



**HAL**  
open science

# **Sturm's theorem on zeros of linear combinations of eigenfunctions**

Pierre Bérard, Bernard Helffer

► **To cite this version:**

Pierre Bérard, Bernard Helffer. Sturm's theorem on zeros of linear combinations of eigenfunctions. *Expositiones Mathematicae*, 2020, 38 (1), 27–50 / <http://dx.doi.org/10.1016/j.exmath.2018.10.002>. 10.1016/j.exmath.2018.10.002 . hal-01541645v4

**HAL Id: hal-01541645**

**<https://hal.science/hal-01541645v4>**

Submitted on 14 Oct 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# STURM'S THEOREM ON ZEROS OF LINEAR COMBINATIONS OF EIGENFUNCTIONS

PIERRE BÉRARD AND BERNARD HELFFER

ABSTRACT. Motivated by recent questions about the extension of Courant's nodal domain theorem, we revisit a theorem published by C. Sturm in 1836, which deals with zeros of linear combination of eigenfunctions of Sturm-Liouville problems. Although well known in the nineteenth century, this theorem seems to have been ignored or forgotten by some of the specialists in spectral theory since the second half of the twentieth-century. Although not specialists in History of Sciences, we have tried to put this theorem into the context of nineteenth century mathematics.

*To appear in Expositiones Mathematicae 2018,  
except for the Appendices C to E (in blue).*

## 1. INTRODUCTION

In this paper, we are interested in the following one-dimensional eigenvalue problem, where  $r$  denotes the spectral parameter.

$$(1.1) \quad \frac{d}{dx} \left( K \frac{dV}{dx} \right) + (rG - L)V = 0, \text{ for } x \in ]\alpha, \beta[,$$

$$(1.2) \quad \left( K \frac{dV}{dx} - hV \right) (\alpha) = 0,$$

$$(1.3) \quad \left( K \frac{dV}{dx} + HV \right) (\beta) = 0.$$

Here,

$$(1.4) \quad K, G, L : [\alpha, \beta] \rightarrow \mathbb{R} \text{ are positive functions,}$$

$$(1.5) \quad h, H \in [0, \infty] \text{ are non negative constants, possibly infinite.}$$

*Remark 1.1.* When  $h = \infty$  (resp.  $H = \infty$ ), the boundary condition should be understood as the Dirichlet boundary condition  $V(\alpha) = 0$  (resp. as the Dirichlet boundary condition  $V(\beta) = 0$ ).

---

*Date:* October 3, 2017. Revised July 31, 2018.

*2010 Mathematics Subject Classification.* 34B24, 34L10, 34L99.

*Key words and phrases.* Sturm-Liouville eigenvalue problem, Sturm's theorems.

Precise assumptions on  $K, G, L$  are given below.

Note that when  $K = G \equiv 1$ , (1.1)–(1.3) is an eigenvalue problem for the classical operator  $-\frac{d^2V}{dx^2} + LV$ .

This eigenvalue problem, in the above generality ( $K, G, L$  functions of  $x$ ), was first studied by Charles Sturm in a Memoir presented to the Paris Academy of sciences in September 1833, summarized in [37, 38], and published in [39, 40].

*Remark 1.2.* In this paper, we have mainly retained the notation of [39], except that we use  $[\alpha, \beta]$  for the interval, instead of Sturm’s notation  $[x, X]$ . We otherwise use today notation and vocabulary. Note that in [40], Sturm uses lower case letters for the functions  $K, G, L$ , the same notation as Joseph Fourier in [13].

As far as the eigenvalue problem (1.1)–(1.3) is concerned, Sturm’s results can be roughly summarized in the following theorems.

**Theorem 1.3** (Sturm, 1836). *Under the assumptions (1.4)–(1.5), the eigenvalue problem (1.1)–(1.3) admits an increasing infinite sequence  $\{\rho_i, i \geq 1\}$  of positive simple eigenvalues, tending to infinity. Furthermore, the associated eigenfunctions  $V_i$  have the following remarkable property: the function  $V_i$  vanishes, and changes sign, precisely  $(i - 1)$  times in the open interval  $] \alpha, \beta [$ .*

**Theorem 1.4** (Sturm, 1836). *Let  $Y = A_m V_m + \dots + A_n V_n$  be a non trivial linear combination of eigenfunctions of the eigenvalue problem (1.1)–(1.3), with  $1 \leq m \leq n$ , and  $\{A_j, m \leq j \leq n\}$  real constants such that  $A_m^2 + \dots + A_n^2 \neq 0$ . Then, the function  $Y$  has at least  $(m - 1)$ , and at most  $(n - 1)$  zeros in the open interval  $] \alpha, \beta [$ .*

The first theorem today appears in most textbooks on Sturm-Liouville theory. Although well known in the nineteenth century, the second theorem (as well as the more precise Theorem 2.15) seems to have been ignored or forgotten by some of the specialists in spectral theory since the second half of the twentieth-century, as the following chronology indicates.

**1833:** Sturm’s Memoir presented to the Paris Academy of sciences in September, summarized in [37, 38].

**1836:** Sturm’s papers [39, 40] published. Joseph Liouville summarizes Sturm’s results in [23, § III, p. 257], and uses them to study the expansion of a given function  $f$  into a series of eigenfunctions of (1.1).

**1877:** Lord Rayleigh writes “*a beautiful theorem has been discovered by Sturm*” as he mentions Theorem 1.4 in [34, Section 142].

**1891:** F. Pockels [30, pp. 68-73] gives a summary of Sturm’s results, including Theorem 1.4, and mentions the different proofs

provided by Sturm, Liouville and Rayleigh. On the basis of a note of Sturm in Férussac's Bulletin [36], Pockels (p. 71, lines 12-17) also suggests that Sturm may have looked for a statement in higher dimension as well, without success. Sturm indeed mentions studying an example with spherical symmetry in dimension 3 (leading to an ordinary differential equation with singularity), to which he may have applied Theorem 1.4.

**1903:** Hurwitz [19] gives a lower bound for the number of zeros of the sum of a trigonometric series with a spectral gap and refers, somewhat inaccurately, to Sturm's Theorems. This result, known as the Sturm-Hurwitz theorem, already appears in a more general framework in Liouville's paper [23].

See [12, § 2] for a generalization of the Sturm-Hurwitz theorem to Fourier integrals with a spectral gap, [28] for geometric applications, and the recent paper [35] which quantifies the Sturm-Hurwitz theorem.

**1931:** Courant & Hilbert [10, 11] extensively mention the Sturm-Liouville problem. They do not refer to the original papers of Sturm, but to Bôcher's book [8] which does not include Theorem 1.4. They then state an extension of the so-called Courant's nodal domain theorem to linear combination of eigenfunctions, [10, footnote, p. 394] and [11, footnote, p. 454], and refer to the dissertation of H. Herrmann [18]. It turns out that neither Herrmann's dissertation, nor his later papers, consider this extension of Courant's Theorem.

**1950:** The book [15] by F. Gantmacher and M. Krein contains several notes on Sturm's contributions. One result (Corollary, Chap. III.5, p. 138), stated in the context of Chebyshev systems, is stronger than Theorem 1.4, yet weaker than Theorem 2.15. The book does however not mention [40].

**1956:** Pleijel mentions Sturm's Theorem 1.4, somewhat inaccurately, in [29, p. 543 and 550].

**1973:** V. Arnold [1] points out that an extension of Courant's theorem to linear combinations of eigenfunctions cannot be true in general. Counterexamples were first given by O. Viro for the 3-sphere (with the canonical metric) [41] and, more recently in the papers [4, 5], see also [17].

It seems to us that Arnold may have not been aware of Theorem 1.4. Indeed, in [3], see also the Supplementary problem 9 in [2, p. 327], he mentions a proof, suggested by I. Gelfand, of the upper bound in Theorem 1.4. Gelfand's idea is to "use fermions rather than bosons", and to apply Courant's nodal domain theorem in the fermionic context. However, Arnold concludes by writing [3, p. 30], "*the arguments [given by Gelfand] do not yet provide a proof*". It is interesting to note that Liouville's and

Rayleigh's proofs of the lower bound in Theorem 1.4 use an idea similar to Gelfand's, see the proof of Claim 3.5.

As far as we know, the first implementation of Gelfand's idea into a complete proof of Theorem 1.4 is given in [5, 6].

*Remark 1.5.* In [40], Theorem 1.4 first appears as a corollary to a much deeper theorem [40, § XXIV], in which Sturm describes the time evolution of the  $x$ -zeros of a solution  $u(x, t)$  of the heat equation. We shall not consider this topic here, and we refer to [14, 26] for modern formulations and a historical analysis.

Our interest in Theorem 1.4 arose from reading [20], and investigating Courant's nodal domain theorem and its extension to linear combination of eigenfunctions.

The main purpose of this paper is to popularize Theorems 1.4 and 2.15, as well as Sturm's originality and ideas. Sturm's results are clearly stated in the summaries [37, 38]. Unfortunately, Sturm's detailed papers [39, 40] are written linearly, and contain very few tagged statements. Our second purpose is to provide an accessible proof of Theorems 1.4 and 2.15, meeting today standards of rigor. We in particular state precise assumptions, clarify some technical points, and provide some alternative proofs. We otherwise closely follow the original proofs, and we provide precise cross-references to Sturm's papers.

In this paper, we make the following strong assumptions.

$$(1.6) \quad \begin{cases} [\alpha, \beta] \subset ]\alpha_0, \beta_0[, \\ K, G, L \in C^\infty(] \alpha_0, \beta_0[), \\ K, G, L > 0 \text{ on } ] \alpha_0, \beta_0[. \end{cases}$$

*Remark 1.6.* Neither Sturm nor Liouville make any explicit regularity assumptions, see Subsection 4.3 and Remark 3.7 for more details.

**Organization of the paper.** In Section 2, we prove Theorem 2.15, Sturm's refined version of Theorem 1.4, following the ideas of [40, § XXVI]. In Section 3, we prove Theorem 3.2, Liouville's version of Theorem 1.4, following [23, 24]. In Section 4, we describe the context of Sturm's papers and his ideas. Appendix A provides the detailed proof of a technical argument. Appendix B contains the citations from Sturm's papers in their original French formulation. Appendix C considers Sturm's theorem under weaker assumptions. Appendices D and E provide cross-references between our paper and the papers of Sturm and Liouville.

**Acknowledgements.** The authors would like to thank N. Kuznetsov and J. Lützen for their comments on an earlier version of this paper. They also thank the anonymous referees for their constructive remarks.

## 2. STURM'S O.D.E. PROOF OF THEOREM 1.4

## 2.1. Preliminary lemmas and notation.

2.1.1. Recall that  $\{(\rho_j, V_j), j \geq 1\}$  are the eigenvalues and eigenfunctions of the eigenvalue problem (1.1)–(1.3).

By our assumption  $L > 0$ , the eigenvalues are positive,  $\rho_j > 0$ . Under the Assumptions (1.6), the functions  $V_j$  are  $C^\infty$  on  $]\alpha_0, \beta_0[$ . This follows from Cauchy's existence and uniqueness theorem, or from Liouville's existence proof [23]. Note that the assumption  $L > 0$  is convenient, but not necessary. It actually suffices that  $\frac{L}{G}$  be bounded from below.

In this section, we fix

$$(2.1) \quad Y = \sum_{j=m}^n A_j V_j,$$

a linear combination of eigenfunctions of the eigenvalue problem (1.1)–(1.3), where  $1 \leq m \leq n$ , and where the  $A_j$  are real constants.

