A story about Sturm, Courant, Gelfand, and Arnold's last published paper

Pierre Bérard
Université Grenoble Alpes

Conference in honor of Bernard Helffer
Nantes, April 2019

Introduction

Work in collaboration with Bernard Helffer, motivated by two papers,

- N. Kuznetsov. On delusive nodal sets of free oscillations. Newsletter of the European Mathematical Society 96 (2015), 34–40.
- V. Arnold. Topological properties of eigenoscillations in mathematical physics. Proc. Steklov Inst. Math. 273 (2011), 25–34. *(Paper submitted by Arnold in December 2009, six months before his death on June 6, 2010).*

The first two sections of Arnold's paper are entitled:

1. Correct and Incorrect Theorems of Courant
2. Courant–Gelfand Theorem

Theorem (Courant, 1923)

Consider the eigenvalue problem,

$$\begin{cases} -\Delta u & = \mu\, u \textit{ in } \Omega \subset \mathbb{R}^2 \\ u|_{\partial\Omega} & = 0 \end{cases}$$

with eigenvalues and eigenfunctions,

$$\begin{cases} \mu_1 < \mu_2 \leq \mu_3 \leq \cdots \nearrow \infty \\ \phi_1, \phi_2, \phi_3, \ldots \end{cases}$$

Then,

$$\forall \phi \in \mathcal{E}(\mu), \quad \beta_0(\phi) \leq \kappa(\mu)\,,$$

where

$$\begin{aligned} \mathcal{E}(\mu) \quad &:= \textit{ eigenspace associated with } \mu. \\ \beta_0(\phi) \quad &:= \#\{\textit{nodal domains of } \phi\}. \\ \kappa(\mu) \quad &:= \min\{k \mid \mu_k = \mu\}. \end{aligned}$$

Motivated by Hilbert's 16th problem, Arnold was interested in the following

Statement ("Courant's generalized theorem", in Arnold's words)

Any linear combination of the first n eigenfunctions divides the domain, by means of its nodes, into no more than n subdomains.

This statement appears in Courant-Hilbert, Methods of mathematical physics, Volume I, footnote page 454 (1953 English translation), with a cross-reference to the Göttingen dissertation of H. Herrmann, Beiträge zur Theorie der Eigenwerte und Eigenfunktionen (1932).

It turns out that neither Herrmann's thesis, nor his published papers contain this statement, let alone its proof.

According to Arnold, "Courant's generalized theorem" is true for $\mathbb{RP}^2$, the real projective space in dimension 2, and false for $\mathbb{RP}^3$.

The proof in dimension 2 (folklore ?, J. Leydold, 1996), and the construction of counterexamples in dimension 3 (O. Viro, 1979), rely on real algebraic geometry.

$\rightsquigarrow$ It turns out that there are counterexamples to "Courant's generalized theorem" for domains in $\mathbb{R}^n$, $n \geq 2$ (including convex domains), with the Dirichlet or Neumann boundary conditions (B. & Helffer, 2018).

Arnold and Gelfand (imaginary dialog)

(Gelfand) *I thought that, except for me, nobody paid attention to Courant's remarkable assertion. But I was so surprised that I delved into it and found a proof.*

[Arnold is quite surprised, but does not have time to mention Viro's counterexamples before Gelfand continues.]

However, I could prove this theorem of Courant only for oscillations of one-dimensional media.

(Arnold) *Where could I read it?*

(Gelfand) *I never write proofs. I just discover new interesting things. Finding proofs (and writing articles) is up to my students.*

Arnold sketches Gelfand's ideas to prove "Courant's generalized theorem", in dimension 1, and refers to the result as the "Courant-Gelfand theorem".

Theorem (Courant-Gelfand)

The zeros of a linear combination of the n first eigenfunctions of the Sturm-Liouville problem

$$(1) \qquad \begin{cases} -y''(x) + q(x)\,y(x) = \lambda\,y(x) \ \text{in }]0,1[\,, \\ y(0) = y(1) = 0\,, \end{cases}$$

divide the interval into at most n sub-intervals.

Quotations from Arnold's Section 2

- Unfortunately, [Gelfand's hints] do not yet provide a *proof* of [Courant's] generalized theorem: many facts are still to be proved.

- Gelfand did not publish anything concerning this: he only told me that he hoped his students would correct this drawback of his theory.

- Viktor Borisovich Lidskii told me that "he knows how to prove all this". . . . Although [Lidskii's] arguments look convincing, the lack of a published formal text with a proof of the Courant-Gelfand theorem is still distressing.

According to Kuznetsov, Gelfand's approach appealed so much to Arnold that he included this theorem, together with Gelfand's hint of proof, in the third Russian edition of his book "Ordinary Differential Equations" (English translation, Springer 1992, Problem 9 in the section "Supplementary problems").

9. Prove that the zeros of a linear combination of the first n eigenfunctions of the Sturm-Liouville problem

$$u_{xx} + q(x)u = \lambda u, \quad u(0) = u(l) = 0, \quad q > 0$$

divide the interval $[0, l]$ into at most n parts.

