COLLECTION OF MONOGRAPHS ON THE THEORY OF FUNCTIONS PUBLISHED UNDER THE DIRECTION OF ÉMILE BOREL

LECTURES

ON

STURM'S METHOD

IN THE THEORY OF LINEAR DIFFERENTIAL EQUATIONS AND THEIR MODERN DEVELOPMENTS

TAUGHT AT THE SORBONNE IN 1913-1914

By MAXIME BÔCHER

PROFESSOR AT HARVARD UNIVERSITY APPROVED PROFESSOR AT THE UNIVERSITY OF PARIS

Compiled and edited by Gaston JULIA

AT L'ÉCOLE NORMALE SUPERIEUR

Translated by **D. H. DELPHENICH**

PARIS

GAUTHEIR-VILLARS AND CO., EDITORS LIBRARY OF THE BUREAU OF LONGITUDES OF L'ÉCOLE POLYTECHNIQUE Quai des Grands-Augustins, 55

1917

PREFACE

For some years now, there has existed an annual exchange of professors between the University of Paris and Harvard University. I have had the honor of representing the ones from Paris since the beginning of November 1913 up to the end of January 1914, and it was during the three months that I taught these lectures at the Sorbonne that Borel was kind enough to propose that I should submit them to his series of monographs. Editing them was the work of Gaston Julia, who was then a third-year student at l'École Normale supérieure: I would like to express my warmest thanks to him for the excellent manner by which he accomplished that task.

I have not cited all of the important work that is attached to my subject, but only the ones that the reader should consult to begin with in order to go further into the theory. On the other hand, those who would like to learn about the historical development of the subject can consult:

1. The article II.A, 7*a* in the *Encyclopädie der mathematischen Wissenschaften* (which is already a bit old-fashioned).

2. The talk that I gave at the Fifth International Congress (Cambridge, 1912), which was published in the first volume of the *Proceedings*, page 163.

I regret that lack of time has prevented me from addressing the beautiful work of Liouville that is followed so closely here and represents a very important complement to Sturm's research. That is why questions of asymptotic values and developments into series of arbitrary functions are excluded from this volume. Meanwhile, even in this incomplete form, I hope that this little book might be regarded as a token of my admiration for the work of that French geometer, which is work that the rest of us in America were proud to be able to develop in several directions.

TABLE OF CONTENTS

]
PREF	ACE	
	CHAPTER I	
	SOME EXISTENCE THEOREMS	
1.	Fundamental theorem	
2.	Various extensions of the fundamental existence theorem	
3.	Review of some known facts	
	CHAPTER II	
	THE ANALOGIES BETWEEN LINEAR DIFFERENTIAL SYSTEMS AND LINEAR ALGEBRAIC SYSTEMS	
4.	Algebraic systems	
5.	Differential systems	
6.	The adjoint equation	
7.	Bilinear forms	
8.	Adjoint systems	
9.	Some properties of adjoint systems	
10.	Quasi-differential equations. Second-order equations	
11.	Characteristic numbers	
	CHAPTER III	
	REAL SOLUTIONS AND THEIR ZEROES IN THE SIMPLEST CASES	
12.	Solutions for one invariable equation	
13.	Effect produced on the solutions by a change of coefficients in the equation	
14.	The comparison theorems	
15.	Sturm's oscillation theorems	
16.	Study of the characteristic values from the standpoint of their reality and order of multiplicity	
	CHAPTER IV	
	CHARACTERISTIC FUNCTIONS AND THEIR ZEROES IN SOME MORE- GENERAL CASES	
17.	The reality of characteristic numbers	
18.	Systems with periodic conditions	
19.	Second extension of the Sturm problem. Klein's oscillation theorem	
	CHAPTER V	
	GREEN FUNCTIONS AND THEIR APPLICATIONS	
20	Existence and fundamental properties of Green functions	
20. 21	The relationship between the theory of differential systems and that of integral equations	
21.	The method of successive approximations for differential systems	

CHAPTER I SOME EXISTENCE THEOREMS (¹)

1. Fundamental theorem. – In this course, we will often have to study linear second-order differential equations. We shall write them in the form:

$$\frac{d^2u}{dx^2} + p\frac{du}{dx} + qu = r,$$

in which p, q, r are functions of the independent variable x.

Above all, it is important to examine the conditions for the existence of solutions to the preceding equation. From that viewpoint, we prove the following theorem:

Theorem:

Let x be a real variable that can vary over the closed, finite interval $A \le x \le B$. p, q, r are continuous functions of that variable x (which can have the form $p_1 + i p_2$, $q_1 + i q_2$, $r_1 + i r_2$, in which p_1 , p_2 , q_1 , q_2 , r_1 , r_2 are real and continuous functions of x, and i is the symbol for the imaginary unit). Let c be an arbitrary value in the interval (A, B). The equation:

$$\frac{d^2u}{dx^2} + p\frac{du}{dx} + qu = r$$

admits a unique solution u that verifies the conditions:

(1)
$$\begin{cases} u(c) = \gamma, \\ u'(c) = \gamma', \end{cases}$$

in which γ and γ' are two arbitrary given constants.

We shall call the combination of the equation and the preceding conditions that the solution must verify at *c* a *differential system*. Our theorem can then be stated as:

Any linear differential system of the preceding type has a unique solution. We shall use the method of successive approximations.

^{(&}lt;sup>1</sup>) For the method of successive approximations in the theory of nonlinear equations, one can consult Picard's *Traité d'Analyse*, Tome 2. For the linear equations, see:

PEANO, Math. Ann. 32 (1888), pp. 450.

BÔCHER, Amer. J. Math. 24 (1902), pp. 311.

First consider the particular equation:

$$\frac{d^2u}{dx^2}=\varphi(x)\,,$$

in which $\varphi(x)$ is continuous in (A, B).

Upon integrating by parts, one will find that any solution of that equation is given by the formula:

$$u = \int_c^x (x-\xi)\varphi(\xi)d\xi - C(x-c) - D .$$

One can choose C and D such that u takes the value γ at c and u' takes the value γ' . In order to do that, take $C = \gamma'$ and $D = \gamma$. The solution thus-obtained is unique, and our theorem is true for this particular case.

If one has, for example, $\gamma = \gamma' = 0$ then one will find that:

$$u = \int_c^x (x - \xi) \varphi(\xi) d\xi, \qquad u' = \int_c^x \varphi(\xi) d\xi$$

Now imagine the general case.

Choose a function $u_0(x)$ that is subject to just the condition that it is continuous and has a continuous derivative in (A, B), and let u_1 be the solution to the equation:

$$u_1'' = -p \, u_0' - q \, u_0 + r ,$$

$$u_1 \, (c) = \gamma , \qquad u_1' \, (c) = \gamma' .$$

From the preceding special case, u_1 is unique and well-defined. It will be a function with a continuous derivative in (A, B). We can then find a function u_2 such that:

$$u_{2}'' = -p u_{1}' - q u_{1} + r \qquad [u_{2}(c) = \gamma, u_{2}'(c) = \gamma']$$

Upon continuing this process indefinitely, we will define an infinite sequence of functions u_0 , u_1 , u_2 , ..., u_n , ... We shall show that u_n tends to a limiting function that is the solution of the proposed linear differential system.

It amounts to the same thing for us to show that the series:

$$u_1 + (u_2 - u_1) + \ldots + (u_n - u_{n-1}) + \ldots,$$

which we shall call:

which will also verify:

$$v_1+v_2+\ldots+v_n+\ldots,$$

is convergent, and that its sum is the desired solution.