*Remark 2.1.* We shall always assume that  $Y \not\equiv 0$ , which is equivalent to assuming that  $\sum_m^n A_j^2 \neq 0$ . As far as the statement of Theorem 1.4 is concerned, and without loss of generality, it is simpler to assume that  $A_m A_n \neq 0$ .

We also introduce the associated family of functions,  $\{Y_k, k \in \mathbb{Z}\}$ , where

$$(2.2) \quad Y_k = (-1)^k \sum_{j=m}^n \rho_j^k A_j V_j.$$

Note that  $Y_0$  is the original linear combination  $Y$ , and that  $Y_k \equiv 0$  if and only if  $Y \equiv 0$ .

Roughly speaking, Sturm's idea is to show that the number of zeros of  $Y_k$ , in the interval  $]\alpha, \beta[$ , is non-decreasing with respect to  $k$ , and then to take the limit when  $k$  tends to infinity, see Subsection 2.3. Up to changing the constants  $A_j$ , it suffices to compare the numbers of zeros of  $Y$  and  $Y_1$ . For this purpose, Sturm compares the signs of  $Y$  and  $Y_1$  near the zeros of  $Y$  (Lemma 2.4), and at the non-zero local extrema of  $Y$  (Lemmas 2.7 and 2.9). The main ingredient for this purpose is the differential relation (2.6). In the sequel, we indicate the pages in Sturm's papers corresponding to the different steps of the proof.

2.1.2. For  $m \leq p \leq n$ , write the equations satisfied by the eigenfunction  $V_p$ ,

$$(2.3) \quad \frac{d}{dx} \left( K \frac{dV_p}{dx} \right) + (\rho_p G - L)V_p = 0,$$

$$(2.4) \quad \left( K \frac{dV_p}{dx} - hV_p \right) (\alpha) = 0,$$

$$(2.5) \quad \left( K \frac{dV_p}{dx} + HV_p \right) (\beta) = 0,$$

and multiply the  $p$ -th equation by  $\rho_p^k A_p$ . Summing up from  $p = m$  to  $n$ , yields the following lemma.

**Lemma 2.2.** *Assume that (1.6) holds. Let  $k \in \mathbb{Z}$ .*

- (1) *The function  $Y_k$  satisfies the boundary conditions (1.2) and (1.3).*
- (2) *The functions  $Y_k$  and  $Y_{k+1}$  satisfy the differential relation*

$$(2.6) \quad GY_{k+1} = K \frac{d^2 Y_k}{dx^2} + \frac{dK}{dx} \frac{dY_k}{dx} - LY_k.$$

- (3) *Under the Assumptions (1.6), the function  $Y_k$  cannot vanish at infinite order at a point  $\xi \in [\alpha, \beta]$ , unless  $Y \equiv 0$ .*

**Proof.** [40, p. 437] Assertions (1) and (2) are clear by linearity.

For Assertion (3), assume that  $Y_k \not\equiv 0$ , and that it vanishes at infinite order at some  $\xi$ . Then, according to (2.6) and its successive derivatives, the function  $Y_{k+1}$  also vanishes at infinite order at  $\xi$ , and so does  $Y_\ell$  for any  $\ell \geq k$ . Assume, as indicated in Remark 2.1, that  $A_n \neq 0$ . Fixing some  $p \geq 0$ , we can write, for any  $\ell \geq k$ ,

$$\frac{d^p V_n}{dx^p}(\xi) + \sum_{j=m}^{n-1} \left( \frac{\rho_j}{\rho_n} \right)^\ell \frac{A_j}{A_n} \frac{d^p V_j}{dx^p}(\xi) = 0.$$

Since  $\rho_n > \rho_j$  for  $m \leq j \leq n-1$ , letting  $\ell$  tend to infinity, we conclude that  $\frac{d^p V_n}{dx^p}(\xi) = 0$ . This would be true for all  $p$ , which is impossible by Cauchy's uniqueness theorem, or by Sturm's argument [39, § II].  $\square$

*Remark 2.3.* Assertion (3), and the fact that the zeros of  $Y$  are isolated, with finite multiplicities, are implicit in [40].

**Lemma 2.4.** *Assume that (1.6) holds. Let  $U$  denote any  $Y_k$ , and  $U_1 = Y_{k+1}$ . Let  $\xi \in [\alpha, \beta]$  be a zero of  $U$ , of order  $p \geq 2$ . Then, there exist constants  $B_\xi$  and  $B_{1,\xi}$ , and smooth functions  $R_\xi$  and  $R_{1,\xi}$ , such that*

$$(2.7) \quad \begin{cases} U(x) = B_\xi(x - \xi)^p + (x - \xi)^{p+1}R_\xi(x), \\ U_1(x) = B_{1,\xi}(x - \xi)^{p-2} + (x - \xi)^{p-1}R_{1,\xi}(x), \\ \text{with } B_\xi B_{1,\xi} > 0. \end{cases}$$

**Proof.** [40, p. 439] Assume that  $\xi$  is a zero of order  $p \geq 2$  of  $U$ , so that

$$U(\xi) = \cdots = \frac{d^{p-1}U}{dx^{p-1}}(\xi) = 0$$

and

$$\frac{d^p U}{dx^p}(\xi) \neq 0.$$

Taylor's formula with integral remainder term, see Laplace [21, p. 179] (in Livre premier, Partie 2, Chap. 3, § 44), gives the existence of some function  $R_\xi$  such that

$$U(x) = B_\xi(x - \xi)^p + (x - \xi)^{p+1}R_\xi(x),$$

where

$$B_\xi = \frac{1}{p!} \frac{d^p U}{dx^p}(\xi) \neq 0.$$

Equation (2.6) implies that

$$(GU_1)(x) = p(p-1)B_\xi(x - \xi)^{p-2}K(x) + (x - \xi)^{p-1}S_\xi(x),$$

for some smooth function  $S_\xi$ . It follows that

$$U_1(x) = B_{1,\xi}(x - \xi)^{p-2} + (x - \xi)^{p-1}R_{1,\xi}(x),$$

for some function  $R_{1,\xi}$ , with  $B_{1,\xi} = p(p-1)\frac{K(\xi)}{G(\xi)}B_\xi$ .

In particular,  $B_{1,\xi}B_\xi > 0$  and this proves the lemma.  $\square$

**Lemma 2.5.** *Assume that (1.6) holds. Assume that  $h \in [0, \infty[$ , i.e., that the boundary condition at  $\alpha$  is not the Dirichlet boundary condition. Let  $U$  denote any  $Y_k$ ,  $U_1 = Y_{k+1}$ , and assume that  $U(\alpha) = 0$ . Then,  $\alpha$  is a zero of  $U$  of even order, i.e., there exists  $n_U \in \mathbb{N} \setminus \{0\}$  such that  $\frac{d^p U}{dx^p}(\alpha) = 0$  for  $0 \leq p \leq 2n_U - 1$  and  $\neq 0$  for  $p = 2n_U$ .*

*When  $H \in [0, \infty[$ , a similar statement holds at the boundary  $\beta$ .*

**Proof.** [40, p. 440-441] Assume that  $U(\alpha) = 0$ . By Lemma 2.2,  $U$  does not vanish at infinite order at  $\alpha$ , so that there exists  $p \geq 1$  with

$$U(\alpha) = \cdots = \frac{d^{p-1}Y_k}{dx^{p-1}}(\alpha) = 0$$

and

$$\frac{d^p U}{dx^p}(\alpha) \neq 0.$$

Taylor's formula with integral remainder term gives

$$U(x) = B_\alpha(x - \alpha)^p + (x - \alpha)^{p+1}R_\alpha(x),$$

where  $B_\alpha = \frac{1}{p!} \frac{d^p U}{dx^p}(\alpha) \neq 0$ .

The boundary condition at  $\alpha$  implies that  $\frac{dU}{dx}(\alpha) = 0$ , and hence that  $p \geq 2$ . By Lemma 2.4, we can write

$$U_1(x) = B_{1,\alpha}(x - \alpha)^{p-2} + (x - \alpha)^{p-1}R_{1,\alpha}(x),$$



with  $B_{1,\alpha} B_\alpha > 0$ .

If  $p = 2$ , then  $U_1(\alpha) \neq 0$ . If  $p > 2$ , one can continue.

If  $p = 2q$ , one arrives at

$$Y_{k+q}(x) = B_{k+q,\alpha} + (x - \alpha)R_{k+q,\alpha}(x),$$

with  $Y_{k+q}(\alpha) = B_{k+q,\alpha}$  and  $B_{k+q,\alpha} B_{k,\alpha} > 0$ .

If  $p = 2q + 1$ , one arrives at

$$Y_{k+q}(x) = B_{k+q,\alpha}(x - \alpha) + (x - \alpha)^2 R_{k+q,\alpha}(x),$$

with  $B_{k+q,\alpha} B_{k,\alpha} > 0$  and  $\frac{dY_{k+q}}{dx}(\alpha) = B_{k+q,\alpha} \neq 0$ . On the other-hand, since  $Y_{k+q}$  satisfies (1.2) and  $Y_{k+q}(\alpha) = 0$ , we must have  $\frac{dY_{k+q}}{dx}(\alpha) = 0$ , because  $h > 0$ . This yields a contradiction and proves that the case  $p = 2q + 1$  cannot occur. The lemma is proved.  $\square$

**2.2. Counting zeros.** Assume that (1.6) holds. Let  $U$  denote any  $Y_k$ , and  $U_1 = Y_{k+1}$ . They satisfy the relation (2.6).

From Lemma 2.2, we know that  $U$  cannot vanish at infinite order at a point  $\xi \in ]\alpha, \beta[$ . If  $\xi \in ]\alpha, \beta[$  and  $U(\xi) = 0$ , we define the *multiplicity*  $m(U, \xi)$  of the zero  $\xi$  by

$$(2.8) \quad m(U, \xi) = \min\{p \mid \frac{d^p U}{dx^p}(\xi) \neq 0\}.$$

From Lemma 2.5, we know that the multiplicity  $m(U, \alpha)$  is even if  $h \in [0, \infty[$ , and that the multiplicity  $m(U, \beta)$  is even if  $H \in [0, \infty[$ . We define the *reduced multiplicity* of  $\alpha$  by

$$(2.9) \quad \overline{m}(U, \alpha) = \begin{cases} \frac{1}{2}m(U, \alpha) & \text{if } h \in [0, \infty[, \\ 0 & \text{if } h = \infty, \end{cases}$$

and a similar formula for the reduced multiplicity of  $\beta$ .

By Lemma 2.2, the function  $U$  has finitely many distinct zeros  $\xi_1(U) < \xi_2(U) < \dots < \xi_p(U)$  in the interval  $]\alpha, \beta[$ . We define the *number of zeros of  $U$  in  $]\alpha, \beta[$ , counted with multiplicities*, by

$$(2.10) \quad N_m(U, ]\alpha, \beta[) = \sum_{j=1}^p m(U, \xi_j(U)),$$

and we use the notation  $N_m(U)$  whenever the interval is clear.

We define the *number of zeros of  $U$  in  $[\alpha, \beta]$ , counted with multiplicities*, by

$$(2.11) \quad \overline{N}_m(U, [\alpha, \beta]) = \sum_{j=1}^p m(U, \xi_j(U)) + \overline{m}(U, \alpha) + \overline{m}(U, \beta),$$

and we use the notation  $\overline{N}_m(U)$  whenever the interval is clear.

We define the *number of zeros of  $U$  in  $] \alpha, \beta[$*  (multiplicities not accounted for) by

$$(2.12) \quad N(U, ] \alpha, \beta[) = p,$$

and we use the notation  $N(U)$  whenever the interval is clear.