Hint. (I. M. Gel'fand). Convert to fermions, i.e., to skew-symmetric solutions of the equation $\sum u_{x_i x_j} + \sum q(x_i)u = \lambda u$ and use the fact that the first eigenfunction of this equation has no zeros inside the fundamental simplex $0 < x_1 < \cdots < x_n < l$.

Arnold, Ordinary differential equations (Springer, 1992)

It turns out that Arnold's "Courant-Gelfand theorem" is actually a weak form of a theorem which goes back to Charles François Sturm (1833).

The purpose of this talk is to explain a complete proof, à la Gelfand, of Sturm's theorem.

P. B. & B. Helffer. Sturm's theorem on the zeros of sums of eigenfunctions: Gelfand's strategy implemented. arXiv:1807.03990. To appear in Moscow Math. Journal.

Theorem (Sturm, 1833)

For $n \geq 1$, let $\phi = \sum_{j=1}^{n} a_j h_j$ be a nontrivial linear combination of the n first eigenfunctions, $h_1, \ldots, h_n$, of the Sturm-Liouville problem

$$(1) \qquad \begin{cases} -y''(x) + q(x)\,y(x) = \lambda\,y(x) \text{ dans }]0,1[\,, \\ y(0) = y(1) = 0\,, \end{cases}$$

1. *The function ϕ has at most $(n-1)$ distinct zeros, counted with multiplicities, in the open interval $]0,1[$.*
2. *If $a_1 = \cdots = a_k = 0$, the function ϕ changes sign at least k times in the interval $]0,1[$.*

Remark. The second assertion is frequently referred to as the "Sturm-Hurwitz theorem".

Proof of Sturm's theorem, following Gelfand's idea

Let q be a real function, C^∞ in a neighbourhood of $\mathrm{I} :=]0,1[$.
Consider the 1-particle operator

$$\mathfrak{h}^{(1)} := -\frac{d^2}{dx^2} + q(x)\,,$$

and more precisely its Dirichlet realization in I, i.e., the Dirichlet eigenvalue problem

$$\begin{cases} -\frac{d^2y}{dx^2} + q\,y = \lambda\,y\,, \\ y(0) = y(1) = 0\,. \end{cases}$$

Let $\{(\lambda_j, h_j), j \geq 1\}$ be the eigenpairs of $\mathfrak{h}^{(1)}$, with

$$\lambda_1 < \lambda_2 < \lambda_3 < \cdots\,,$$

and $\{h_j, j \geq 1\}$ an orthonormal basis of associated eigenfunctions.

Following Gelfand's idea, consider the Dirichlet realization $\mathfrak{h}^{(n)}$ of the n-particle operator in I^n,

$$\mathfrak{h}^{(n)} := \sum_{j=1}^{n} \Big(-\frac{\partial^2}{\partial x_j^2} + q(x_j) \Big) = -\Delta + Q \,,$$

where $Q(x_1, \ldots, x_n) = q(x_1) + \cdots + q(x_n)$.

Let $\vec{k} = (k_1, \cdots, k_n)$ denote a vector with integer components, and $\vec{x} = (x_1, \cdots, x_n)$ a vector in I^n. A complete set of eigenpairs of $\mathfrak{h}^{(n)}$ is given by the $(\Lambda_{\vec{k}}, H_{\vec{k}})$, with

$$\begin{cases} \Lambda_{\vec{k}} = \lambda_{k_1} + \cdots + \lambda_{k_n} \,, \text{ and} \\ H_{\vec{k}}(\vec{x}) = h_{k_1}(x_1) \cdots h_{k_n}(x_n) \,. \end{cases}$$

The symmetric group $\mathfrak{s}_n$ acts on I^n by $\sigma(\vec{x}) = (x_{\sigma(1)}, \cdots, x_{\sigma(n)})$, if $\vec{x} = (x_1, \cdots, x_n)$. It also acts on $L^2(\mathrm{I}^n)$, and on the functions $H_{\vec{k}}$. A fundamental domain of the action of $\mathfrak{s}_n$ on I^n is the n-simplex

$$\Omega_n^{\mathrm{I}} := \{0 < x_1 < x_2 < \cdots < x_n < 1\}.$$

Following Gelfand, we restrict the operator to functions which are anti-symmetric under the action of $\mathfrak{s}_n$ (Fermions), and consider in particular, the Slater determinant $\mathfrak{S}_n$, defined by,

$$\mathfrak{S}_n(x_1, \ldots, x_n) = \begin{vmatrix} h_1(x_1) & h_1(x_2) & \ldots & h_1(x_n) \\ h_2(x_1) & h_2(x_2) & \ldots & h_2(x_n) \\ \vdots & \vdots & & \vdots \\ h_n(x_1) & h_n(x_2) & \ldots & h_n(x_n) \end{vmatrix}.$$

Let $\vec{h}_{[n]}(x)$ denote the vector

$$\vec{h}_{[n]}(x) = {}^T(h_1(x), \ldots, h_n(x)) \ .$$

With this notation, we write the Slater determinant as the determinant of its column vectors,

$$\mathfrak{S}_n(x_1, \ldots, x_n) = \left| \vec{h}_{[n]}(x_1) \ldots \vec{h}_{[n]}(x_n) \right| \ .$$

