Differential invariants by transvection

Jan A. Sanders¹ and Gloria Marí Beffa²

(1) vrije Universiteit amsterdam

(2) University of Wisconsin at Madison

ISLAND 3

Transvectants

The goal of this talk is to show that

transvectants

give us a natural language in which to describe the process of computing the **Differential Invariants** in many geometries.

classical invariant theory

Transvectants were invented in the middle of the 19th century to compute new covariants (or invariants) from old ones in classical invariant theory.

determinant

The simplest example is given by two linear forms. Their first transvectant is the determinant of the coefficients.

discriminant

Another example is the discriminant of a quadratic form, which is the second transvectant of the quadratic form with itself.

classical invariant theory

Transvectants were invented in the middle of the 19th century to compute new covariants (or invariants) from old ones in classical invariant theory.

determinant

The simplest example is given by two linear forms. Their first transvectant is the determinant of the coefficients.

discriminant

Another example is the discriminant of a quadratic form, which is the second transvectant of the quadratic form with itself.

classical invariant theory

Transvectants were invented in the middle of the 19th century to compute new covariants (or invariants) from old ones in classical invariant theory.

determinant

The simplest example is given by two linear forms. Their first transvectant is the determinant of the coefficients.

discriminant

Another example is the discriminant of a quadratic form, which is the second transvectant of the quadratic form with itself.

finite number

Gordan (1868) proved that every covariant of one or more binary forms (the groundforms) could be expressed as a polynomial of transvectants of the basic forms, and that there only exists a finite number of algebraically independent ones.

Emmy Noether

To carry out the algorithm in particular cases is not an easy task. It motivated one student of Gordan to great levels of abstraction.

finite number

Gordan (1868) proved that every covariant of one or more binary forms (the groundforms) could be expressed as a polynomial of transvectants of the basic forms, and that there only exists a finite number of algebraically independent ones.

Emmy Noether

To carry out the algorithm in particular cases is not an easy task. It motivated one student of Gordan to great levels of abstraction.

Hilbert 1890

Hilbert (1890) reproved Gordan's result, simplifying and generalizing the whole thing tremendously and laying the foundations for commutative algebra in the process.

nonconstructive proof

The only criticism one could have was that the proof was nonconstructive.

Hilbert 1893

In 1893 Hilbert gave a constructive proof.

Hilbert 1890

Hilbert (1890) reproved Gordan's result, simplifying and generalizing the whole thing tremendously and laying the foundations for commutative algebra in the process.

nonconstructive proof

The only criticism one could have was that the proof was nonconstructive.

Hilbert 1893

In 1893 Hilbert gave a constructive proof.

Hilbert 1890

Hilbert (1890) reproved Gordan's result, simplifying and generalizing the whole thing tremendously and laying the foundations for commutative algebra in the process.

nonconstructive proof

The only criticism one could have was that the proof was nonconstructive.

Hilbert 1893

In 1893 Hilbert gave a constructive proof.

Differential invariants

 \mathfrak{sl}_2

Let us define

$$F = \sum_{\alpha=1}^{n} \sum_{k=1}^{\infty} k(k-1) u_{k-1}^{\alpha} \frac{\partial}{\partial u_{k}^{\alpha}}$$
$$H = \sum_{\alpha=1}^{n} \sum_{k=0}^{\infty} 2k u_{k}^{\alpha} \frac{\partial}{\partial u_{k}^{\alpha}}$$
$$D = \sum_{\alpha=1}^{n} \sum_{k=0}^{\infty} u_{k+1}^{\alpha} \frac{\partial}{\partial u_{k}^{\alpha}}.$$

 $u = (u^1, \dots, u^n)$: curve, x: curve parameter u^{α} : coordinates on an *n*-manifold. u_k^{α} : kth (invariant) derivative. with respect to x.

Jan A. Sanders (1)

Differential invariants by transvection

The operators F, D and H form an \mathfrak{sl}_2 and all commute with the prolongation of vector fields of the form

$$\mathbf{v} = \frac{\partial}{\partial x} + \sum_{\alpha=1}^{n} \phi^{\alpha} \frac{\partial}{\partial u^{\alpha}}$$

This implies that there exists a representation of \mathfrak{sl}_2 in the space of differential invariants of curves.