Finally, we define the *number of sign changes of  $U$  in the interval  $] \alpha, \beta[$*  by

$$(2.13) \quad N_v(U, ] \alpha, \beta[) = \sum_{j=1}^p \frac{1}{2} \left[ 1 - (-1)^{m(U, \xi_j(U))} \right].$$

*Remark 2.6.* Note that sign changes of the function  $Y$  correspond to zeros with odd multiplicity.

**2.3. Comparing the numbers of zeros of  $Y_k$  and  $Y_{k+1}$ .** Assume that (1.6) holds. Let  $U$  be some  $Y_k$  and  $U_1 = Y_{k+1}$ . In this subsection, we show that the number of zeros of  $U_1$  is not smaller than the number of zeros of  $U$ .

**Lemma 2.7.** *Let  $\xi < \eta$  be two zeros of  $U$  in  $[\alpha, \beta]$ . Then, there exists some  $a_{\xi, \eta} \in ] \xi, \eta[$  such that  $U(a_{\xi, \eta})U_1(a_{\xi, \eta}) < 0$ .*

*Remark 2.8.* We do not assume that  $\xi, \eta$  are consecutive zeros.

**Proof.** [40, p. 437] Since  $U$  cannot vanish identically in  $] \xi, \eta[$  (see Lemma 2.2), there exists some  $x_0 \in ] \xi, \eta[$  such that  $U(x_0) \neq 0$ . Let  $\varepsilon_0 = \text{sign}(U(x_0))$ . Then  $\varepsilon_0 U$  takes a positive value at  $x_0$ , and hence  $M := \sup\{\varepsilon_0 U(x) \mid x \in [\xi, \eta]\}$  is positive and achieved at some  $a_{\xi, \eta} \in ] \xi, \eta[$ . Denote this point by  $a$  for short, then,

$$\varepsilon_0 U(a) > 0, \quad \frac{dU}{dx}(a) = 0 \quad \text{and} \quad \varepsilon_0 \frac{d^2U}{dx^2}(a) \leq 0.$$

It follows from (2.6) that  $\varepsilon_0 U_1(a) < 0$ , or equivalently, that  $U(a)U_1(a) < 0$ . The lemma is proved.  $\square$

**Lemma 2.9.** *Let  $\xi \in ] \alpha, \beta[$ . Assume that  $U(\xi) = 0$ , and that  $U$  does not change sign in  $] \alpha, \xi[$ . Then, there exists some  $a_\xi \in [\alpha, \xi[$  such that  $U(a_\xi)U_1(a_\xi) < 0$ .*

*Let  $\eta \in ] \alpha, \beta[$ . Assume that  $U(\eta) = 0$ , and that  $U$  does not change sign in  $] \eta, \beta[$ . Then, there exists some  $b_\eta \in ] \eta, \beta[$  such that  $U(b_\eta)U_1(b_\eta) < 0$ .*

**Proof.** [40, p. 438] Since  $U$  cannot vanish identically in  $] \alpha, \xi[$  (see Lemma 2.2), there exists  $x_0 \in ] \alpha, \xi[$  such that  $U(x_0) \neq 0$ . Let  $\varepsilon_\xi = \text{sign}(U(x_0))$ . Since  $U$  does not change sign in  $] \alpha, \xi[$ ,  $\varepsilon_\xi U(x) \geq 0$  in  $] \alpha, \xi[$ . Then,

$$M_\xi := \sup\{\varepsilon_\xi U(x) \mid x \in [\alpha, \xi]\} > 0.$$

Let

$$a_\xi := \inf\{x \in [\alpha, \xi] \mid \varepsilon_\xi U(x) = M_\xi\}.$$

Then  $a_\xi \in ]\alpha, \xi[$ .

If  $a_\xi \in ]\alpha, \xi[$ , then  $\varepsilon_\xi U(a_\xi) > 0$ ,  $\frac{dU}{dx}(a_\xi) = 0$ , and  $\varepsilon_\xi \frac{d^2U}{dx^2}(a_\xi) \leq 0$ . By (2.6), this implies that  $\varepsilon_\xi U_1(a_\xi) < 0$ . Equivalently,  $U(a_\xi)U_1(a_\xi) < 0$ .

**Claim 2.10.** *If  $a_\xi = \alpha$ , then  $\varepsilon_\xi U(\alpha) > 0$ ,  $h = 0$ ,  $\frac{dU}{dx}(\alpha) = 0$ , and  $\varepsilon_\xi \frac{d^2U}{dx^2}(\alpha) \leq 0$ .*

*Proof of the claim.* Assume that  $a_\xi = \alpha$ , then  $\varepsilon_\xi U(\alpha) > 0$ , and hence  $h \neq \infty$ . If  $h$  were in  $]0, \infty[$ , we would have  $\varepsilon_\xi \frac{dU}{dx}(\alpha) = h\varepsilon_\xi U(\alpha) > 0$ , and hence  $a_\xi > \alpha$ . It follows that the assumption  $a_\xi = \alpha$  implies that  $h = 0$  and  $\frac{dU}{dx}(\alpha) = 0$ . If  $\varepsilon_\xi \frac{d^2U}{dx^2}(\alpha)$  were positive, we would have  $a_\xi > \alpha$ . Therefore, the assumption  $a_\xi = \alpha$  also implies that  $\varepsilon_\xi \frac{d^2U}{dx^2}(\alpha) \leq 0$ . The claim is proved.

If  $a_\xi = \alpha$ , then by Claim 2.10 and (2.6), we have  $\varepsilon_\xi U_1(\alpha) < 0$ . Equivalently,  $U(\alpha)U_1(\alpha) < 0$ . The first assertion of the lemma is proved. The proof of the second assertion is similar.  $\square$

**Proposition 2.11.** *Assume that (1.6) holds, and let  $k \in \mathbb{Z}$ . Then,*

$$(2.14) \quad N_v(Y_{k+1}, ]\alpha, \beta[) \geq N_v(Y_k, ]\alpha, \beta[),$$

*i.e., in the interval  $] \alpha, \beta [$ , the function  $Y_{k+1}$  changes sign at least as many times as the function  $Y_k$ .*

**Proof.** [40, p. 437-439] We keep the notation  $U = Y_k$  and  $U_1 = Y_{k+1}$ . By Lemma 2.2, the functions  $U$  and  $U_1$  have finitely many zeros in  $] \alpha, \beta [$ , with finite multiplicities. Since  $\alpha$  and  $\beta$  are fixed, we skip the mention to the interval  $] \alpha, \beta [$  in the proof, and we examine several cases.

*Case 1.* If  $N_v(U) = 0$ , there is nothing to prove.

*Case 2.* Assume that  $N_v(U) = 1$ . Then  $U$  admits a unique zero  $\xi \in ] \alpha, \beta [$  having odd multiplicity. Without loss of generality, we may assume that  $U \geq 0$  in  $] \alpha, \xi [$  and  $U \leq 0$  in  $] \xi, \beta [$ . By Lemma 2.9, there exist  $a \in ] \alpha, \xi [$  and  $b \in ] \xi, \beta [$  such that  $U_1(a) < 0$  and  $U_1(b) > 0$ .

It follows that the function  $U_1$  vanishes and changes sign at least once in  $] \alpha, \beta [$ , so that  $N_v(U_1) \geq 1 = N_v(U)$ , which proves the lemma in Case 2.

*Case 3.* If  $N_v(U) = 2$ , the function  $U$  has exactly two zeros, having odd multiplicities,  $\xi$  and  $\eta$  in  $] \alpha, \beta [$ ,  $\alpha < \xi < \eta < \beta$ , and we may assume that  $U|_{] \alpha, \xi [} \geq 0$ ,  $U|_{] \xi, \eta [} \leq 0$ , and  $U|_{] \eta, \beta [} \geq 0$ . The arguments given in Case 2 imply that there exist  $a \in ] \alpha, \xi [$  such that  $U_1(a) < 0$  and  $b \in ] \eta, \beta [$  such that  $U_1(b) < 0$ . In  $] \xi, \eta [$  the function  $U$  does not vanish identically and therefore achieves a global minimum at a point  $c$  such that  $U(c) < 0$ ,  $\frac{dU}{dx}(c) = 0$ , and  $\frac{d^2U}{dx^2}(c) \geq 0$ . Equation (2.6) then implies that  $U_1(c) > 0$ .

We can conclude that the function  $U_1$  vanishes and changes sign at least twice in  $] \alpha, \beta [$ , so that  $N_v(U_1) \geq 2 = N_v(U)$ .

*Case 4.* Assume that  $N_v(U) = p \geq 3$ . Then,  $U$  has exactly  $p$  zeros, with odd multiplicities, in  $] \alpha, \beta [$ ,  $\alpha < \xi_1 < \xi_2 < \dots < \xi_p < \beta$ , and one can assume that

$$U|_{] \alpha, \xi_1 [} \geq 0, \quad (-1)^p U|_{] \xi_p, \beta [} \geq 0, \quad \text{and} \\ (-1)^i U|_{] \xi_i, \xi_{i+1} [} \geq 0 \text{ for } 1 \leq i \leq p-1.$$

One can repeat the arguments given in the Cases 2 and 3, and conclude that there exist  $a_0, \dots, a_p$  with  $a_0 \in [\alpha, \xi_1[$ ,  $a_i \in ] \xi_i, \xi_{i+1} [$  for  $1 \leq i \leq p-1$ , and  $a_p \in ] \xi_p, \beta [$  such that  $(-1)^i U_1(a_i) < 0$ .

We can then conclude that the function  $U_1$  vanishes and changes sign at least  $p$  times in  $] \alpha, \beta [$ , i.e. that  $N_v(U_1) \geq p = N_v(U)$ .

This concludes the proof of Proposition 2.11.  $\square$

**Proposition 2.12.** *Assume that (1.6) holds. For any  $k \in \mathbb{Z}$ ,*

$$(2.15) \quad N_m(Y_{k+1}, ] \alpha, \beta [) \geq N_m(Y_k, ] \alpha, \beta [),$$

*i.e., in the interval  $] \alpha, \beta [$ , counting multiplicities of zeros, the function  $Y_{k+1}$  vanishes at least as many times as the function  $Y_k$ .*

**Proof.** [40, p. 439-442] Let  $U = Y_k$  and  $U_1 = Y_{k+1}$ . If  $U$  does not vanish in  $] \alpha, \beta [$ , there is nothing to prove. We now assume that  $U$  has at least one zero in  $] \alpha, \beta [$ . By Lemma 2.2,  $U$  and  $U_1$  have finitely many zeros in  $] \alpha, \beta [$ . Let

$$\alpha < \xi_1 < \dots < \xi_k < \beta$$

be the distinct zeros of  $U$ , with multiplicities  $p_i = m(U, \xi_i)$  for  $1 \leq i \leq k$ . Let  $\sigma_0$  be the sign of  $U$  in  $] \alpha, \xi_1 [$ ,  $\sigma_i$  the sign of  $U$  in  $] \xi_i, \xi_{i+1} [$  for  $1 \leq i \leq k-1$ , and  $\sigma_k$  the sign of  $U$  in  $] \xi_k, \beta [$ . Note that

$$\sigma_i = \text{sign} \left( \frac{d^{p_i} U}{dx^{p_i}}(\xi_i) \right) \text{ for } 1 \leq i \leq k.$$

By Lemma 2.9, there exist  $a_0 \in [\alpha, \xi_1[$  and  $a_k \in ] \xi_k, \beta [$  such that  $U(a_0)U_1(a_0) < 0$  and  $U(a_k)U_1(a_k) < 0$ . By Lemma 2.7, there exists  $a_i \in ] \xi_i, \xi_{i+1} [$ ,  $1 \leq i \leq k-1$ , such that  $U(a_i)U_1(a_i) < 0$ .