A linear combination of the n first eigenfunctions can be written as

$$\langle \vec{b}_{[n]}, \vec{h}_{[n]} \rangle = b_1 h_1 + \cdots + b_n h_n \, .$$

Lemma

For all $n \geq 1$,

1. $\mathfrak{S}_n \not\equiv 0$ *on* I^n.
2. $\mathfrak{S}_n$ *is the first Dirichlet eigenfunction of* $\mathfrak{h}^{(n)}$ *in* Ω_n^{I}. *In particular, for all* $0 < x_1 < \cdots < x_n < 1$, *one has* $\mathfrak{S}_n(x_1, \ldots, x_n) \neq 0$.
3. *For all* $0 < x_1 < \cdots < x_n < 1$, *the* n *vectors* $\vec{h}_{[n]}(x_i), 1 \leq i \leq n$ *are linearly independent in* $\mathbb{R}^n$.

Remark. The second assertion is Gelfand's Observation **A** in Arnold's paper.

Proof of the lemma

The first assertion is proved by induction, and follows from the linear independence of the eigenfunctions h_j, and from the fact that h_1 does not vanish in $]0,1[$.

The second assertion follows from the relation

$$\mathfrak{h}^{(n)}\mathfrak{S}_n := (-\Delta + Q)\,\mathfrak{S}_n = (\lambda_1 + \cdots + \lambda_n)\,\mathfrak{S}_n\,,$$

from the fact that the least eigenvalue of the operator $\mathfrak{h}^{(n)}$ acting on Fermions is precisely $\Lambda_n := \lambda_1 + \cdots + \lambda_n$, and from the standard argument to prove that an eigenfunction associated with the least eigenvalue does not change sign.

The third assertion is an immediate consequence of the second.

$\square$

Proposition (Zeros without multiplicities)

Let $n \geq 1$ and let $\phi = \langle \vec{b}_{[n]}, \vec{h}_{[n]} \rangle$ be a nontrivial linear combination of the eigenfunctions $h_1, \ldots, h_n$.

1. *The function ϕ has at most $(n-1)$ distinct zeros in $]0,1[$.*
2. *If the function ϕ has exactly $(n-1)$ distinct zeros $c_1 < \cdots < c_{n-1}$ in $]0,1[$ then, there exists a nonzero constant C such that*

$$\phi(x) = C\, \mathfrak{S}_n(c_1, \ldots, c_{n-1}, x), \quad \forall x \in]0,1[\,.$$

Remarks. The first assertion is precisely the Courant-Gelfand theorem. The second assertion is an amended version of Gelfand's Observation **B** in Arnold's paper.

Proof of the proposition (zeros without multiplicities)

• Assume that ϕ has at least n distinct zeros, $c_1 < \cdots < c_n$. One then has the linear system,

$$\begin{aligned} &b_1 h_1(c_1) + \cdots b_n h_n(c_1) = 0 \\ &\ldots \\ &b_1 h_1(c_n) + \cdots b_n h_n(c_n) = 0 \end{aligned}$$

whose determinant is $\mathfrak{S}_n(c_1, \ldots, c_n) \neq 0$, and hence the linear combination is trivial.

• The coefficient of $h_n(x)$ in the linear combination $\mathfrak{S}_n(c_1, \ldots, c_{n-1}, x)$ is precisely $\mathfrak{S}_{n-1}(c_1, \ldots, c_{n-1}) \neq 0$. It is nontrivial, and of the form $\langle \vec{s}_{[n]}, \vec{h}_{[n]} \rangle$.
From the assumptions, the two vectors $\vec{b}_{[n]}$ and $\vec{s}_{[n]}$ are orthogonal to the $(n-1)$ vectors $\vec{h}_{[n]}(c_1)$, ..., $\vec{h}_{[n]}(c_{n-1})$ which are linear independent because $\mathfrak{S}_{n-1}(c_1, \ldots, c_{n-1}) \neq 0$. $\square$

Proposition (Zeros without multiplicities)

Let $n \geq 1$ and let $\phi = \langle \vec{b}_{[n]}, \vec{h}_{[n]} \rangle$ be a nontrivial linear combination of the eigenfunctions $h_1, \ldots, h_n$.

1. *The function ϕ has at most $(n-1)$ distinct zeros in $]0,1[$.*
2. *If the function ϕ has exactly $(n-1)$ distinct zeros $c_1 < \cdots < c_{n-1}$ in $]0,1[$ then, there exists a nonzero constant C such that*

$$\phi(x) = C\,\mathfrak{S}_n(c_1, \ldots, c_{n-1}, x), \quad \forall x \in]0,1[\,.$$

Victor Kleptsyn recently sent us an unpublished draft, dated June 27 2011, with a proof of this proposition, along Gelfand's ideas.

However, the proposition is not satisfactory in view of Sturm's theorem. Indeed, we would like to take the multiplicities of zeros into account.

Notation

Let $n, p \geq 1$ with $n \geq p$.

Let $0 < \bar{c}_1 < \cdots < \bar{c}_p < 1$ be distinct points, with multiplicities, $\mathrm{mult}(\bar{c}_j) = k_j \geq 1$, such that $k_1 + \cdots + k_p = n$.

