x_1,\dots,x_{n-2}).$$

Reichstein's simple computational argument will be given later in this talk. Slinky gives a general cubic form in n variables as a sum of $\frac{n(n+1)}{2}$ cubes; Reichstein needs roughly $\frac{n^2}{4}$. The true minimum, by Alexander-Hirschowitz, is $\lceil\frac{(n+1)(n+2)}{6}\rceil$ for $n\neq 5$, and one more if $n=5$. Later, we'll give a constructive proof of *every* cubic form in n variables as a sum of $\frac{n(n+1)}{2}$ cubes, but this isn't canonical.

5. Basic Definitions

Let $H_d(\mathbb{C}^n)$ denote the set of forms $p(x_1, \dots, x_n)$ of degree d with coefficients in $\mathbb{C}$. The dimension of the vector space $H_d(\mathbb{C}^n)$ is $N(n,d) := \binom{n+d-1}{d}$. Let $\mathcal{I}(n,d)$ be the index set of monomials:

$$\mathcal{I}(n,d) = \left\{ (i_1, \dots, i_n) : 0 \le i_k \in \mathbb{Z}, \quad \sum_k i_k = d \right\}.$$

5. Basic Definitions

Let $H_d(\mathbb{C}^n)$ denote the set of forms $p(x_1, \dots, x_n)$ of degree d with coefficients in $\mathbb{C}$. The dimension of the vector space $H_d(\mathbb{C}^n)$ is $N(n,d) := \binom{n+d-1}{d}$. Let $\mathcal{I}(n,d)$ be the index set of monomials:

$$\mathcal{I}(n,d) = \left\{ (i_1, \dots, i_n) : 0 \le i_k \in \mathbb{Z}, \quad \sum_k i_k = d \right\}.$$

Let $x^i = x_1^{i_1} \cdots x_n^{i_n}$ and $c(i) = \frac{d!}{\prod i_k!}$ denote the multinomial coefficient. If $p \in H_d(\mathbb{C}^n)$, then we can write

$$p(x_1, \dots, x_n) = \sum_{i \in \mathcal{I}(n,d)} c(i) a(p;i) x^i.$$

5. Basic Definitions

The following fundamental theorem sieves for algebraic geometers.
The only accessible proof I know is in Ehrenborg-Rota.

5. Basic Definitions

The following fundamental theorem sieves for algebraic geometers. The only accessible proof I know is in Ehrenborg-Rota.

Theorem

Suppose $F : \mathbb{C}^N \to \mathbb{C}^N$ *is a polynomial map; that is,*

$$F(t_1, \dots, t_N) = (f_1(t_1, \dots, t_N), \dots, f_N(t_1, \dots, t_N))$$

where each $f_j \in \mathbb{C}[t_1, \dots, t_N]$*. Then either (i) or (ii) holds:*
(i) The N *polynomials* $\{f_j : 1 \le j \le N\}$ *are algebraically dependent and* $F(\mathbb{C}^N)$ *lies in some non-trivial* $\{P = 0\}$ *in* $\mathbb{C}^N$*.*
(ii) The N *polynomials* $\{f_j : 1 \le j \le N\}$ *are algebraically independent and* $F(\mathbb{C}^N)$ *is (at least) dense in* $\mathbb{C}^N$*.*
Furthermore, the second case occurs if and only there is a point $u \in \mathbb{C}^N$ *at which the Jacobian matrix* $\left[\frac{\partial f_i}{\partial t_j}(u)\right]$ *has full rank.*

5. Basic Definitions

When $N = N(n, d)$, we may interpret such an F as a map from $\mathbb{C}^N$ to $H_d(\mathbb{C}^n)$ by indexing $\mathcal{I}(n, d)$ as $\{i_j : 1 \le j \le N\}$ and making the interpretation in an abuse of notation that

$$F(\{t_j\})(x) = \sum_{j=1}^{N} f_j(t_1 \dots, t_N)x^{i_j}.$$

5. Basic Definitions

When $N = N(n,d)$, we may interpret such an F as a map from $\mathbb{C}^N$ to $H_d(\mathbb{C}^n)$ by indexing $\mathcal{I}(n,d)$ as $\{i_j : 1 \le j \le N\}$ and making the interpretation in an abuse of notation that

$$F(\{t_j\})(x) = \sum_{j=1}^{N} f_j(t_1 \dots, t_N) x^{i_j}.$$

Definition

A **canonical form** for $H_d(\mathbb{C}^n)$ is any polynomial map F from $\mathbb{C}^N$ to $H_d(\mathbb{C}^n)$ so that almost every $p \in H_d(\mathbb{C}^n)$ is in the range of F.

5. Basic Definitions

When $N = N(n,d)$, we may interpret such an F as a map from $\mathbb{C}^N$ to $H_d(\mathbb{C}^n)$ by indexing $\mathcal{I}(n,d)$ as $\{i_j : 1 \le j \le N\}$ and making the interpretation in an abuse of notation that

$$F(\{t_j\})(x) = \sum_{j=1}^{N} f_j(t_1 \dots, t_N) x^{i_j}.$$

Definition

A **canonical form** for $H_d(\mathbb{C}^n)$ is any polynomial map F from $\mathbb{C}^N$ to $H_d(\mathbb{C}^n)$ so that almost every $p \in H_d(\mathbb{C}^n)$ is in the range of F.

5. Basic Definitions

When $N = N(n,d)$, we may interpret such an F as a map from $\mathbb{C}^N$ to $H_d(\mathbb{C}^n)$ by indexing $\mathcal{I}(n,d)$ as $\{i_j : 1 \le j \le N\}$ and making the interpretation in an abuse of notation that

$$F(\{t_j\})(x) = \sum_{j=1}^{N} f_j(t_1 \dots, t_N) x^{i_j}.$$

Definition

A **canonical form** for $H_d(\mathbb{C}^n)$ is any polynomial map F from $\mathbb{C}^N$ to $H_d(\mathbb{C}^n)$ so that almost every $p \in H_d(\mathbb{C}^n)$ is in the range of F.

If $\{\phi_j\}$ is a basis of $H_d(\mathbb{C}^n)$, e.g. $\phi_j = c(i_j)x^{i_j}$, then

$$F(\{t_j\})(x) = \sum_{j=1}^{N(n,d)} t_j \phi_j$$

is technically (though not traditionally) a canonical form.

5. Basic Definitions

One gets the impression from the literature that canonical forms are extremely rare. Actually, a “general” polynomial map from $\mathbb{C}^N$ to itself is canonical. But *interesting* canonical forms, those with natural interpretations, still command our attention.

Let's return to the Main Example.
The partials of $(t_1x_1 + t_2x_2)^2 + (t_3x_2)^2$ with respect to the t_j's are:

$$2x_1(t_1x_1 + t_2x_2), \quad 2x_2(t_1x_1 + t_2x_2), \quad 2x_2(t_3x_2).$$

5. Basic Definitions

One gets the impression from the literature that canonical forms are extremely rare. Actually, a "general" polynomial map from $\mathbb{C}^N$ to itself is canonical. But *interesting* canonical forms, those with natural interpretations, still command our attention.

Let's return to the Main Example.
The partials of $(t_1x_1 + t_2x_2)^2 + (t_3x_2)^2$ with respect to the t_j's are:

$$2x_1(t_1x_1 + t_2x_2), \quad 2x_2(t_1x_1 + t_2x_2), \quad 2x_2(t_3x_2).$$

If we specialize at $(t_1, t_2, t_3) = (1, 0, 1)$, so $t_1x_1 + t_2x_2 = x_1$ and $t_3x_2 = x_2$, then this becomes $2x_1^2, 2x_1x_2, 2x_2^2$ and these do span $H_2(\mathbb{C}^2)$, providing an abstract existential proof that you can complete the square for binary quadratic forms!

5. Basic Definitions

We can prove the sextic theorem in a similar fashion. Suppose

$$p(x,y) = f^2(x,y) + g^3(x,y) =$$
$$(t_1x^3 + t_2x^2y + t_3xy^2 + t_4y^3)^2 + (t_5x^2 + t_6xy + t_7y^2)^3 :$$
$$f(x,y) = t_1x^3 + t_2x^2y + t_3xy^2 + t_4y^3,$$
$$g(x,y) = t_5x^2 + t_6xy + t_7y^2.$$

5. Basic Definitions

We can prove the sextic theorem in a similar fashion. Suppose

$$p(x,y) = f^2(x,y) + g^3(x,y) =$$
$$(t_1x^3 + t_2x^2y + t_3xy^2 + t_4y^3)^2 + (t_5x^2 + t_6xy + t_7y^2)^3 :$$
$$f(x,y) = t_1x^3 + t_2x^2y + t_3xy^2 + t_4y^3,$$
$$g(x,y) = t_5x^2 + t_6xy + t_7y^2.$$

Then the partials with respect to the t_j's are:

$$2x^3f,\ 2x^2yf,\ 2xy^2f,\ 2y^3f; \quad 3x^2g^2,\ 3xyg^2,\ 3y^2g^2.$$

5. Basic Definitions

We can prove the sextic theorem in a similar fashion. Suppose

$$p(x,y) = f^2(x,y) + g^3(x,y) =$$
$$(t_1x^3 + t_2x^2y + t_3xy^2 + t_4y^3)^2 + (t_5x^2 + t_6xy + t_7y^2)^3 :$$
$$f(x,y) = t_1x^3 + t_2x^2y + t_3xy^2 + t_4y^3,$$
$$g(x,y) = t_5x^2 + t_6xy + t_7y^2.$$

Then the partials with respect to the t_j's are:

$$2x^3f,\ 2x^2yf,\ 2xy^2f,\ 2y^3f;\quad 3x^2g^2,\ 3xyg^2,\ 3y^2g^2.$$

If we specialize at $f = x^3, g = y^2$, then these partials become:

$$2x^6,\ 2x^5y,\ 2x^4y^2,\ 2x^3y^3;\quad 3x^2y^4,\ 3xy^5\ ,3y^6.$$

5. Basic Definitions

We can prove the sextic theorem in a similar fashion. Suppose

$$p(x,y) = f^2(x,y) + g^3(x,y) =$$
$$(t_1x^3 + t_2x^2y + t_3xy^2 + t_4y^3)^2 + (t_5x^2 + t_6xy + t_7y^2)^3 :$$
$$f(x,y) = t_1x^3 + t_2x^2y + t_3xy^2 + t_4y^3,$$
$$g(x,y) = t_5x^2 + t_6xy + t_7y^2.$$

Then the partials with respect to the t_j's are:

$$2x^3f,\ 2x^2yf,\ 2xy^2f,\ 2y^3f;\quad 3x^2g^2,\ 3xyg^2,\ 3y^2g^2.$$

If we specialize at $f = x^3, g = y^2$, then these partials become:

$$2x^6,\ 2x^5y,\ 2x^4y^2,\ 2x^3y^3;\quad 3x^2y^4,\ 3xy^5\ ,3y^6.$$

These trivially span $H_6(\mathbb{C}^2)$.
I haven't been able to find an algorithmic proof yet.

5. Basic Definitions

There are two ways to show that F is a canonical form. One way is to use the Theorem and find a single point at which the Jacobian has full rank, or, equivalently, look for a particular representation $F(u)$ at which $\left\{\frac{\partial F}{\partial t_j}(u)\right\}$ spans $H_d(\mathbb{C}^n)$.