From the process of their formation, u_n and u_{n-1} have continuous derivatives in (A, B). Therefore, v_n is also continuous and has continuous derivatives v'_n .

We now prove that the two series:

$$(2) v_1 + v_2 + \dots + v_n + \dots,$$

$$v_1'+v_2'+\cdots+v_n'+\cdots$$

are uniformly convergent in (A, B). We can start with v_2 . v_2 and v'_1 are continuous in (A, B), so their absolute values have a maximum C:

$$\begin{vmatrix} |v_2| \\ |v_2'| \end{vmatrix} = C .$$

p and q are continuous, so |p| + |q| is a continuous function that admits a maximum M in (A, B):

$$|p|+|q| \leq M.$$

Finally, let *L* be the greater of the two quantities 1 and B - A. We shall show that for any *n*, we have:

$$\frac{|v_2|}{|v_2'|} = \frac{C L^{n-2} M^{n-2} |x-c|^{n-2}}{(n-2)!}.$$

For n = 2, those inequalities will reduce to the ones that led us to choose *C*. Suppose that they are true up to the index n - 1, and show that they are true for *n*.

One has:

$$u''_{n} = -p u'_{n-1} - q u_{n-1} + r,$$

$$u''_{n-1} = -p u'_{n-2} - q u_{n-2} + r.$$

Therefore, upon subtracting the two, one will have:

$$v_n'' = -p v_{n-1}' - q v_{n-1} + r$$

 v_n then verifies an equation of the special type that was studied, and one will have:

$$v_n(c) = u_n(c) - u_{n-1}(c) = 0,$$

 $v'_n(c) = 0.$

One then concludes that one has:

$$v_{n} = \int_{c}^{x} (x - \xi) [-p(\xi) v_{n-1}'(\xi) - q(\xi) v_{n-1}(\xi)] d\xi,$$

$$v_{n}' = \int_{c}^{x} [-p(\xi) v_{n-1}'(\xi) - q(\xi) v_{n-1}(\xi)] d\xi.$$

We remark that $|x - \xi|$ and 1 are $\leq L$, so we will have:

$$\begin{vmatrix} v_n \\ v'_n \end{vmatrix} \\ \leq \int_c^x LM \frac{C L^{n-3} M^{n-3} |x-c|^{n-3}}{(n-3)!} |\xi-c|^{n-3} |d\xi| .$$

Upon performing the integration, we will get the inequality that was to be proved.

We now point out that |x - c| < L, so we will have:

$$\frac{|v_n|}{|v'_n|} \le \frac{C L^{2(n-2)} M^{n-2}}{(n-2)!}.$$

The series whose general term is:

$$\frac{CL^{2(n-2)}M^{n-2}}{(n-2)!}$$

is obviously convergent for any C, L, M. Hence, the series (2) and (3) are uniformly convergent in all of the interval (A, B). Moreover, one notes that v_n is continuous, along with v'_n , since u_n and u_{n-1} are.

If we let *u* denote the sum $v_1 + v_2 + ...$ then the function *u* will be continuous, along with its derivative *u'*, which is equal to $v'_1 + v'_2 + ...$

It remains to be shown that the function $u = v_1 + v_2 + ...$ is a solution of the differential system:

1. First of all, one has:

$$v_1(c) = u_1(c) = \gamma, \quad v'_1(c) = u'_1(c) = \gamma',$$

 $v_n(c) = u_n(c) - u_{n-1}(c) = 0 \text{ and } v'_n(c) = 0 \text{ for } n \ge 2.$

Therefore:

$$u(c) = g, \qquad u'(c) = \gamma'.$$

2. I say that u verifies the differential equation at any point of (A, B).

In order to show that, define the function -pu'-qu. The two series u and u' are convergent, so one has:

$$p u' - q u = (-p v_1' - q v_1) + (-p v_2' - q v_2) + \cdots$$

However, one has:

$$-p v'_{n-1} - q v_{n-1} = v''_n$$
 for $n \ge 3$,

$$-p v_1' - q v_1 = v_1'' + v_2'' - r.$$

Hence:

$$-pu' - qu = -r + v_1'' + v_2'' + \cdots$$

From the way that we combined the two series u and u', it is clear that the series $v_1'' + v_2'' + \cdots$ is uniformly convergent, so it will follow that u has a second derivative that is equal to $v_1'' + v_2'' + \cdots$

That permits us to write:

$$-pu'-qu=-r+u'',$$

and we see that u indeed verifies the differential equation.

It is thus proved that the differential system has a solution u that is continuous and has a continuous derivative in the entire interval (A, B).

That solution is unique. If there were two of them then their difference would be a solution to the system:

$$u'' + pu' + qu = 0,$$
 $u(c) = 0,$
 $u'(c) = 0,$

and would not be identically zero.

Now, that system, which we call the *system with no right-hand side or homogeneous system* corresponding to the proposed system, has only one solution, which is zero.

Indeed, one immediately establishes Abel's identity between two solutions u_1 and u_2 of the homogeneous equation:

$$u_1' u_2 - u_2' u_1 = k \exp\left(-\int_0^x p \, dx\right)$$

in which *k* is a certain constant.

Therefore, suppose that the homogeneous system admits a solution u_1 that is not identically zero. That would say that one would find a point c_1 in (A, B) where $u_1(c_1) \neq 0$.

Hence, define a solution u_2 to the homogeneous equation such that:

$$u_2(c_1) = 0$$
, $u'_2(c_1) = 1$,

which is possible.

Construct Abel's identity with those two solutions u_1 and u_2 . Upon substituting the value c for x, one will have:

$$u_1(c) = 0$$
, $u'_1(c) = 0$.

Hence k = 0, i.e.:

$$u_1' u_2 - u_2' u_1 = 0$$

Now, it one takes x to be equal to c_1 then the left-hand side will reduce to $-u'_1(c_1)$, which is not zero. The contradiction is obvious. The homogeneous system will then have only the zero solution.

Therefore, the proposed system has a unique and well-defined solution.

2. Various extensions of the fundamental existence theorem. – The proof that we gave of the existence theorem is valid when x is real and the functions p(x), q(x), r(x) are continuous. We shall indicate some cases in which the proof that we just gave can be applied with slight modifications that are either more general than the preceding case or constitute some applications of the same method to some similar questions.

1. Suppose, as always, that x is real, $A \le x \le B$, and suppose that p, q, r are finite, but they can take on a finite number of discontinuities in (A, B). We must specify what we mean by a solution of the differential equation here. It is obvious that it will be a function that must verify the equation at any point where p, q, r are continuous. Moreover, we shall demand that this function is everywhere-continuous and has a continuous derivative. Since the hypothesis of the continuity of p and q will be introduced only when we wish to fix the maximum M of |p| + |q|, and since p and q are finite here, we can also fix a number M that is greater than |p| + |q|, so our argument will be immediately valid. There is therefore one and only one solution to a differential system like the ones that we are considering.

2. Suppose that x is a complex variable and p, q, r are analytic functions that are defined in a continuum D in the Weierstrass sense. That domain, which is called a *Weierstrass domain*, is assumed to be simply connected. Its characteristic property is that one can describe a circle with its center at any point in the domain that is small enough that all of the points that are interior to that circle or situated on its circumference will belong to the domain. In certain cases, that domain can extend to infinity, but it will always be supposed to be simply connected. There would be no inconvenience in supposing that this domain can overlap itself at certain points in the manner of a Riemann surface. The reader will easily understand why one has needs only to make some very slight modifications to the preceding analysis in order to prove that the system (1) has one and only one solution in the present case.