Denote by $\mathfrak{S}_n^*(\bar{c}_1, k_1; \ldots; \bar{c}_p, k_p)$ the Generalized Slater determinant

$$\left| \vec{h}_{[n]}(\bar{c}_1) \ldots \vec{h}_{[n]}^{(k_1-1)}(\bar{c}_1) \ \ldots \ \vec{h}_{[n]}(\bar{c}_p) \ldots \vec{h}_{[n]}^{(k_p-1)}(\bar{c}_p) \right| ,$$

whose column vectors are the $\vec{h}_{[n]}$ and their successive derivatives, at the points $\bar{c}_j$, with the multiplicities k_j.

Proposition (Zeros with multiplicities)

Let $n \geq 1$ and let $\phi = \langle \vec{b}_{[n]}, \vec{h}_{[n]} \rangle$ be a nontrivial linear combination of the eigenfunctions $h_1, \ldots, h_n$.

1. *The function ϕ has at most $(n-1)$ zeros, counted with multiplicities, in $]0,1[$.*
2. *If the function ϕ has exactly p distinct zeros $0 < \bar{c}_1 < \cdots < \bar{c}_p < 1$, with multiplicities $\mathrm{mult}(\bar{c}_j) = k_j \geq 1$, such that $k_1 + \cdots + k_p = n-1$ then, there exists a nonzero constant C such that*
$$\phi(x) = C\, \mathfrak{S}_n^*(\bar{c}_1, k_1; \ldots; \bar{c}_p, k_p; x) \quad \forall x \in]0,1[\,.$$
3. *Furthermore, if $b_1 = \cdots = b_k = 0$, the function ϕ changes sign at least k times in $]0,1[$.*

Proof of the proposition (zeros with multiplicities)

Assume that the number of zeros of ϕ, counted with multiplicities, is at least n. This means that there is a linear system of n equations with n unknowns,

$$\begin{aligned}
&b_1 h_1(\bar{c}_1) + \cdots + b_n h_n(\bar{c}_1) = 0 \\
&\ldots \\
&b_1 h_1^{(k_1-1)}(\bar{c}_1) + \cdots + b_n h_n^{(k_1-1)}(\bar{c}_1) = 0 \\
&\ldots \\
&b_1 h_1(\bar{c}_p) + \cdots + b_n h_n(\bar{c}_p) = 0 \\
&\ldots \\
&b_1 h_1^{(k_p-1)}(\bar{c}_p) + \cdots + b_n h_n^{(k_p-1)}(\bar{c}_p) = 0
\end{aligned}$$

We need to study the determinant of this system: this is the subject of the next lemma.

The third assertion is a consequence of the second one (Liouville's remark).

Lemma (Zeros with multiplicities)

Let $n \geq 1$ and $n \geq p \geq 1$. Let $0 < \bar{c}_1 < \cdots < \bar{c}_p < 1$ be distinct points with multiplicities $\mathrm{mult}(\bar{c}_j) = k_j \geq 1$, such that $k_1 + \cdots + k_p = n$.

1. *Then, $\mathfrak{S}_n^*(\bar{c}_1, k_1; \ldots; \bar{c}_p, k_p) \neq 0$.*
2. *In particular, the n vectors*

$$\vec{h}_{[n]}(\bar{c}_1), \ldots, \vec{h}_{[n]}^{(k_1-1)}(\bar{c}_1), \ \ldots, \ \vec{h}_{[n]}(\bar{c}_p), \ldots, \vec{h}_{[n]}^{(k_p-1)}(\bar{c}_p)$$

are linearly independent in $\mathbb{R}^n$.

Proof of the lemma (zeros with multiplicities)

It suffices to consider the case $p \leq n-1$, in which at least one multiplicity is bigger than or equal to 2.

Consider the point

$$\vec{c}_{[n]} = (\bar{c}_1, \ldots, \bar{c}_1, \ldots, \bar{c}_p, \ldots, \bar{c}_p)$$

where, for $1 \leq j \leq p$, $\bar{c}_j$, is repeated k_j times.

The point $\vec{c}_{[n]}$ belongs to the boundary of the simplex Ω_n^{I}, and is a zero of $\mathfrak{S}_n$.

Since

$$\mathfrak{h}^{(n)}(\mathfrak{S}_n) = \Lambda_n \, \mathfrak{S}_n$$

in I^n, one can apply Bers' theorem at $\vec{c}_{[n]}$.

Application of Bers' theorem

Let $\vec{x}_{[n]} = \vec{c}_{[n]} + \vec{\xi}_{[n]}$. Then, there exists a harmonic homogeneous polynomial $\widehat{P}_k$, of degree k, in $\mathbb{R}^n$ such that

$$\mathfrak{S}_n(\vec{c}_{[n]} + \vec{\xi}_{[n]}) = \widehat{P}_k(\vec{\xi}_{[n]}) + \omega_{k+1}(\vec{\xi}_{[n]}),$$

where $\omega_{k+1}(t\vec{\xi}_{[n]}) = O(t^{k+1})$ when t tends to 0.

We do not a priori know the degree k.

It turns out, in the present situation, that one can determine $\widehat{P}_k$ up to a multiplicative constant.