5. Basic Definitions

There are two ways to show that F is a canonical form. One way is to use the Theorem and find a single point at which the Jacobian has full rank, or, equivalently, look for a particular representation $F(u)$ at which $\left\{\frac{\partial F}{\partial t_j}(u)\right\}$ spans $H_d(\mathbb{C}^n)$.

This has often be done via apolarity, see below. With the apolarity interpretation, this is known classically as the Lasker-Wakeford Theorem. A beautiful modern version is given in Ehrenborg-Rota.

5. Basic Definitions

There are two ways to show that F is a canonical form. One way is to use the Theorem and find a single point at which the Jacobian has full rank, or, equivalently, look for a particular representation $F(u)$ at which $\left\{\frac{\partial F}{\partial t_j}(u)\right\}$ spans $H_d(\mathbb{C}^n)$.

This has often be done via apolarity, see below. With the apolarity interpretation, this is known classically as the Lasker-Wakeford Theorem. A beautiful modern version is given in Ehrenborg-Rota.

The second (and better way, if possible) is to give a constructive algorithm for writing a general form in $H_d(\mathbb{C}^n)$ in the shape $F(u)$.

6. Apolarity

I sometimes call the following the "obvious inner product"; it can be found in ancient invariant theory, and analysts call it the "Fisher inner product". Either way, it's at least 120 years old.

6. Apolarity

I sometimes call the following the "obvious inner product"; it can be found in ancient invariant theory, and analysts call it the "Fisher inner product". Either way, it's at least 120 years old. Recall that

$$p(x_1 \dots, x_n) = \sum_{i \in \mathcal{I}(n,d)} c(i) a(p;i) x^i.$$

We now define, for $p, q \in H_d(\mathbb{C}^n)$:

$$[p, q] = \sum_{i \in \mathcal{I}(n,d)} c(i) a(p;i) a(q;i).$$

(Ehrenborg-Rota use this extensively, but never explicitly.)

6. Apolarity

I sometimes call the following the "obvious inner product"; it can be found in ancient invariant theory, and analysts call it the "Fisher inner product". Either way, it's at least 120 years old. Recall that

$$p(x_1 \dots, x_n) = \sum_{i \in \mathcal{I}(n,d)} c(i) a(p; i) x^i.$$

We now define, for $p, q \in H_d(\mathbb{C}^n)$:

$$[p, q] = \sum_{i \in \mathcal{I}(n,d)} c(i) a(p; i) a(q; i).$$

(Ehrenborg-Rota use this extensively, but never explicitly.)
This is only an inner product for real forms; for complex forms you need $\overline{a(q; i)}$. The conjugate actually only makes our expressions more complicated, $[p, q]$ is really just a bilinear form on $H_d(\mathbb{C}^n)$.

6. Apolarity

Definition

$p, q \in H_d(\mathbb{C}^n)$ are **apolar** if $[p, q] = 0$.

6. Apolarity

Definition

$p, q \in H_d(\mathbb{C}^n)$ are **apolar** if $[p, q] = 0$.

6. Apolarity

Definition

$p, q \in H_d(\mathbb{C}^n)$ are **apolar** if $[p, q] = 0$.

For $\alpha \in \mathbb{C}^n$, define $(\alpha\cdot)^d \in H_d(\mathbb{C}^n)$ by

$$(\alpha\cdot)^d(x) = (\alpha \cdot x)^d = \left(\sum_{j=1}^{n} \alpha_j x_j\right)^d = \sum_{i \in \mathcal{I}(n,d)} c(i)\alpha^i x^i,$$

where the usual multinomial conventions apply. We define the differential operator $q(D)$ for $q \in H_e(\mathbb{C}^n)$ in the usual way by

$$q(D) = \sum_{i \in \mathcal{I}(n,e)} c(i)a(q;i)\left(\frac{\partial}{\partial x_1}\right)^{i_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{i_n}.$$

The reason the obvious inner product is so useful is that it has many nice properties; all can be verified formally.

6. Apolarity

The inner product satisfies these properties:

- $[p, q] = [q, p]$.

6. Apolarity

The inner product satisfies these properties:

- $[p, q] = [q, p]$.
- $p(\alpha) = [p, (\alpha \cdot)^d]$.

6. Apolarity

The inner product satisfies these properties:

- $[p, q] = [q, p]$.
- $p(\alpha) = [p, (\alpha \cdot)^d]$.
- If $e = d$, then $p(D)q = d![p, q] = q(D)p$.

6. Apolarity

The inner product satisfies these properties:

- $[p, q] = [q, p]$.
- $p(\alpha) = [p, (\alpha\cdot)^d]$.
- If $e = d$, then $p(D)q = d![p, q] = q(D)p$.
- If $f \in H_e(\mathbb{C}^n)$ and $g \in H_{d-e}(\mathbb{C}^n)$, then

$$d![fg, p] = (fg)(D)p = f(D)g(D)p = e![f, g(D)p]$$

6. Apolarity

The inner product satisfies these properties:

- $[p, q] = [q, p]$.
- $p(\alpha) = [p, (\alpha\cdot)^d]$.
- If $e = d$, then $p(D)q = d![p, q] = q(D)p$.
- If $f \in H_e(\mathbb{C}^n)$ and $g \in H_{d-e}(\mathbb{C}^n)$, then

$$d![fg, p] = (fg)(D)p = f(D)g(D)p = e![f, g(D)p]$$

- $\frac{1}{d}\frac{\partial p}{\partial x_j}(\alpha) = [p, x_j(\alpha\cdot)^{d-1}]$, etc.

6. Apolarity

The inner product satisfies these properties:

- $[p, q] = [q, p]$.
- $p(\alpha) = [p, (\alpha\cdot)^d]$.
- If $e = d$, then $p(D)q = d![p, q] = q(D)p$.
- If $f \in H_e(\mathbb{C}^n)$ and $g \in H_{d-e}(\mathbb{C}^n)$, then

$$d![fg, p] = (fg)(D)p = f(D)g(D)p = e![f, g(D)p]$$

- $\frac{1}{d}\frac{\partial p}{\partial x_j}(\alpha) = [p, x_j(\alpha\cdot)^{d-1}]$, etc.
- If $e \le d$ and $g \in H_{d-e}(\mathbb{C}^n)$, then

$$g(D)(\alpha\cdot)^d = \frac{d!}{e!} g(\alpha)(\alpha\cdot)^e.$$

6. Apolarity

The consequences of this definition include:

- p is apolar to $(\alpha\cdot)^d$ iff $p(\alpha) = 0$.

6. Apolarity

The consequences of this definition include:

- p is apolar to $(\alpha\cdot)^d$ iff $p(\alpha) = 0$.
- p is apolar to $g(x)(\alpha\cdot)^{d-e}$ for every $g \in H_e(\mathbb{C}^n)$ iff all e-th order derivatives of p vanish at α iff p vanishes to e-th order at α.

6. Apolarity

The consequences of this definition include:

- p is apolar to $(\alpha\cdot)^d$ iff $p(\alpha) = 0$.
- p is apolar to $g(x)(\alpha\cdot)^{d-e}$ for every $g \in H_e(\mathbb{C}^n)$ iff all e-th order derivatives of p vanish at α iff p vanishes to e-th order at α.
- If $\deg q \le \deg p$ and $q(D)p = 0$, then all multiples of q in $H_d(\mathbb{C}^n)$ are apolar to p.

6. Apolarity

The consequences of this definition include:

- p is apolar to $(\alpha\cdot)^d$ iff $p(\alpha) = 0$.
- p is apolar to $g(x)(\alpha\cdot)^{d-e}$ for every $g \in H_e(\mathbb{C}^n)$ iff all e-th order derivatives of p vanish at α iff p vanishes to e-th order at α.
- If $\deg q \le \deg p$ and $q(D)p = 0$, then all multiples of q in $H_d(\mathbb{C}^n)$ are apolar to p.
- Classically, if p and q are forms of possibly different degree, p is apolar to q if $p(D)q = 0$. The definitions coincide when the degrees are equal, but not otherwise. In that case, the definition is not symmetric: if $\deg p > \deg q$, then $p(D)q$ will always equal 0.

7. Why doesn't constant-counting work?

Why are canonical forms even an issue? The main reason is that maps which one would think have full range don't. Apart from sums of squares, where the orthogonal group plays a role, the simplest example occurs in $H_4(\mathbb{C}^3)$. Since $N(3,4) = \binom{6}{2} = 15$, one expects that a general ternary quartic could be written as

$$p(x_1, x_2, x_3) = \sum_{k=1}^{5} (\alpha_{k1}x_1 + \alpha_{k2}x_2 + \alpha_{k3}x_3)^4$$

7. Why doesn't constant-counting work?

Why are canonical forms even an issue? The main reason is that maps which one would think have full range don't. Apart from sums of squares, where the orthogonal group plays a role, the simplest example occurs in $H_4(\mathbb{C}^3)$. Since $N(3,4) = \binom{6}{2} = 15$, one expects that a general ternary quartic could be written as

$$p(x_1, x_2, x_3) = \sum_{k=1}^{5} (\alpha_{k1}x_1 + \alpha_{k2}x_2 + \alpha_{k3}x_3)^4$$

This would be a canonical form, if the partials with respect to the α_{kj}'s at some chosen value would span $H_4(\mathbb{C}^3)$. By apolarity, this means that there should be no non-zero quartic which is singular at the five points $\alpha_k = (\alpha_{k1}, \alpha_{k2}, \alpha_{k3})$.

7. Why doesn't constant-counting work?

Why are canonical forms even an issue? The main reason is that maps which one would think have full range don't. Apart from sums of squares, where the orthogonal group plays a role, the simplest example occurs in $H_4(\mathbb{C}^3)$. Since $N(3,4) = \binom{6}{2} = 15$, one expects that a general ternary quartic could be written as

$$p(x_1, x_2, x_3) = \sum_{k=1}^{5} (\alpha_{k1}x_1 + \alpha_{k2}x_2 + \alpha_{k3}x_3)^4$$

This would be a canonical form, if the partials with respect to the α_{kj}'s at some chosen value would span $H_4(\mathbb{C}^3)$. By apolarity, this means that there should be no non-zero quartic which is singular at the five points $\alpha_k = (\alpha_{k1}, \alpha_{k2}, \alpha_{k3})$.

However, as Clebsch argued in the 1860's, since $N(3,2) = 6$, any choice of five α_k's pass through a non-zero quadratic $h(x_1, x_2, x_3)$, and so h^2 will be apolar to all the partials and a sum of five 4th powers is not a canonical form.

