Apolarity, steampunk canonical forms and the obvious inner product

Bruce Reznick
University of Illinois at Urbana-Champaign

Geometric Commutative Algebra and Applications
2011 Spring Central Section Meeting of AMS
Iowa City, March 19, 2011

1. Introduction

Corrected version of this talk.

1. Introduction

Corrected version of this talk.

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

1. Introduction

Corrected version of this talk.

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

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

1. Introduction

Corrected version of this talk.

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

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

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

1. Introduction

Corrected version of this talk.
Webpage for this talk: www.math.uiuc.edu/∼reznick/iowa.html
What is steampunk? It is a style based on combining 19th century Victorian culture with elements of modern life like computers.

What are canonical forms? Before I give a definition, here's the **Basic Example** for this talk.
Any binary quadratic form over $\mathbb{C}$ can be written as

$$p(x, y) = ax^2 + 2bxy + cy^2. \tag{1}$$

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

$$p(x, y) = (\alpha_1 x + \alpha_2 y)^2 + (\alpha_3 y)^2, \tag{2}$$

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

1. Introduction

Corrected version of this talk.
Webpage for this talk: www.math.uiuc.edu/∼reznick/iowa.html
What is steampunk? It is a style based on combining 19th century Victorian culture with elements of modern life like computers.

What are canonical forms? Before I give a definition, here's the **Basic Example** for this talk.
Any binary quadratic form over $\mathbb{C}$ can be written as

$$p(x, y) = ax^2 + 2bxy + cy^2. \tag{1}$$

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

$$p(x, y) = (\alpha_1 x + \alpha_2 y)^2 + (\alpha_3 y)^2, \tag{2}$$

at least unless $a = 0$ and $b \neq 0$.
That is, a "general" binary quadratic form can be written with the square completed.

1. Introduction

Corrected version of this talk.
Webpage for this talk: www.math.uiuc.edu/∼reznick/iowa.html
What is steampunk? It is a style based on combining 19th century Victorian culture with elements of modern life like computers.

What are canonical forms? Before I give a definition, here's the **Basic Example** for this talk.
Any binary quadratic form over $\mathbb{C}$ can be written as

$$p(x, y) = ax^2 + 2bxy + cy^2. \quad (1)$$

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

$$p(x, y) = (\alpha_1 x + \alpha_2 y)^2 + (\alpha_3 y)^2, \quad (2)$$

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.

Note also that there are three parameters in each of these representations.

Note also that there are three parameters in each of these representations.

This equality is necessary, but not sufficient. There are three parameters in

$$(\alpha_1 x + \alpha_2 y)^2 + (i\alpha_1 x + \alpha_3 y)^2, \tag{3}$$

but the coefficient of x^2 is 0, so the sum (3) must have a factor of y, which a general binary quadratic form does not. This example may look silly, but it isn't.

Note also that there are three parameters in each of these representations.

This equality is necessary, but not sufficient. There are three parameters in

$$(\alpha_1 x + \alpha_2 y)^2 + (i\alpha_1 x + \alpha_3 y)^2, \tag{3}$$

but the coefficient of x^2 is 0, so the sum (3) must have a factor of y, which a general binary quadratic form does not. This example may look silly, but it isn't.

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 something of interest in the algebraic geometry of binary forms.

I freely acknowledge that any result not specifically credited to someone else may nevetheless be old. (References from the audience welcome.)

I freely acknowledge that any result not specifically credited to someone else may nevetheless be old. (References from the audience welcome.)

There is very serious mathematics in this area being done today.

There is a strong connection between this talk and the wonderful 1993 paper of Richard Ehrenborg and Gian-Carlo Rota connecting apolarity and canonical forms using matroids. There is also recent work on these questions from a more sophisticated point of view by Tony Geramita, Tony Iarrobino, J. M. Landsberg, Zach Teitler, Enrico Carlini, Giorgio Ottaviani and others I might have missed.

And of course the deep theorem of Alexander-Hirschowitz hovers over all.

I freely acknowledge that any result not specifically credited to someone else may nevetheless be old. (References from the audience welcome.)

There is very serious mathematics in this area being done today. There is a strong connection between this talk and the wonderful 1993 paper of Richard Ehrenborg and Gian-Carlo Rota connecting apolarity and canonical forms using matroids. There is also recent work on these questions from a more sophisticated point of view by Tony Geramita, Tony Iarrobino, J. M. Landsberg, Zach Teitler, Enrico Carlini, Giorgio Ottaviani and others I might have missed. And of course the deep theorem of Alexander-Hirschowitz hovers over all.

This talk has been constructed to avoid profundity, especially in the proofs.

Here are two 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.

Here are two 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.

Here are two 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 cubic form $p(x_1, \dots, x_n)$ has a unique representation in the form

$$p(x_1, \dots, x_n) = \sum_{1 \le i \le j \le n} (\alpha_{\{i,j\},i} x_i + \cdots + \alpha_{\{i,j\},j} x_j)^3. \tag{4}$$

Here are two 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 cubic form $p(x_1, \dots, x_n)$ has a unique representation in the form

$$p(x_1, \dots, x_n) = \sum_{1 \le i \le j \le n} (\alpha_{\{i,j\},i} x_i + \cdots + \alpha_{\{i,j\},j} x_j)^3. \qquad (4)$$

Here are two 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 cubic form $p(x_1, \ldots, x_n)$ has a unique representation in the form

$$p(x_1, \ldots, x_n) = \sum_{1 \le i \le j \le n} (\alpha_{\{i,j\},i} x_i + \cdots + \alpha_{\{i,j\},j} x_j)^3. \qquad (4)$$

Even though (4) is actually a canonical form, it is not a representation of p as a minimal **number** of cubes, so its application to, for example, the rank of tensors is unclear.