Notation

We work in a neighbourhood of the point

$$\vec{c}_{[n]} = (\bar{c}_1, \ldots, \bar{c}_1, \ \ldots, \ \bar{c}_p, \ldots, \bar{c}_p)$$

We group the components of the vectors $\vec{x}$ et $\vec{\xi}$ in p groups of lengths k_1, ..., k_p, and write these vectors as

$$\vec{x} = \left(x^{(1)}, \ldots, x^{(p)}\right) \quad \text{and} \quad \vec{\xi} = \left(\xi^{(1)}, \ldots, \xi^{(p)}\right) .$$

For $m \geq 1$, call P_m the Vandermonde polynomial defined by $P_1(x_1) = 1$ and, for $m \geq 2$,

$$P_m(x_1, \ldots, x_m) = (x_1 - x_2) \cdots (x_1 - x_m) \ldots (x_{m-1} - x_m) .$$

With these notation, we have the following lemma.

Lemma (Local form of $\mathfrak{S}_n$)

The polynomial $\widehat{P}_k$ *has the following properties.*

1. *For any permutation* $\sigma = (\sigma_1, \ldots, \sigma_p) \in \mathfrak{s}_{k_1} \times \cdots \times \mathfrak{s}_{k_p} \subset \mathfrak{s}_n$,
$$\widehat{P}_k(\sigma \cdot \vec{\xi}\,) = \varepsilon(\sigma)\, \widehat{P}_k(\vec{\xi}\,)\,.$$
2. *The set of real zeros of* $\widehat{P}_k$ *is characterized by*
$$\widehat{P}_k(\vec{\xi}\,) = 0 \Leftrightarrow \prod_{j=1}^{p} P_{k_j}(\xi^{(j)}) = 0\,.$$
3. *There exists a nonzero constant* $\rho(\vec{c}\,)$ *such that*
$$\widehat{P}_k(\vec{\xi}\,) = \rho(\vec{c})\, P_{k_1}(\xi^{(1)}) \ldots P_{k_p}(\xi^{(p)})\,.$$

In particular, $\widehat{P}_k$ has degree $k = \sum_j \frac{k_j(k_j-1)}{2}$, and

$$\mathfrak{S}_n(\vec{c} + \vec{\xi}\,) = \rho(\vec{c}\,)\, P_{k_1}(\xi^{(1)}) \ldots P_{k_p}(\xi^{(p)}) + \omega_{k+1}(\vec{c}, \vec{\xi}\,)\,.$$

Examples in dimension 5, for the harmonic oscillator

One can in particular apply Bers' theorem to the Vandermonde polynomial which is harmonic and, up to multiplication by an exponential, the Slater determinant of the harmonic oscillator.

Exemple 1.
Take $\vec{c}_{[5]} = (\bar{c}_1, \bar{c}_1, \bar{c}_2, \bar{c}_2, \bar{c}_3)$, with $\bar{c}_1 < \bar{c}_2 < \bar{c}_3$. Then,

$$\widehat{P}_k(\vec{\xi}) = \rho_2 \, (\xi_1 - \xi_2)(\xi_3 - \xi_4) = \rho_2 \, P_2(\xi_1, \xi_2) P_2(\xi_3, \xi_4) \, .$$

Exemple 2.
Take $\vec{c}_{[5]} = (\bar{c}_1, \bar{c}_1, \bar{c}_1, \bar{c}_2, \bar{c}_3)$, with $\bar{c}_1 < \bar{c}_2 < \bar{c}_3$. Then,

$$\widehat{P}_k(\vec{\xi}) = \rho_3 \, (\xi_1 - \xi_2)(\xi_1 - \xi_3)(\xi_2 - \xi_3) = \rho_3 \, P_3(\xi_1, \xi_2, \xi_3) \, .$$

Let us come back to our lemma ...

Lemma (Local form of $\mathfrak{S}_n$)

The polynomial $\widehat{P}_k$ has the following properties.

1. *For any permutation $\sigma = (\sigma_1, \dots, \sigma_p) \in \mathfrak{s}_{k_1} \times \cdots \times \mathfrak{s}_{k_p} \subset \mathfrak{s}_n$,*

$$\widehat{P}_k(\sigma \cdot \vec{\xi}\,) = \varepsilon(\sigma)\, \widehat{P}_k(\vec{\xi}\,)\,.$$

2. *The set of real zeros of $\widehat{P}_k$ is characterized by*

$$\widehat{P}_k(\vec{\xi}\,) = 0 \Leftrightarrow \prod_{j=1}^{p} P_{k_j}(\xi^{(j)}) = 0\,.$$

3. *There exists a nonzero constant $\rho(\vec{c}\,)$ such that*

$$\widehat{P}_k(\vec{\xi}\,) = \rho(\vec{c})\, P_{k_1}(\xi^{(1)}) \dots P_{k_p}(\xi^{(p)})\,.$$

Proof of the lemma (local form of $\mathfrak{S}_n$)

We have $\mathfrak{S}_n(\vec{c} + \vec{\xi}\,) = \widehat{P}_k(\vec{\xi}\,) + \omega_{k+1}(\vec{\xi}\,)$. Look at the action of permutations, and use the fact that $\mathfrak{S}_n$ is anti-symmetric.

• The first assertion in the lemma, and the $(\Leftarrow)$ part of the second assertion, follow from the local analysis.