The case in which p, q, r are analytic functions of the real variable x in (A, B) reduces, in turn, to the preceding one because one can put (A, B) in a two-dimensional domain in which p, q, r are analytic at x. The solution to the system (1) will be necessarily analytic.

3. Suppose that x is real, and p, q, r depend upon the real parameter λ . (One can suppose that p, q, r depend upon k real parameters $\lambda_1, \lambda_2, ..., \lambda_k$. To simplify, we shall take just one of them, although our results will still be valid for k parameters.) We suppose that p, q, r are continuous in the two real variables x and λ . [x is in the closed interval (A, B), while λ is in the interval (A₁, A₂), which can be open.]

One can likewise suppose that γ and γ' are continuous functions of λ . That solution will be a continuous function of x and λ under our hypotheses. In order to ensure that, it will suffice to take the C that is used in our calculus of inequalities to have a positive value that is greater than $|v_2|$ and $|v'_2|$ for any x in (A, B) and λ in an arbitrary closed interval that is interior to (A_1, A_2) , and for

any *M* with a value that is $\geq |p| + |q|$ under the same conditions.

The argument proceeds as follows: The series:

$$C+\frac{CL^2M}{1!}+\cdots,$$

which majorizes the series $\sum v_n$ and $\sum v'_n$, is independent of λ . Hence, $\sum v_n$ and $\sum v'_n$ are uniformly convergent for x in (A, B) and λ in a closed interval that is interior to (A₁, A₂). Hence, their sums u and u' are continuous in x and λ , and since u'' = -pu'-qu+r will also be continuous in x and λ .

4. Suppose that x is real, while p, q, r are continuous in x and λ and analytic in λ (or more generally, analytic in $\lambda_1, \lambda_2, ..., \lambda_k$), γ and γ' are also analytic in λ when x is in an interval (A, B) and λ is in a Weierstrass domain D. One can once more conclude that the solution u, as well as u', are continuous in x and λ and analytic in λ .

5. What is the effect that is produced on the solution by arbitrary small variations of the coefficients p, q, r, and γ , γ' ? x is supposed to be real.

In order to see what that effect would be, it is important to understand how the solutions depends upon p, q, r, γ, γ' .

It is determined when one knows those five functions of x (γ and γ' are constants, but we shall include them within the concept of *functions* of x), and we then write $u = \mathcal{F}(p, q, r, \gamma, \gamma')$.

However, \mathcal{F} is not an ordinary function of the five arguments p, q, r, γ, γ' because if that were true then the value of u would not vary at a point x_0 when one changes p, q, r, γ, γ' arbitrarily, provided that the values of those five functions remain the same. Some simple examples will show that, on the contrary, the value of u will change under those conditions, so u is not an ordinary function of p, q, r, γ, γ' . It is what one calls a *functional* of those five arguments, i.e., a function of x that is determined in (A, B) by three functions p, q, r that are defined in (A, B) and the two constants γ and γ' . We shall confine ourselves to the case where p, q, r are continuous here.

Under those conditions, I say that the functional:

$$u = \mathcal{F}(p, q, r, \gamma, \gamma')$$

is a *continuous functional* of its five arguments, i.e., when one is given a positive number δ that is as small as one desires, one can determine a positive number ε such that arbitrary variations of the functions p, q, r whose absolute values are everywhere less than ε in the interval (A, B) and arbitrary variations of the constants γ and γ' whose absolute values are likewise less than ε will imply a variation of the functional \mathcal{F} whose absolute value is less than δ everywhere in (A, B).

One can then answer the question that was posed to begin with by proving that $u = \mathcal{F}(p, q, r, \gamma, \gamma')$ is a continuous functional.

Chapter I

Indeed, one first proves that the sum and product of several continuous functionals are continuous functionals, just as one has for continuous functions. One also sees very easily that the integral:

$$\psi = \int_c^x \Phi(p,q,r,\gamma,\gamma') dx$$

will be a continuous functional when Φ is also one.

It follows from those elementary propositions that the terms in the two series $\sum v_n$, $\sum v'_n$ are continuous functions of p, q, r, γ , γ' . Those two series of continuous functionals have terms that are respectively lower in modulus than the ones in a series of constants that one can easily calculate from the known expressions for v_n and v'_n when one supposes that one restricts oneself to the domain of functions that is defined by functions p, q, r that are continuous in x in (A, B) and finite along with γ and γ' :

$$\begin{array}{l} \mid p \mid < N \\ \mid q \mid < N \\ \mid r \mid < N \end{array} \quad \begin{array}{l} \mid \gamma \mid < N, \\ \mid \gamma' \mid < N, \end{array}$$

in which N is a certain constant.

One chooses $C \ge |v_1|$, $|v_1'|$ for any p, q, r, γ, γ' that belong to that domain, and one will soon see that this is possible. Similarly, one chooses $M \ge |p| + |q|$; for example, M = 2N.

The two series $\sum v_n$ and $\sum v'_n$ are then uniformly convergent in the preceding domain, i.e., one can determine and index N_1 such that for $n_1 \ge N_1$:

$$\left|\sum_{n=n_1}^{+\infty} v_n\right|$$
 and $\left|\sum_{n=n_1}^{+\infty} v'_n\right|$

are $< \varepsilon$ for any p, q, r, γ, γ' that belong to the domain envisioned. Each functional v_n, v'_n is continuous, so one can prove, just as one does for series whose terms are continuous functions, that the sums $\sum v_n$ and $\sum v'_n$, i.e., u and u', are continuous functionals of p, q, r, γ, γ'

Moreover, we remark that nothing in the preceding argument compels us to suppose that p, q, r have a finite number of discontinuities while they all remain finite.

Finally, this method of variations still applies when p, q, r are analytic with respect to the complex variable x in a domain D. Any small variation of p, q, r that leaves those functions analytic will produce a variation of u, u' that is uniformly small in any region that it interior to D.

6. Quasi-differential equations:

Let an equation have the form:

$$\frac{d}{dx}\left[a_1\frac{d}{dx}(au)+bu\right]+l\frac{dy}{dx}(au)+qu=r.$$

If a, a_1 , b are differentiable then one will have an ordinary differential equation upon performing the differentiations.

If a, a_1 , b, l, q, r are continuous, but not differentiable, functions then one will no longer be dealing with an ordinary equation. We suppose that a and a_1 are everywhere non-zero in (A, B).

A solution of that equation will be a function u that permits one to perform the differentiations and that also verifies the preceding equation.

u will have no derivative when a has none, but a u does. Hence, the preceding relation is not a relation between the values of u and its first and second derivatives, as would be the case with ordinary differential equations. That is what we call a quasi-differential equation. Our approximation method still applies. We start from a function u_0 that we substitute in $-l\frac{d}{dr}(au)-qu+r$, and is such that a u has a continuous derivative. Step by step, we will have

the desired solution u in the form of a series that is unique and verifies the conditions:

$$u(c) = \gamma,$$
$$\left[\frac{d}{dx}(au)\right]_{c} = \gamma'$$

Moreover, those equations reduce to first-order differential systems by setting:

$$y = a u,$$

$$z = b u + a_1 y',$$

thanks to the hypotheses that $a \neq 0, a_1 \neq 0$.