7. Why doesn't constant-counting work?

A few years later, Sylvester gave another proof. Given

$$p(x_1, x_2, x_3) = \sum_{r+s+t=4} \frac{4!}{r!s!t!} a_{rst} x_1^r x_2^s x_3^t,$$

define the catalecticant H_p as a quadratic form in 6 variables (or a 6×6 symmetric matrix defined linearly in terms of p).

$$H_p = \begin{pmatrix} a_{400} & a_{220} & a_{202} & a_{310} & a_{301} & a_{211} \\ a_{220} & a_{040} & a_{022} & a_{130} & a_{121} & a_{031} \\ a_{202} & a_{022} & a_{004} & a_{112} & a_{103} & a_{013} \\ a_{310} & a_{130} & a_{112} & a_{220} & a_{211} & a_{121} \\ a_{301} & a_{121} & a_{103} & a_{211} & a_{202} & a_{112} \\ a_{211} & a_{031} & a_{013} & a_{121} & a_{112} & a_{022} \end{pmatrix}$$

7. Why doesn't constant-counting work?

Under this definition, $H_{(\alpha\cdot)^4}$ is a perfect square. Thus if p is a sum of five fourth powers, then $rank(H_p) \le 5$, so H_p is singular. This can't happen for a general ternary quartic, for which the determinant is non-zero. This gives the algebraic relation of the coefficients in Clebsch's proof.

7. Why doesn't constant-counting work?

Under this definition, $H_{(\alpha\cdot)^4}$ is a perfect square. Thus if p is a sum of five fourth powers, then $rank(H_p) \le 5$, so H_p is singular. This can't happen for a general ternary quartic, for which the determinant is non-zero. This gives the algebraic relation of the coefficients in Clebsch's proof.
Clebsch's proof and Sylvester's proof are really the same, because

$$H_p(t_1, \dots, t_6) =$$
$$[(t_1x_1^2 + t_2x_2^2 + t_3x_3^2 + t_4x_1x_2 + t_5x_1x_3 + t_6x_2x_3)^2(D)]p.$$

7. Why doesn't constant-counting work?

Under this definition, $H_{(\alpha\cdot)^4}$ is a perfect square. Thus if p is a sum of five fourth powers, then $rank(H_p) \le 5$, so H_p is singular. This can't happen for a general ternary quartic, for which the determinant is non-zero. This gives the algebraic relation of the coefficients in Clebsch's proof.
Clebsch's proof and Sylvester's proof are really the same, because

$$H_p(t_1, \dots, t_6) =$$
$$[(t_1 x_1^2 + t_2 x_2^2 + t_3 x_3^2 + t_4 x_1 x_2 + t_5 x_1 x_3 + t_6 x_2 x_3)^2(D)]p.$$

Our 19th century ancestors saw that funny things happen when $(n, d) = (3, 4), (4, 4), (5, 4), (5, 3)$. In the early 1990s, Alexander and Hirschowitz proved that these are the only cases these funny things can happen. And since this starts to become serious algebraic geometry, it's time to change topics.

And now, a biographical interlude.

Emanuel Lasker (1868-1941) received his Ph.D. under Max Noether at Göttingen in 1902. He first developed the concept of a primary ideal and proved the primary decomposition theorem for an ideal of a polynomial ring in terms of primary ideals.

8. The Lasker-Wakeford Theorem

And now, a biographical interlude.

Emanuel Lasker (1868-1941) received his Ph.D. under Max Noether at Göttingen in 1902. He first developed the concept of a primary ideal and proved the primary decomposition theorem for an ideal of a polynomial ring in terms of primary ideals.

He is probably better known for being the world chess champion for 27 years (1894-1921), which spanned the life of ...

8. The Lasker-Wakeford Theorem

And now, a biographical interlude.

Emanuel Lasker (1868-1941) received his Ph.D. under Max Noether at Göttingen in 1902. He first developed the concept of a primary ideal and proved the primary decomposition theorem for an ideal of a polynomial ring in terms of primary ideals.

He is probably better known for being the world chess champion for 27 years (1894-1921), which spanned the life of ...

Edward Kingsley Wakeford (1894-1916).

8. The Lasker-Wakeford Theorem

And now, a biographical interlude.

Emanuel Lasker (1868-1941) received his Ph.D. under Max Noether at Göttingen in 1902. He first developed the concept of a primary ideal and proved the primary decomposition theorem for an ideal of a polynomial ring in terms of primary ideals.

He is probably better known for being the world chess champion for 27 years (1894-1921), which spanned the life of ...

Edward Kingsley Wakeford (1894-1916).

If you remember your European history, those dates will give you pause.

8. The Lasker-Wakeford Theorem

The memorial article by J. H. Grace about Wakeford in the *Proceedings of the London Mathematical Society* may be the angriest obituary I've ever read in a scholarly journal, and it can be found in its entirety on my webpage:

http://www.math.uiuc.edu/∼reznick/wakeford.pdf

8. The Lasker-Wakeford Theorem

The memorial article by J. H. Grace about Wakeford in the *Proceedings of the London Mathematical Society* may be the angriest obituary I've ever read in a scholarly journal, and it can be found in its entirety on my webpage:

http://www.math.uiuc.edu/∼reznick/wakeford.pdf

"He [EKW] was slightly wounded early in 1916, and soon after coming home was busy again with Canonical Forms.... [H]e discovered a paper of Hilbert's which contained the very theorem he had long been in want of – first vaguely, and later quite definitely. This was in March; April found him, full of the most joyous and reverential admiration for the great German master, working away in fearful haste to finish the dissertation ...

8. The Lasker-Wakeford Theorem

The memorial article by J. H. Grace about Wakeford in the *Proceedings of the London Mathematical Society* may be the angriest obituary I've ever read in a scholarly journal, and it can be found in its entirety on my webpage:

http://www.math.uiuc.edu/~reznick/wakeford.pdf

"He [EKW] was slightly wounded early in 1916, and soon after coming home was busy again with Canonical Forms.... [H]e discovered a paper of Hilbert's which contained the very theorem he had long been in want of – first vaguely, and later quite definitely. This was in March; April found him, full of the most joyous and reverential admiration for the great German master, working away in fearful haste to finish the dissertation ...

He returned to the front in June and was killed in July.... He only needed a chance, and he never got it."

8. The Lasker-Wakeford Theorem

I found this terminology in the text *The theory of determinants, matrices and invariants* by H. W. Turnbull, who was by some accounts the last old-style invariant theorist. Turnbull was a good mathematician and his book is a real Rosetta Stone for understanding 19th century algebra. He described the theorem as "paradoxical and very curious". I will rephrase it for simplicity.

Theorem (Lasker-Wakeford)

If $F : \mathbb{C}^M \to H_d(\mathbb{C}^n)$ $(M \geq N(n,d))$ then F is a canonical form if and only if there is a point $u \in \mathbb{C}^M$ so that there is no non-zero form q which is apolar to all N forms $\{\frac{\partial F}{\partial t_1}(u), \ldots, \frac{\partial F}{\partial t_N}(u)\}$.

Most writers restrict themselves to $M = N(n,d)$, as do I.

8. The Lasker-Wakeford Theorem

The point is simply that $\{\frac{\partial F}{\partial t_1}(u), \ldots, \frac{\partial F}{\partial t_N}(u)\}$ spans $H_d(\mathbb{C}^n)$ if and only if its perp is $\{0\}$. To a large extent, an appeal to apolarity is unnecessary. Except that, for binary forms, zeros correspond to linear factors and counting is often all we need to do.

The classical "Fundamental Theorem of Apolarity" can now be easily stated and proved. I don't know how it was understood before the Nullstellensatz. The theorem applies even when $e > d$.

Theorem (FTA)

Suppose $q \in H_e(\mathbb{C}^n)$ is irreducible and $p \in H_d(\mathbb{C}^n)$. Then $q(D)p = 0$ iff there exist $\alpha_k \subset \{\alpha : q(\alpha) = 0\}$ and $\lambda_k \in \mathbb{C}$ such that

$$p(x) = \sum_{k=1}^{m} \lambda_k (\alpha_k \cdot x)^d.$$

8. The Lasker-Wakeford Theorem

About 15 years ago, I proved a fairly obvious generalization of this theorem when q is not irreducible. The proof is, in spirit, the same as the one given above, both are in the Iowa beamer notes on my webpage: http://www.math.uiuc.edu/∼reznick/iowa-31911F.pdf

8. The Lasker-Wakeford Theorem

About 15 years ago, I proved a fairly obvious generalization of this theorem when q is not irreducible. The proof is, in spirit, the same as the one given above, both are in the Iowa beamer notes on my webpage: http://www.math.uiuc.edu/~reznick/iowa-31911F.pdf

Theorem

Suppose $q \in H_e(\mathbb{C}^n)$ *factors as* $\prod_{j=1}^r q_j^{m_j}$ *into a product of distinct irreducible factors and suppose* $p \in H_d(\mathbb{C}^n)$*. Then* $q(D)p = 0$ *iff there exist* $\alpha_{jk} \subset \{q_j(\alpha) = 0\}$*, and* $\phi_{jk} \in H_{m_j - 1}(\mathbb{C}^n)$ *such that*

$$p(x) = \sum_{j=1}^{r} \left(\sum_{k=1}^{n_j} \phi_{jk}(x)(\alpha_{kj} \cdot x)^{d-(m_j-1)} \right).$$

The *Mathematical Reviews* comment on this paper included the sentence: "The proof is remarkably elementary." That's my all-time favorite review!

8. The Lasker-Wakeford Theorem

In the case of binary forms, where zeros correspond to linear factors, the FTA is even simpler, and can be made equivalent to Gundelfinger's generalization of Sylvester's canonical forms. This next theorem also appears in Ehrenborg-Rota.

Theorem

Suppose $\sum m_j = d + 1$ and let $\ell_i(x, y) = \alpha_i x + \beta_i y$. Suppose further that ℓ_i and ℓ_j are pairwise linearly independent for $i \neq j$. Then the following set is a basis for $H_d(\mathbb{C}^2)$:

$$\left\{ x^{(m_j-1)-k} y^k (\beta_j x - \alpha_j y)^{d-(m_j-1)} : 0 \le k \le m_j - 1 \right\}$$

8. The Lasker-Wakeford Theorem

Remember canonical forms? Suppose $t_1, \dots, t_n$ appear in a canonical form as

$$(t_1 x_1 + \cdots + t_n x_n)^d = \ell^d.$$

Then $\frac{\partial F}{\partial t_j} = d x_j \ell^{d-1}$, and in applying Lasker-Wakeford, note that a form is apolar to each of these if and only if it is singular at $(t_1, \dots, t_n)$. Start thinking about general forms which are singular at general sets of points and you enter the context in which Alexander-Hirschowitz comes in.

8. The Lasker-Wakeford Theorem

Remember canonical forms? Suppose $t_1, \ldots, t_n$ appear in a canonical form as

$$(t_1 x_1 + \cdots + t_n x_n)^d = \ell^d.$$

Then $\frac{\partial F}{\partial t_j} = d x_j \ell^{d-1}$, and in applying Lasker-Wakeford, note that a form is apolar to each of these if and only if it is singular at $(t_1, \ldots, t_n)$. Start thinking about general forms which are singular at general sets of points and you enter the context in which Alexander-Hirschowitz comes in.

The rest of this talk will consider first new canonical forms for binary forms, and then for cubic forms.

9. Sylvester's canonical Forms

In 1869, Sylvester (1814-1897) reflected on the discovery of his canonical forms, done while he was working as an actuary in the office next to Cayley's.

9. Sylvester's canonical Forms

In 1869, Sylvester (1814-1897) reflected on the discovery of his canonical forms, done while he was working as an actuary in the office next to Cayley's.