Summarizing, we have obtained:

$$(2.16) \quad \text{For } 0 \leq i \leq k, \quad U(a_i)U_1(a_i) < 0.$$

We have the relation

$$(2.17) \quad N_m(U, ] \alpha, \beta [) = \sum_{i=1}^k N_m(U, ] a_{i-1}, a_i [) = \sum_{i=1}^k p_i.$$

Indeed, for  $1 \leq i \leq k$ , the interval  $] a_{i-1}, a_i [$  contains precisely one zero  $\xi_i$  of  $U$ , with multiplicity  $p_i$ .

For  $U_1$ , we have the inequality

$$(2.18) \quad N_m(U_1, ]\alpha, \beta]) \geq \sum_{i=1}^k N_m(U_1, ]a_{i-1}, a_i]),$$

because  $U_1$  might have zeros in the interval  $]a_0, \alpha[$  if  $a_0 > \alpha$  (resp. in the interval  $]a_k, \beta[$  if  $a_k < \beta$ ).

**Claim 2.13.** For  $1 \leq i \leq k$ ,

$$N_m(U_1, ]a_{i-1}, a_i]) \geq N_m(U, ]a_{i-1}, a_i]) = p_i.$$

To prove the claim, we consider several cases.

• If  $p_i = 1$ , then  $U(a_{i-1})U(a_i) < 0$  and, by (2.16),  $U_1(a_{i-1})U_1(a_i) < 0$ , so that  $N_m(U_1, ]a_{i-1}, a_i]) \geq 1$ .

• If  $p_i \geq 2$ , we apply Lemma (2.4) at  $\xi_i$ : there exist real numbers  $B, B_1$  and smooth functions  $R$  and  $R_1$ , such that, in a neighborhood of  $\xi_i$ ,

$$(2.19) \quad \begin{cases} U(x) &= B(x - \xi_i)^{p_i} + (x - \xi_i)^{p_i+1}R(x), \\ U_1(x) &= B_1(x - \xi_i)^{p_i-2} + (x - \xi_i)^{p_i-1}R_1(x), \end{cases}$$

where  $\text{sign}(B) = \text{sign}(B_1) = \sigma_i$ .

We now use (2.16) and the fact that  $\text{sign}(U(a_i)) = \sigma_i$ .

◊ If  $p_i \geq 2$  is odd, then  $\sigma_{i-1}\sigma_i = -1$ . It follows that

$$\sigma_i U_1(a_i) < 0 \text{ and } \sigma_i U_1(a_{i-1}) > 0.$$

By (2.19), for  $\varepsilon$  small enough, we also have

$$\sigma_i U_1(\xi_i + \varepsilon) > 0 \text{ and } \sigma_i U_1(\xi_i - \varepsilon) < 0.$$

This means that  $U_1$  vanishes at order  $p_i - 2$  at  $\xi_i$ , and at least once in the intervals  $]a_{i-1}, \xi_i - \varepsilon[$  and  $] \xi_i + \varepsilon, a_i[$ , so that

$$N_m(U_1, ]a_{i-1}, a_i]) \geq p_i - 2 + 2 = p_i = N_m(U, ]a_{i-1}, a_i]).$$

◊ If  $p_i \geq 2$  is even, then  $\sigma_{i-1}\sigma_i = 1$ . It follows that

$$\sigma_i U_1(a_i) < 0 \text{ and } \sigma_i U_1(a_{i-1}) < 0.$$

By (2.19), for  $\varepsilon$  small enough, we also have

$$\sigma_i U_1(\xi_i + \varepsilon) > 0 \text{ and } \sigma_i U_1(\xi_i - \varepsilon) > 0.$$

This means that  $U_1$  vanishes at order  $p_i - 2$  at  $\xi_i$ , and at least once in the intervals  $]a_{i-1}, \xi_i - \varepsilon[$  and  $] \xi_i + \varepsilon, a_i[$ , so that

$$N_m(U_1, ]a_{i-1}, a_i]) \geq p_i - 2 + 2 = p_i = N_m(U, ]a_{i-1}, a_i]).$$

The claim is proved, and the proposition as well.  $\square$

**Proposition 2.14.** Assume that (1.6) holds. For any  $k \in \mathbb{Z}$ ,

$$(2.20) \quad \overline{N}_m(Y_{k+1}, [\alpha, \beta]) \geq \overline{N}_m(Y_k, [\alpha, \beta]),$$

i.e., in the interval  $[\alpha, \beta]$ , counting multiplicities of interior zeros, and reduced multiplicities of  $\alpha$  and  $\beta$ , the function  $Y_{k+1}$  vanishes at least as many times as the function  $Y_k$ .

**Proof.** [40, p. 440-442] Recall that the reduced multiplicity of  $\alpha$  (resp.  $\beta$ ) is zero if the Dirichlet condition holds at  $\alpha$  (resp. at  $\beta$ ) or if  $U(\alpha) \neq 0$  (resp.  $U(\beta) \neq 0$ ). Furthermore, according to Lemma 2.5, if  $h \in [0, \infty[$  and  $U(\alpha) = 0$  (resp. if  $H \in [0, \infty[$  and  $U(\beta) = 0$ ), then  $m(U, \alpha) = 2p$  (resp.  $m(U, \beta) = 2q$ ).

*Case 1.* Assume that  $N_m(U, ]\alpha, \beta[) = 0$ . Without loss of generality, we may assume that  $U > 0$  in  $] \alpha, \beta [$ .

- If  $U(\alpha) \neq 0$  and  $U(\beta) \neq 0$ , there is nothing to prove.
- Assume that  $U(\alpha) = U(\beta) = 0$ . Then, there exists  $a \in ]\alpha, \beta[$  such that

$$U(a) = \sup\{U(x) \mid x \in [\alpha, \beta]\},$$

with

$$U(a) > 0, \quad \frac{dU}{dx}(a) = 0, \quad \text{and} \quad \frac{d^2U}{dx^2}(a) \leq 0.$$

It follows from (2.6) that  $U_1(a) < 0$ , and that

$$(2.21) \quad \overline{N}_m(U_1, [\alpha, \beta]) = \overline{N}(U_1, [\alpha, a]) + \overline{N}(U_1, [a, \beta]).$$

It now suffices to look separately at the intervals  $[\alpha, a]$  and  $[a, \beta]$ .

◊ Interval  $[\alpha, a]$ . If the Dirichlet condition holds at  $\alpha$ , there is nothing to prove. If  $h \in [0, \infty[$ ,  $m(U, \alpha) = 2p \geq 2$  and, by Lemma 2.4,

$$(2.22) \quad \begin{aligned} U(x) &= B(x - \alpha)^{2p} + (x - \alpha)^{2p+1}R(x), \\ U_1(x) &= B_1(x - \alpha)^{2p-2} + (x - \alpha)^{2p-1}R_1(x), \end{aligned}$$

with  $B > 0$  and  $B_1 > 0$ .

It follows that  $U_1(\alpha + \varepsilon) > 0$  for any positive  $\varepsilon$  small enough so that  $N_m(U_1, ]\alpha, a[) \geq 1$ . It follows that

$$(2.23) \quad \begin{aligned} \overline{N}_m(U_1, [\alpha, a]) &= \overline{m}(U_1, \alpha) + N_m(U_1, ]\alpha, a[) \geq p - 1 + 1, \\ \text{i.e.} \end{aligned}$$

$$\overline{N}_m(U_1, [\alpha, a]) \geq \overline{N}(U, [\alpha, a]).$$

◊ Interval  $[a, \beta]$ . The proof is similar.

- Assume that  $U(\alpha) = 0$  and  $U(\beta) \neq 0$ . The proof is similar to the previous one with  $a \in ]a, \beta]$ .
- Assume that  $U(\alpha) \neq 0$  and  $U(\beta) = 0$ . The proof is similar to the previous one with  $a \in [\alpha, a[$ .

*Case 2.* Assume that  $N_m(U, ]\alpha, \beta[) \geq 1$ .

- If  $U(\alpha) \neq 0$  (resp.  $U(\beta) \neq 0$ ), there is nothing to prove for the boundary  $\alpha$  (resp.  $\beta$ ).
- If  $U(\alpha) = 0$  (resp.  $U(\beta) = 0$ ), the number  $a_0$  (resp.  $a_k$ ) which appears in the proof of Proposition 2.12 belongs to the open interval  $] \alpha, \xi_1 [$  (resp. to the open interval  $] \xi_k, \beta [$ ), where  $\xi_1$  (resp.  $\xi_k$ ) is the smallest (resp. largest) zero of  $U$  in  $] \alpha, \beta [$ . We can then apply the

proof of Step. 1 to the interval  $[\alpha, a_0]$  (resp. to the interval  $[a_k, \beta]$ ) to prove that  $\overline{N}_m(U_1, [\alpha, a_0]) \geq \overline{N}_m(U, [\alpha, a_0])$  (resp. to prove that  $\overline{N}_m(U_1, [a_k, \beta]) \geq \overline{N}_m(U, [a_k, \beta])$ ). This proves Proposition 2.14.  $\square$

We can now state Sturm's refined version of Theorem 1.4.

**Theorem 2.15.** *Assume that (1.6) holds, and let  $Y$  be the non trivial linear combination*

$$(2.24) \quad Y = \sum_{p=m}^n A_p V_p,$$

where  $1 \leq m \leq n$ , and where  $\{A_p, m \leq p \leq n\}$  are real constants such that  $A_m^2 + \dots + A_n^2 \neq 0$ . Then, with the notation of Subsection 2.2,

$$(2.25) \quad N_v(Y, ]\alpha, \beta[) \leq N_m(Y, ]\alpha, \beta[) \leq \overline{N}_m(Y, [\alpha, \beta]),$$

$$(2.26) \quad (m-1) \leq N_v(Y, ]\alpha, \beta[) \text{ and } \overline{N}_m(Y, [\alpha, \beta]) \leq (n-1).$$

**Proof.** [40, p. 442] Let  $N(V)$  be any of the above functions. We may of course assume that  $A_m \neq 0$  and  $A_n \neq 0$ . In the preceding lemmas, we have proved that  $N(Y_{k+1}) \geq N(Y_k)$  for any  $k \in \mathbb{Z}$ . This inequality can also be rewritten as

$$(2.27) \quad N(Y_{(-k)}) \leq N(Y) \leq N(Y_k) \text{ for any } k \geq 1.$$

Letting  $k$  tend to infinity, we conclude that

$$(2.28) \quad N(V_m) \leq N(Y) \leq N(V_n),$$

and we can apply Theorem 1.3.  $\square$

**Remark.** For a complete proof of the limiting argument when  $k$  tends to infinity, we refer to Appendix A.

### 3. LIOUVILLE'S APPROACH TO THEOREM 1.4

**3.1. Main statement.** We keep the notation of Section 2. Starting from a linear combination  $Y$  as in (2.1), Liouville also considers the family  $Y_k$  given by (2.2), and shows that the number of zeros of  $Y_{k+1}$  is not smaller than the number of zeros of  $Y_k$ . His proof is based on a generalization of Rolle's theorem.

*Remark 3.1.* In his proof, Liouville [24] only considers the zeros in the open interval  $] \alpha, \beta [$ .

As in Section 2, for  $1 \leq m \leq n$ , we fix  $Y = \sum_{j=m}^n A_j V_j$ , a linear combination of eigenfunctions of the eigenvalue problem (1.1)–(1.3), and we assume that  $A_m A_n \neq 0$ , see Remark 2.1.