7. What we have said about the second-order equation applies with no essential modification to the n^{th} -order equation and to the differential system:

$$\frac{d^{n}u}{dx^{n}} + p_{1}\frac{d^{n-1}u}{dx^{n-1}} + \dots + p_{n}u = r,$$

$$u(c) = \gamma_{1},$$

$$u'(c) = \gamma',$$

$$\dots$$

$$u^{(n-1)}(c) = \gamma^{(n-1)}.$$

3. Review of some known facts. – Let an equation be given:

$$\frac{d^n u}{dx^n} + p_1 \frac{d^{n-1} u}{dx^{n-1}} + \dots + p_n u = r$$

Let c be a point in the interval (A, B) in which the coefficients are continuous, or else a point in the Weierstrass domain where they are analytic. We shall say *principal solutions* for the point c to mean the solutions $u_1(x), u_2(x), ..., u_n(x)$ of the homogeneous equation:

$$\frac{d^n u}{dx^n} + p_1 \frac{d^{n-1} u}{dx^{n-1}} + \dots + p_n u = 0$$

that verify the conditions:

$u_1(x)=1,$	$u_1'(c)=0,$,	$u_1^{(n-1)}(c) = 0$,
$u_2(x)=0,$	$u_2'(c)=1,$,	$u_2^{(n-1)}(c) = 0$,
		,	,
$u_n(x)=0,$	$u_n'(c)=0,$,	$u_n^{(n-1)}(c) = 1$.

Each of those *n* functions $u_1(x), ..., u_n(x)$ is well-defined and unique.

.

 $u_1(x), ..., u_n(x)$ constitute a fundamental system of integrals for the homogeneous equation since any integral of that equation will be a linear and homogeneous combination with constant coefficients of $u_1, u_2, ..., u_n$, and conversely.

Recall that the necessary and sufficient condition for the n solutions $u_1, u_2, ..., u_n$ of the homogeneous equation to form a fundamental system is that their Wronskian:

$$\Delta = \begin{vmatrix} u_1 & u_2 & \cdots & u_n \\ u'_1 & u'_2 & \cdots & u'_n \\ \vdots & \vdots & \cdots & \vdots \\ u_1^{(n-1)} & u_2^{(n-1)} & \cdots & u_n^{(n-1)} \end{vmatrix}$$
 must be $\neq 0$.

.

That Wronskian is calculated by Abel's formula:

$$\Delta(x) = \Delta(c) \exp\left(-\int_0^c p_1 dx\right) \,.$$

It cannot be zero at a point without being identically zero.

Another necessary and sufficient condition for the system of solutions $u_1, u_2, ..., u_n$, to be fundamental is that $u_1, u_2, ..., u_n$ must be linearly independent, i.e., there can exist no constants that are not all zero $c_1, c_2, ..., c_n$ such that one has:

$$c_1 u_1 + c_2 u_2 + \ldots + c_n u_n = 0$$

identically.

The condition $\Delta \neq 0$ is the condition for the *n* solutions $u_1, u_2, ..., u_n$ to be linearly independent. If one returns to the inhomogeneous equation:

$$\frac{d^n u}{dx^n} + p_1 \frac{d^{n-1} u}{dx^{n-1}} + \dots + p_n u = r$$

then its general solution will be obtained by adding a particular solution u_0 to the general solution of the homogeneous equation:

$$c_1 \eta_1 + c_2 \eta_2 + \ldots + c_n \eta_n$$
.

($c_1, c_2, ..., c_n$ are constants, while $\eta_1, ..., \eta_n$ is a fundamental system of solutions of the homogeneous equation.)

CHAPTER II

THE ANALOGIES BETWEEN LINEAR DIFFERENTIAL SYSTEMS AND LINEAR ALGEBRAIC SYSTEMS (²)

4. Algebraic systems. – In his research concerning differential equations, Sturm, like many of his predecessors, was led to envision them as the limits of finite difference equations.

For example, consider the equation:

$$\frac{d^2u}{dx^2} + p\frac{du}{dx} + qu = r,$$

in which p, q, r are continuous functions of x in the interval (A, B).

$$\begin{array}{c|c} A & x_1 & x_{k-1} & B \\ \hline Figure 1. \end{array}$$

Imagine that (A, B) has been divided into k equal parts by the points $x_1, x_2, ..., x_{k-1}$. Δx denotes one of the differences:

$$x_i - x_{i-1},$$
 $\begin{pmatrix} x_0 = A, \\ x_k = B \end{pmatrix}$

One has:

$$\Delta u (x_i) = u (x_{i+1}) - u (x_i) ,$$

$$\Delta^2 u (x_i) = u (x_{i+2}) - 2u (x_{i+1}) + u (x_i) .$$

Consider the equation then:

$$\frac{\Delta^2 u(x_i)}{\Delta x^2} + p(x_i) \frac{\Delta u(x_i)}{\Delta x} + q(x_i) u(x_i) = r(x_i),$$

which must be true for all values i = 0, 1, 2, ..., k - 2.

If one multiplies both sides by Δx^2 then the equation can be written:

⁽²⁾ For the purely-algebraic facts, see, for example:

BÔCHER, *Introduction to Higher Algebra*, New York, 1907, or the German edition in 1909, Teubner. For the theory of the *adjoint equation*, see:

DARBOUX, Théorie des Surfaces, t. II, Chap. V.

For the differential systems, see:

BÔCHER, Trans. Amer. Math. Soc. 14 (1913), pp. 403.

(1)
$$P_i u(x_i) + Q_i u(x_{i+1}) + R_i u(x_{i+2}) = r_i, \quad r_i = r(x_i) \Delta x^2 \quad (i = 0, 1, 2, ..., k-2).$$

Equation (1), which couples the values of u at three consecutive dividing points, must be its values for i = 0, 1, ..., k - 2, as one has seen. It thus gives rise to k - 1 algebraic equations between the k + 1 unknowns $u(x_0), u(x_1), ..., u(x_{k-1}), u(x_k)$. If k tends to infinity then the initial differential equation will appear to be the limit of a system of linear algebraic equations in which the number of unknowns, which always exceeds the number of equations by two units, will tend to infinity.

The system (1) admits an infinitude of solutions, in general, which depend linearly upon two arbitrary constants. Without going into the details, one might expect that the differential equation will have an infinitude of solutions that depend linearly upon two arbitrary constants.

If one has a linear differential equation of order *n* then one can regard it as the limit of a finitedifference equation, i.e., a linear algebraic system, each equation of which links the values of the solution *u* at n + 1 consecutive points that are chosen from $x_0, x_1, ..., x_k$, and the number of unknowns in them exceeds the number of equations by *n*.

We shall first imagine the homogeneous algebraic system:

(2)
$$\begin{cases} a_{11} \xi_1 + a_{12} \xi_2 + \dots + a_{1N} \xi_N = 0, \\ \dots \\ a_{M1} \xi_1 + a_{M2} \xi_2 + \dots + a_{MN} \xi_N = 0. \end{cases}$$

One knows that:

1. If the system (2) has only the solution:

$$\xi_1=\xi_2=\ldots=\xi_N=0$$

then it is called *incompatible*.