The canonical binary forms generalize a couple of classical theorems of Sylvester from 1851. Sylvester also gave algorithms in the process of giving his proofs.

The canonical binary forms generalize a couple of classical theorems of Sylvester from 1851. Sylvester also gave algorithms in the process of giving his proofs.

The given canonical cubic form was suggested by a much more profound result of Boris Reichstein from 1987:

Theorem (Reichstein)

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

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

The canonical binary forms generalize a couple of classical theorems of Sylvester from 1851. Sylvester also gave algorithms in the process of giving his proofs.

The given canonical cubic form was suggested by a much more profound result of Boris Reichstein from 1987:

Theorem (Reichstein)

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

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

The canonical binary forms generalize a couple of classical theorems of Sylvester from 1851. Sylvester also gave algorithms in the process of giving his proofs.
The given canonical cubic form was suggested by a much more profound result of Boris Reichstein from 1987:

Theorem (Reichstein)

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

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

Anybody here know this?

2. Preliminaries

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

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

Here is an alternative that is amazing to non-algebraic geometers, but not, apparently, to the algebraic geometers I've talked to. 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 $a \in \mathbb{C}^N$ at which the Jacobian matrix $\left[\frac{\partial f_i}{\partial t_j}(a)\right]$ has full rank.

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 $\{i_k : 1 \le k \le N\}$ and making the interpretation in an abuse of notation that

$$F(t_1, \dots, t_N) = \sum_{k=1}^{N} c(i_k) f_k(t_1 \dots, t_N) x^{i_k}$$

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 $\{i_k : 1 \le k \le N\}$ and making the interpretation in an abuse of notation that

$$F(t_1, \dots, t_N) = \sum_{k=1}^{N} c(i_k) f_k(t_1 \dots, t_N) x^{i_k}$$

To retrieve our Main Example, let $i_1 = (2,0)$, $i_2 = (1,1)$, $i_3 = (0,2)$, and suppose the original polynomial map is $F(t_1,t_2,t_3) = (t_1^2, t_1 t_2, t_2^2 + t_3^2)$. Then

$$\begin{aligned} F(t_1,t_2,t_3) &= t_1^2 x_1^2 + 2 t_1 t_2 x_1 x_2 + (t_2^2 + t_3^2)^2 x_2^2 \\ &= (t_1 x_1 + t_2 x_2)^2 + (t_3 x_2)^2. \end{aligned}$$

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.

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.

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.

I am somewhat puzzled that the simplest canonical forms were often not considered so.

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.

I am somewhat puzzled that the simplest canonical forms were often not considered so.

Suppose $\{p_j : 1 \leq j \leq N\}$ is a basis for $H_d(\mathbb{C}^n)$. Then for every $p \in H_d(\mathbb{C}^n)$, there exist $\alpha_j \in \mathbb{C}$ so that $p = \sum \alpha_j p_j$. This must not have been considered "sporting" in the 19th century.

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.

I am somewhat puzzled that the simplest canonical forms were often not considered so.
Suppose $\{p_j : 1 \le j \le N\}$ is a basis for $H_d(\mathbb{C}^n)$. Then for every $p \in H_d(\mathbb{C}^n)$, there exist $\alpha_j \in \mathbb{C}$ so that $p = \sum \alpha_j p_j$. This must not have been considered "sporting" in the 19th century.

The interpretation of the Jacobian having full rank is that, for some $a \in \mathbb{C}^n$, the forms $\{\frac{\partial F}{\partial t_j}(a)\}$ span $H_d(\mathbb{C}^n)$.

Let's see how this works in 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:

$$2x_1(t_1x_1 + t_2x_2), \quad 2x_2(t_1x_1 + t_2x_2), \quad 2x_2(t_3x_2). \tag{5}$$

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 (5) 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!

Let's see how this works in 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:

$$2x_1(t_1x_1 + t_2x_2), \quad 2x_2(t_1x_1 + t_2x_2), \quad 2x_2(t_3x_2). \tag{5}$$

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 (5) 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!

The classical version of the Jacobian argument is called the *Lasker-Wakeford Theorem*. It's worthwhile to spend a few minutes on who these people are.

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.

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). I quote from the memorial article by J. H. Grace in the *Proceedings of the London Mathematical Society*. It may be the angriest obituary I've ever read, and it can be found in its entirety on my webpage.

Edward Kingsley Wakeford (1894-1916). I quote from the memorial article by J. H. Grace in the *Proceedings of the London Mathematical Society*. It may be the angriest obituary I've ever read, and it can be found in its entirety on my webpage.

"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 ...

Edward Kingsley Wakeford (1894-1916). I quote from the memorial article by J. H. Grace in the *Proceedings of the London Mathematical Society*. It may be the angriest obituary I've ever read, and it can be found in its entirety on my webpage.

"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."

I found the Lasker-Wakeford Theorem in *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 a so that there is no non-zero form q which is apolar to all N forms $\{\frac{\partial F}{\partial t_1}(a), \dots, \frac{\partial F}{\partial t_N}(a)\}$.