"I discovered and developed the whole theory of canonical binary forms for odd degrees, and, as far as yet made out, for even degrees too, at one evening sitting, with a decanter of port wine to sustain nature's flagging energies, in a back office in Lincoln's Inn Fields. The work was done, and well done, but at the usual cost of racking thought — a brain on fire, and feet feeling, or feelingless, as if plunged in an ice-pail. That night we slept no more"

9. Sylvester's canonical Forms

In 1869, Sylvester (1814-1897) reflected on the discovery of his canonical forms, done while he was working as an actuary in the office next to Cayley's.

"I discovered and developed the whole theory of canonical binary forms for odd degrees, and, as far as yet made out, for even degrees too, at one evening sitting, with a decanter of port wine to sustain nature's flagging energies, in a back office in Lincoln's Inn Fields. The work was done, and well done, but at the usual cost of racking thought — a brain on fire, and feet feeling, or feelingless, as if plunged in an ice-pail. That night we slept no more"

The "$\lambda\ell^{2k}$" must be what Sylvester meant by "as far as yet made out". As if anticipating modern mathematical preferences, Sylvester proved his theory with one brilliant algorithm.

9. Sylvester's canonical forms

Theorem (Sylvester)

Suppose $p(x,y) = \sum_{j=0}^{d} \binom{d}{j} a_j x^{d-j} y^j$ *and* $h(x,y) = \sum_{t=0}^{r} c_t x^{r-t} y^t = \prod_{j=1}^{r} (\beta_j x - \alpha_j y)$ *is a product of pairwise distinct linear factors. Then there exist* $\lambda_k \in \mathbb{C}$ *so that*

$$p(x,y) = \sum_{k=1}^{r} \lambda_k (\alpha_k x + \beta_k y)^d$$

if and only if

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_r \\ a_1 & a_2 & \cdots & a_{r+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d-r} & a_{d-r+1} & \cdots & a_d \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

9. Sylvester's canonical forms

Here is an example of Sylvester's algorithm in action. Let

$$p(x,y) = 3x^5 - 20x^3y^2 + 10xy^4 =$$

$$\binom{5}{0}\cdot 3\ x^5 + \binom{5}{1}\cdot 0\ x^4y + \binom{5}{2}\cdot(-2)\ x^3y^2$$

$$+\binom{5}{3}\cdot 0\ x^2y^3 + \binom{5}{4}\cdot 2\ xy^4 + \binom{5}{5}\cdot 0\ y^5;$$

$$\begin{pmatrix} 3 & 0 & -2 & 0 \\ 0 & -2 & 0 & 2 \\ -2 & 0 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

We have $h(x,y) = y(x^2+y^2) = y(y-ix)(y+ix)$.

9. Sylvester's canonical forms

Here is an example of Sylvester's algorithm in action. Let

$$p(x,y) = 3x^5 - 20x^3y^2 + 10xy^4 =$$

$$\binom{5}{0}\cdot 3\ x^5 + \binom{5}{1}\cdot 0\ x^4y + \binom{5}{2}\cdot(-2)\ x^3y^2$$

$$+\binom{5}{3}\cdot 0\ x^2y^3 + \binom{5}{4}\cdot 2\ xy^4 + \binom{5}{5}\cdot 0\ y^5;$$

$$\begin{pmatrix} 3 & 0 & -2 & 0 \\ 0 & -2 & 0 & 2 \\ -2 & 0 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

We have $h(x,y) = y(x^2+y^2) = y(y-ix)(y+ix)$. Accordingly, there exist $\lambda_k \in \mathbb{C}$ so that

$$p(x,y) = \lambda_1 x^5 + \lambda_2(x+iy)^5 + \lambda_3(x-iy)^5.$$

Indeed, $\lambda_1 = \lambda_2 = \lambda_3 = 1$, as may be checked.

9. Sylvester's canonical forms

A few remarks about Sylvester's algorithm

- If $h(D) = \prod_{j=1}^{r}(\beta_j \frac{\partial}{\partial x} - \alpha_j \frac{\partial}{\partial y}) = \sum_{t=0}^{r} c_t \frac{\partial^r}{\partial x^{r-t} \partial y^t}$, then

$$h(D)p = \sum_{m=0}^{d-r} \frac{d!}{(d-r-m)!m!} \left(\sum_{i=0}^{d-r} a_{i+m} c_i \right) x^{d-r-m} y^m$$

  The coefficients of $h(D)p$ are, up to multiple, the rows in the matrix product, so the matrix condition is $h(D)p = 0$. The algorithm is the FTA configured for products of linear factors.

9. Sylvester's canonical forms

A few remarks about Sylvester's algorithm

- If $h(D) = \prod_{j=1}^{r}(\beta_j \frac{\partial}{\partial x} - \alpha_j \frac{\partial}{\partial y}) = \sum_{t=0}^{r} c_t \frac{\partial^r}{\partial x^{r-t} \partial y^t}$, then

$$h(D)p = \sum_{m=0}^{d-r} \frac{d!}{(d-r-m)!m!} \left(\sum_{i=0}^{d-r} a_{i+m} c_i \right) x^{d-r-m} y^m$$

  The coefficients of $h(D)p$ are, up to multiple, the rows in the matrix product, so the matrix condition is $h(D)p = 0$. The algorithm is the FTA configured for products of linear factors.
- An alternate proof of Sylvester's Theorem is basically equivalent to computing the solution of constant-coefficient linear recurrence sequences.

9. Sylvester's canonical forms

A few remarks about Sylvester's algorithm

- If $h(D) = \prod_{j=1}^{r}(\beta_j \frac{\partial}{\partial x} - \alpha_j \frac{\partial}{\partial y}) = \sum_{t=0}^{r} c_t \frac{\partial^r}{\partial x^{r-t} \partial y^t}$, then
$$h(D)p = \sum_{m=0}^{d-r} \frac{d!}{(d-r-m)!m!} \left(\sum_{i=0}^{d-r} a_{i+m} c_i \right) x^{d-r-m} y^m$$
The coefficients of $h(D)p$ are, up to multiple, the rows in the matrix product, so the matrix condition is $h(D)p = 0$. The algorithm is the FTA configured for products of linear factors.
- An alternate proof of Sylvester's Theorem is basically equivalent to computing the solution of constant-coefficient linear recurrence sequences.
- If $d = 2s - 1$ and $r = s$, then the matrix is $s \times (s+1)$ and has a non-trivial null-vector. The corresponding h (which can be given in terms of the coefficients of p) has distinct factors unless its discriminant vanishes. This gives the canonical form in odd degree.

- If $d = 2s$ and $r = s$, then the matrix is square, and in general, there exists λ so that $p(x, y) - \lambda x^{2s}$ has a matrix with a non-trivial null-vector as above. This gives the canonical form in even degree. That extra wobble is what Sylvester must have meant by "as far as yet made out".

9. Sylvester's canonical forms

- If $d = 2s$ and $r = s$, then the matrix is square, and in general, there exists λ so that $p(x, y) - \lambda x^{2s}$ has a matrix with a non-trivial null-vector as above. This gives the canonical form in even degree. That extra wobble is what Sylvester must have meant by "as far as yet made out".
- In this even case, the determinant of the square matrix is the *catalecticant*. Sylvester apologized for introducing this term: "Meicatalecticizant would more completely express the meaning of that which, for the sake of brevity, I denominate the catalecticant." Sylvester was very interested in the technical aspects of poetry and a "catalectic" verse is one in which the last line is missing a foot.

9. Sylvester's canonical forms

- If $d = 2s$ and $r = s$, then the matrix is square, and in general, there exists λ so that $p(x, y) - \lambda x^{2s}$ has a matrix with a non-trivial null-vector as above. This gives the canonical form in even degree. That extra wobble is what Sylvester must have meant by "as far as yet made out".
- In this even case, the determinant of the square matrix is the *catalecticant*. Sylvester apologized for introducing this term: "Meicatalecticizant would more completely express the meaning of that which, for the sake of brevity, I denominate the catalecticant." Sylvester was very interested in the technical aspects of poetry and a "catalectic" verse is one in which the last line is missing a foot.
- To his credit, in the same paper, Sylvester introduced the term "unimodular" in its current meaning.

10. New steampunk canonical forms

Constant-counting works for a wide range of binary forms:

10. New steampunk canonical forms

Constant-counting works for a wide range of binary forms:

Theorem

Suppose $d \geq 1$, $\ell_j(x,y) = \beta_j x + \gamma_j y$, $1 \leq j \leq m$, *are fixed pairwise non-proportional linear forms and suppose* $e_k \mid d$, $1 \leq k \leq r$ *and* $m + \sum_{k=1}^{r}(e_k + 1) = d + 1$. *Then a general binary form of degree* d *can be written as*

$$p(x,y) = \sum_{j=1}^{m} c_j \ell_j^d(x,y) + \sum_{k=1}^{r} f_k^{d/e_k}(x,y),$$

where $c_j \in \mathbb{C}$ *and* f_k *is a form of degree* e_k.

10. New steampunk canonical forms

Constant-counting works for a wide range of binary forms:

Theorem

Suppose $d \geq 1$, $\ell_j(x,y) = \beta_j x + \gamma_j y$, $1 \leq j \leq m$, are fixed pairwise non-proportional linear forms and suppose $e_k \mid d$, $1 \leq k \leq r$ and $m + \sum_{k=1}^{r}(e_k + 1) = d + 1$. Then a general binary form of degree d can be written as

$$p(x,y) = \sum_{j=1}^{m} c_j \ell_j^d(x,y) + \sum_{k=1}^{r} f_k^{d/e_k}(x,y),$$

where $c_j \in \mathbb{C}$ and f_k is a form of degree e_k.

10. New steampunk canonical forms

Constant-counting works for a wide range of binary forms:

Theorem

Suppose $d \geq 1$, $\ell_j(x,y) = \beta_j x + \gamma_j y$, $1 \leq j \leq m$, *are fixed pairwise non-proportional linear forms and suppose* $e_k \mid d$, $1 \leq k \leq r$ *and* $m + \sum_{k=1}^{r}(e_k + 1) = d + 1$. *Then a general binary form of degree* d *can be written as*

$$p(x,y) = \sum_{j=1}^{m} c_j \ell_j^d(x,y) + \sum_{k=1}^{r} f_k^{d/e_k}(x,y),$$

where $c_j \in \mathbb{C}$ *and* f_k *is a form of degree* e_k.

This recovers Sylvester's canonical form, upon taking $r = \lfloor d/2 \rfloor$ and $e_k \equiv 1$, so that $m = 0$ if d is odd and $m = 1$ if d is even.

10. New steampunk canonical forms

Constant-counting works for a wide range of binary forms:

Theorem

Suppose $d \geq 1$, $\ell_j(x,y) = \beta_j x + \gamma_j y$, $1 \leq j \leq m$, are fixed pairwise non-proportional linear forms and suppose $e_k \mid d$, $1 \leq k \leq r$ and $m + \sum_{k=1}^{r}(e_k + 1) = d + 1$. Then a general binary form of degree d can be written as

$$p(x,y) = \sum_{j=1}^{m} c_j \ell_j^d(x,y) + \sum_{k=1}^{r} f_k^{d/e_k}(x,y),$$

where $c_j \in \mathbb{C}$ and f_k is a form of degree e_k.