2. If the system (2) has several solutions:

$$\xi_1', \xi_2', \dots, \xi_N', \\ \xi_1'', \xi_2'', \dots, \xi_N'',$$

then one will have a more general solution that is given by the formulas:

 $c_1 \xi_1' + c_2 \xi_1'' + \cdots, \qquad c_1 \xi_2' + c_2 \xi_2'' + \cdots, \qquad \dots, \qquad c_1 \xi_N' + c_2 \xi_N'' + \cdots,$

in which $c_1, c_2, ...$ are some arbitrary constants, and one can always find a finite number of solutions such that the latter formula gives the general solution to the system (2).

If a finite number of those solutions are linearly independent then one says that they form a *fundamental system of solutions of system* (2).

Recall that the solutions:

 $\xi_1', \xi_2', \dots, \xi_N', \\ \xi_1'', \xi_2'', \dots, \xi_N'',$

are called *linearly independent* if it is impossible to find constants k', k'', ... that are not all zero such that one has, at the same time:

```
\begin{aligned} &k'\,\xi_1'+k''\,\xi_1''+\cdots\,=0\;,\\ &k'\,\xi_2'+k''\,\xi_2''+\cdots\,=0\;,\\ &\ldots \end{aligned}
```

The rank of the matrix:

 $\left(\begin{array}{cccc}a_{11}&a_{12}&\cdots&a_{1N}\\ \vdots&\vdots&\cdots&\vdots\\ a_{M1}&\cdots&\cdots&a_{MN}\end{array}\right)$

is the maximum order of a non-zero determinant that is drawn from that matrix. If that rank is *p* then one will have the following theorem:

The number of solutions to a fundamental system is always N - p. That number will be called the **index of compatibility of the system** (2), or simply its **index**.

The reader has already observed the striking analogy between the preceding and the features that were recalled at the end of the last chapter in relation to the homogeneous differential equation.

If one envisions an inhomogeneous system:

(3)
$$\begin{cases} a_{11} \xi_1 + a_{12} \xi_2 + \dots + a_{1N} \xi_N = b_1, \\ \dots \\ a_{M1} \xi_1 + a_{M2} \xi_2 + \dots + a_{MN} \xi_N = b_M \end{cases}$$

then one will be led to consider, in parallel with the matrix:

$$\left(egin{array}{cccc} a_{11} & \cdots & a_{1N} \ dots & \cdots & dots \ a_{M1} & \cdots & a_{MN} \end{array}
ight),$$

the augmented matrix:

$$\left(egin{array}{ccccc} a_{11}&\cdots&a_{1N}&b_1\dots&\cdots&dots&dots\ a_{M1}&\cdots&a_{MN}&b_M \end{array}
ight),$$

and one knows that the necessary and sufficient condition for the system (3) to be compatible is that the rank of the matrix of a is equal to the rank of the augmented matrix.

Under those conditions, the general solution of the system (3) will be obtained by adding a particular solution to the general solution of the homogeneous system (2).

5. Differential systems. – In the case where M = N - n, the system (3) is the analogue of an inhomogeneous linear equation of order n. In order to get the analogue of the algebraic system where M = N, one must add n supplementary conditions that the solution must verify to the differential equation of order n. For example, one can take:

$$u(c) = \gamma,$$

$$u'(c) = \gamma',$$

$$\dots,$$

$$u^{(n-1)}(c) = \gamma^{(n-1)},$$

in which c is a point of (A, B), $u^{(i)}(c)$ is the value of $\frac{d^i u}{dx^i}$ at the point c, and γ , γ' , ..., $\gamma^{(n-1)}$ are n arbitrary constants. (If one passes to the finite-difference equation then those n conditions will indeed give n relations between the values of u at n - 1 consecutive points after c, so n linear equations that must be added to the N - n equations that produce the equation itself.)

In what follows, we shall envision some more-general supplementary conditions:

1. One can take two points a, b in the interval (A, B) or in the Weierstrass domain, where one restricts x to remain, and one imposes the following condition on a solution:

$$\alpha u(a) + \alpha' u'(a) + \dots + \alpha^{(n-1)} u^{(n-1)}(a) + \beta u(b) + \beta' u'(b) + \dots + \beta^{(n-1)} u^{(n-1)}(b) = \gamma,$$

in which α , α' , ..., β , β' , ..., γ are constants. (It would be pointless to introduce the n^{th} derivative of u since the differential equation will permit one to express it as a function of u, u', ..., $u^{(n-1)}$.)

We shall let U(u) denote the left-hand side of the preceding condition. In this general case, we take *m* conditions of the preceding form:

$$U_i(u) = \gamma_i$$
 (*i* = 1, 2, ..., *m*).

One can take h points a, b, ... in (A, B), instead of only two, and impose a linear relation between the values of u and its (n - 1) first derivatives at the points a, b, ...

2. One can consider a second type of condition:

$$U_i(u) = \int_A^B [f_i^{(0)}(x)u(x) + f_i^{(1)}(x)u'(x) + \dots + f_i^{(n-1)}(x)u^{(n-1)}(x)]dx = \gamma_i,$$

Chapter II

in which $f_i^{(0)}$, $f_i^{(1)}$, ..., $f_i^{(n-1)}$ are given functions of x that we suppose to be continuous in order to avoid complicating matters.

3. Finally, upon taking $U_i(u)$ to be the sum of an expression of type 1. and an integral of type 2. above, one can take the more-general conditions:

$$U_i(u) = \gamma_i$$
.

Those conditions are *linear* from two standpoints:

1. One can imagine that they are limits of linear relations between the values of u at the k division points in the interval (A, B) when k increases indefinitely.

2. If one has several functions $u_1, u_2, ..., and$ one forms the functional U_i for the function $c_1 u_1 + c_2 u_2 + ..., in$ which $c_1, c_2, ...$ are arbitrary constants, then one will have:

$$U_i(c_1u_1+c_2u_2+...)=c_1U_i(u_1)+c_2U_i(u_2)+...$$

That functional $U_i(u)$ is distributive, and one knows that distributivity is the essential characteristic of algebraic linear expressions. Any expression that possesses that distributive property can be called *linear*.

Having said that, we can envision differential systems of the form:

(1)
$$L(u) = 0,$$

 $U_i(u) = 0$ $(i = 1, 2, ..., m),$

in which:

$$L(u) = \frac{d^{n}u}{dx^{n}} + p_{1}\frac{d^{n-1}u}{dx^{n-1}} + \dots + p_{n}u.$$

The conditions $U_i(u) = 0$ can be linear conditions of the most-general type 3. that was envisioned above.

If m is arbitrary then that system will be the analogue of the linear algebraic system of M equations in N unknowns that we have exhibited in no. **4**.

The analogue of the system [**§ 4**, (3)] will be:

(2)
$$\begin{array}{c} L(u) = r, \\ U_i(u) = \gamma_i \quad (i = 1, 2, ..., m). \end{array}$$

We shall develop a theory of compatibility for those two types of systems – viz., homogeneous and inhomogeneous ones – that is analogous to the one that we recalled for algebraic systems (2) and (3) in no. 4.

Take a homogeneous system (1) that:

1. Might admit only the solution u = 0. It will then be called *incompatible*.

2. Might admit linearly-independent solutions $\eta_1, \eta_2, ..., \eta_k$. One will always have $k \le n$ since the equation L(u) = 0 will have at most n linearly-independent solutions. Under those conditions:

$$c_1 \eta_1 + c_2 \eta_2 + \ldots + c_k \eta_k$$

will be a solution of the system that depends upon k arbitrary constants $c_1, c_2, ..., c_k$. We say *fundamental system of solutions* to mean a system of linearly-independent solutions $\eta_1, \eta_2, ..., \eta_k$ that give the general solution to the differential system by the formula:

$$c_1 \eta_1 + c_2 \eta_2 + \ldots + c_k \eta_k$$

If one considers the system with the right-hand side (2) then:

1. It might not have a solution because it might happen that no solution of L(u) = r verifies the supplementary conditions.