I found the Lasker-Wakeford Theorem in *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 a so that there is no non-zero form q which is apolar to all N forms $\{\frac{\partial F}{\partial t_1}(a), \dots, \frac{\partial F}{\partial t_N}(a)\}$.

I found the Lasker-Wakeford Theorem in *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 $F : \mathbb{C}^M \to H_d(\mathbb{C}^n)$ $(M \ge N(n,d))$ then F is a canonical form if and only if there is a point a so that there is no non-zero form q which is apolar to all N forms $\{\frac{\partial F}{\partial t_1}(a), \dots, \frac{\partial F}{\partial t_N}(a)\}$.

Most writers restrict themselves to $M = N(n,d)$ as do I.
The question for us now is, what does "apolarity" mean in this context?

Apolarity

I 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.

Apolarity

I 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.)

Apolarity

I 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 $H_d(\mathbb{C}^n)$.

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. Proofs are available in many of the papers online.

The inner product satisfies these properties:

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

The inner product satisfies these properties:

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

The inner product satisfies these properties:

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

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

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

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.

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

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

The consequences of this definition include:

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

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.

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.

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.

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.

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

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

Proof.

Fix q. Define the two subspaces

$$A = \{p \in H_d(\mathbb{C}^n) : q(D)p = 0\},$$
$$B = \{p \in H_d(\mathbb{C}^n) : p = \sum \lambda_k(\alpha_k \cdot)^d, \quad q(\alpha_k) = 0\}.$$

We want to show that $B \subseteq A$ and $B^\perp \subseteq A^\perp$; if so, then since $H_d(\mathbb{C}^n)$ is finite dimensional, it will follow that $A = B$.

Proof.

Fix q. Define the two subspaces

$$A = \{p \in H_d(\mathbb{C}^n) : q(D)p = 0\},$$
$$B = \{p \in H_d(\mathbb{C}^n) : p = \sum \lambda_k(\alpha_k \cdot)^d, \quad q(\alpha_k) = 0\}.$$

We want to show that $B \subseteq A$ and $B^\perp \subseteq A^\perp$; if so, then since $H_d(\mathbb{C}^n)$ is finite dimensional, it will follow that $A = B$.

Proof.

Fix q. Define the two subspaces

$$A = \{p \in H_d(\mathbb{C}^n) : q(D)p = 0\},$$
$$B = \{p \in H_d(\mathbb{C}^n) : p = \sum \lambda_k(\alpha_k \cdot)^d, \quad q(\alpha_k) = 0\}.$$

We want to show that $B \subseteq A$ and $B^\perp \subseteq A^\perp$; if so, then since $H_d(\mathbb{C}^n)$ is finite dimensional, it will follow that $A = B$.

First, if $e > d$, then $A = H_d(\mathbb{C}^n)$, so $B \subseteq A$. If $e \le d$ and $p \in B$, then $q(D)p = \frac{d!}{(d-e)!} \sum \lambda_k q(\alpha_k)(\alpha_k \cdot)^{d-e} = 0$, so $p \in A$.

Proof.

Fix q. Define the two subspaces

$$A = \{p \in H_d(\mathbb{C}^n) : q(D)p = 0\},$$
$$B = \{p \in H_d(\mathbb{C}^n) : p = \sum \lambda_k(\alpha_k\cdot)^d, \quad q(\alpha_k) = 0\}.$$

We want to show that $B \subseteq A$ and $B^\perp \subseteq A^\perp$; if so, then since $H_d(\mathbb{C}^n)$ is finite dimensional, it will follow that $A = B$.

First, if $e > d$, then $A = H_d(\mathbb{C}^n)$, so $B \subseteq A$. If $e \le d$ and $p \in B$, then $q(D)p = \frac{d!}{(d-e)!}\sum \lambda_k q(\alpha_k)(\alpha_k\cdot)^{d-e} = 0$, so $p \in A$.

Observe that $f \in B^\perp$ iff $q(\alpha) = 0$ implies $[f,(\alpha\cdot)^d] = f(\alpha) = 0$. Since q is irreducible, the Nullstellensatz implies that $q \mid f$. If $e > d$, this is impossible unless $f = 0$, so $B^\perp = \{0\} \subseteq A^\perp$. If $e \le d$, then $f = gq$ where $g \in H_{d-e}(\mathbb{C}^n)$. But $p \in A \implies q(D)p = 0 \implies [p,f] = [p,gq] = \frac{d!}{(d-e)!}[q(D)p,g] = 0$. It follows that $f \in A^\perp$, completing the proof. □

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.

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.

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

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

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.

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*

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

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.

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

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

In the case of binary forms, where zeros correspond to linear factors, the theorem 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 that* ℓ_i *and* ℓ_j *are pairwise linearly independent for* $i \neq j$. *Then every* $p \in H_d(\mathbb{C}^2)$ *can be written as* $\sum_j \phi_j(x, y)(\beta_j x - \alpha_j y)^{d-(m_j-1)}$, *where* $\phi_j \in H_{d-(m_j-1)}(\mathbb{C}^2)$.

Theorem

Suppose $\sum m_j = d+1$ *and let* $\ell_i(x,y) = \alpha_i x + \beta_i y$. *Suppose that* ℓ_i *and* ℓ_j *are pairwise linearly independent for* $i \neq j$. *Then every* $p \in H_d(\mathbb{C}^2)$ *can be written as* $\sum_j \phi_j(x,y)(\beta_j x - \alpha_j y)^{d-(m_j-1)}$, *where* $\phi_j \in H_{d-(m_j-1)}(\mathbb{C}^2)$.