If ψ is a differential invariant, so is $D\psi$.

The operators F, D and H form an \mathfrak{sl}_2 and all commute with the prolongation of vector fields of the form

$$\mathbf{v} = \frac{\partial}{\partial x} + \sum_{\alpha=1}^{n} \phi^{\alpha} \frac{\partial}{\partial u^{\alpha}}$$

This implies that there exists a representation of \mathfrak{sl}_2 in the space of differential invariants of curves.

If ψ is a differential invariant, so is $D\psi$.

The operators F, D and H form an \mathfrak{sl}_2 and all commute with the prolongation of vector fields of the form

$$\mathbf{v} = \frac{\partial}{\partial x} + \sum_{\alpha=1}^{n} \phi^{\alpha} \frac{\partial}{\partial u^{\alpha}}$$

This implies that there exists a representation of \mathfrak{sl}_2 in the space of differential invariants of curves.

If ψ is a differential invariant, so is $D\psi$.

The operators F, D and H form an \mathfrak{sl}_2 and all commute with the prolongation of vector fields of the form

$$\mathbf{v} = \frac{\partial}{\partial x} + \sum_{\alpha=1}^{n} \phi^{\alpha} \frac{\partial}{\partial u^{\alpha}}$$

This implies that there exists a representation of \mathfrak{sl}_2 in the space of differential invariants of curves.

If ψ is a differential invariant, so is $D\psi$.

This happens in particular if F is nilpotent, that is to say, for a given ψ there exists a k such that $F^k \psi = 0$.

For instance, polynomials in *u* and its derivatives.

This happens in particular if F is nilpotent, that is to say, for a given ψ there exists a k such that $F^k \psi = 0$.

For instance, polynomials in *u* and its derivatives.

This happens in particular if F is nilpotent, that is to say, for a given ψ there exists a k such that $F^k \psi = 0$.

For instance, polynomials in u and its derivatives.

This happens in particular if F is nilpotent, that is to say, for a given ψ there exists a k such that $F^k \psi = 0$.

For instance, polynomials in u and its derivatives.

Let U and V be \mathfrak{sl}_2 -modules. We define the *n*-transvectant $\tau^{(n)}: U \otimes V \to U \otimes V$ as follows.

Let $f \in U$ and $g \in V$, with *H*-eigenvalues $w_f = \omega(f)$ and $w_g = \omega(g)$, respectively, and we denote $f_i = D^i f, g_i = D^i g$.

Then

$$au^{(m)}f\otimes g=\sum_{i+j=m}(-1)^{i}inom{\omega_{f}+m-1}{j}f_{i}\otimesinom{\omega_{g}+m-1}{i}g_{j}$$

If the geometry is determined by a bilinear form, we denote the result of transvection followed by contraction by $(f,g)^{(m)}$.

< □ > < 同 > < 回 > < Ξ > < Ξ

Let U and V be \mathfrak{sl}_2 -modules. We define the *n*-transvectant $\tau^{(n)}: U \otimes V \to U \otimes V$ as follows.

Let $f \in U$ and $g \in V$, with *H*-eigenvalues $w_f = \omega(f)$ and $w_g = \omega(g)$, respectively, and we denote $f_i = D^i f, g_i = D^i g$.

Then

$$au^{(m)}f\otimes g=\sum_{i+j=m}(-1)^{i}inom{\omega_{f}+m-1}{j}f_{i}\otimesinom{\omega_{g}+m-1}{i}g_{j}$$

If the geometry is determined by a bilinear form, we denote the result of transvection followed by contraction by $(f,g)^{(m)}$.

Image: A math a math

Let U and V be \mathfrak{sl}_2 -modules. We define the *n*-transvectant $\tau^{(n)}: U \otimes V \to U \otimes V$ as follows.

Let $f \in U$ and $g \in V$, with *H*-eigenvalues $w_f = \omega(f)$ and $w_g = \omega(g)$, respectively, and we denote $f_i = D^i f, g_i = D^i g$.