2. It might admit one solution u_0 . If $\eta_1, \eta_2, ..., \eta_k$ constitute a fundamental system of solutions of the homogeneous system (2) that corresponds to (1) then the general solution of that system (2) will be:

$$u_0 + c_1 \eta_1 + c_2 \eta_2 + \ldots + c_k \eta_k$$

in which $c_1, c_2, ..., c_k$ are k arbitrary constants. One has seen that the necessary and sufficient condition for a homogeneous algebraic system [§ 4, (2)] to have k independent solutions is that the

rank of the matrix $\begin{pmatrix} a_{11} - a_{1N} \\ a_{M1} - a_{MN} \end{pmatrix}$ should be N - k.

Let us study an analogous condition for the homogeneous differential system (1).

Let $u_1, u_2, ..., u_n$ be a fundamental system of solutions of L(u) = 0. Choose the constants in such a way that:

$$u = c_1 u_1 + \ldots + c_n u_n$$

verifies the conditions:

$$U_i(u) = 0$$
 (*i* = 1, 2, ..., *m*).

One has the following conditions for determining the c_i :

$$c_1 U_1(u_1) + c_2 U_1(u_2) + \ldots + c_n U_1(u_n) = 0$$
,

$$c_1 U_2 (u_1) + c_2 U_2 (u_2) + \dots + c_n U_2 (u_n) = 0,$$

....
$$c_1 U_m (u_1) + c_2 U_m (u_2) + \dots + c_n U_m (u_n) = 0.$$

Those are linear and homogeneous equations in $c_1, ..., c_n$. One must consider the matrix:

$$\mathcal{T} = \begin{pmatrix} U_{1}(u_{1}) & \cdots & U_{1}(u_{n}) \\ U_{2}(u_{1}) & \cdots & U_{2}(u_{n}) \\ \vdots & \cdots & \vdots \\ U_{m}(u_{1}) & \cdots & U_{m}(u_{n}) \end{pmatrix}$$

If its rank is k then one will have n - k linearly-independent solutions $c_1, ..., c_n$. Each of those solutions gives a function $u = c_1 u_1 + ... + c_n u_n$. We then have n - k linearly-independent solutions for the system (1). [Those solutions are linearly-independent since $u_1, ..., u_n$ is a fundamental system of L(u) = 0.]

One will then have this result:

The necessary and sufficient condition for the homogeneous system (1) to have p linearlyindependent solutions is that the rank of the matrix T must be n - p.

That result is theoretical. Suppose that one has solved the equation L(u) = 0.

Let u_0 be a particular solution of L(u) = r, and let $u_1, u_2, ..., u_n$ be a fundamental system of integrals of L(u) = 0.

 $u = u_0 + c_1 u_1 + ... + c_n u_n$ will then be the general solution to L(u) = r. In order for such a solution to verify:

$$U_i(u) = \gamma_i$$
 (*i* = 1, 2, ..., *m*),

it is necessary and sufficient that the c_i should verify the equations:

$$c_1 U_1(u_1) + c_2 U_1(u_2) + \dots + c_n U_1(u_n) = \gamma_1 - U_1(u_0),$$

$$c_1 U_m(u_1) + c_2 U_m(u_2) + \dots + c_n U_m(u_n) = \gamma_m - U_m(u_0).$$

Conforming to the result that was found for the algebraic system $[\S 4, (3)]$, the two matrices:

$$\mathcal{T} = \begin{pmatrix} U_1(u_1) & \cdots & U_1(u_n) \\ \vdots & \cdots & \vdots \\ U_m(u_1) & \cdots & U_m(u_n) \end{pmatrix}$$

and

$$\mathcal{T}_{1} = \begin{pmatrix} U_{1}(u_{1}) & U_{1}(u_{2}) & \cdots & U_{1}(u_{n}) & \gamma_{1} - U_{1}(u_{0}) \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ U_{m}(u_{1}) & U_{m}(u_{n}) & \cdots & U_{m}(u_{n}) & \gamma_{m} - U_{m}(u_{0}) \end{pmatrix}.$$

The differential system (2) will have solutions if T and T_1 have the same rank.

If n - p is that common rank then the system (2) will have an infinitude of solutions that depend linearly upon p arbitrary constants since that is true for the system that determines the c_i .

That entirely-theoretical result supposes that the equation L(u) = r has been solved.

Meanwhile, if the rank of T is n, and one supposes that m = n, in addition, then the homogeneous system (1) will have no solution, but the system that determines the c_i will always have a unique solution. One will then have the following statement in a form that is independent of T and T_1 , which is a statement that has its analogue for linear algebraic systems:

If m = n and the homogeneous system (1) is incompatible then the system (2) will always have one and only one solution.

The parallelism of the linear algebraic systems and the differential ones can be pursued further. In order to avoid some very great complications, in what follows, we shall not envision limit conditions of the form $U_i(u) = \gamma_i$ that include integrals of the second type that was indicated. We shall even confine ourselves to "two-point conditions at *a* and *b*" on *u*, i.e., ones with the form:

$$U_{i}(u) = \alpha u(a) + \alpha' u'(a) + \dots + \alpha^{(n-1)} u^{(n-1)}(a) + \beta u(b) + \beta' u'(b) + \dots + \beta^{(n-1)} u^{(n-1)}(b) = \gamma_{i}$$

so the one-point conditions will be only one particular case.

The theory of the adjoint system to a given differential system likewise exhibits the greatest analogy with the parallel theory in the algebraic equations.

6. The adjoint equation. – First, recall some properties of the Lagrange adjoint equation. If one considers a differential linear expression:

$$L(u) = l_n \frac{d^n u}{dx^n} + l_{n-1} \frac{d^{n-1} u}{dx^{n-1}} + \dots + l_0 u,$$

and if one seeks to determine a multiplier v such that v L(u) is the derivative of a linear expression in $u, u', ..., u^{(n-1)}$ then one will find that v must satisfy the differential equation:

$$(-1)^{n} \frac{d^{n}(l_{n}v)}{dx^{n}} + (-1)^{n-1} \frac{d^{n-1}(l_{n-1}v)}{dx^{n-1}} + \dots + l_{0}v = 0,$$

which one calls the *Lagrange adjoint equation*. One obviously supposes in that statement that any coefficient l_i admits continuous derivatives up to order *i*. In that way, one makes the function l_i

more specific. Later on, that restriction shall be dropped. Finally, one supposes that $l_n \neq 0$ in all of the interval (A, B) considered in order for the equation L(u) = r to have continuous coefficients at every point in the interval when one divides both sides of it by the coefficient of $\frac{d^n u}{dx^n}$.

If we let M(v) denote the left-hand side of the adjoint equation:

$$M(v) = (-1)^{n} \frac{d^{n}(l_{n}v)}{dx^{n}} + (-1)^{n-1} \frac{d^{n-1}(l_{n-1}v)}{dx^{n-1}} + \dots + l_{0}v$$

then we will easily prove the identity:

(1)
$$v L(u) - u M(v) = \frac{d}{dx} P(u, v)$$

for any u and v. P(u, v) is a bilinear form relative to the variables $u, u', ..., u^{(n-1)}, v, v', ..., v^{(n-1)}$.