• For the part $(\Rightarrow)$, suppose that there exists $\vec{\eta}$ such that $\widehat{P}(\vec{\eta}) = 0$ and $\prod_{j=1}^{p} P_{k_j}(\eta^{(j)}) \neq 0$. Since $\widehat{P}$ is harmonic, there exist $\vec{\eta}^{\,\pm}$ as close to $\vec{\eta}$ as we want, such that $\widehat{P}(\vec{\eta}^{\,+})\,\widehat{P}(\vec{\eta}^{\,-}) < 0$. Looking at $\mathfrak{S}_n(\vec{c} + t\vec{\eta}^{\,\pm})$, we arrive at a contradiction with the fact that $\mathfrak{S}_n$ does not change sign in the simplex Ω_n^{I}.

We have proved that $\widehat{P}_k$ and the product of Vandermonde polynomials, have the same set of real zeros.

• The third assertion follows from a result on the divisibility of harmonic real polynomials. The statement we need here goes back to Brelot et Choquet (1954), see also Murdoch (1964) et Logunov-Mallinikova (2015).

Theorem

Let P et Q be two real polynomials in n variables. Assume that P is harmonic, and that the sets of real zeros of the polynomials P and Q satisfy $\mathcal{Z}_{\mathbb{R}}(P) \subset \mathcal{Z}_{\mathbb{R}}(Q)$. Then, there exists a real polynomial R such that $Q = P\,R$.

$\square$

Proof of the lemma (zeros with multiplicities)

The formula

$$\mathfrak{S}_n(\vec{c}+\vec{\xi}\,) = \rho(\vec{c}\,)\, P_{k_1}(\xi^{(1)}) \dots P_{k_p}(\xi^{(p)}) + \omega_{k+1}(\vec{c}, \vec{\xi}\,)\,.$$

allows us to evaluate the successive derivatives of the function $\mathfrak{S}_n(x_1, \dots, x_n)$ at the point $\vec{c}$.

Observe that

$$\partial_{x_m}^{m-1}\, \partial_{x_{m-1}}^{m-2} \cdots \partial_{x_3}^{2}\, \partial_{x_2}\, P_m = (-1)^{\frac{m(m-1)}{2}}\, (m-1)!\, (m-2)!\, \dots\, 2!\,,$$

and obtain

$$\mathfrak{S}_n^*(\bar{c}_1, k_1; \dots; \bar{c}_p, k_p) \neq 0$$

which is precisely what the lemma “zeros with multiplicities” asserts. $\square$

The proof of the proposition follows easily.

Topological classification of linear combinations of eigenfunctions

Let $\mathcal{L}(n)$ be the vector space of linear combinations of the first n eigenfunctions $h_1, \ldots, h_n$. Let $0 \neq \phi \in \mathcal{L}(n)$. Let

$$\mathcal{Z}(\phi) = \begin{cases} \emptyset \text{ if } \phi \text{ does not vanish in }]0,1[\,, \\ \{(\bar{c}_1, k_1), \ldots (\bar{c}_p, k_p)\} \text{ otherwise}\,, \end{cases}$$

be the zero set of ϕ, where $0 < \bar{c}_1 < \cdots < \bar{c}_p < 1$ are the distinct zeros of ϕ, and where the positive integers k_j are the corresponding multiplicities.

Define $m\left(\mathcal{Z}(\phi)\right) := k_1 + \cdots + k_p + 1$. Then, according to Sturm's theorem, $m(\mathcal{Z}) \leq n$.

Given a set $\mathcal{Z}_0$ as above, define the function

$$\mathfrak{S}^*(\mathcal{Z}_0;x) := \varepsilon(\mathcal{Z}_0) \left| \vec{h}_{[m]}(\bar{c}_1) \ldots \vec{h}_{[m]}^{(k_1-1)}(\bar{c}_1) \ldots \vec{h}_{[m]}(\bar{c}_p) \ldots \vec{h}_{[m]}^{(k_p-1)}(\bar{c}_p) \vec{h}_{[m]}(x) \right|,$$

where $m = m(\mathcal{Z}_0)$, $\vec{h}_{[m](x)}$ is the column vector $(h_1(x), \ldots, h_m(x))$, and $\varepsilon(\mathcal{Z}_0) = \pm 1$ is such that $\mathfrak{S}^*(\mathcal{Z}_0; x)$ is positive in $]0, \bar{c}_1[$, with the convention that $\mathfrak{S}^*(\emptyset; x) = h_1(x) > 0$.

Define

$$\mathcal{L}(n; \mathcal{Z}_0) := \{0 \neq \phi \in \mathcal{L}(n) \mid \mathcal{Z}(\phi) = \mathcal{Z}_0\}.$$

Denote by $\mathcal{L}(n, \mathcal{Z}_0, +)$ the subset of ϕ's in $\mathcal{L}(n; \mathcal{Z}_0)$ such that $\phi|_{]0,\bar{c}_1[} > 0$. If $\phi \in \mathcal{L}(n, \mathcal{Z}_0)$, then ϕ or $-\phi$ is in $\mathcal{L}(n, \mathcal{Z}_0, +)$.