Then

$$au^{(m)}f\otimes g=\sum_{i+j=m}(-1)^{i}\binom{\omega_{f}+m-1}{j}f_{i}\otimes\binom{\omega_{g}+m-1}{i}g_{j}$$

If the geometry is determined by a bilinear form, we denote the result of transvection followed by contraction by $(f,g)^{(m)}$.

• • • • • • • • • • • • •

Let U and V be \mathfrak{sl}_2 -modules. We define the *n*-transvectant $\tau^{(n)}: U \otimes V \to U \otimes V$ as follows.

Let $f \in U$ and $g \in V$, with *H*-eigenvalues $w_f = \omega(f)$ and $w_g = \omega(g)$, respectively, and we denote $f_i = D^i f, g_i = D^i g$.

Then

$$au^{(m)}f\otimes g=\sum_{i+j=m}(-1)^iinom{\omega_f+m-1}{j}f_i\otimesinom{\omega_g+m-1}{i}g_j$$

If the geometry is determined by a bilinear form, we denote the result of transvection followed by contraction by $(f,g)^{(m)}$.

< ロ > < 同 > < 三 > < 三

Lemma

If
$$f, g \in \ker F$$
, then $\tau^{(m)} f \otimes g \in \ker F$.

Corollary

If $f, g \in \ker F$, then $(f, g)^{(m)} \in \ker F$.

イロト イヨト イヨト イヨト

Lemma

If
$$f, g \in \ker F$$
, then $\tau^{(m)} f \otimes g \in \ker F$.

Corollary

If $f,g \in \ker F$, then $(f,g)^{(m)} \in \ker F$.

< ∃ > <

We now show how such a representation implies that transvection can take the role of differentiation in the process of finding differential invariants of parametrized curves.

By doing so we are capable of finding bases of differential invariants which are always in the kernel of *F*, and so in the complement of the image of *D*.

In many cases one can perform the process with appropriate weights to ensure the result is also invariant under reparametrizations.

We now show how such a representation implies that transvection can take the role of differentiation in the process of finding differential invariants of parametrized curves.

By doing so we are capable of finding bases of differential invariants which are always in the kernel of F, and so in the complement of the image of D.

In many cases one can perform the process with appropriate weights to ensure the result is also invariant under reparametrizations.

We now show how such a representation implies that transvection can take the role of differentiation in the process of finding differential invariants of parametrized curves.

By doing so we are capable of finding bases of differential invariants which are always in the kernel of F, and so in the complement of the image of D.

In many cases one can perform the process with appropriate weights to ensure the result is also invariant under reparametrizations.

Wylczinski

In the case of differential invariants of projective curves Wylczinski proved that one could lift a curve in $\mathbb{R}P^n$ to a curve in \mathbb{R}^{n+1} the standard way, multiply by a factor the lift and define a different vector $\mu \in \mathbb{R}^{n+1}$.

He then proceeded to recurrently differentiate this vector and to produce a basis of differential invariants for projective curves by taking determinants of the derivatives.

It just so happens that the vector μ is a relative invariant of the prolonged projective action with constant weight and the Wylczinski process mirrors the transvection process.

Wylczinski

In the case of differential invariants of projective curves Wylczinski proved that one could lift a curve in $\mathbb{R}P^n$ to a curve in \mathbb{R}^{n+1} the standard way, multiply by a factor the lift and define a different vector $\mu \in \mathbb{R}^{n+1}$.

He then proceeded to recurrently differentiate this vector and to produce a basis of differential invariants for projective curves by taking determinants of the derivatives.

It just so happens that the vector μ is a relative invariant of the prolonged projective action with constant weight and the Wylczinski process mirrors the transvection process.

A D > A A P >

Wylczinski

In the case of differential invariants of projective curves Wylczinski proved that one could lift a curve in $\mathbb{R}P^n$ to a curve in \mathbb{R}^{n+1} the standard way, multiply by a factor the lift and define a different vector $\mu \in \mathbb{R}^{n+1}$.

He then proceeded to recurrently differentiate this vector and to produce a basis of differential invariants for projective curves by taking determinants of the derivatives.

It just so happens that the vector μ is a relative invariant of the prolonged projective action with constant weight and the Wylczinski process mirrors the transvection process.

Transvectation can substitute differentiation in Wylczinski's original method.