**Theorem 3.2.** *Counting zeros with multiplicities in the interval  $] \alpha, \beta [$ , the function  $Y$  (1) has at most  $(n-1)$  zeros and, (2) has at least  $(m-1)$  zeros.*

**Proof.** Liouville uses the following version of Rolle's theorem (Michel Rolle (1652-1719) was a French mathematician). This version of Rolle's theorem seems to go back to Cauchy and Lagrange.

**Lemma 3.3.** *Let  $f$  be a function in  $] \alpha_0, \beta_0[$ . Assume that*

$$f(x') = f(x'') = 0 \text{ for some } x', x'', \alpha_0 < x' < x'' < \beta_0.$$

- (1) *If the function  $f$  is differentiable, and has  $\nu - 1$  distinct zeros in the interval  $]x', x''[$ , then the derivative  $f'$  has at least  $\nu$  distinct zeros in  $]x', x''[$ .*
- (2) *If the function  $f$  is smooth, and has  $\mu - 1$  zeros counted with multiplicities in the interval  $]x', x''[$ , then the derivative  $f'$  has at least  $\mu$  zeros counted with multiplicities in  $]x', x''[$ .*

**Proof of the lemma.** Call  $x_1 < x_2 < \dots < x_{\nu-1}$  the distinct zeros of  $f$  in  $]x', x''[$ . Since  $f(x') = f(x'') = 0$ , by Rolle's theorem [32], the function  $f'$  vanishes at least once in each open interval determined by the  $x_j$ ,  $1 \leq j \leq \nu - 1$ , as well as in the intervals  $]x', x_1[$  and  $]x_{\nu-1}, x''[$ . It follows that  $f'$  has at least  $\nu$  distinct zeros in  $]x', x''[$ , which proves the first assertion.

Call  $m_j$  the multiplicity of the zero  $x_j$ ,  $1 \leq j \leq \nu - 1$ . Then  $f'$  has at least  $\nu$  zeros, one in each of the open intervals determined by  $x', x''$  and the  $x_j$ 's, and has a zero at each  $x_j$  with multiplicity  $m_j - 1$ , provided that  $m_j > 1$ . It follows that the number of zeros of  $f'$  in  $]x', x''[$ , counting multiplicities, is at least

$$\sum_{j=1}^{\nu-1} (m_j - 1) + \nu = \sum_{j=1}^{\nu-1} m_j + 1,$$

which proves the second assertion. □

### 3.2. Proof of the assertion “ $Y$ has at most $(n - 1)$ zeros in $] \alpha, \beta[$ , counting multiplicities”.

Write (1.1) for  $V_1$  and for  $V_p$ , for some  $m \leq p \leq n$ . Multiply the first equation by  $-V_p$ , the second by  $V_1$ , and add the resulting equations. Then

$$(3.1) \quad V_1 \frac{d}{dx} \left( K \frac{dV_p}{dx} \right) - V_p \frac{d}{dx} \left( K \frac{dV_1}{dx} \right) + (\rho_p - \rho_1) G V_1 V_p = 0.$$

Use the identity

$$(3.2) \quad V_1 \frac{d}{dx} \left( K \frac{dV_p}{dx} \right) - V_p \frac{d}{dx} \left( K \frac{dV_1}{dx} \right) = \frac{d}{dx} \left( V_1 K \frac{dV_p}{dx} - V_p K \frac{dV_1}{dx} \right),$$

and integrate from  $\alpha$  to  $t$  to get the identity

$$(3.3) \quad (\rho_1 - \rho_p) \int_{\alpha}^t G V_1 V_p dx = K(t) \left( V_1(t) \frac{dV_p}{dx}(t) - V_p(t) \frac{dV_1}{dx}(t) \right).$$



Here we have used the boundary condition (1.2) which implies that

$$\left( V_1(\alpha) \frac{dV_p}{dx}(\alpha) - V_p(\alpha) \frac{dV_1}{dx}(\alpha) \right) = 0.$$

Multiplying the identity (3.3) by  $A_p$ , and summing for  $p$  from  $m$  to  $n$ , we obtain

$$(3.4) \quad \int_{\alpha}^t GV_1 \sum_{p=m}^n (\rho_1 - \rho_p) A_p V_p dx = K(t) \left( V_1 \frac{dY}{dx} - Y \frac{dV_1}{dx} \right) (t).$$

or

$$(3.5) \quad \int_{\alpha}^t GV_1 \sum_{p=m}^n (\rho_1 - \rho_p) A_p V_p dx = K(t) V_1^2(t) \frac{d}{dt} \left( \frac{Y}{V_1} \right) (t),$$

where we have used the fact that the function  $V_1$  does not vanish in the interval  $] \alpha, \beta [$ .

Let  $\Psi(x) = \frac{Y}{V_1}(x)$ . The zeros of  $Y$  in  $] \alpha, \beta [$  are the same as the zeros of  $\Psi$ , with the same multiplicities. Let  $\mu$  be the number of zeros of  $Y$ , counted with multiplicities. Using Lemma 3.3, Assertion (2), one can show that  $\frac{d\Psi}{dx}$  has at least  $\mu - 1$  zeros in  $] \alpha, \beta [$ , and hence so does the left-hand side of (3.5),

$$\int_{\alpha}^t GV_1 \sum_{p=m}^n (\rho_1 - \rho_p) A_p V_p dx.$$

On the other hand, this function vanishes at  $\alpha$  and  $\beta$  (because of the boundary condition (1.3) or orthogonality). By Lemma 3.3, its derivative,

$$(3.6) \quad V_1 \sum_{p=m}^n (\rho_1 - \rho_p) A_p V_p$$

has at least  $\mu$  zeros counted with multiplicities in  $] \alpha, \beta [$ . We have proved the following

**Lemma 3.4.** *If the function  $Y = \sum_{p=m}^n A_p V_p$  has at least  $\mu$  zeros counted with multiplicities in the interval  $] \alpha, \beta [$ , then the function  $Y_1 = \sum_{p=m}^n (\rho_1 - \rho_p) A_p V_p$  has at least  $\mu$  zeros, counted with multiplicities, in  $] \alpha, \beta [$ .*

Applying this lemma iteratively, we deduce that if  $Y$  has at least  $\mu$  zeros counted with multiplicities in  $] \alpha, \beta [$ , then, for any  $k \geq 1$ , the function

$$(3.7) \quad Y_k = \sum_{p=m}^n (\rho_1 - \rho_p)^k A_p V_p$$

has at least  $\mu$  zeros, counted with multiplicities, in  $] \alpha, \beta [$ .

We may of course assume that the coefficient  $A_n$  is non-zero. The above assertion can be rewritten as the statement:

For all  $k \geq 0$ , the equation

$$(3.8) \quad A_m \left( \frac{\rho_m - \rho_1}{\rho_n - \rho_1} \right)^k V_m + \cdots + A_{n-1} \left( \frac{\rho_{n-1} - \rho_1}{\rho_n - \rho_1} \right)^k V_{n-1} + A_n V_n = 0$$

has at least  $\mu$  solutions in  $] \alpha, \beta[$ , counting multiplicities.

Letting  $k$  tend to infinity, and using the fact that  $V_n$  has exactly  $(n-1)$  zeros in  $] \alpha, \beta[$ , this implies that  $\mu \leq (n-1)$ . This proves the first assertion.  $\square$

### 3.3. Proof of the assertion “ $Y$ has at least $(m-1)$ zeros in $] \alpha, \beta[$ , counting multiplicities”.

We have seen that the number of zeros of  $Y_k$  is less than or equal to the number of zeros of the function  $Y_{k+1}$ . This assertion actually holds for any  $k \in \mathbb{Z}$ , and can also be rewritten as,

$$(3.9) \quad N_m(Y_{-k}) \leq N_m(Y),$$

for any  $k \geq 0$ , where

$$(3.10) \quad Y_{-k} = A_m(\rho_m - \rho_1)^{-k} V_m + \cdots + A_n(\rho_n - \rho_1)^{-k} V_n,$$

and we can again let  $k$  tend to infinity. The second assertion is proved and Theorem 3.2 as well.  $\square$

### 3.4. Liouville’s 2nd approach to the 2nd part of Theorem 3.2.

If the function  $Y$  has  $\mu_1$  distinct zeros, and  $\mu \leq \mu_1$  sign changes, we call  $a_i$ ,  $\alpha < a_1 < \cdots < a_\mu < \beta$ , the points at which  $Y$  changes sign.

**Claim 3.5.** *The function  $Y$  changes sign at least  $(m-1)$  times in the interval  $] \alpha, \beta[$ .*

**Proof of the claim.** Assume, by contradiction, that  $\mu \leq (m-2)$ . Consider the function

$$(3.11) \quad x \mapsto W(x) := \Delta(a_1, \dots, a_\mu; x),$$

where the function  $\Delta$  is defined as the determinant

$$(3.12) \quad \begin{vmatrix} V_1(a_1) & V_1(a_2) & \cdots & V_1(a_\mu) & V_1(x) \\ V_2(a_1) & V_2(a_2) & \cdots & V_2(a_\mu) & V_2(x) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ V_{\mu+1}(a_1) & V_{\mu+1}(a_2) & \cdots & V_{\mu+1}(a_\mu) & V_{\mu+1}(x) \end{vmatrix}.$$

The function  $W$  vanishes at the points  $a_i$ ,  $1 \leq i \leq \mu$ . According to the first part in Theorem 3.2,  $W$  being a linear combination of the first  $\mu+1$  eigenfunctions, vanishes at most  $\mu$  times in  $] \alpha, \beta[$ , counting multiplicities. This implies that each zero  $a_i$  of  $W$  has order one, and that  $W$  does not have any other zero in  $] \alpha, \beta[$ . It follows that the function  $YW$  vanishes only at the points  $\{a_i\}$ ,  $1 \leq i \leq \mu$ , and that it

does not change sign. We can assume that  $YW \geq 0$ . On the other hand, we have

$$(3.13) \quad \int_{\alpha}^{\beta} GYW \, dx = 0,$$

because  $Y$  involves the functions  $V_p$  with  $p \geq m$  and  $W$  the functions  $V_q$  with  $q \leq \mu + 1 \leq m - 1$ . This gives a contradiction.  $\square$

*Remark 3.6.* Liouville does actually not use the determinant (3.12), but a similar approach, see [23, p. 259], Lemme 1<sup>er</sup>. The determinant  $\Delta$  appears in [34, Section 142]. The paper [6] is based on a careful analysis of this determinant.

*Remark 3.7.* The arguments in Subsection 3.2, using Assertion (1) of Lemma 3.3, instead of Assertion (2), yield an upper bound on the number of zeros of  $Y$ , multiplicities not accounted for. This estimate holds under weaker regularity assumptions, namely only assuming that the functions  $G, L$  are continuous, and that the function  $K$  is  $C^1$ , see Appendix C, and compare with [15], Chap. III.5.

#### 4. MATHEMATICAL CONTEXT OF STURM'S PAPERS. STURM'S MOTIVATIONS AND IDEAS

**4.1. On Sturm's style.** Sturm's papers [39, 40] are written in French, and quite long, about 80 pages each. One difficulty in reading them is the lack of layout structure. The papers are written linearly, and divided into sequences of sections, without any title. Most results are stated without tags, "Theorem" and the like, and only appear in the body of the text. For example, [39] only contains one theorem stated as such, see § XII, p. 125. In order to have an overview of the results contained in [39], the reader should look at the announcement [37]. Theorem 1.4 is stated in [38].

For a more thorough analysis of Sturm's papers on differential equations, we refer to [26, 14]. We refer to [7, 33] for the relationships between Theorem 1.3 and Sturm's theorem on the number of real roots of real polynomials.