This recovers Sylvester's canonical form, upon taking $r = \lfloor d/2 \rfloor$ and $e_k \equiv 1$, so that $m = 0$ if d is odd and $m = 1$ if d is even.

The novelty in these canonical forms is the existence of forms of intermediate degree taken to intermediate powers.

Proof.

The parameters are the m constants c_j and, for each k, the $e_k + 1$ coefficients of $f_k(x, y) = \sum_{u=0}^{e_k} \alpha_{ku} x^{e_k - u} y^u$.
The partials with respect to the c_j's are simply $\{\ell_1^d, \cdots, \ell_m^d\}$, and the partial with respect to α_{ku} is

$$(d/e_k) x^{e_k - u} y^u f_k^{d/e_k - 1}.$$

Proof.

The parameters are the m constants c_j and, for each k, the $e_k + 1$ coefficients of $f_k(x,y) = \sum_{u=0}^{e_k} \alpha_{ku} x^{e_k - u} y^u$.

The partials with respect to the c_j's are simply $\{\ell_1^d, \cdots, \ell_m^d\}$, and the partial with respect to α_{ku} is

$$(d/e_k) x^{e_k - u} y^u f_k^{d/e_k - 1}.$$

10. New steampunk canonical forms

Proof.

The parameters are the m constants c_j and, for each k, the $e_k + 1$ coefficients of $f_k(x, y) = \sum_{u=0}^{e_k} \alpha_{ku} x^{e_k - u} y^u$.
The partials with respect to the c_j's are simply $\{\ell_1^d, \cdots, \ell_m^d\}$, and the partial with respect to α_{ku} is

$$(d/e_k) x^{e_k - u} y^u f_k^{d/e_k - 1}.$$

Now evaluate the Jacobian at a choice of parameters so that $f_k(x, y) = \tilde{\ell}_k^{e_k}$, where the linear forms $\tilde{\ell}_k$ are chosen so that the combined set $\{\ell_j, \tilde{\ell}_k\}$ is pairwise linearly independent. Then $f_k^{d/e_k - 1} = \tilde{\ell}_k^{d - e_k}$, and it is taken times a basis of $H_{e_k}(\mathbb{C}^2)$. By an earlier theorem, this set, taken all together, is a basis for $H_d(\mathbb{C}^2)$ and so this is a canonical form. □

10. New steampunk canonical forms

Clearly, the most interesting canonical forms of this kind occur when $m = 0$, so that $e_k \mid d$ and $\sum_{k=1}^{r}(e_k + 1) = d + 1$. We wish to exclude the case $r = 1$, $e_1 = d$, because then the theorem is vacuous. Otherwise, write $e_k m_k = d$, where $m_k > 1$ and observe that $\sum_{k=1}^{r}(e_k + 1) = d + 1 \implies r - 1 + \sum_{k=1}^{r} \frac{d}{m_k} = d \implies 1 = \sum_{k=1}^{r} \frac{1}{m_k} + \sum_{j=1}^{r-1} \frac{1}{d}$. It is well-known, and not hard to prove, that for any fixed r there are only finitely many such "Egyptian fraction" decompositions, hence for any r there are only finitely new canonical forms with r terms.

10. New steampunk canonical forms

Clearly, the most interesting canonical forms of this kind occur when $m = 0$, so that $e_k \mid d$ and $\sum_{k=1}^{r}(e_k + 1) = d + 1$. We wish to exclude the case $r = 1$, $e_1 = d$, because then the theorem is vacuous. Otherwise, write $e_k m_k = d$, where $m_k > 1$ and observe that $\sum_{k=1}^{r}(e_k + 1) = d + 1 \implies r - 1 + \sum_{k=1}^{r} \frac{d}{m_k} = d \implies 1 = \sum_{k=1}^{r} \frac{1}{m_k} + \sum_{j=1}^{r-1} \frac{1}{d}$. It is well-known, and not hard to prove, that for any fixed r there are only finitely many such "Egyptian fraction" decompositions, hence for any r there are only finitely new canonical forms with r terms.

For $r = 2$, there are three: $(d, e_1, e_2) = (3, 1, 1), (4, 2, 1), (6, 3, 2)$.

10. New steampunk canonical forms

Clearly, the most interesting canonical forms of this kind occur when $m = 0$, so that $e_k \mid d$ and $\sum_{k=1}^{r}(e_k + 1) = d + 1$. We wish to exclude the case $r = 1$, $e_1 = d$, because then the theorem is vacuous. Otherwise, write $e_k m_k = d$, where $m_k > 1$ and observe that $\sum_{k=1}^{r}(e_k + 1) = d + 1 \implies r - 1 + \sum_{k=1}^{r} \frac{d}{m_k} = d \implies 1 = \sum_{k=1}^{r} \frac{1}{m_k} + \sum_{j=1}^{r-1} \frac{1}{d}$. It is well-known, and not hard to prove, that for any fixed r there are only finitely many such "Egyptian fraction" decompositions, hence for any r there are only finitely new canonical forms with r terms.

For $r = 2$, there are three: $(d, e_1, e_2) = (3, 1, 1), (4, 2, 1), (6, 3, 2)$. For $r = 3$, there are 22, the most exotic of which is that a general binary form of degree 84 can be written as the square of a form of degree 42 plus the cube of a form of degree 28 plus the seventh power of a form of degree 12: $43 + 29 + 13 = 85$. I won't be looking for an algorithm very soon.

10. New steampunk canonical forms

If we further assume that all e_k's are equal, then we get a simple situation; $e \mid d$ and $e+1 \mid d+1$. Let $f(d)$ denote the number of $e < d$ with this property. We see that

$$d \equiv 0 \pmod{e}, \qquad d \equiv -1 \pmod{(e+1)}$$
$$\implies d \equiv e \pmod{(e^2+e)}$$

by the Chinese Remainder Theorem, hence $d = e + te(e+1)$ for $t \geq 1$, and standard generating function techniques give

$$\sum_{d=1}^{\infty} f(d)x^d = \sum_{e=1}^{\infty}\sum_{t=1}^{\infty} x^{e+te(e+1)} = \sum_{e=1}^{\infty} \frac{x^{e^2+2e}}{1-x^{e^2+e}}.$$

10. New steampunk canonical forms

If we further assume that all e_k's are equal, then we get a simple situation; $e \mid d$ and $e+1 \mid d+1$. Let $f(d)$ denote the number of $e < d$ with this property. We see that

$$d \equiv 0 \pmod{e}, \qquad d \equiv -1 \pmod{(e+1)}$$
$$\implies d \equiv e \pmod{(e^2+e)}$$

by the Chinese Remainder Theorem, hence $d = e + te(e+1)$ for $t \geq 1$, and standard generating function techniques give

$$\sum_{d=1}^{\infty} f(d)x^d = \sum_{e=1}^{\infty}\sum_{t=1}^{\infty} x^{e+te(e+1)} = \sum_{e=1}^{\infty} \frac{x^{e^2+2e}}{1-x^{e^2+e}}.$$

It follows that if $F(N) = \sum_{d=1}^{N} f(d)$, then

$$F(N) = \sum_{e=1}^{\infty} \left\lfloor \frac{N-e}{e^2+e} \right\rfloor \implies F(N) = N + \mathcal{O}(N^{1/2}).$$

10. New steampunk canonical forms

That is, the "average" number of these really nice canonical forms is about one per degree. Half of them are just the Sylvester canonical forms for odd degree $2k-1$, with $e=1$ and $r=k$. The first new one is $d=8$, $e=2$, $r=3$: a general binary octic is a sum of three quadratics to the fourth power. The first time $f(d)=2$ is $d=15$; the first time $f(d)=3$ is $d=99$.

10. New steampunk canonical forms

That is, the "average" number of these really nice canonical forms is about one per degree. Half of them are just the Sylvester canonical forms for odd degree $2k-1$, with $e=1$ and $r=k$. The first new one is $d=8$, $e=2$, $r=3$: a general binary octic is a sum of three quadratics to the fourth power. The first time $f(d)=2$ is $d=15$; the first time $f(d)=3$ is $d=99$.

How large can $f(d)$ get? It's not too hard to show that $f(2^m-1)=d(m)-1$, where $d(m)$ denotes the usual divisor function; if $d=2^m-1$, then $e=2^j-1$, where $1\le j<m$ and $j\mid m$. This means that $(f(d))$ is unbounded and, taking $d=2^{2^u-1}$ we see that $f(d)$ is of order $\log\log d$ infinitely often. Empirical evidence suggests that this is too modest and $\log d$ might be possible. For $d\le 10^8$, the maximum value is

$$f(7316000)=12.$$

The data suggest that $(f(d))$ has an underlying distribution.

10. New steampunk canonical forms

We now look at specific instances of this new theorem.

- If $r = 0$ and $m = d + 1$, we get the canonical form for a basis.

10. New steampunk canonical forms

We now look at specific instances of this new theorem.

- If $r = 0$ and $m = d + 1$, we get the canonical form for a basis.
- If $e_k \equiv 1$, then Sylvester's algorithm can be adapted to show uniqueness.

10. New steampunk canonical forms

We now look at specific instances of this new theorem.

- If $r = 0$ and $m = d + 1$, we get the canonical form for a basis.
- If $e_k \equiv 1$, then Sylvester's algorithm can be adapted to show uniqueness.
- If $e_k \equiv 2$, an analogue to Sylvester's canonical forms occurs for general forms of even degree $d = 2k$: they are the sum of the k-th power of $\lfloor (d+1)/3 \rfloor$ quadratics plus a linear combination of any pre-specified $d - 3\lfloor (d+1)/3 \rfloor$ $2k$-th powers of linear forms. We don't have an algorithm for this. We want one. One problem is that it's easy to kill ℓ^d with a linear differential operator; $q^{d/2}$, not so much.

10. New steampunk canonical forms

We now look at specific instances of this new theorem.

- If $r = 0$ and $m = d + 1$, we get the canonical form for a basis.
- If $e_k \equiv 1$, then Sylvester's algorithm can be adapted to show uniqueness.
- If $e_k \equiv 2$, an analogue to Sylvester's canonical forms occurs for general forms of even degree $d = 2k$: they are the sum of the k-th power of $\lfloor (d+1)/3 \rfloor$ quadratics plus a linear combination of any pre-specified $d - 3\lfloor (d+1)/3 \rfloor$ $2k$-th powers of linear forms. We don't have an algorithm for this. We want one. One problem is that it's easy to kill ℓ^d with a linear differential operator; $q^{d/2}$, not so much.
- If $d = 4$, $m = 0$, $e_1 = 2$ and $e_2 = 1$, a general binary quartic can be written as the sum of the square of a quadratic form and the fourth power of a linear form. (More later.)

10. New steampunk canonical forms

If $d = 6$, $m = 0$, $e_1 = 3$ and $e_2 = 2$, then $4 + 3 = 7$ implies that a general binary sextic form can be written as the sum of the square of a cubic form and the cube of a quadratic form. We don't have an algorithm for doing this and we $(\text{really})^2$ want one! Mathematica computations suggest that sextic $p = f^2 + g^3$ for 40 different choices of $\{f^2, g^3\}$. Proofs are welcomed.