If one integrates the two side of the identity (1) between a and b, which are two arbitrary points in the interval (A, B), then one will have the identity:

(2)
$$\int_{a}^{b} [v L(u) - u M(v)] dx = [P(u,v)]_{a}^{b}.$$

Those two forms (1) and (2) are equivalent because if one supposes that (2) is true for any b in (A, B) then upon differentiating the two sides of (2) with respect to the upper limit b, which one can call x, one will get back to (1).

We shall call (1) the *Lagrange identity* and (2), *Green's formula*, although (1) and (2) were given by Lagrange, and Green did not give the formula (2). However, (2) is the one-dimensional analogue of Green's famous formula that permits one to reduce the calculation of a double integral to that of a curvilinear integral in certain cases. That explains our choice of terminology.

When one establishes the identity (1) or (2), one can then, in turn, deduce the property of the multiplier because if v is the solution of the adjoint equation then one will have M(v) = 0, so:

$$v L(u) = \frac{d}{dx} P(u, v)$$

If one integrates both sides between *A* and *B* then one will have:

(3)
$$\int_{A}^{B} v L(u) dx = \left[P(u,v)\right]_{A}^{B}$$

The integral $\int_{A}^{B} v L(u) dx$ will then be expressed by a linear function with values in $u, u', ..., u^{(n-1)}$ at the two points A and B.

Let us find the analogy for algebraic equations. To simplify the problem, suppose that L(u) is a second-order differential expression, so the right-hand side of (3) will be a linear function in the values of u and u' at A and B.

What corresponds to the expression L(u) here is M linear forms in $\xi_1, \xi_2, ..., \xi_N$ (M = N - 2):

 $\xi_1, ..., \xi_N$ are the values that *u* takes at the division point along the interval (A, B) [using the finitedifference equation that one substitutes for L(u) = r]. Multiplying L(u) by a function *v* and integrating from *A* to *B* will have the following analogue here: Multiply the preceding forms by constants $y_1, y_2, ..., y_M$, respectively, and add the results. Furthermore, $\int_A^B v L(u) dx$ is expressed linearly in terms of just the values of *u* and *u'* at *A* and *B*. Now, the value *u* (*A*) corresponds to ξ_1 , and u(B) corresponds to ξ_N . In the finite-difference equation, u'(A) corresponds to a linear expression in ξ_1 and ξ_2 , while u'(B) is a linear expression in ξ_{N-1} and ξ_N . We will get the analogue of a multiplier *v* by choosing the constants $y_1, ..., y_M$ in such a way that the expression:

$$y_1(a_{11}\xi_1 + \ldots + a_{1N}\xi_N) + \ldots + y_M(a_{M1}\xi_1 + \ldots + a_{MN}\xi_N)$$

includes only the values of ξ_1 , ξ_2 , ξ_{N-1} , ξ_N .

One must then annul the coefficients of $\xi_2, ..., \xi_{N-2}$, which will give the following equations that must be satisfied by $y_1, y_2, ..., y_M$:

(4)
$$\begin{cases} a_{13} \quad y_1 + a_{23} \quad y_2 + \dots + a_{M3} \quad y_M = 0, \\ a_{14} \quad y_1 + a_{24} \quad y_2 + \dots + a_{M4} \quad y_M = 0, \\ \dots \\ a_{1,N-2} \quad y_1 + a_{2,N-2} \quad y_2 + \dots + a_{M,N-2} \quad y_M = 0. \end{cases}$$

The matrix of coefficients in those equations is deduced from the matrix of coefficients of the forms above in $\xi_1, ..., \xi_N$:

$$\left(\begin{array}{cccc}a_{11}&a_{12}&\cdots&a_{1N}\\ \vdots&\vdots&\cdots&\vdots\\ a_{M1}&\cdots&\cdots&a_{MN}\end{array}\right)$$

by permuting the rows and columns and then suppressing the first two and last two rows.

A system such as (4) is the analogue of the adjoint equation, and in the case that we are currently considering, M = N - 2, there will be N - 4 equations in N - 2 unknowns in (4), which one would have to expect since the adjoint equation has order two here.

We shall return to this analogy after we have defined the adjoint system to a given differential system. Indeed, one has an even closer analogy with algebraic systems if one considers differential systems, instead of differential equations.

In the meantime, we shall recall some known facts from the theory of bilinear forms.

7. Bilinear forms. – Suppose that one has two systems of N variables:

```
x_1, x_2, \dots, x_N, y_1, y_2, \dots, y_N.
```

A bilinear form with respect to those variables is an expression such as:

 $a_{11} x_1 y_1 + a_{12} x_1 y_2 + \dots + a_{1N} x_1 y_N$ $+ a_{21} x_2 y_1 + a_{22} x_2 y_2 + \dots + a_{2N} x_2 y_N$ $+ \dots$ $+ a_{N1} x_N y_1 + a_{N2} x_N y_2 + \dots + a_{NN} x_N y_N$

that is linear with respect to x and with respect to y separately.

The determinant:

$$A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{21} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{vmatrix}$$

is called the *determinant of the form*.

If $A \neq 0$ then the form is ordinary. If A = 0 then the form is singular.

Make a substitution of the x_i that replaces them with X_i such that:

$$X_1 = c_{11} x_1 + c_{12} x_2 + \dots + c_{1N} x_N,$$

....
$$X_N = c_{N1} x_1 + c_{N2} x_2 + \dots + c_{NN} x_N.$$

We suppose that this substitution is ordinary, i.e., that the determinant $C = |c_{ik}|$ is not zero.

One can then replace the x_i as functions of X_i in the given bilinear form, which will then take the form $\sum d_{ii} X_i v_i$.

The transformed form will be singular or ordinary according to whether the proposed form is singular or ordinary, respectively.

Imagine that transformed form. One can write it as:

$$X_1 Y_1 + X_2 Y_2 + \ldots + X_N Y_N$$

on the condition that one must set:

$$Y_1 = d_{11} y_1 + d_{12} y_2 + \dots + d_{1N} y_N,$$

$$Y_N = d_{N1} y_1 + d_{N1} y_N + \dots + d_{NN} y_N.$$

That substitution replaces the y_i with the Y_i . If the proposed form is ordinary then that substitution will be ordinary.

Therefore, upon supposing that the initial form is ordinary, one can reduce the initial form $\sum_{i,j=1}^{N} a_{ij} x_i y_j$ to the canonical form $\sum_{i=1}^{N} X_i Y_i$ by two convenient ordinary substitutions that take the x_i to the X_i and the y_i to the Y_i .

The determinant *C* is $\neq 0$, so the forms $X_1, X_2, ..., X_N$ in $x_1, ..., x_N$ are independent, one can give them arbitrarily. The forms $Y_1, ..., Y_N$ will then be well-defined by the choice of X_i , and they will be independent since their determinant $D \neq 0$.

If we leave *p* of the forms X_i unchanged and change the remaining q = N - p then what effect will that have on the forms Y_i ?