**4.2. Sturm's motivations.** Sturm's motivations come from mathematical physics, and more precisely, from the problem of heat diffusion in a non-homogeneous bar. He considers the heat equation,

$$(4.1) \quad G \frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left( K \frac{\partial u}{\partial x} \right) - Lu, \text{ for } (x, t) \in ]\alpha, \beta[ \times \mathbb{R}_+,$$

with boundary conditions

$$(4.2) \quad \begin{cases} K(\alpha) \frac{\partial u}{\partial x}(\alpha, t) - h u(\alpha, t) = 0, \\ K(\beta) \frac{\partial u}{\partial x}(\beta, t) + H u(\beta, t) = 0, \end{cases}$$

for all  $t > 0$ , and with the initial condition

$$(4.3) \quad u(x, 0) = f(x), \text{ for } x \in ]\alpha, \beta[,$$

where  $f$  is a given function.

The functions  $K, G, L$  and the constants  $h, H$  describe the physical properties of the bar, see [40, Introduction, p. 376]. Sturm refers to the book of Siméon Denis Poisson [31], rather than to Fourier's book [13], because Poisson's equations are more general, see [33, Chap. III].

The boundary conditions (1.2)-(1.3) and (4.2) first appeared in the work of Fourier [13] but are called "Robin's condition" in the recent literature. Victor Gustave Robin (1855-1897) was a French mathematician.

As was popularized by Fourier and Poisson, in order to solve (4.1), Sturm uses the method of separation of variables, and is therefore led to the eigenvalue problem (1.1)-(1.3).

**4.3. Sturm's assumptions.** In [39, 40], Sturm implicitly assumes that the functions  $K, G, L$  are  $C^\infty$  and, explicitly, that  $K$  is positive, see [39, p. 108]. For the eigenvalue problem, he also assumes that  $G, L$  are positive, see [40, p. 381]. In [40, p. 394], he mentions that  $L$  could take negative values, and implicitly assumes, in this case, that  $\frac{L}{G}$  is bounded from below.

In [24], Liouville does not mention any regularity assumption on the functions  $G, K, L$ . He however indicates a regularity assumption (piecewise  $C^2$  functions) in a previous paper, [23, Footnote (\*), p. 256].

**4.4. Sturm's originality.** Before explaining Sturm's proofs, we would like to insist on the *originality* of his approach. Indeed, unlike his predecessors, Sturm does not look for explicit solutions of the differential equation (4.4) (i.e., solutions in closed form, or given as sums of series or as integrals), but he rather looks for *qualitative properties* of the solutions, properties which can be deduced directly from the differential equation itself. The following excerpts are translated from [39, Introduction]<sup>1</sup>.

*One only knows how to integrate these equations in a very small number of particular cases, and one can otherwise not even obtain a first integral; even when one knows the expression of the function which satisfies such an equation, in finite form, as a series, as integrals either definite or indefinite, it is most generally difficult to recognize in this expression the behaviour and the characteristic properties of this function. . . .*

*Although it is important to be able to determine the value of the unknown function for an isolated value of the variable it depends upon, it*

---

<sup>1</sup>See Appendix B for the original citations in French.

*is not less necessary to discuss the behaviour of this function, or otherwise stated, the form and the twists and turns of the curve whose ordinate would be the function, and the abscissa the independent variable. It turns out that one can achieve this goal by the sole consideration of the differential equation themselves, without having to integrate them. This is the purpose of the present memoir. . . .*

**4.5. Sturm and the existence and uniqueness theorem for ordinary differential equation.** In [39, p. 108], Sturm considers the differential equation

$$\frac{d}{dx} \left( K \frac{dV}{dx} \right) + GV = 0, \quad (I)$$

and takes the existence and uniqueness theorem for granted. More precisely, he claims [39, p. 108], without any reference whatsoever, *The complete integral of equation (I) must contain two arbitrary constants, for which one can take the values of  $V$  and of  $\frac{dV}{dx}$  corresponding to some particular value of  $x$ . Once these values are fixed, the function  $V$  is fully determined by equation (I), it has a uniquely determined value for each value of  $x$ .*

On the other hand, he gives two arguments for the fact that a solution of (I) and its derivative cannot vanish simultaneously at a point without vanishing identically, see [39, § II]. When the coefficients  $K, G$  of the differential equation depend upon a parameter  $m$ , e.g. continuously, Sturm also takes for granted the fact that the solution  $V(x, m)$ , and its zeros, depend continuously on  $m$ .

In [40, § II], Sturm mentions the existence proof given by Liouville in [23], see also [22]. According to [16], Augustin-Louis Cauchy may have presented the existence and uniqueness theorem for ordinary differential equations in his course at École polytechnique as early as in the year 1817-1818. Following a recommendation of the administration of the school, Cauchy delivered the notes of his lectures in 1824, see [9] and, in particular, the introduction by Christian Gilain who discovered these notes in 1974. These notes apparently had a limited distribution. Liouville entered the École polytechnique in 1825, and there attended the mathematics course given by Ampère<sup>2</sup> (as a matter of fact Ampère and Cauchy gave the course every other year, alternatively). Liouville's proof of the existence theorem for differential equations in [22], à la Picard but before Picard, though limited to the particular case of 2nd order linear equations, might be the first well circulated proof of an existence theorem for differential equations, see [25, § 34]. Cauchy's theorem was later popularized in the second volume of Moigno's book, published in 1844, see [27], "Vingt-sixième Leçon" § 159, pp. 385–396.

---

<sup>2</sup>We are grateful to J. Lützen for providing this information.

4.6. **Sturm's proof of Theorem 1.3.** Theorem 1.3 is proved in [40]. For the first assertion, see § III (p. 384) to VII; for the second assertion, see § VIII (p. 396) to X.

The proof is based on the paper [39] in which Sturm studies the zeros of the solution of the initial value problem,

$$(4.4) \quad \frac{d}{dx} \left( K(x, m) \frac{dV}{dx}(x, m) \right) + G(x, m)V(x, m) = 0,$$

$$(4.5) \quad \left( K \frac{dV}{dx} - hV \right) (\alpha, m) = 0.$$

Here  $K, G$  are assumed to be functions of  $x$  depending on a real parameter  $m$ , with  $K$  positive (the constants  $h$  and  $H$  may also depend on the parameter  $m$ ). The solution  $V(x, m)$  is well defined up to a scaling factor. The main part of [39] is devoted to studying how the zeros of the function  $V(x, m)$  (and other related functions) depend on the parameter  $m$ , see [39, § XII, p. 125]. While developing this program, Sturm proves the *oscillation, separation and comparison theorems* which nowadays bear his name, [39, § XV, XVI and XXXVII].

The eigenvalue problem (1.1)–(1.3) itself is studied in [40]. For this purpose, Sturm considers the functions

$$K(x, r) \equiv K(x) \text{ and } G(x, r) = rG(x) - L(x),$$

the solution  $V(x, r)$  of the corresponding initial value problem (4.4)–(4.5), and applies the results and methods of [39].

The spectral data of the eigenvalue problem (1.1)–(1.3) are determined by the following transcendental equation in the spectral parameter  $r$ ,

$$(4.6) \quad K(\beta) \frac{dV}{dx}(\beta, r) + HV(\beta, r) = 0,$$

see, [40], §III, page 383, line 8 from bottom.

4.7. **Sturm's two proofs of Theorem 1.4.** Theorem 1.4 appears in [40, § XXV, p. 431], see also the announcement [38].

Sturm's general motivation, see the introductions to [39] and [40], was the investigation of heat diffusion in a (non-homogeneous) bar, whose physical properties are described by the functions  $K, G, L$ . He first obtained Theorem 1.4 as a corollary of a much deeper theorem which describes the behaviour, as time varies, of the  $x$ -zeros of a solution  $u(x, t)$  of the heat equation (4.1)–(4.3). When the initial temperature  $u(x, 0)$  is given by a linear combination of simple states,

$$(4.7) \quad u(x, 0) = Y(x) = \sum_{j=m}^n A_j V_j$$

the function  $u(x, t)$  is given by

$$(4.8) \quad u(x, t) = \sum_{j=m}^n e^{-t\rho_j} A_j V_j.$$

When  $t$  tends to infinity, the  $x$ -zeros of  $u(x, t)$  approach those of  $V_p$ , where  $p$  is the least integer  $j$ ,  $m \leq j \leq n$  such that  $A_j \neq 0$ .

J. Liouville, who was aware of Theorem 1.4, made use of it in [23], and provided a purely “ordinary differential equation” proof in [24], a few months before the actual publication of [40]. This induced Sturm to provide two proofs of Theorem 1.4 in [40], his initial proof using the heat equation, and another proof based on the sole ordinary differential equation. The proofs of Sturm actually give a more precise result. In [40, p. 379], Sturm writes,

*M. Liouville gave a direct proof of this theorem, which for me was a mere corollary of the preceding one, without taking care of the particular case in which the function vanishes at one of the extremities of the bar. I have also found, after him, another direct proof which I give in this memoir. M. Liouville made use of the same theorem in a very nice memoir which he published in the July issue of his journal, and which deals with the expansion of an arbitrary function into a series made of the functions  $V$  which we have considered.*

The time independent analog to studying the behaviour of the  $x$ -zeros of (4.8) is to study the behaviour of the zeros of the family of functions  $\{Y_k\}_{k \in \mathbb{Z}}$ , where

$$(4.9) \quad Y_k(x) = \sum_{j=m}^n \rho_j^k A_j V_j,$$

as  $k$  tends to infinity.

#### APPENDIX A. THE LIMITING ARGUMENT IN (3.8)

Recall that we assume that  $A_n \neq 0$ . Define

$$(A.1) \quad \omega = \left( \frac{\rho_{n-1} - \rho_1}{\rho_n - \rho_1} \right)^k.$$

One can rewrite (3.8) as

$$V_n(x) + \omega \Pi(x) = 0,$$

where

$$(A.2) \quad \Pi(x) = \sum_{p=m}^{n-1} \left( \frac{\rho_p - \rho_1}{\rho_{n-1} - \rho_1} \right)^k \frac{A_p}{A_n} V_p.$$

It follows that  $\Pi$  is uniformly bounded by

$$(A.3) \quad |\Pi(x)| \leq M := n \max_p \left| \frac{A_p}{A_n} \right| \max_p \sup_{[\alpha, \beta]} |V_p|.$$

Similarly,

$$(A.4) \quad \left| \frac{d\Pi}{dx}(x) \right| \leq N := n \max_p \left| \frac{A_p}{A_n} \right| \max_p \sup_{[\alpha, \beta]} \left| \frac{dV_p}{dx} \right|.$$

Call  $\xi_1 < \xi_2 < \dots < \xi_{n-1}$  the zeros of the function  $V_n$  in the interval  $] \alpha, \beta [$ .

• Assume that  $V_n(\alpha) \neq 0$  and  $V_n(\beta) \neq 0$ .

Since  $\frac{dV_n}{dx}(\xi_i) \neq 0$ , there exist  $\delta_1, \varepsilon_1 > 0$  such that  $|\frac{dV_n}{dx}(x)| \geq \varepsilon_1$  for  $x \in [\xi_i - \delta_1, \xi_i + \delta_1]$ , and  $|V_n(x)| \geq \varepsilon_1$  in  $[\alpha, \beta] \setminus \cup ]\xi_i - \delta_1, \xi_i + \delta_1 [$ .