10. New steampunk canonical forms

If $d = 6$, $m = 0$, $e_1 = 3$ and $e_2 = 2$, then $4 + 3 = 7$ implies that a general binary sextic form can be written as the sum of the square of a cubic form and the cube of a quadratic form. We don't have an algorithm for doing this and we $(\text{really})^2$ want one! Mathematica computations suggest that sextic $p = f^2 + g^3$ for 40 different choices of $\{f^2, g^3\}$. Proofs are welcomed.

We have proved a general canonical form for even degree which Sylvester might have enjoyed.

Theorem

A general binary form of degree $d = 2k$, $k \geq 1$, can be written as

$$(\alpha_0 x^2 + \beta_0 xy + \gamma_0 y^2)^k + \sum_{j=1}^{k-1} (\alpha_j x + \beta_j y)^{2k}.$$

10. New steampunk canonical forms

How many ways can this be done?

10. New steampunk canonical forms

How many ways can this be done?

If $d = 2k = 4$, there are six different pairs (f^2, ℓ^4). Write $p = f^2 + g^2$, which can be done in essentially three ways. But the orthogonal group acts on a sum of two squares: $f^2 + g^2 = (uf + vg)^2 + (vf - ug)^2$ if $u^2 + v^2 = 1$, and for two choices, $vf - ug = \ell^2$ is itself a square: $3 \times 2 = 6$.

10. New steampunk canonical forms

How many ways can this be done?
If $d = 2k = 4$, there are six different pairs (f^2, ℓ^4). Write $p = f^2 + g^2$, which can be done in essentially three ways. But the orthogonal group acts on a sum of two squares:
$f^2 + g^2 = (uf + vg)^2 + (vf - ug)^2$ if $u^2 + v^2 = 1$, and for two choices, $vf - ug = \ell^2$ is itself a square: $3 \times 2 = 6$.

This theorem says that for $p \in F_{2k}(\mathbb{C}^n)$, there exists quadratic q so that $p - q^k$ is a sum of $k - 1$ $2k$-th powers, and by Sylvester's algorithm, this means that the $(k+2) \times k$ catalecticant matrix of $p - q^k$ has rank $k - 1$. Mathematica tests show that for $d = 2k = 6$ this happens in 22 ways, for $d = 2k = 8$ in 62 ways, for $d = 2k = 10$, in 147 ways, and one test (which took two days of computation) showed that for $d = 2k = 12$, the number is 308. There actually is a pattern: so far, the number is $2\binom{k+3}{5} - \binom{k+2}{3}$. No proof yet.

11. Quadratic forms

Every quadratic form $p \in H_2(\mathbb{C}^n)$ is a sum of n squares, but since the naive number of coefficients, $n \times n$, is $> N(n, 2) = \frac{n(n+1)}{2}$, a sum of n squares is not, *per se*, a canonical form. However, the standard "upper triangular" representation is a canonical form.

11. Quadratic forms

Every quadratic form $p \in H_2(\mathbb{C}^n)$ is a sum of n squares, but since the naive number of coefficients, $n \times n$, is $> N(n,2) = \frac{n(n+1)}{2}$, a sum of n squares is not, *per se*, a canonical form. However, the standard “upper triangular” representation is a canonical form. Let $\{t_{ij} : 1 \le i \le j \le n\}$ be the $N(n,2)$ parameters, and let

$$F(\{t_{ij}\})(x) := \sum_{i=1}^{n} L_i^2; \qquad L_i = t_{ii}x_i + \cdots + t_{in}x_n.$$

11. Quadratic forms

Every quadratic form $p \in H_2(\mathbb{C}^n)$ is a sum of n squares, but since the naive number of coefficients, $n \times n$, is $> N(n,2) = \frac{n(n+1)}{2}$, a sum of n squares is not, *per se*, a canonical form. However, the standard "upper triangular" representation is a canonical form. Let $\{t_{ij} : 1 \le i \le j \le n\}$ be the $N(n,2)$ parameters, and let

$$F(\{t_{ij}\})(x) := \sum_{i=1}^{n} L_i^2; \qquad L_i = t_{ii}x_i + \cdots + t_{in}x_n.$$

Then $\frac{\partial F}{\partial t_{ij}} = 2L_i x_j$, and if we specialize to $L_i = x_i$, then the set of partials is literally $\{2x_i x_j : 1 \le i \le j \le n\}$, which spans $H_2(\mathbb{C}^n)$.

11. Quadratic forms

A constructive proof is better, of course. Suppose $p \in H_2(\mathbb{C}^n)$ and $p(x) = \sum_i a_{ii} x_i^2 + 2 \sum_{i<j} a_{ij} x_i x_j$. Then

$$\frac{\partial p}{\partial x_1} = 2 \sum_{j=1}^{n} a_{1j} x_j.$$

11. Quadratic forms

A constructive proof is better, of course. Suppose $p \in H_2(\mathbb{C}^n)$ and $p(x) = \sum_i a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j$. Then

$$\frac{\partial p}{\partial x_1} = 2\sum_{j=1}^{n} a_{1j}x_j.$$

If $p(1,0,\dots,0) = a_{11} \neq 0$, which is generally true, define

$$q(x_1,\dots,x_n) = p(x_1,\dots,x_n) - \frac{1}{a_{11}}\left(\sum_{j=1}^{n} a_{1j}x_j\right)^2.$$

A constructive proof is better, of course. Suppose $p \in H_2(\mathbb{C}^n)$ and $p(x) = \sum_i a_{ii} x_i^2 + 2\sum_{i<j} a_{ij} x_i x_j$. Then

$$\frac{\partial p}{\partial x_1} = 2\sum_{j=1}^{n} a_{1j} x_j.$$

If $p(1,0,\ldots,0) = a_{11} \neq 0$, which is generally true, define

$$q(x_1,\ldots,x_n) = p(x_1,\ldots,x_n) - \frac{1}{a_{11}}\left(\sum_{j=1}^{n} a_{1j} x_j\right)^2.$$

Then

$$\frac{\partial q}{\partial x_1} = \frac{\partial p}{\partial x_1} - \frac{\partial p}{\partial x_1} = 0 \implies q = q(x_2,\ldots,x_n).$$

11. Quadratic forms

A constructive proof is better, of course. Suppose $p \in H_2(\mathbb{C}^n)$ and $p(x) = \sum_i a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j$. Then

$$\frac{\partial p}{\partial x_1} = 2\sum_{j=1}^{n} a_{1j}x_j.$$

If $p(1,0,\dots,0) = a_{11} \neq 0$, which is generally true, define

$$q(x_1,\dots,x_n) = p(x_1,\dots,x_n) - \frac{1}{a_{11}}\left(\sum_{j=1}^{n} a_{1j}x_j\right)^2.$$

Then

$$\frac{\partial q}{\partial x_1} = \frac{\partial p}{\partial x_1} - \frac{\partial p}{\partial x_1} = 0 \implies q = q(x_2,\dots,x_n).$$

Now just iterate this, losing one variable at a time, to get the traditional upper triangular sum of squares.

11. Quadratic forms

It's worth noting that *every* quadratic form in $H_2(\mathbb{C}^n)$ is a sum of n squares, and this can also be made algorithmic. Begin with

Theorem (Biermann's Theorem)

If $p \in H_d(\mathbb{C}^n)$ and $p(i) = 0$ for every $i \in \mathcal{I}(n, d)$, then $p = 0$.

11. Quadratic forms

It's worth noting that *every* quadratic form in $H_2(\mathbb{C}^n)$ is a sum of n squares, and this can also be made algorithmic. Begin with

Theorem (Biermann's Theorem)

If $p \in H_d(\mathbb{C}^n)$ and $p(i) = 0$ for every $i \in \mathcal{I}(n,d)$, then $p = 0$.

11. Quadratic forms

It's worth noting that *every* quadratic form in $H_2(\mathbb{C}^n)$ is a sum of n squares, and this can also be made algorithmic. Begin with

Theorem (Biermann's Theorem)

If $p \in H_d(\mathbb{C}^n)$ and $p(i) = 0$ for every $i \in \mathcal{I}(n,d)$, then $p = 0$.

This gives a finite set of $N(n,2)$ points to check for quadratic forms. Here's the algorithm. Given $p \in H_2(\mathbb{C}^n)$, index $\mathcal{I}(n,2)$ as you wish and look at $p(i)$. If this is always zero, then $p = 0$ and there's nothing to prove. Otherwise, take the first i at which $p(i) \neq 0$, and make an invertible linear change of variables taking $i \mapsto (1,0,\dots,0)$. Do the argument of the last slide, and get p as a square plus a quadratic form in $n-1$ variables. Iterate to get p as a sum of n squares.

12. Reichstein and canonically completing the cube

There is a wonderful non-trivial way to complete the cube, but almost nobody knows it. It appears in a paper by Boris Reichstein from 1987 which according to MathSciNet has had no citations. It is a truly beautiful theorem, though it was not transparently presented and was framed in the context of trilinear forms.

12. Reichstein and canonically completing the cube

There is a wonderful non-trivial way to complete the cube, but almost nobody knows it. It appears in a paper by Boris Reichstein from 1987 which according to MathSciNet has had no citations. It is a truly beautiful theorem, though it was not transparently presented and was framed in the context of trilinear forms.

I will prove this by giving Reichstein's constructive proof. First, here is some numerology about the number of cubes. By Alexander-Hirschowitz, for $n \neq 5$, a general cubic form in n variables can be written as a sum of $\lceil \frac{1}{n} N(n,3) \rceil = \left\lceil \frac{1}{n}\binom{n+2}{3} \right\rceil = \left\lceil \frac{(n+1)(n+2)}{6} \right\rceil$ cubes. (For $n = 5$, you need $\lceil \frac{6\cdot 7}{6} \rceil + 1$ cubes.)

12. Reichstein and canonically completing the cube

Reichstein's Theorem writes a general cubic form as

$$\sum_{k=1}^{n}(\alpha_{k1}x_1 + \cdots + \alpha_{kn}x_n)^3 + q(x_3, \ldots, x_n).$$

This is a sum of $\sum_{0 \le k \le n/2}(n-2k) = \lfloor \frac{(n+1)^2}{4} \rfloor$ cubes, which is, on average, about 50% larger than what is necessary.

12. Reichstein and canonically completing the cube

Reichstein's Theorem writes a general cubic form as

$$\sum_{k=1}^{n} (\alpha_{k1} x_1 + \cdots + \alpha_{kn} x_n)^3 + q(x_3, \dots, x_n).$$

This is a sum of $\sum_{0 \le k \le n/2} (n - 2k) = \lfloor \frac{(n+1)^2}{4} \rfloor$ cubes, which is, on average, about 50% larger than what is necessary.
But $N(n,3) - N(n-2,3) = \frac{n^3+3n^2+2n}{6} - \frac{n^3-3n^2+2n}{6} = n^2$, so that the total number of coefficients is

$$\sum_{0 \le k \le n/2} (n-2k)^2 = N(n,3),$$

showing that this is a potential canonical form.

12. Reichstein and canonically completing the cube

The validity can be verified by Lasker-Wakeford, specializing at $x_1, x_2, x_1 + kx_2 + x_k$ (for $k \geq 3$) for linear forms in $(x_1, \ldots, x_n)$, etc., but Reichstein's constructive proof is better.