Proof.

This is a direct consequence of the last theorem. More directly,

$$A = \left\{ \sum_j \phi_j(\beta_j x - \alpha_j y)^{d-(m_j-1)} : \phi_j \in H_{m_j-1}(\mathbb{C}^2) \right\}.$$

Observe that $p \in A^\perp \iff p$ is apolar to each possible $x^r y^{m_j-1-r}(\beta_j x - \alpha_j y)^{d-(m_j-1)} \iff p$ vanishes to $m_j - 1$-st order at $(\beta_j, -\alpha_j) \iff p$ is divisible by each $\ell_j^{m_j}$. Since $\sum_j m_j = d+1$, this implies $p = 0$, hence $A = \{0\}^\perp = H_d(\mathbb{C}^2)$. □

Remember canonical forms? Suppose $\alpha_1, \ldots, \alpha_n$ appear in a canonical form as

$$(\alpha_1 x_1 + \cdots + \alpha_n x_n)^d = \ell^d.$$

Then $\frac{\partial F}{\partial \alpha_j} = d x_j \ell^{d-1}$, and in applying Lasker-Wakeford, note that a form is apolar to each of these if and only if it is singular at $(\alpha_1, \ldots, \alpha_n)$. Start thinking about general forms which are singular at general sets of points and you enter the context in which Alexander-Hirschowitz comes in, and we change topics.

Some canonical forms for binary forms

In 1869, J. J. Sylvester (1814-1897) reflected on the discovery of some of his most famous research in 1851, done while he was working as an actuary.

Some canonical forms for binary forms

In 1869, J. J. Sylvester (1814-1897) reflected on the discovery of some of his most famous research in 1851, done while he was working as an actuary.

"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"

Theorem (Sylvester)

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

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

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

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

Note that the constant counts work out: $2 \cdot k = (2k - 1) + 1$ and $1 + 2 \cdot k = 2k + 1$.

These are immediate consequences of the results I just described, but Sylvester did even better: these were a consequence of his algorithm for computing the representations.

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

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

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

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

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

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

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

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

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

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

Since

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

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

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

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

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

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

Since

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

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.

A few remarks about Sylvester's algorithm

- If $h(D) = \prod_{j=1}^{r}(\beta_j \frac{\partial}{\partial x} - \alpha_j \frac{\partial}{\partial y})$, 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 $h(D)p = 0$. This is FTA configured for products of linear factors.

A few remarks about Sylvester's algorithm

- If $h(D) = \prod_{j=1}^{r}(\beta_j \frac{\partial}{\partial x} - \alpha_j \frac{\partial}{\partial y})$, 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 $h(D)p = 0$. This is 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.

A few remarks about Sylvester's algorithm

- If $h(D) = \prod_{j=1}^{r}(\beta_j \frac{\partial}{\partial x} - \alpha_j \frac{\partial}{\partial y})$, 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 $h(D)p = 0$. This is 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".

- 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.

- 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.

Possibly new results

It is easy to see that if $\ell_j(x, y) = \alpha_j x + \beta_j y, 1 \le j \le d + 1$ is a set of $d + 1$ pairwise distinct linear forms, then $\{\ell_j^d\}$ is a basis for $H_d(\mathbb{C}^2)$: the representation of $\{\ell_j^d\}$ with respect to the basis $\{\binom{d}{j} x^{d-j} y^j\}$ has Vandermonde determinant

$$\prod_{1 \le i < j \le n} (\alpha_i \beta_j - \alpha_j \beta_i) \neq 0.$$

Possibly new results

It is easy to see that if $\ell_j(x,y) = \alpha_j x + \beta_j y, 1 \le j \le d+1$ is a set of $d+1$ pairwise distinct linear forms, then $\{\ell_j^d\}$ is a basis for $H_d(\mathbb{C}^2)$: the representation of $\{\ell_j^d\}$ with respect to the basis $\{\binom{d}{j} x^{d-j} y^j\}$ has Vandermonde determinant

$$\prod_{1 \le i < j \le n} (\alpha_i \beta_j - \alpha_j \beta_i) \ne 0.$$

We now give our generalization which includes this result, Sylvester's theorem and quite a bit more.

Theorem

Suppose $m, n \geq 0$ *and* $d \geq 1$ *are integers so that* $m + n = d + 1$. *Suppose* $\ell_j(x, y) = \beta_j x + \gamma_j y$, $1 \leq j \leq m$, *are fixed pairwise linearly independent linear forms and suppose* $e_k \mid d$, $1 \leq k \leq r$ *and* $\sum_{k=1}^{r}(e_k + 1) = n$. *Then a general binary form of degree* d *can be written as*

$$p(x, y) = \sum_{j=1}^{m} c_j \ell_j^d(x, y) + \sum_{k=1}^{r} f_k^{d/e_k}(x, y), \qquad (6)$$

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

Theorem

Suppose $m, n \geq 0$ and $d \geq 1$ are integers so that $m + n = d + 1$. Suppose $\ell_j(x, y) = \beta_j x + \gamma_j y$, $1 \leq j \leq m$, are fixed pairwise linearly independent linear forms and suppose $e_k \mid d$, $1 \leq k \leq r$ and $\sum_{k=1}^{r}(e_k + 1) = n$. Then a general binary form of degree d can be written as

$$p(x, y) = \sum_{j=1}^{m} c_j \ell_j^d(x, y) + \sum_{k=1}^{r} f_k^{d/e_k}(x, y), \tag{6}$$

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

The novelty here is the existence of forms of intermediate degree taken to higher powers. Multiples of fixed d-th powers are thrown in to make the constant count work out right.