In many other geometric flat manifolds G/H the method works identically for both differentiation and transvection.

That is, one can lift the curve to a vector in a higher dimensional \mathbb{R}^m where the group acts linearly.

We can modify the vector to produce relative differential invariants and we can then apply recurrent transvection of it with a properly chosen differential invariant, requiring the eigenvalue to be 2.

Transvectation can substitute differentiation in Wylczinski's original method.

In many other geometric flat manifolds G/H the method works identically for both differentiation and transvection.

That is, one can lift the curve to a vector in a higher dimensional \mathbb{R}^m where the group acts linearly.

We can modify the vector to produce relative differential invariants and we can then apply recurrent transvection of it with a properly chosen differential invariant, requiring the eigenvalue to be 2.

Transvectation can substitute differentiation in Wylczinski's original method.

In many other geometric flat manifolds G/H the method works identically for both differentiation and transvection.

That is, one can lift the curve to a vector in a higher dimensional \mathbb{R}^m where the group acts linearly.

We can modify the vector to produce relative differential invariants and we can then apply recurrent transvection of it with a properly chosen differential invariant, requiring the eigenvalue to be 2.

Transvectation can substitute differentiation in Wylczinski's original method.

In many other geometric flat manifolds G/H the method works identically for both differentiation and transvection.

That is, one can lift the curve to a vector in a higher dimensional \mathbb{R}^m where the group acts linearly.

We can modify the vector to produce relative differential invariants and we can then apply recurrent transvection of it with a properly chosen differential invariant, requiring the eigenvalue to be 2.

Transvectation can substitute differentiation in Wylczinski's original method.

In many other geometric flat manifolds G/H the method works identically for both differentiation and transvection.

That is, one can lift the curve to a vector in a higher dimensional \mathbb{R}^m where the group acts linearly.

We can modify the vector to produce relative differential invariants and we can then apply recurrent transvection of it with a properly chosen differential invariant, requiring the eigenvalue to be 2.

We have shown explicitly the projective, conformal and Grasmannian Lagrangian cases.

Thus, it is not at all surprising that Schwarzian and Euclidean curvature lie in the kernel of the operator F and can be written as transvectants of the derivative u_1 .

In fact, being differential invariants of lowest order within their corresponding geometrical background (projective and Euclidean respectively) they needed to be.

We have shown explicitly the projective, conformal and Grasmannian Lagrangian cases.

Thus, it is not at all surprising that Schwarzian and Euclidean curvature lie in the kernel of the operator F and can be written as transvectants of the derivative u_1 .

In fact, being differential invariants of lowest order within their corresponding geometrical background (projective and Euclidean respectively) they needed to be.

We have shown explicitly the projective, conformal and Grasmannian Lagrangian cases.

Thus, it is not at all surprising that Schwarzian and Euclidean curvature lie in the kernel of the operator F and can be written as transvectants of the derivative u_1 .

In fact, being differential invariants of lowest order within their corresponding geometrical background (projective and Euclidean respectively) they needed to be.

A question

The following construction is an answer to a question that was asked in 1999 by John McKay.

$\mathcal{M} \cong PSL(2)/H \cong \mathbb{R}P^1$

Assume that $\mathcal{M} \cong PSL(2)/H \cong \mathbb{R}P^1$. It is known that any differential invariant for curves in $\mathbb{R}P^1$ can be written as a function of the *Schwarzian derivative* of u,

$$S(u) = \frac{u_3}{u_1} - \frac{3}{2} \left(\frac{u_2}{u_1}\right)^2$$

and its derivatives with respect to x.

$S(u) \in \ker F$

It is also trivial to check that S(u) is in the kernel of F.

A question

The following construction is an answer to a question that was asked in 1999 by John McKay.

$\mathcal{M} \cong PSL(2)/H \cong \mathbb{R}\mathrm{P}^1$

Assume that $\mathcal{M} \cong PSL(2)/H \cong \mathbb{R}P^1$. It is known that any differential invariant for curves in $\mathbb{R}P^1$ can be written as a function of the *Schwarzian derivative* of u,

$$S(u) = \frac{u_3}{u_1} - \frac{3}{2} \left(\frac{u_2}{u_1}\right)^2$$

and its derivatives with respect to x.