For  $k$  large enough, we have  $\omega M, \omega N \leq \varepsilon_1/2$ . It follows that in the interval  $[\xi_i - \delta_1, \xi_i + \delta_1]$ ,

$$\left| \frac{d}{dx}(V_n + \omega \Pi) \right| \geq \left| \frac{dV_n}{dx} \right| - \omega N \geq \varepsilon_1/2.$$

Furthermore,

$$V_n(\xi_i \pm \delta_1) + \omega \Pi(\xi_i \pm \delta_1) \geq |V_n(\xi_i \pm \delta_1)| - \omega M \geq \varepsilon_1/2.$$

Since  $V_n(\xi_i + \delta_1)V_n(\xi_i - \delta_1) < 0$ , we can conclude that the function  $V_n + \omega \Pi$  has exactly one zero in each interval  $]\xi_i - \delta_1, \xi_i + \delta_1 [$ .

In  $[\alpha, \beta] \setminus \cup ]\xi_i - \delta_1, \xi_i + \delta_1 [$ , we have

$$|V_n(x) + \omega \Pi(x)| \geq |V_n(x)| - \omega M \geq \varepsilon_1/2,$$

which implies that  $V_n(x) + \omega \Pi(x) \neq 0$ .

• Assume that  $V_n(\alpha) = 0$  and  $V_n(\beta) \neq 0$ . This corresponds to the case  $h = +\infty$  and  $H \neq +\infty$ . Hence the  $V_j$  verify Dirichlet at  $\alpha$  and  $\Pi$  verifies Dirichlet at  $\alpha$ . Observing that  $V_n'(\alpha) \neq 0$ , it is immediate to see that there exists  $\delta_1 > 0$ , such that, for  $k$  large enough,  $V_n(x) + \omega \Pi(x)$  has only  $\alpha$  as zero in  $[\alpha, \alpha + \delta_1]$ .

• The other cases are treated in the same way. □



APPENDIX B. CITATIONS FROM STURM'S PAPERS  
FRENCH ORIGINAL AND ENGLISH TRANSLATION

**Citation from** [39, Introduction].

On ne sait [ces équations] les intégrer que dans un très petit nombre de cas particuliers hors desquels on ne peut pas même en obtenir une intégrale première ; et lors même qu'on possède l'expression de la fonction qui vérifie une telle équation, soit sous forme finie, soit en série, soit en intégrales définies ou indéfinies, il est le plus souvent difficile de reconnaître dans cette expression la marche et les propriétés caractéristiques de cette fonction. ...

S'il importe de pouvoir déterminer la valeur de la fonction inconnue pour une valeur isolée quelconque de la variable dont elle dépend, il n'est pas moins nécessaire de discuter la marche de cette fonction, ou en d'autres termes, d'examiner la forme et les sinuosités de la courbe dont cette fonction serait l'ordonnée variable, en prenant pour abscisse la variable indépendante. Or on peut arriver à ce but par la seule considération des équations différentielles elles-mêmes, sans qu'on ait besoin de leur intégration. Tel est l'objet du présent mémoire. ...

One only knows how to integrate these equations in a very small number of particular cases, and one can otherwise not even obtain a first integral; even when one knows the expression of the function which satisfies such an equation, in finite form, as a series, as integrals either definite or indefinite, it is most generally difficult to recognize in this expression the behaviour and the characteristic properties of this function. ...

Although it is important to be able to determine the value of the unknown function for an isolated value of the variable it depends upon, it is not less necessary to discuss the behaviour of this function, or otherwise stated, the form and the twists and turns of the curve whose ordinate would be the function, and the abscissa the independent variable. It turns out that one can achieve this goal by the sole consideration of the differential equation themselves, without having to integrate them. This is the purpose of the present memoir. ...

**Citation from** [39, p. 108].

L'intégrale complète de l'équation (I) doit contenir deux constantes arbitraires, pour lesquelles on peut prendre les valeurs de  $V$  et de  $\frac{dV}{dx}$  correspondantes à une valeur particulière de  $x$ . Lorsque ces valeurs sont fixées, la fonction  $V$  est entièrement définie par l'équation (I), elle a une valeur déterminée et unique pour chaque valeur de  $x$ .

The complete integral of equation (I) must contain two arbitrary constants, for which one can take the values of  $V$  and of  $\frac{dV}{dx}$  corresponding to some particular value of  $x$ . Once these values are fixed, the function  $V$  is fully determined by equation (I), it has a uniquely determined value for each value of  $x$ .

**Citation from** [40, p. 379].

M. Liouville a démontré directement ce théorème, qui n'était pour moi qu'un corollaire du précédent, sans s'occuper du cas particulier où la fonction serait nulle à l'une des extrémités de la barre. J'en ai aussi trouvé après lui une autre démonstration directe que je donne dans ce mémoire. M. Liouville a fait usage du même théorème dans un très beau Mémoire qu'il a publié dans le numéro de juillet de son journal et qui a pour objet le développement d'une fonction arbitraire en une série composée de fonctions  $V$  que nous avons considérées.

M. Liouville gave a direct proof of this theorem, which for me was a mere corollary of the preceding one, without taking care of the particular case in which the function vanishes at one of the extremities of the bar. I have also found, after him, another direct proof which I give in this memoir. M. Liouville made use of the same theorem in a very nice memoir which he published in the July issue of his journal, and which deals with the expansion of an arbitrary function into a series made of the functions  $V$  which we have considered.

## APPENDIX C. STURM'S RESULTS UNDER WEAKER ASSUMPTIONS

We proved Theorems 2.15 and 3.2 under the Assumptions (1.6). In this section, we consider the weaker assumptions

$$(C.1) \quad \begin{cases} [\alpha, \beta] \subset ]\alpha_0, \beta_0[, \\ K \in C^1(]\alpha_0, \beta_0[), \\ G, L \in C^0(]\alpha_0, \beta_0[), \\ K, G, L > 0 \text{ on } ]\alpha_0, \beta_0[. \end{cases}$$

Under these assumptions, the functions  $V_j$  are  $C^2$  on  $]\alpha_0, \beta_0[$ . This follows easily for example from Liouville's existence proof [23], and we have the following lemma, whose proof is analogous to the proof of Lemma 2.2

**Lemma C.1.** *Let  $k \in \mathbb{Z}$ .*

- (1) *The function  $Y_k$  satisfies the boundary conditions (1.2) and (1.3).*
- (2) *The functions  $Y_k$  and  $Y_{k+1}$  satisfy the relation*

$$(C.2) \quad GY_{k+1} = K \frac{d^2 Y_k}{dx^2} + \frac{dK}{dx} \frac{dY_k}{dx} - LY_k.$$

- (3) *Under the Assumptions (C.1), the function  $Y_k$  cannot vanish identically on an open interval  $]\alpha_1, \beta_1[ \subset ]\alpha_0, \beta_0[$ , unless  $Y \equiv 0$ .*

In Subsection 3.2, we have used Lemma 3.3 (2) which relies on the fact that the functions  $V_j$  are  $C^\infty$ . If the functions  $V_j$  are only  $C^2$ , we can apply Lemma 3.3 (1). It is easy to conclude that Liouville's proofs in Subsection 3.2 and 3.3 go through, under the weaker Assumptions (C.1), if we only count distinct zeros, see (2.11). More precisely, we can prove the following claim.

**Claim C.2.** *Under the Assertions (C.1), for any  $k \in \mathbb{Z}$ , if the function  $Y_k$  has at least  $\mu$  distinct zeros in the interval  $]\alpha, \beta[$ , then the function  $Y_{k+1}$  has at least  $\mu$  distinct zeros in the interval  $]\alpha, \beta[$ .*

We can then deduce from this claim, as in Section 3, that a linear combination  $Y = \sum_{j=m}^n A_j V_j$  has at most  $(n-1)$  distinct zeros (in particular it has finitely many zeros).

Once this result is secured, we can define zeros at which  $Y$  changes sign (without using the multiplicity), and apply Sturm's lower bound argument to conclude that the function  $Y$  must change sign at least  $(m-1)$  times.

## APPENDIX D. STURM'S ORIGINAL O.D.E PROOF

The first proof of Theorem 1.4 appears in [40, § XXV, p. 431], as a corollary of a more profound theorem (§ XXIV) which describes the behaviour, as  $t$  grows from 0 to infinity, of the zeros of  $x \mapsto u(x, t)$ , where  $u$  is a solution of the heat (4.1)-(4.3).

Sturm proves that the number  $N(t)$  of zeros of the function  $x \mapsto u(x, t)$  is piecewise constant, non-increasing in  $t$ , and that jumps occur precisely for values of  $t$  such that  $u(x, t)$  and  $\frac{\partial u}{\partial t}(x, t)$  have common zeros. We refer to [14] for an analysis of this aspect of Sturm's paper [40].

The second proof, purely o.d.e., is developed in [40, § XXVI, p. 436 ff]. In this section, we give the main steps of this proof (with page numbers and number of line from top  $\ell \downarrow$ , resp. from bottom  $\ell \uparrow$ ).

**p. 436**  $\ell \uparrow 13$ , Sturm mentions Liouville's proof [24].

*M. Liouville a démontré directement le théorème du numéro précédent (dans le cahier d'août de son journal) sans employer la considération de la variable auxiliaire  $t$  qui entre dans la fonction  $u$  (42) dont j'ai fait usage. Il n'a pas tenu compte toutefois de la racine  $x$  ou  $X^3$  lorsqu'elle existe. Je vais donner ici une autre démonstration directe du même théorème, indépendante de celui du n° XXIV.*

He introduces the linear combination

$$Y = C_i V_i + C_{i+1} V_{i+1} + \cdots + C_p V_p. \quad (43)$$

and, **p. 436**  $\ell \uparrow 1$ , its companion

$$Y_1 = -(C_i \rho_i V_i + C_{i+1} \rho_{i+1} V_{i+1} + \cdots + C_p \rho_p V_p) .$$

**p. 437**, Sturm establishes the differential relation

$$gY_1 = k \frac{d^2 Y}{dx^2} + \frac{dk}{dx} \frac{dY}{dx} - \ell Y. \quad (44)$$

He also notes  $\ell \downarrow 5$ , that the function  $Y$  satisfies the boundary conditions (1.2)-(1.3). Sturm' idea *Je vais prouver . . .*, is to prove that the function  $Y_1$  has at least as many zeros in  $]\alpha, \beta[$ , counted with multiplicities, as the function  $Y$  in the same circumstances.

**p. 437**  $\ell \uparrow 10$ , Sturm makes the *implicit assumption* that the zeros of  $Y$  are isolated.

**p. 439**  $\ell \downarrow 5$ , Sturm states that the number of sign changes of  $Y_1$  in  $]\alpha, \beta[$  is not smaller than the number of sign changes of  $Y$ . He then considers the zeros with multiplicities, and *implicitly assumes* that the function  $Y$  (assumed not to be identically zero) does not vanish at infinite order at some point.

---

<sup>3</sup>Respectively  $\alpha$  and  $\beta$  with our notation.

**p. 440**  $\ell \uparrow 13$ , Sturm states that the number of zeros of  $Y_1$  in  $] \alpha, \beta[$ , counted with multiplicities, is not smaller than the number of zeros of  $Y$ . He then examines ( $\ell \uparrow 6$ ) the possible zeros of  $Y$  at  $\alpha$  or  $\beta$ .

**p. 442**  $\ell \downarrow 7$ , Sturm states that the number of zeros of  $Y_1$  in  $[\alpha, \beta]$ , counted with multiplicities (with a special rule for counting multiplicities at  $\alpha, \beta$ ), is not smaller than the number of zeros of  $Y$ .

**p. 442**  $\ell \downarrow 13$ , Sturm iterates the procedure (with  $Y_k$ ), and uses a limiting argument to conclude that the number of zeros of  $Y$  in  $[\alpha, \beta]$ , counting multiplicities, is at most  $p - 1$ .