12. Reichstein and canonically completing the cube

The validity can be verified by Lasker-Wakeford, specializing at $x_1, x_2, x_1 + kx_2 + x_k$ (for $k \geq 3$) for linear forms in $(x_1, \dots, x_n)$, etc., but Reichstein's constructive proof is better.
The proof requires a well-known fact: A general pair of quadratic forms can be simultaneously diagonalized. That is, if general $f, g \in H_2(\mathbb{C}^n)$ are given, then there exist n linearly independent forms $L_i(x) = \sum_{j=1}^n \alpha_{ij} x_j$ and $c_i \in \mathbb{C}$ so that

$$f = \sum_{i=1}^{n} L_i^2, \qquad g = \sum_{i=1}^{n} c_i L_i^2.$$

12. Reichstein and canonically completing the cube

The validity can be verified by Lasker-Wakeford, specializing at $x_1, x_2, x_1 + kx_2 + x_k$ (for $k \geq 3$) for linear forms in $(x_1, \ldots, x_n)$, etc., but Reichstein's constructive proof is better.

The proof requires a well-known fact: A general pair of quadratic forms can be simultaneously diagonalized. That is, if general $f, g \in H_2(\mathbb{C}^n)$ are given, then there exist n linearly independent forms $L_i(x) = \sum_{j=1}^n \alpha_{ij} x_j$ and $c_i \in \mathbb{C}$ so that

$$f = \sum_{i=1}^n L_i^2, \qquad g = \sum_{i=1}^n c_i L_i^2.$$

This can be made constructive. If $rank(f) = n$ and the determinant of the symmetric matrix associated with the pencil $f - \lambda g$ has n distinct roots $\{c_i\}$, then each $f - c_i g$ is singular. Routine methods can then be used to find the L_i's.

12. Reichstein and canonically completing the cube

We now prove Reichstein's Theorem. Suppose $p \in H_3(\mathbb{C}^n)$. We can generally simultaneously diagonalize $\frac{\partial p}{\partial x_1}$ and $\frac{\partial p}{\partial x_2}$: there exist linearly independent $L_i(x) = \sum_{j=1}^n \alpha_{ij} x_j$ and $c_i \in \mathbb{C}$ so that

$$\frac{\partial p}{\partial x_1} = \sum_{i=1}^n L_i^2, \qquad \frac{\partial p}{\partial x_2} = \sum_{i=1}^n c_i L_i^2.$$

12. Reichstein and canonically completing the cube

We now prove Reichstein's Theorem. Suppose $p \in H_3(\mathbb{C}^n)$. We can generally simultaneously diagonalize $\frac{\partial p}{\partial x_1}$ and $\frac{\partial p}{\partial x_2}$: there exist linearly independent $L_i(x) = \sum_{j=1}^n \alpha_{ij} x_j$ and $c_i \in \mathbb{C}$ so that

$$\frac{\partial p}{\partial x_1} = \sum_{i=1}^n L_i^2, \qquad \frac{\partial p}{\partial x_2} = \sum_{i=1}^n c_i L_i^2.$$

Since mixed partials are equal, we obtain the equation

$$\sum_{i=1}^n 2\alpha_{i2} L_i = \sum_{i=1}^n 2c_i \alpha_{i1} L_i,$$

and since the L_i's are linearly independent, $\alpha_{i2} = c_i \alpha_{i1}$. (This is important!)

12. Reichstein and canonically completing the cube

As before, it is generally true that $\alpha_{i1} \neq 0$ and we can let

$$q(x_1, \dots, x_n) = p(x_1, \dots, x_n) - \sum_{i=1}^{n} \frac{1}{3\alpha_{i1}} L_i^3$$

12. Reichstein and canonically completing the cube

As before, it is generally true that $\alpha_{i1} \neq 0$ and we can let

$$q(x_1, \dots, x_n) = p(x_1, \dots, x_n) - \sum_{i=1}^{n} \tfrac{1}{3\alpha_{i1}} L_i^3$$

$$\implies \frac{\partial q}{\partial x_1} = \frac{\partial p}{\partial x_1} - \sum_{i=1}^{n} \frac{3\alpha_{i1}}{3\alpha_{i1}} L_i^2 = 0,$$

12. Reichstein and canonically completing the cube

As before, it is generally true that $\alpha_{i1} \neq 0$ and we can let

$$q(x_1, \ldots, x_n) = p(x_1, \ldots, x_n) - \sum_{i=1}^{n} \tfrac{1}{3\alpha_{i1}} L_i^3$$

$$\implies \frac{\partial q}{\partial x_1} = \frac{\partial p}{\partial x_1} - \sum_{i=1}^{n} \frac{3\alpha_{i1}}{3\alpha_{i1}} L_i^2 = 0,$$

$$\frac{\partial q}{\partial x_2} = \frac{\partial p}{\partial x_2} - \sum_{i=1}^{n} \frac{3\alpha_{i2}}{3\alpha_{i1}} L_i^2 = \frac{\partial p}{\partial x_2} - \sum_{i=1}^{n} c_i L_i^2 = 0$$

12. Reichstein and canonically completing the cube

As before, it is generally true that $\alpha_{i1} \neq 0$ and we can let

$$q(x_1, \ldots, x_n) = p(x_1, \ldots, x_n) - \sum_{i=1}^{n} \tfrac{1}{3\alpha_{i1}} L_i^3$$

$$\implies \frac{\partial q}{\partial x_1} = \frac{\partial p}{\partial x_1} - \sum_{i=1}^{n} \frac{3\alpha_{i1}}{3\alpha_{i1}} L_i^2 = 0,$$

$$\frac{\partial q}{\partial x_2} = \frac{\partial p}{\partial x_2} - \sum_{i=1}^{n} \frac{3\alpha_{i2}}{3\alpha_{i1}} L_i^2 = \frac{\partial p}{\partial x_2} - \sum_{i=1}^{n} c_i L_i^2 = 0$$

$$\implies q = q(x_3, \ldots, x_n).$$

12. Reichstein and canonically completing the cube

As before, it is generally true that $\alpha_{i1} \neq 0$ and we can let

$$q(x_1,\dots,x_n) = p(x_1,\dots,x_n) - \sum_{i=1}^{n} \tfrac{1}{3\alpha_{i1}} L_i^3$$

$$\implies \frac{\partial q}{\partial x_1} = \frac{\partial p}{\partial x_1} - \sum_{i=1}^{n} \frac{3\alpha_{i1}}{3\alpha_{i1}} L_i^2 = 0,$$

$$\frac{\partial q}{\partial x_2} = \frac{\partial p}{\partial x_2} - \sum_{i=1}^{n} \frac{3\alpha_{i2}}{3\alpha_{i1}} L_i^2 = \frac{\partial p}{\partial x_2} - \sum_{i=1}^{n} c_i L_i^2 = 0$$

$$\implies q = q(x_3,\dots,x_n).$$

By iterating, we obtain Reichstein's form for cubics:

$$p(x_1,\dots,x_n) = \sum_{i=0}^{\lfloor (n-1)/2 \rfloor} \sum_{j=1}^{n-2i} \ell_{ij}^3(x_{1+2i},\dots,x_n).$$

13. Slinky

Recall Slinky:

$$p(x_1, \dots, x_n) = \sum_{1 \le i \le j \le n} (\alpha_{\{i,j\},i} x_i + \cdots + \alpha_{\{i,j\},j} x_j)^3.$$

13. Slinky

Recall Slinky:

$$p(x_1, \ldots, x_n) = \sum_{1 \le i \le j \le n} (\alpha_{\{i,j\},i} x_i + \cdots + \alpha_{\{i,j\},j} x_j)^3.$$

Slinky is potentially canonical, because $\sum_{k=1}^{n} k(n+1-k) = \binom{n+2}{3}$. You can probably guess by now how it's going to be proved.

13. Slinky

Recall Slinky:

$$p(x_1, \ldots, x_n) = \sum_{1 \le i \le j \le n} (\alpha_{\{i,j\},i} x_i + \cdots + \alpha_{\{i,j\},j} x_j)^3.$$

Slinky is potentially canonical, because $\sum_{k=1}^{n} k(n+1-k) = \binom{n+2}{3}$. You can probably guess by now how it's going to be proved. Given $p \in H_3(\mathbb{C}^n)$, $\frac{\partial p}{\partial x_n}$ is a quadratic form, so we can generally complete the square in the upper triangular way:

$$\frac{\partial p}{\partial x_n} = \sum_{j=1}^{n} (\alpha_{jj} x_j + \cdots + \alpha_{jn} x_n)^2.$$

13. Slinky

Let

$$q(x_1, \ldots, x_n) = p(x_1, \ldots, x_n) - \sum_{j=1}^{n} \frac{1}{3\alpha_{jn}}(\alpha_{jj}x_j + \cdots + \alpha_{jn}x_n)^3.$$

13. Slinky

Let

$$q(x_1,\ldots,x_n) = p(x_1,\ldots,x_n) - \sum_{j=1}^{n} \tfrac{1}{3\alpha_{jn}}(\alpha_{jj}x_j + \cdots + \alpha_{jn}x_n)^3.$$

Then

$$\frac{\partial q}{\partial x_n} = \frac{\partial p}{\partial x_n} - \frac{\partial p}{\partial x_n} = 0 \implies q = q(x_1,\ldots,x_{n-1}).$$

and repeat. We assume $\alpha_{jn} \neq 0$, etc., which is generally true. In this way, for each pair (i,j) with $1 \le i \le j \le n$, we get exactly one summand using only the x_k's with $i \le k \le j$.

13. Slinky

This last construction worked because in the upper diagonal sum of squares for quadratic forms, there is a variable, x_n, which appears in every summand. This is not the case for the cubic version, so there is no obvious way to bump it up to quartics.

13. Slinky

This last construction worked because in the upper diagonal sum of squares for quadratic forms, there is a variable, x_n, which appears in every summand. This is not the case for the cubic version, so there is no obvious way to bump it up to quartics.

The Reichstein form, on the other hand, **can** be generalized to quartics, in the same way, by integrating on the coefficient of x_n. One gets a general $p \in H_4(\mathbb{C}^n)$ as a sum of $\sum_{j=0}^{n} \frac{(n+1-j)^2}{4} \approx \frac{1}{12} n^3$ fourth powers, which is about twice the minimal number. But this quartic version has no universally-used variable, so it can't be bumped up to the fifth power.

14. A number theoretic interlude

Imagine a general canonical form for quartics of "Reichstein-type"

$$p(x_1, \dots, x_n) = \sum_{k=1}^{r} (\alpha_{k1} x_1 + \cdots + \alpha_{kn} x_n)^4 + q(x_1, \dots, x_m).$$

14. A number theoretic interlude

Imagine a general canonical form for quartics of "Reichstein-type"

$$p(x_1, \dots, x_n) = \sum_{k=1}^{r} (\alpha_{k1} x_1 + \cdots + \alpha_{kn} x_n)^4 + q(x_1, \dots, x_m).$$

We would need $N(n, 4) = r \times n + N(m, 4)$.

14. A number theoretic interlude

Imagine a general canonical form for quartics of "Reichstein-type"

$$p(x_1, \dots, x_n) = \sum_{k=1}^{r} (\alpha_{k1}x_1 + \cdots + \alpha_{kn}x_n)^4 + q(x_1, \dots, x_m).$$

We would need $N(n, 4) = r \times n + N(m, 4)$.