$S(u) \in \ker F$

It is also trivial to check that S(u) is in the kernel of F.

A question

The following construction is an answer to a question that was asked in 1999 by John McKay.

$\mathcal{M} \cong PSL(2)/H \cong \mathbb{R}\mathrm{P}^1$

Assume that $\mathcal{M} \cong PSL(2)/H \cong \mathbb{R}P^1$. It is known that any differential invariant for curves in $\mathbb{R}P^1$ can be written as a function of the *Schwarzian derivative* of u,

$$S(u) = \frac{u_3}{u_1} - \frac{3}{2} \left(\frac{u_2}{u_1}\right)^2$$

and its derivatives with respect to x.

$S(u) \in \ker F$

It is also trivial to check that S(u) is in the kernel of F.

The Schwarzian as a transvectant

Schwarzian

We compute, with u_1 and $\frac{1}{u_1} \in \ker F$,

$$\begin{aligned} \tau^{(2)} u_1 \otimes \frac{1}{u_1} &= \sum_{i+j=2} {\binom{3}{j}} u_{i+1} \otimes D^j \frac{1}{u_1} \\ &= 3u_1 \otimes (2\frac{u_2^2}{u_1^3} - \frac{u_3}{u_1^2}) - 3u_2 \otimes \frac{u_2}{u_1^2} + u_3 \otimes \frac{1}{u_1} \end{aligned}$$

and this contracts by symmetrization to

$$3\frac{u_2^2}{u_1^2}-2\frac{u_3}{u_1}=-2S(u).$$

This shows that the Schwarzian is a transvectant, and a differential invariant by construction.

Jan A. Sanders (1)

The Schwarzian as a transvectant

Schwarzian

We compute, with u_1 and $\frac{1}{u_1} \in \ker F$,

$$\begin{aligned} \tau^{(2)} u_1 \otimes \frac{1}{u_1} &= \sum_{i+j=2} \binom{3}{j} u_{i+1} \otimes D^j \frac{1}{u_1} \\ &= 3u_1 \otimes (2\frac{u_2^2}{u_1^3} - \frac{u_3}{u_1^2}) - 3u_2 \otimes \frac{u_2}{u_1^2} + u_3 \otimes \frac{1}{u_1} \end{aligned}$$

and this contracts by symmetrization to

$$3\frac{u_2^2}{u_1^2}-2\frac{u_3}{u_1}=-2S(u).$$

This shows that the Schwarzian is a transvectant, and a differential invariant by construction.

Jan A. Sanders (1)

July 2007 16 / 23

The set of differential invariants of curves in this manifold under the action of Sp(n) has been classified.

Some of these invariants are projectively invariant and had been previously found.

The set of differential invariants of curves in this manifold under the action of Sp(n) has been classified.

Some of these invariants are projectively invariant and had been previously found.

The set of differential invariants of curves in this manifold under the action of Sp(n) has been classified.

Some of these invariants are projectively invariant and had been previously found.

The set of differential invariants of curves in this manifold under the action of Sp(n) has been classified.

Some of these invariants are projectively invariant and had been previously found.

symmetric matrix

A Lagrangian plane in \mathbb{R}^{2n} can be identified with a symmetric matrix and, in fact, the quotient $\operatorname{Sp}(n)/H$ can be locally identified with matrices of the form $\begin{pmatrix} I & u \\ 0 & I \end{pmatrix}$

where u is $n \times n$ and symmetric and where l represents the unit $n \times n$ matrix.

The subgroup H can thus be represented by matrices of the form

$$\left(\begin{array}{cc}I&0\\S&I\end{array}\right)\left(\begin{array}{cc}g&0\\0&g^{-T}\end{array}\right)$$

with $g \in \operatorname{GL}(n, \mathbb{R})$ and S symmetric.

symmetric matrix

A Lagrangian plane in \mathbb{R}^{2n} can be identified with a symmetric matrix and, in fact, the quotient $\operatorname{Sp}(n)/H$ can be locally identified with matrices of the form $\begin{pmatrix} I & u \\ 0 & I \end{pmatrix}$

where u is $n \times n$ and symmetric and where I represents the unit $n \times n$ matrix.