Proposition

For any $\phi \in \mathcal{L}(n, \mathcal{Z}_0, +)$, *and any* $j \in \{1, \ldots, p\}$,

$$\begin{cases} \operatorname{sign}\left(\phi^{(k_j)}(\bar{c}_j)\right) = (-1)^{k_1+\cdots+k_j}\,, \\ \operatorname{sign}\left(\phi|_{]\bar{c}_j, \bar{c}_{j+1}[}\right) = (-1)^{k_1+\cdots+k_j}\,, \end{cases}$$

where $\bar{c}_{p+1} = 1$.

In particular, the set $\mathcal{L}(n, \mathcal{Z}_0, +)$ *retracts continuously to the function* $\mathfrak{S}^*(\mathcal{Z}_0; \cdot)$. *More precisely, for any* $\phi \in \mathcal{L}(n, \mathcal{Z}_0, +)$, *the map* $\phi_t = (1-t)\,\mathfrak{S}^*(\mathcal{Z}_0; \cdot) + t\,\phi$, *for* $t \in [0, 1]$, *is a continuous curve from* $\phi_0 = \mathfrak{S}^*(\mathcal{Z}_0; \cdot)$ *to* $\phi_1 = \phi$, *entirely contained in* $\mathcal{L}(n, \mathcal{Z}_0, +)$.

Some historical remarks

Arnold's "Courant-Gelfand theorem" is a weak form of a theorem of Charles François Sturm. Sturm's results were presented as a Memoir to the Paris Academy of Sciences in September 1833, and published in the first volume of Liouville's journal in 1836 (two papers, pp. 106–186 and pp. 373–444).

Sturm's first proof of the above theorem (1833) used the heat equation. In 1836, J. Liouville published a direct (ode) proof replacing the heat flow by a discrete family of functions $\{U_\ell, \ell \geq 0\}$. In the second paper of 1836, Sturm gives two proofs, one using the heat equation, the second being a direct (ode) proof, different from Liouville's proof. Where Liouville uses Rolle's theorem, Sturm studies sign variations. Liouville also remarks that the second assertion in Sturm's theorem, follows from the first one.

Go to the proof

In his book “The theory of sound” (1877), Lord Rayleigh proves Sturm's “beautiful theorem”, following Liouville's ideas and, on this occasion, introduces a “Slater determinant”.

Sturm's theorem (on zeros of sums of eigenfunctions) is explained in F. Pockels (1891), and then seems to sink into oblivion. M. Bôcher (“Leçons sur les méthodes de Sturm”, 1917) does not mention it, neither do Courant-Hilbert, Arnold, Kuznetsov, etc. .

▸ Go to Liouville's proof

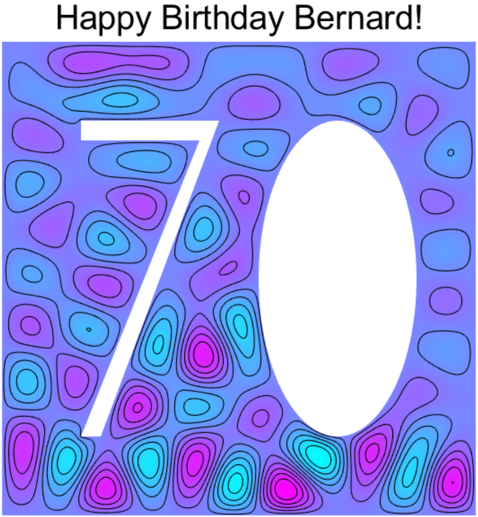
Heat equation approach, see: Galaktionov (2004) and Galaktionov-Harwin (2005).

P. B. & B. Helffer. Sturm's theorem on zeros of linear combinations of eigenfunctions. Expositiones Math. 23 (2018), or arXiv:1706.08247.

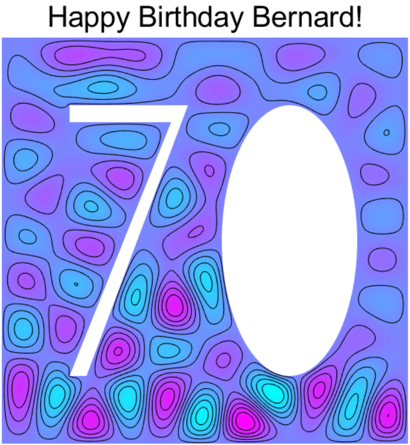
The non-vanishing of determinants such as $\mathfrak{S}_n$ appear in a 1916-paper by O. Kellogg, who investigates when an orthonormal family of continuous real functions is oscillatory. A. Haar (1917) relates the non-vanishing property of the determinants to the approximation of functions (Chebyshev systems).

Kellogg (1918) determines conditions on the kernel of an integral operator under which the associated eigenfunctions form an oscillatory family. This led to the theory of "oscillating matrices and kernels" which was developed by the Russian school, see in particular the book by F. Gantmacher and M. Krein ("Oscillatory matrices and kernels"). In this book, Gantmacher and Krein prove (a version of) Sturm's theorem via the associated integral equation.

Note that anti-symmetric functions (Fermions) also appear in this more general framework. Gelfand's approach of Sturm's theorem is more direct, but the ideas are very similar.