Proof.

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

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

Now evaluate the Jacobian at a choice of parameters so that the forms of degree e_k are powers of linear forms; specifically, let $f_k(x, y) = \tilde{\ell}_k^{e_k}$, 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 earlier theorems, this set, taken all together, is a basis for $H_d(\mathbb{C}^2)$ and so (6) is a canonical form. □

A few remarks:

- This result follows easily from the end of the Ehrenborg-Rota paper, but does not seem to appear there explicitly.

A few remarks:

- This result follows easily from the end of the Ehrenborg-Rota paper, but does not seem to appear there explicitly.
- If $e_k \equiv 1$, $m = d + 1$ and $n = 0$, this is the "Vandermonde basis" canonical form.

A few remarks:

- This result follows easily from the end of the Ehrenborg-Rota paper, but does not seem to appear there explicitly.
- If $e_k \equiv 1$, $m = d + 1$ and $n = 0$, this is the "Vandermonde basis" canonical form.
- If $e_k \equiv 1$, $d = 2r - 1$, $m = 0$ and $n = r$, this recovers Sylvester's canonical form for binary forms of odd degree.

A few remarks:

- This result follows easily from the end of the Ehrenborg-Rota paper, but does not seem to appear there explicitly.
- If $e_k \equiv 1$, $m = d + 1$ and $n = 0$, this is the "Vandermonde basis" canonical form.
- If $e_k \equiv 1$, $d = 2r - 1$, $m = 0$ and $n = r$, this recovers Sylvester's canonical form for binary forms of odd degree.
- If $e_k \equiv 1$, $d = 2r$, $m = 1$, $\ell_1(x, y) = x$ and $n = r$, this recovers Sylvester's canonical form for binary forms of even degree.

A few remarks:

- This result follows easily from the end of the Ehrenborg-Rota paper, but does not seem to appear there explicitly.
- If $e_k \equiv 1$, $m = d + 1$ and $n = 0$, this is the "Vandermonde basis" canonical form.
- If $e_k \equiv 1$, $d = 2r - 1$, $m = 0$ and $n = r$, this recovers Sylvester's canonical form for binary forms of odd degree.
- If $e_k \equiv 1$, $d = 2r$, $m = 1$, $\ell_1(x, y) = x$ and $n = r$, this recovers Sylvester's canonical form for binary forms of even degree.
- If $e_k \equiv 1$, so that $n = 2r$ and $m = n - 2r$, this provides a kind of "interpolation" between the first two examples. Further, Sylvester's algorithm can still be used. Let $q = \prod_{j=1}^{m} \ell_j$. Then $q(D)p$ has degree $d - m = 2r - 1$. Write $q(D)p$ as a sum of r $2r - 1$-st powers and then "integrate" to find an explicit representation.

- If $e_k \equiv 2$, then 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 $d - 3\lfloor (d+1)/3 \rfloor$ $2k$-th power of specified linear forms, say x^{2k} and y^{2k}. We don't have an algorithm for this. We want one.

- If $e_k \equiv 2$, then 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 $d - 3\lfloor (d+1)/3 \rfloor$ $2k$-th power of specified linear forms, say x^{2k} and y^{2k}. We don't have an algorithm for this. We want one.
- 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 and the fourth power of a linear form. (We have an algorithm for this which shows that it can be done in six different ways.)

- If $e_k \equiv 2$, then 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 $d - 3\lfloor (d+1)/3 \rfloor$ $2k$-th power of specified linear forms, say x^{2k} and y^{2k}. We don't have an algorithm for this. We want one.
- 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 and the fourth power of a linear form. (We have an algorithm for this which shows that it can be done in six different ways.)
- If $d = 6$, $m = 0$, $e_1 = 3$ and $e_2 = 2$, then $(\frac{6}{3} + 1) + (\frac{6}{2} + 1) = 6 + 1$ implies that a general binary sextic can be written as the sum of the square of a cubic and the cube of a quadratic form. We don't have an algorithm for this. We want one.

To illustrate Lasker-Wakeford in action, consider the representation

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

To illustrate Lasker-Wakeford in action, consider the representation

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

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

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

To illustrate Lasker-Wakeford in action, consider the representation

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

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 distinct linear powers for $\{f,g\}$, say $f = x^3, g = y^2$, then these partials become:

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

To illustrate Lasker-Wakeford in action, consider the representation

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

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 distinct linear powers for $\{f,g\}$, say $f = x^3, g = y^2$, 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 $H_6(\mathbb{C}^2)$ and there is no non-zero form apolar to them, so we have a genuine steampunk canonical form!