The subgroup H can thus be represented by matrices of the form

$$\left(\begin{array}{cc}I&0\\S&I\end{array}\right)\left(\begin{array}{cc}g&0\\0&g^{-T}\end{array}\right)$$

with $g \in GL(n, \mathbb{R})$ and S symmetric.

Let
$$v_k = u_1^{-1/2} u_k u_1^{-1/2}$$
.

With this description, it is known that a basis for the space of differential invariants of Lagrangian curves is given by the eigenvalues of the *Lagrangian Schwarzian* derivative of *u*

$$S(u) = v_3 - \frac{3}{2}v_2^2$$

together with the off-diagonal entries of the matrix of differential invariants

$$I_4 = v_4 - 2v_3v_2 - 2v_2v_3 + 3v_2^3$$

It was shown that I_4 contains in its diagonal the derivative of the eigenvalues of S(u).

Let
$$v_k = u_1^{-1/2} u_k u_1^{-1/2}$$
.

With this description, it is known that a basis for the space of differential invariants of Lagrangian curves is given by the eigenvalues of the Lagrangian Schwarzian derivative of u

$$S(u) = v_3 - \frac{3}{2}v_2^2$$

together with the off-diagonal entries of the matrix of differential invariants

$$I_4 = v_4 - 2v_3v_2 - 2v_2v_3 + 3v_2^3$$

It was shown that I_4 contains in its diagonal the derivative of the eigenvalues of S(u).

Let
$$v_k = u_1^{-1/2} u_k u_1^{-1/2}$$
.

With this description, it is known that a basis for the space of differential invariants of Lagrangian curves is given by the eigenvalues of the Lagrangian Schwarzian derivative of u

$$S(u) = v_3 - \frac{3}{2}v_2^2$$

together with the off-diagonal entries of the matrix of differential invariants

$$I_4 = v_4 - 2v_3v_2 - 2v_2v_3 + 3v_2^3$$

It was shown that I_4 contains in its diagonal the derivative of the eigenvalues of S(u).

July 2007 19 / 23

The Lagrangian Schwarzian generates projective differential invariants for Lagrangian planes.

One can show that, if $g = \Theta u_1^{-1/2}$, where $\Theta \in O(n)$ is such that $\Theta S(u)\Theta^T$ is diagonal (uniquely determined by Gramm-Schmidt), then

$$\mu = \left(\begin{array}{c} u \\ I \end{array}\right) g^{\mathsf{T}}$$

is a relative invariant for the action of Sp(n) on the space of Lagrangian planes.

We denote by $\mathbf{D} = \Theta S(u) \Theta^T$ the diagonalization of S(u). These will be our initial differential invariants. They are generated by a certain contraction of a transvectant of u_1 . One can show that, if $g = \Theta u_1^{-1/2}$, where $\Theta \in O(n)$ is such that $\Theta S(u)\Theta^T$ is diagonal (uniquely determined by Gramm-Schmidt), then

$$\mu = \left(\begin{array}{c} u \\ I \end{array}\right) g^{\mathsf{T}}$$

is a relative invariant for the action of Sp(n) on the space of Lagrangian planes.

We denote by $\mathbf{D} = \Theta S(u) \Theta^T$ the diagonalization of S(u). These will be our initial differential invariants. They are generated by a certain contraction of a transvectant of u_1 .

Let us consider the group-invariant contraction

$$C: M_{2n \times n} \otimes M_{2n \times n} \to M_{n \times n}$$

given by $C(v, w) = v^T J w$, where

$$J = \left(\begin{array}{cc} 0 & I \\ -I & 0 \end{array}\right).$$

Jan A. Sanders (1)

Theorem

Let us call $V_0 = \mu$ and define $V_{i+1} = (\mathbf{D}, V_i)^{(1)}$, i = 1, 2, ... Then, the entries of $\mathbf{D} = \mathbf{\Theta}\mathbf{S}(\mathbf{u})\mathbf{\Theta}^{\mathsf{T}}$ and $C(V_2, V_0)^{(0)} = V_2^{\mathsf{T}}JV_0$ generate all differential invariants for curves of Lagrangian planes under the action of the symplectic group.