Level sets of the 70th Dirichlet eigenfunction of the birthday cake



Level sets of the 70th Dirichlet eigenfunction of the birthday cake

Thank you for your attention

Thank you to the organizers for this nice celebration!

Proof of Sturm's theorem following Liouville and Rayleigh

Given any $1 \leq m \leq n$, let $U = \sum_{k=m}^{n} a_k \, h_k$ be any nontrivial real linear combination of eigenfunctions of (1).

$$(1) \qquad \begin{cases} -\,y''(x) + q(x)\,y(x) = \lambda\,y(x) \text{ dans }]0,1[\,, \\ y(0) = y(1) = 0\,, \end{cases}$$

Proof that U has at most $(n-1)$ zeros, counting multiplicities (Liouville's proof)

Write equation (1) for h_1 and for h_k, and get the relation

$$\big(h_1\, h_k' - h_1'\, h_k\big)' = (\lambda_1 - \lambda_k)\, h_1\, h_k\,.$$

Multiply by a_k, and sum from $k = m$ to $k = n$ to obtain

$$(2) \qquad \big(h_1\, U' - h_1'\, U\big)' = h_1\, U_1\,,$$

where $U_1 = \sum_{k=m}^{n} (\lambda_1 - \lambda_k)\, a_k\, h_k$. Equivalently,

$$U_1 = U'' + (\lambda_1 - q)\, U\,.$$

More generally,

$$\left(h_1\, U_\ell' - h_1'\, U_\ell\right)' = h_1\, U_{\ell+1}\,,$$

where $U_\ell = \sum_{k=m}^{n} (\lambda_1 - \lambda_k)^\ell\, a_k\, h_k$. Equivalently,

$$U_{\ell+1} = U_\ell'' + (\lambda_1 - q)\, U_\ell\,.$$

Note that this implies that U cannot vanish at infinite order at a point.

Integrate the relation $\left(h_1\, U' - h_1'\, U\right)' = h_1\, U_1$ from 0 to x, use the Dirichlet boundary condition, and obtain

$$h_1(x)\, U'(x) - h_1'(x)\, U(x) = \int_0^x h_1(t)\, U_1(t)\, dt\,.$$

or

$$h_1^2(x)\, \frac{d}{dx} \frac{U}{h_1}(x) = \int_0^x h_1(t)\, U_1(t)\, dt\,.$$

Count zeros with multiplicities. Assume that U has N zeros in $]0,1[$. Then so does $\frac{U}{h_1}$.

By Rolle's theorem, $\frac{d}{dx}\frac{U}{h_1}$ has *at least* $(N-1)$ zeros in $]0,1[$, and so does the function $x \mapsto \int_0^x h_1(t)\, U_1(t)\, dt$.

The function $x \mapsto \int_0^x h_1(t)\, U_1(t)\, dt$ has at least $(N-1)$ zeros in $]0,1[$. It vanishes at both 0 and 1 because the h_j form an orthonormal family.

By Rolle's theorem, its derivative, $h_1\, U_1$, has at least N zeros in $]0,1[$.

Repeat the argument (Sturm's idea): for any $\ell \geq 1$, $U_\ell = \sum_{k=m}^{n} (\lambda_1 - \lambda_k)^\ell\, a_k\, h_k$ has *at least* N zeros in $]0,1[$.

Let ℓ tend to infinity, use the fact that the eigenvalues λ_k are simple, and the fact that h_n has $(n-1)$ zeros in $]0,1[$, to conclude that $N \leq (n-1)$. $\square$

Proof that U changes sign at least $(m-1)$ times, if $a_1 = \cdots = a_{m-1} = 0$ (Liouville's proof revisited by Rayleigh)

It is easy to prove that the following two assertions are equivalent:

(i) For any $n \geq 1$, any nontrivial real linear combination $\sum_{j=1}^{n} c_j h_j$ has at most $(n-1)$ distinct zeros in $]0,1[$.

(ii) For any $n \geq 1$, and any $x_1 < \cdots < x_n$ in $]0,1[$, $\det\left(h_i(x_j)\right)_{1\leq i,j\leq n} \neq 0$.

Assume that U changes sign exactly p times at the points $z_1 < \cdots < z_p$ in the interval $]0,1[$, and that $p < (m-1)$, i.e., $p+1 \leq m-1$. Consider the function,

$$V(x) := \begin{vmatrix} h_1(z_1) & \dots & h_1(z_p) & h_1(x) \\ \vdots & & \vdots & \vdots \\ h_{p+1}(z_1) & \dots & h_{p+1}(z_p) & h_{p+1}(x) \end{vmatrix}.$$

From the above assertions, the function V is not identically zero. It vanishes at the points $z_j, 1 \leq j \leq p$, and it is a linear combination of the eigenfunctions $h_1, \ldots, h_{p+1}$.

The first assertion in Sturm's theorem tells us that V does not have any other zero, and that each z_j has order 1, so that V changes sign precisely at the points z_j.

Since $p + 1 \leq m - 1$, the functions U and V are orthogonal, and since they vanish and change sign at the same points, their product $U\,V$ does not change sign in $]0, 1[$. It follows that $U\,V$ vanishes identically, a contradiction.

Back to Historical remarks