More variables

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

$$p(x_1, x_2, x_3) = \sum_{k=1}^{5} (\alpha_{k1} x_1 + \alpha_{k2} x_2 + \alpha_{k3} x_3)^4 \tag{7}$$

More variables

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

$$p(x_1, x_2, x_3) = \sum_{k=1}^{5} (\alpha_{k1}x_1 + \alpha_{k2}x_2 + \alpha_{k3}x_3)^4 \qquad (7)$$

This would be a canonical form, provided the partials with respect to the α_{kj}'s at some value would span $H_4(\mathbb{C}^3)$. Using apolarity, there should be no non-zero quartic apolar to $\{x_i \ell_k^3 : 1 \le k \le 5\}$; that is, no quartic which is singular at the five points $(\alpha_{k1}, \alpha_{k2}, \alpha_{k3})$. However, as Clebsch argued in the 1860's, any choice of five α_k's pass through a non-zero quadratic $h(x_1, x_2, x_3)$ (since $N(3,2) = 6$), and so h^2 will be apolar to all the partials and so it's not a canonical form.

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.

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

It happens that $H_{(\alpha\cdot)^4}$ 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 (7).

It happens that $H_{(\alpha\cdot)^4}$ 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 (7). These are really the same proof, because

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

It happens that $H_{(\alpha\cdot)^4}$ 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 (7). These are really the same proof, because

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

This becomes a serious topic in algebraic geometry, so I won't talk about it here this morning. 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.

Sums of two squares

Suppose $p \in H_{2k}(\mathbb{C}^2)$, has even degree. Then the FTA (algebra this time) says that p can be written as a product of $2k$ linear forms, which in general are pairwise linearly independent, so by pairing them off, we have $p = gh$, where $f, g \in H_2(\mathbb{C}^2)$ in essentially $\binom{2k-1}{k}$ different ways. Further, by the old algebraist's trick:

$$p = gh = \left(\frac{g+h}{2}\right)^2 + \left(\frac{g-h}{2i}\right)^2.$$

Sums of two squares

Suppose $p \in H_{2k}(\mathbb{C}^2)$, has even degree. Then the FTA (algebra this time) says that p can be written as a product of $2k$ linear forms, which in general are pairwise linearly independent, so by pairing them off, we have $p = gh$, where $f, g \in H_2(\mathbb{C}^2)$ in essentially $\binom{2k-1}{k}$ different ways. Further, by the old algebraist's trick:

$$p = gh = \left(\frac{g+h}{2}\right)^2 + \left(\frac{g-h}{2i}\right)^2.$$

A sum of two squares is not a canonical form because the right-hand side has $2(k+1) = (2k+1)+1$ parameters. However, if $(u,v) \in \mathbb{C}^2$, $u^2 + v^2 = 1$, then

$$a^2 + b^2 = (ua + vb)^2 + (va - ub)^2 \tag{8}$$

Sums of two squares

Suppose $p \in H_{2k}(\mathbb{C}^2)$, has even degree. Then the FTA (algebra this time) says that p can be written as a product of $2k$ linear forms, which in general are pairwise linearly independent, so by pairing them off, we have $p = gh$, where $f, g \in H_2(\mathbb{C}^2)$ in essentially $\binom{2k-1}{k}$ different ways. Further, by the old algebraist's trick:

$$p = gh = \left(\frac{g+h}{2}\right)^2 + \left(\frac{g-h}{2i}\right)^2.$$

A sum of two squares is not a canonical form because the right-hand side has $2(k+1) = (2k+1)+1$ parameters. However, if $(u,v) \in \mathbb{C}^2$, $u^2 + v^2 = 1$, then

$$a^2 + b^2 = (ua+vb)^2 + (va-ub)^2 \tag{8}$$

(These correspond to $gh = (\zeta g)(\zeta^{-1} h)$ where $\zeta = u + iv$; in terms of sos forms, when $u, v \in \mathbb{R}$, all the representations in (8) have the same Gram matrix.)

Accordingly, for a general p, we can choose (u, v) so that the coefficient of x^k, say, disappears in the second sum, and

$$p(x, y) = \left(\sum_{k=0}^{d} \alpha_k x^{d-k} y^k\right)^2 + \left(\sum_{k=1}^{d} \beta_k x^{d-k} y^k\right)^2 \qquad (9)$$

is another old canonical form for binary forms of even degree.

Accordingly, for a general p, we can choose (u, v) so that the coefficient of x^k, say, disappears in the second sum, and

$$p(x,y) = \left(\sum_{k=0}^{d} \alpha_k x^{d-k} y^k\right)^2 + \left(\sum_{k=1}^{d} \beta_k x^{d-k} y^k\right)^2 \qquad (9)$$

is another old canonical form for binary forms of even degree. When $d = 1$, this gives yet another proof of the Basic Example.

Accordingly, for a general p, we can choose (u, v) so that the coefficient of x^k, say, disappears in the second sum, and

$$p(x, y) = \left(\sum_{k=0}^{d} \alpha_k x^{d-k} y^k\right)^2 + \left(\sum_{k=1}^{d} \beta_k x^{d-k} y^k\right)^2 \qquad (9)$$

is another old canonical form for binary forms of even degree. When $d = 1$, this gives yet another proof of the Basic Example.

What about the counterexample (3) in the beginning?:

$$(\alpha_1 x + \alpha_2 y)^2 + (i\alpha_1 x + \alpha_3 y)^2.$$

This can be put in a more general context. Suppose we don't just look at maps $\mathbb{C}^N \to H_d(\mathbb{C}^n)$, but also maps from N dimensional subspaces of $\mathbb{C}^M$ for $M > N$. Here is the simplest non-trivial case:

For which $0 \neq r = (r_1, \ldots, r_4) \in \mathbb{C}^4$ is it true that

$$(\alpha_1 x + \alpha_2 y)^2 + (\alpha_3 x + \alpha_4 y)^2, \qquad \sum_{k=1}^{4} r_k \alpha_k = 0 \tag{10}$$

is a canonical form for binary quadratic forms?