It turns out that if $n = 12$, there does **not** exist $m < 12$ so that

$$12 \;\Big|\; \binom{15}{4} - \binom{m+3}{4},$$

so number theory rules out universal Reichstein-type canonical forms for quartics, at least in 12 variables.

14. A number theoretic interlude

Imagine a general canonical form for quartics of “Reichstein-type”

$$p(x_1, \dots, x_n) = \sum_{k=1}^{r} (\alpha_{k1}x_1 + \cdots + \alpha_{kn}x_n)^4 + q(x_1, \dots, x_m).$$

We would need $N(n, 4) = r \times n + N(m, 4)$.

It turns out that if $n = 12$, there does **not** exist $m < 12$ so that

$$12 \;\Big|\; \binom{15}{4} - \binom{m+3}{4},$$

so number theory rules out universal Reichstein-type canonical forms for quartics, at least in 12 variables.

14. A number theoretic interlude

Now define

$$A_d = \left\{ n : 0 \le m < n \implies n \nmid \binom{n+d-1}{d} - \binom{m+d-1}{d} \right\}.$$

We have a few partial results.

- If $3 \nmid k$, then $n = 2^{2k} \cdot 3 \in A_4$.

14. A number theoretic interlude

Now define

$$A_d = \left\{ n : 0 \le m < n \implies n \nmid \binom{n+d-1}{d} - \binom{m+d-1}{d} \right\}.$$

We have a few partial results.

- If $3 \nmid k$, then $n = 2^{2k} \cdot 3 \in A_4$.
- If $p \equiv 1 \pmod{144}$ is prime, then $12p \in A_4$.

14. A number theoretic interlude

Now define

$$A_d = \left\{ n : 0 \le m < n \implies n \not| \binom{n+d-1}{d} - \binom{m+d-1}{d} \right\}.$$

We have a few partial results.

- If $3 \not| k$, then $n = 2^{2k} \cdot 3 \in A_4$.
- If $p \equiv 1 \pmod{144}$ is prime, then $12p \in A_4$.
- If p is prime, then $p \mid \binom{n+p-1}{p} - \binom{n}{p}$, hence A_p is empty for prime p (such as $p = 2, 3$.)

14. A number theoretic interlude

Now define

$$A_d = \left\{ n : 0 \le m < n \implies n \nmid \binom{n+d-1}{d} - \binom{m+d-1}{d} \right\}.$$

We have a few partial results.

- If $3 \nmid k$, then $n = 2^{2k} \cdot 3 \in A_4$.
- If $p \equiv 1 \pmod{144}$ is prime, then $12p \in A_4$.
- If p is prime, then $p \mid \binom{n+p-1}{p} - \binom{n}{p}$, hence A_p is empty for prime p (such as $p = 2, 3$.)
- The smallest elements of $A_6, A_8, A_{10}, A_{12}, A_{14}$ and A_{15} are 10, 1792, 6, 242, 338 and 273 respectively. If A_9 or A_{16} are non-empty, then their smallest elements are at least 10^5. (Fortunately, steampunk allows Mathematica.)

15. Slowpoke

Our last expression for cubic forms is not canonical: for any $p \in H_3(\mathbb{C}^n)$, there exists an invertible linear change of variables $y_j = \sum \lambda_{jk} x_k$ and n linear forms ℓ_j so that

$$p(x_1, \dots, x_n) = \sum_{j=1}^{n} \ell_j^3(x_1, \dots, x_n) + q(y_2, \dots, y_n).$$

15. Slowpoke

Our last expression for cubic forms is not canonical: for any $p \in H_3(\mathbb{C}^n)$, there exists an invertible linear change of variables $y_j = \sum \lambda_{jk} x_k$ and n linear forms ℓ_j so that

$$p(x_1, \dots, x_n) = \sum_{j=1}^{n} \ell_j^3(x_1, \dots, x_n) + q(y_2, \dots, y_n).$$

The proof of this is constructive. Repeating the argument gives p as a sum of $\frac{n(n+1)}{2}$ cubes, the same number as in Slinky.

15. Slowpoke

Our last expression for cubic forms is not canonical: for any $p \in H_3(\mathbb{C}^n)$, there exists an invertible linear change of variables $y_j = \sum \lambda_{jk} x_k$ and n linear forms ℓ_j so that

$$p(x_1, \dots, x_n) = \sum_{j=1}^{n} \ell_j^3(x_1, \dots, x_n) + q(y_2, \dots, y_n).$$

The proof of this is constructive. Repeating the argument gives p as a sum of $\frac{n(n+1)}{2}$ cubes, the same number as in Slinky.
We need a lemma: for any integer m, there exist $m+1$ linear forms $\ell_{j,m} = \ell_{j,m}(y_1, \dots, y_m)$ so that

$$\sum_{j=1}^{m+1} \ell_{j,m} = 0 \quad \text{and} \quad \sum_{j=1}^{m+1} \ell_{j,m}^2 = \sum_{k=1}^{m} y_k^2.$$

15. Slowpoke

The simplest proof is to set $\ell_{m+1,m} = -\sum_{j=1}^{m} \ell_{j,m}$ and then observe that the quadratic form $\sum_{j=1}^{m} t_j^2 + (\sum_{j=1}^{m} t_j)^2$ has full rank, and so can be written as a sum of m squares. Finally, invert the system.

15. Slowpoke

The simplest proof is to set $\ell_{m+1,m} = -\sum_{j=1}^{m} \ell_{j,m}$ and then observe that the quadratic form $\sum_{j=1}^{m} t_j^2 + (\sum_{j=1}^{m} t_j)^2$ has full rank, and so can be written as a sum of m squares. Finally, invert the system.

As an explicit solution, let $\alpha = \frac{-(m+1)+\sqrt{m+1}}{m(m+1)}$ and define

$$\ell_{j,m}(x_1, \ldots, x_n) = x_j + \alpha \sum_{j=1}^{m} x_j, \qquad 1 \le j \le m,$$

$$\ell_{m+1,m}(x_1, \ldots, x_n) = -(1 + m\alpha) \sum_{j=1}^{m} x_j.$$

15. Slowpoke

The simplest proof is to set $\ell_{m+1,m} = -\sum_{j=1}^{m} \ell_{j,m}$ and then observe that the quadratic form $\sum_{j=1}^{m} t_j^2 + (\sum_{j=1}^{m} t_j)^2$ has full rank, and so can be written as a sum of m squares. Finally, invert the system.

As an explicit solution, let $\alpha = \frac{-(m+1)+\sqrt{m+1}}{m(m+1)}$ and define

$$\ell_{j,m}(x_1, \dots, x_n) = x_j + \alpha \sum_{j=1}^{m} x_j, \qquad 1 \le j \le m,$$

$$\ell_{m+1,m}(x_1, \dots, x_n) = -(1 + m\alpha) \sum_{j=1}^{m} x_j.$$

Now suppose $p \in H_3(\mathbb{C}^n)$. By Biermann's Theorem, there is a finite list to check to find a point u where $p(u) \neq 0$, and after an invertible linear change of variables, taking $\{x_j\} \mapsto \{u_j\}$, we may assume that

15. Slowpoke

$$p = u_1^3 + 3h_1(u_2, \ldots, u_n)u_1^2 + 3h_2(u_2, \ldots, u_n)u_1 + h_3(u_2, \ldots, u_n),$$

15. Slowpoke

$$p = u_1^3 + 3h_1(u_2, \ldots, u_n)u_1^2 + 3h_2(u_2, \ldots, u_n)u_1 + h_3(u_2, \ldots, u_n),$$

We then let $u_1 = y_1 - h_1(u_2, \ldots, u_n)$ to clear the quadratic term :

$$p = y_1^3 + 3y_1\tilde{h}_2(u_2, \ldots, u_n) + \tilde{h}_3(u_2, \ldots, u_n).$$

15. Slowpoke

$$p = u_1^3 + 3h_1(u_2,\ldots,u_n)u_1^2 + 3h_2(u_2,\ldots,u_n)u_1 + h_3(u_2,\ldots,u_n),$$

We then let $u_1 = y_1 - h_1(u_2,\ldots,u_n)$ to clear the quadratic term :

$$p = y_1^3 + 3y_1\tilde{h}_2(u_2,\ldots,u_n) + \tilde{h}_3(u_2,\ldots,u_n).$$

and do a standard diagonalization of $\tilde{h}_2$ as a quadratic form, with the accompanying change of variables, yielding:

$$p = y_1^3 + 3y_1(y_2^2 + \cdots + y_r^2) + k_3(y_2,\ldots,y_n); \quad r \le n.$$

15. Slowpoke

$$p = u_1^3 + 3h_1(u_2, \ldots, u_n)u_1^2 + 3h_2(u_2, \ldots, u_n)u_1 + h_3(u_2, \ldots, u_n),$$

We then let $u_1 = y_1 - h_1(u_2, \ldots, u_n)$ to clear the quadratic term :

$$p = y_1^3 + 3y_1\tilde{h}_2(u_2, \ldots, u_n) + \tilde{h}_3(u_2, \ldots, u_n).$$

and do a standard diagonalization of $\tilde{h}_2$ as a quadratic form, with the accompanying change of variables, yielding:

$$p = y_1^3 + 3y_1(y_2^2 + \cdots + y_r^2) + k_3(y_2, \ldots, y_n); \quad r \le n.$$

Finally, observe that if

$$q = \frac{1}{r}\sum_{j=1}^{r}\left(y_1 + \sqrt{r}\cdot \ell_{j,r-1}(y_2, \ldots, y_r)\right)^3,$$

then the lemma implies that $p - q$ is a cubic form in $(y_2, \ldots, y_n)$, which is what we wanted.

16. Other kinds of canonical forms

It seems obvious to 21st century mathematicians, if not 19th century mathematicians, that one might look at polynomial maps $F : S \mapsto \mathbb{C}^N$, where S is an N-dimensional subspace of $\mathbb{C}^M$ for some $M > N$.

Theorem

For fixed $\{r_k\}$, a general binary quadratic form can be written as

$$(t_1x + t_2y)^2 + (t_3x + t_4y)^2, \quad \textit{where} \quad \sum_{k=1}^{4} r_k t_k = 0$$

unless $r_3 = \pm ir_1$ and $r_2 = \pm ir_4$.

This is proved by parameterizing the plane and looking at the Jacobian.
The choice $(r_1, r_2, r_3, r_4) = (1, 0, i, 0)$ gives the "silly" example.

16. Other kinds of canonical forms

We have proved the following result for $d = 2k = 2, 4, 6, 8$ and believe it is true in general. If true, it would provide another complement to Sylvester's canonical form for even degree:

A general binary form of degree $d = 2k$ can be written as

$$p(x, y) = \sum_{j=1}^{k+1} (\alpha_j x + \beta_j y)^{2k}, \quad \text{where} \quad \sum_{j=1}^{k+1} (\alpha_j + \beta_j) = 0.$$

In the cases given, the Jacobian is non-zero when the parameters are given the values of small positive consecutive integers.

17. To the audience

Thanks for your patience and for coming to the talk!

18. Oh, good. I have some extra time