**p. 443**, Sturm proves the lower bound for the number of zeros and, ( $\ell \uparrow 6$ ), compares the present proof with the heat equation proof, the functions  $Y_k$  are equal to  $\frac{d^k u}{dt^k}(x, 0)$ . Finally, in a footnote, he mentions that  $Y$  cannot vanish identically unless all the coefficients  $C_j$  are zero. He does not mention the fact that  $Y$  can actually not vanish at infinite order at any point.

**p. 444**  $\ell \downarrow 3$ , Sturm explains what to do when no assumption is made on the sign of the function  $\ell$ . Taking  $Y$  as above, and defining

$$Y_1 = - (C_i(\rho_i + c)V_i + C_{i+1}(\rho_{i+1} + c)V_{i+1} + \cdots + C_p(\rho_p + c)V_p) ,$$

where  $c$  is a constant, he obtains

$$gY_1 = k \frac{d^2 Y}{dx^2} + \frac{dk}{dx} \frac{dY}{dx} - (gc + \ell)Y .$$

It suffices to assume that the constant  $c$  is such that  $gc + \ell > 0$  and to follow the previous proof with this new definition of  $Y_1$ .

#### APPENDIX E. CROSS REFERENCES TO STURM'S AND LIOUVILLE'S PAPERS

In this Appendix, we give the references to pages in Sturm's paper [40, § XXVI] for the results in our paper.

- Lemma 2.2: **p. 437**. Note that the third assertion does not appear in Sturm's paper. He indeed implicitly assumes that the zeros of  $Y$  are isolated.
- Lemma 2.4: **p. 439**.
- Lemma 2.5: **p. 440-441**.
- Lemma 2.7: **p. 437**.
- Lemma 2.9: **p. 438**.
- Proposition 2.11: **p. 437-439**.
- Proposition 2.12: **p. 439-442**.
- Proposition 2.14: **p. 440-442**.
- Theorem 2.15: **p. 442**.

Here are the pages in Liouville's paper [24].

- Theorem 3.2: **p. 272**.

- Lemma 3.3: Mentioned **p. 272**. No precise statement, no proof provided by Liouville.
- Proof of first assertion. Lemma 3.4: **p. 274**.
- Proof of second assertion: **p. 276** and reference to [23].  
Claim 3.5: We use the determinant  $\Delta$  to simplify Liouville's [23, Lemme 1<sup>er</sup>, p. 259].

Numbers inserted after a reference indicate the pages where it is cited.

#### REFERENCES

- [1] V. Arnold. Topology of real algebraic curves (works of I.G. Petrovsky and their development)[Russian]. Usp. Mat. Nauk. 28:5 (1973),260–262. Translated by O. Viro, in V.I. Arnold, Collected works, Vol. 2, pp. 251-254. Springer 2014. [3](#)
- [2] V. Arnold. Ordinary differential equations. Translated from the 3rd Russian edition by Roger Cooke. Springer-Verlag 1992. [3](#)
- [3] V. Arnold. Topological properties of eigenoscillations in mathematical physics. Proc. Steklov Inst. Math., 273 (2011), 25–34. [3](#)
- [4] P. Bérard and B. Helffer. On Courant’s nodal domain property for linear combinations of eigenfunctions, Part I. arXiv:1705.03731. [3](#)
- [5] P. Bérard and B. Helffer. On Courant’s nodal domain property for linear combinations of eigenfunctions, Part II. arXiv:1803.00449. [3](#), [4](#)
- [6] P. Bérard and B. Helffer. Sturm’s theorem on the zeros of sums of eigenfunctions: Gelfand’s strategy implemented. arXiv:1807.03990. [4](#), [18](#)
- [7] M. Bôcher. The published and unpublished work of Charles Sturm on algebraic and differential equations. Proc. Amer. Math. Soc., 18 (1911), 1–18. [18](#)
- [8] M. Bôcher. Leçons sur les méthodes de Sturm dans la théorie des équations différentielles linéaires et leurs développements modernes. Gauthier-Villars et Cie, Éditeurs. Paris 1917. [3](#)
- [9] A. L. Cauchy. Équations différentielles ordinaires. Cours inédit (Fragment). Critical edition by Christian Gilain. Paris-Québec : Études vivantes, and New York: Johnson Reprint, p. I-LVI et p. 1–146, 1981. [20](#)
- [10] R. Courant and D. Hilbert. Methoden der mathematischen Physik, Vol. I. Springer 1931. [3](#)
- [11] R. Courant and D. Hilbert. Methods of mathematical physics. Vol. 1. First english edition. Interscience, New York 1953. [3](#)
- [12] A. Eremenko and D. Novikov. Oscillation of Fourier integrals with a spectral gap. Journal de mathématiques pures et appliquées, 83 (2004), 313-365. [3](#)
- [13] J.B. Joseph Fourier. Théorie analytique de la chaleur. Chez Firmin Didot, père et fils, 1822 (639 pages). [2](#), [19](#)
- [14] V.A. Galaktionov and P.J. Harwin. Sturm’s theorems on zero sets in nonlinear parabolic equations. in *Sturm-Liouville theory: Past and present*. W.O. Amrein, A.M. Hinz, D.B. Pearson ed. Birkhäuser Verlag Basel 2005, 173–199. [4](#), [18](#), [27](#)
- [15] F. Gantmacher and M. Krein. Oscillation matrices and kernels and small vibrations of mechanical systems. AMS Chelsea Publishing, 2002. [3](#), [18](#)
- [16] C. Gilain. Cauchy et le cours d’analyse de l’École polytechnique. Revue de la SABIX, 5 (1989), 3–31. [20](#)
- [17] G. Gladwell and H. Zhu. The Courant-Herrmann conjecture. ZAMM - Z. Angew. Math. Mech., 83:4 (2003), 275–281. [3](#)
- [18] H. Herrmann. Beiträge zur Theorie der Eigenwerte und Eigenfunktionen. Göttinger Dissertation 1932. Published by Teubner. [3](#)
- [19] A. Hurwitz. Über die Fourierschen Konstanten integrierbaren Funktionen. Math. Annalen, 57 (1903), 425–446. [3](#)
- [20] N. Kuznetsov. On delusive nodal sets of free oscillations. Newsletter of the European Mathematical Society, 96 (2015), 34–40. [4](#)
- [21] P.-S. Laplace. Théorie analytique des probabilités. 3rd edition. Coursier Paris 1820. [7](#)
- [22] J. Liouville. Analyse appliquée. Mémoire sur la théorie analytique de la chaleur. Annales de Mathématiques Pures et Appliquées, 21 (1830), 131–181. [20](#)
- [23] J. Liouville. Mémoire sur le développement de fonctions ou parties de fonctions en séries dont les divers termes sont assujétis à satisfaire à une même

- équation différentielle du second ordre, contenant un paramètre variable. *Journal de Mathématiques Pures et Appliquées*, 1 (1836), 253–265. [2](#), [3](#), [4](#), [5](#), [18](#), [19](#), [20](#), [22](#), [26](#), [29](#)
- [24] J. Liouville. Démonstration d’un théorème dû à M. Sturm et relatif à une classe de fonctions transcendentes. *Journal de Mathématiques Pures et Appliquées*, 1 (1836), 269–277. [4](#), [14](#), [19](#), [22](#), [27](#), [28](#)
- [25] J. Lützen. Sturm and Liouville’s work on ordinary linear differential equations. The emergence of Sturm-Liouville theory. *Archive History Exact Sciences*, 29:4 (1984), 309–376. [20](#)
- [26] J. Lützen and A. Mingareli. Charles François Sturm and differential equations. in *Collected works of Charles François Sturm*. Jean-Claude Pont (ed.), in coll. with Flavis Padovani. Birkhäuser Verlag Basel, 2009, 25–47. [4](#), [18](#)
- [27] F.N.M. Moigno (Monsieur l’abbé). Leçons de calcul différentiel et de calcul intégral, rédigées principalement d’après les méthodes et les ouvrages de M. A.-L. Cauchy, et étendues aux travaux les plus récents des géomètres. Bachelier, Paris. Vol. 1 (1840) and Vol. 2 (1844). [20](#)
- [28] V. Ovsienko and S. Tabachnikov. Projective differential geometry old and new: from Schwarzian derivative to cohomology of diffeomorphism groups. Cambridge University Press 2005. [3](#)
- [29] Å. Pleijel. Remarks on Courant’s nodal theorem. *Comm. Pure. Appl. Math.*, 9 (1956), 543–550. [3](#)
- [30] F. Pockels. Über die partielle Differentialgleichung  $\Delta u + k^2 u = 0$  und deren Auftreten in der mathematischen Physik. Teubner, Leipzig 1891. [2](#)
- [31] S. D. Poisson. Théorie mathématique de la chaleur. Bachelier, Paris 1835. [19](#)
- [32] M. Rolle. Démonstration d’une méthode pour résoudre les égalités de tous les degrés. Chez Jean Cusson, Paris 1691. [15](#)
- [33] H. Sinaceur. Corps et modèles. Essai sur l’histoire de l’algèbre réelle. Seconde édition corrigée. Vrin, Paris 1999. [18](#), [19](#)
- [34] J.W. Strutt, Baron Rayleigh. The Theory of Sound. Vol. I. Macmillan and Co., London, 1877. [2](#), [18](#)
- [35] S. Steinerberger. Quantitative projections in the Sturm oscillation theorem. arXiv:1804.05779, 18 Apr 2018. [3](#)
- [36] C. Sturm Extrait d’un mémoire de M. Sturm, présenté à l’Académie des sciences, dans sa séance du 1er juin 1829. *Bulletin de Férussac*, XI (1829), 422–425. [3](#)
- [37] C. Sturm. Analyse générale d’un mémoire sur les propriétés générales des fonctions qui dépendent d’équations différentielles linéaires du second ordre, présenté à l’Académie des sciences de Paris, le 30 septembre 1833. L’institut. *Journal général des sociétés et travaux scientifiques de la France et de l’étranger*, 1 (1833), 247–248. [2](#), [4](#), [18](#)
- [38] C. Sturm. Monsieur Sturm nous prie d’insérer la note suivante. L’institut. *Journal général des sociétés et travaux scientifiques de la France et de l’étranger*, 1 (1833), 247–248. [2](#), [4](#), [18](#), [21](#)
- [39] C. Sturm. Mémoire sur les équations différentielles linéaires du second ordre. *Journal de Mathématiques Pures et Appliquées*, 1 (1836), 106–186. [2](#), [4](#), [6](#), [18](#), [19](#), [20](#), [21](#), [24](#), [25](#)
- [40] C. Sturm. Mémoire sur une classe d’équations à différences partielles. *Journal de Mathématiques Pures et Appliquées*, 1 (1836), 373–444. [2](#), [3](#), [4](#), [6](#), [7](#), [9](#), [10](#), [11](#), [13](#), [14](#), [18](#), [19](#), [20](#), [21](#), [22](#), [25](#), [27](#), [28](#)
- [41] O. Viro. Construction of multi-component real algebraic surfaces. *Soviet Math. Dokl.*, 20:5 (1979), 991–995. [3](#)



PB: INSTITUT FOURIER, UNIVERSITÉ GRENOBLE ALPES AND CNRS, B.P.74,  
F38402 SAINT MARTIN D'HÈRES CEDEX, FRANCE.

*E-mail address:* pierreherard@gmail.com

BH: LABORATOIRE JEAN LERAY, UNIVERSITÉ DE NANTES AND CNRS, F44322  
NANTES CEDEX, FRANCE.

*E-mail address:* Bernard.Helffer@univ-nantes.fr