For which $0 \neq r = (r_1, \ldots, r_4) \in \mathbb{C}^4$ is it true that

$$(\alpha_1 x + \alpha_2 y)^2 + (\alpha_3 x + \alpha_4 y)^2, \qquad \sum_{k=1}^{4} r_k \alpha_k = 0 \tag{10}$$

is a canonical form for binary quadratic forms?

Theorem

A general binary quadratic form can be written as in (10), unless $r_3 = \epsilon r_1$ and $r_4 = \epsilon r_2$, where $\epsilon \in \{i, -i\}$.

For which $0 \neq r = (r_1, \dots, r_4) \in \mathbb{C}^4$ is it true that

$$(\alpha_1 x + \alpha_2 y)^2 + (\alpha_3 x + \alpha_4 y)^2, \qquad \sum_{k=1}^{4} r_k \alpha_k = 0 \tag{10}$$

is a canonical form for binary quadratic forms?

Theorem

A general binary quadratic form can be written as in (10), unless $r_3 = \epsilon r_1$ and $r_4 = \epsilon r_2$, where $\epsilon \in \{i, -i\}$.

In other words, (10) is canonically a canonical form.

For which $0 \neq r = (r_1, \dots, r_4) \in \mathbb{C}^4$ is it true that

$$(\alpha_1 x + \alpha_2 y)^2 + (\alpha_3 x + \alpha_4 y)^2, \qquad \sum_{k=1}^{4} r_k \alpha_k = 0 \qquad (10)$$

is a canonical form for binary quadratic forms?

Theorem

A general binary quadratic form can be written as in (10), unless $r_3 = \epsilon r_1$ and $r_4 = \epsilon r_2$, where $\epsilon \in \{i, -i\}$.

In other words, (10) is canonically a canonical form.

In the special case $(r_1, r_2, r_3, r_4) = (1, 0, i, 0)$, observe that $\sum_{k=1}^{4} r_k \alpha_k = 0 \iff r_3 = i r_1$, and we recover the now-not-so-silly counterexample.

Completing the d-th power, all props to Boris Reichstein

Let me give a ridiculous (but constructive) proof of completing the square for a quadratic form. Suppose $p \in H_2(\mathbb{C}^n)$. Then we have:

$$\frac{\partial p}{\partial x_n} = \alpha_1 x_1 + \cdots + \alpha_n x_n.$$

Completing the d-th power, all props to Boris Reichstein

Let me give a ridiculous (but constructive) proof of completing the square for a quadratic form. Suppose $p \in H_2(\mathbb{C}^n)$. Then we have:

$$\frac{\partial p}{\partial x_n} = \alpha_1 x_1 + \cdots + \alpha_n x_n.$$

If we assume $\alpha_n \neq 0$, and let

$$q(x_1, \dots, x_n) = p(x_1, \dots, x_n) - \tfrac{1}{2\alpha_n}(\alpha_1 x_1 + \cdots + \alpha_n x_n)^2,$$

then it is easy to see that

Completing the d-th power, all props to Boris Reichstein

Let me give a ridiculous (but constructive) proof of completing the square for a quadratic form. Suppose $p \in H_2(\mathbb{C}^n)$. Then we have:

$$\frac{\partial p}{\partial x_n} = \alpha_1 x_1 + \cdots + \alpha_n x_n.$$

If we assume $\alpha_n \neq 0$, and let

$$q(x_1, \ldots, x_n) = p(x_1, \ldots, x_n) - \tfrac{1}{2\alpha_n}(\alpha_1 x_1 + \cdots + \alpha_n x_n)^2,$$

then it is easy to see that

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

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

This argument can be repeated. Suppose $p \in H_3(\mathbb{C}^n)$ is cubic; then $\frac{\partial p}{\partial x_n}$ is a quadratic form, so we can complete the square in general, but now lets do it diagonally from left to right:

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

This argument can be repeated. Suppose $p \in H_3(\mathbb{C}^n)$ is cubic; then $\frac{\partial p}{\partial x_n}$ is a quadratic form, so we can complete the square in general, but now lets do it diagonally from left to right:

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

You can guess where this is going. 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

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

This argument can be repeated. Suppose $p \in H_3(\mathbb{C}^n)$ is cubic; then $\frac{\partial p}{\partial x_n}$ is a quadratic form, so we can complete the square in general, but now lets do it diagonally from left to right:

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

You can guess where this is going. 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

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

We repeat this running backwards inductively to get an algebra prelim question in a 19th century graduate math program:

Theorem

A canonical form for the general cubic is

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

Theorem

A canonical form for the general cubic is

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

Proof.

We can do this abstractly with Lasker-Wakeford, or directly: use the previous construction and observe that

$$\binom{n+1}{2} + N(n-1,3) = \binom{n+1}{2} + \binom{n+1}{3}$$
$$= \binom{n+2}{3} = N(n,3).$$

□

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 appears in the context of trilinear forms.

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 appears in the context of trilinear forms.

Theorem (Reichstein)

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

$$\sum_{k=1}^{n} (\alpha_{k1} x_1 + \cdots + \alpha_{kn} x_n)^3 + q(x_1, \dots, x_{n-2}). \qquad (11)$$

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 appears in the context of trilinear forms.

Theorem (Reichstein)

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

$$\sum_{k=1}^{n} (\alpha_{k1}x_1 + \cdots + \alpha_{kn}x_n)^3 + q(x_1, \dots, x_{n-2}). \tag{11}$$

We have $N(n,3) - N(n-2,3) = \frac{n^3+3n^2+2n}{6} - \frac{n^3-3n^2+2n}{6} = n^2$, so if (11) is generally possible, it is a canonical form. This can be verified by Lasker-Wakeford, specializing at $x_1, x_2, x_1 + kx_2 + x_k$ (for $k \geq 3$), but Reichstein's constructive proof is prettier.

Proof.

A general pair of quadratic forms can be simultaneously diagonalized; so that for general p, that there exist n linearly independent forms $L_j(x) = \sum_{k=1}^{n} \alpha_{j,k} x_k$ and complex numbers c_j so that

$$\frac{\partial p}{\partial x_{n-1}} = \sum_{j=1}^{n} L_j^2, \qquad \frac{\partial p}{\partial x_n} = \sum_{j=1}^{n} c_j L_j^2,$$

Mixed partials are equal, so $2\sum_{j=1}^{n} \alpha_{j,n} L_j = 2\sum_{j=1}^{n} c_j \alpha_{j,n-1} L_j$ and the L_j's are linearly independent, so $\alpha_{j,n} = c_j \alpha_{j,n-1}$. Now,

Proof.

A general pair of quadratic forms can be simultaneously diagonalized; so that for general p, that there exist n linearly independent forms $L_j(x) = \sum_{k=1}^n \alpha_{j,k} x_k$ and complex numbers c_j so that

$$\frac{\partial p}{\partial x_{n-1}} = \sum_{j=1}^n L_j^2, \qquad \frac{\partial p}{\partial x_n} = \sum_{j=1}^n c_j L_j^2,$$

Mixed partials are equal, so $2\sum_{j=1}^n \alpha_{j,n} L_j = 2\sum_{j=1}^n c_j \alpha_{j,n-1} L_j$ and the L_j's are linearly independent, so $\alpha_{j,n} = c_j \alpha_{j,n-1}$. Now,

$$q(x_1, \dots, x_n) = p(x_1, \dots, x_n) - \sum_{j=1}^n \tfrac{c_j}{3\alpha_{j,n}} (\alpha_{j,1} x_1 + \cdots + \alpha_{j,n} x_n)^3$$

$$\implies \frac{\partial q}{\partial x_{n-1}} = \frac{\partial q}{\partial x_n} = 0 \implies q = q(x_1, \dots, x_{n-2}).$$

□

Reichstein's construction gives a cubic as a sum of roughly $\frac{n^2}{4}$ cubes, which is more than the actual minimum of roughly $\frac{n^2}{6}$.

Reichstein's construction gives a cubic as a sum of roughly $\frac{n^2}{4}$ cubes, which is more than the actual minimum of roughly $\frac{n^2}{6}$. We conclude the talk by veering off in another direction.

Reichstein's construction gives a cubic as a sum of roughly $\frac{n^2}{4}$ cubes, which is more than the actual minimum of roughly $\frac{n^2}{6}$.

We conclude the talk by veering off in another direction.

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

$$p(x_1, \dots, x_n) = \sum_{k=1}^{r} (\alpha_{k1}x_1 + \cdots + \alpha_{kn}x_n)^4 + q(x_1, \dots, x_m). \quad (12)$$

Reichstein's construction gives a cubic as a sum of roughly $\frac{n^2}{4}$ cubes, which is more than the actual minimum of roughly $\frac{n^2}{6}$.

We conclude the talk by veering off in another direction.

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

$$p(x_1, \dots, x_n) = \sum_{k=1}^{r} (\alpha_{k1} x_1 + \cdots + \alpha_{kn} x_n)^4 + q(x_1, \dots, x_m). \quad (12)$$

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

Reichstein's construction gives a cubic as a sum of roughly $\frac{n^2}{4}$ cubes, which is more than the actual minimum of roughly $\frac{n^2}{6}$.

We conclude the talk by veering off in another direction.

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

$$p(x_1, \dots, x_n) = \sum_{k=1}^{r} (\alpha_{k1}x_1 + \cdots + \alpha_{kn}x_n)^4 + q(x_1, \dots, x_m). \quad (12)$$

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.

Reichstein's construction gives a cubic as a sum of roughly $\frac{n^2}{4}$ cubes, which is more than the actual minimum of roughly $\frac{n^2}{6}$.

We conclude the talk by veering off in another direction.

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

$$p(x_1, \ldots, x_n) = \sum_{k=1}^{r} (\alpha_{k1} x_1 + \cdots + \alpha_{kn} x_n)^4 + q(x_1, \ldots, x_m). \quad (12)$$

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.

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 $n = 2^{2k} \cdot 3 \in A_4$.

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 $12p \in A_4$.

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 $p = 2, 3$.)

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 $p = 2, 3$.)
- 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.)

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 $p = 2, 3$.)
- 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.)

Thus the way is clear for a Reichstein canonical form for quintics. I hope one of you in the audience can find it.

Thanks

Thanks to Dave Anderson and Julianna Tymoczko for organizing this session and the workshop which preceded it, and for inviting me to speak.