When we choose the forms $X_1, ..., X_p, X_{p+1}, ..., X_{p+q}$, the corresponding forms Y will be $Y_1, ..., Y_p, Y_{p+1}, ..., Y_{p+q}$. If we choose the forms $X_1, X_2, ..., X'_{p+1}, ..., X'_{p+q}$, while preserving the preceding first p forms, and make $X'_{p+1}, ..., X'_{p+q}$ different from $X_{p+1}, ..., X_{p+q}$, but still forming a linearly-independent system with $X_1, ..., X_p$, then they will correspond to linearly-independent forms Y'_1 , $Y'_2, ..., Y'_p, Y'_{p+1}, ..., Y'_{p+q}$ for the Y. One obviously has the identity:

(1)
$$X_{1}Y_{1} + \dots + X_{p}Y_{p} + X_{p+1}Y_{p+1} + \dots + X_{p+q}Y_{p+q} = X_{1}Y'_{1} + \dots + X_{p}Y'_{p} + X'_{p+1}Y'_{p+1} + \dots + X'_{p+q}Y'_{p+q},$$

in which one supposes that the X and the X', as well as the Y and Y', are replaced with their expressions as functions of x_i and y_i , resp.

Choose the x_i to be the unique system of values that makes:

$$X_1 = 1$$
, $X_2 = X_3 = \ldots = X_p = X'_{p+1} = \ldots = X'_{p+q} = 0$.

For those values of the x_i , the forms $X_{p+1}, ..., X_{p+q}$ will take the numerical values $A_{p+1}, ..., A_{p+q}$. Upon substituting those values in the identity (1), one will have the identity with respect to the y_i :

$$Y'_{p+1} = A_{p+1} Y_{p+1} + \ldots + A_{p+q} Y_{p+q}$$
.

One will find analogous results for Y'_{p+2} , ..., Y'_{p+q} . Thus, the last q forms Y'_{p+1} , ..., Y'_{p+q} are expressed by linear and homogeneous functions of the last q old ones Y_{p+1} , ..., Y_{p+q} .

Now choose the x_i such that:

$$X_1 = 1$$
, $X_2 = X_3 = \ldots = X_p = X'_{p+1} = \ldots = X'_{p+q} = 0$,

while the $X_{p+1}, ..., X_{p+q}$ take certain numerical values $B_{p+1}, ..., B_{p+q}$, and the identity (1) will give the following identity with respect to the y_i :

$$Y'_1 = Y_1 + B_{p+1} Y_{p+1} + \ldots + B_{p+q} Y_{p+q}$$

One gets the same result for Y'_2 , ..., Y'_p .

The first *p* forms $Y'_1, ..., Y'_p$ are equal to the *p* old ones $Y_1, Y_2, ..., Y_p$, respectively, augmented by homogeneous linear functions of $Y_{p+1}, ..., Y_{p+q}$.

8. Adjoint systems. – Having recalled that, one has seen that if one considers:

$$L(u) = l_n \frac{d^n u}{dx^n} + l_{n-1} \frac{d^{n-1} u}{dx^{n-1}} + \dots + l_0 u,$$

and if M(u) denotes the adjoint expression then one will have the Lagrange identity:

$$v L(u) - u M(v) = \frac{d}{dx} [P(u,v)] ,$$

in which P(u, v) is a bilinear form with respect to the systems of variables:

$$u(x), u'(x), ..., u^{(n-1)}(x),$$

 $v(x), v'(x), ..., v^{(n-1)}(x).$

One can very easily construct P(u, v), and one finds that:

$$P(u, v) = u \left[l_1 v - \frac{d(l_2 v)}{dx} + \dots + (-1)^{n-1} \frac{d^{n-1}(l_n v)}{dx^{n-1}} \right]$$

+ $u' \left[l_2 v - \frac{d(l_3 v)}{dx} + \dots + (-1)^{n-2} \frac{d^{n-2}(l_n v)}{dx^{n-2}} \right]$
+ \dots
+ $u^{(n-1)} l_n v$.

The determinant of the form is written:

	•••	•••	$(-1)^{n-1}l_n$	
	•••	$(-1)^{n-2}l_n$	0	
		··	0	,
	·		•	
l_n	0		0	

in which the elements of the second diagonal are equal to $\pm l_n$ and the elements below them are all equal to zero.

That determinant is equal to $\pm (l_n)^n$ then, and since we suppose that the equation L(u) = r is regular in the interval in question (A, B), we suppose that $l_n \neq 0$ in that interval. Therefore, the preceding determinant is not zero. The form P(u, v) is ordinary for any value of x that is taken in the interval (A, B).

Green's formula then gives us:

$$\int_{a}^{b} [v L(u) - u M(v)] dx = [P(u, v)]_{a}^{b}$$

The quantity $[P(u,v)]_a^b$ is a bilinear form with respect to the following two series of 2n variables:

$$u(a), u'(a), ..., u^{(n-1)}(a), u(b), u'(b), ..., u^{(n-1)}(b),$$

 $v(a), v'(a), ..., v^{(n-1)}(a), v(b), v'(b), ..., v^{(n-1)}(b).$

That bilinear form has a particular character. In reality, it is the sum of two bilinear forms: The first one is with respect to the $u^{(i)}(a)$ and $v^{(i)}(a)$, while the second one is with respect to $u^{(i)}(b)$ and $v^{(i)}(b)$.

The determinant of that form will be:

$$\begin{vmatrix} \cdots & \cdots & \cdots & \mp l_n(a) & 0 & \cdots & \cdots & 0 \\ & \pm l_n(a) & 0 & \vdots & & \vdots \\ & \ddots & \ddots & 0 & \vdots & (\text{zeroes}) & \vdots \\ & \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ \frac{l_n(a) & 0 & \cdots & \cdots & 0}{0 & \cdots & \cdots & 0} & 0 & \cdots & \cdots & 0 \\ \vdots & & & & \ddots & & & \vdots & & & \ddots & \ddots \\ \vdots & & & & & & \ddots & & & \vdots \\ \vdots & & & & & & & \ddots & & \vdots \\ \vdots & & & & & & & \ddots & & \vdots \\ 0 & \cdots & \cdots & 0 & & l_n(b) & 0 & \cdots & 0 \end{vmatrix} = \pm [l_n(a)]^n [l_n(n)]^n .$$

Chapter II

From our hypothesis, it is $\neq 0$, so the form $[P(u,v)]_a^b$ will be ordinary. It will then be reducible to the canonical form.

Take 2n forms $U_1, ..., U_{2n}$ that are linear in $u(a), u'(a), ..., u^{(n-1)}(a), u(b), ..., u^{(n-1)}(b)$, and independent:

$$U_{1} = \alpha_{1} u(a) + \alpha_{1}' u'(a) + \dots + \alpha_{1}^{(n-1)} u^{(n-1)}(a) + \beta_{1} u(b) + \beta_{1}' u'(b) + \dots + \beta_{1}^{(n-1)} u^{(n-1)}(b), \dots \\ U_{2n} = \alpha_{2n} u(a) + \dots + \alpha_{2n}^{(n-1)} u^{(n-1)}(a) + \beta_{2n} u(b) + \dots + \beta_{2n}^{(n-1)} u^{(n-1)}(b).$$

The U_i are chosen to be independent, so that will imply 2n forms $V_1, ..., V_{2n}$ that are linear in $v(a), v'(a), ..., v^{(n-1)}(a), v(b), v'(b), ..., v^{(n-1)}(b)$ and linearly independent, so they are found determined by the choice of U_i and are such that one has:

$$[P(u,v)]_a^b = U_1 V_{2n} + U_2 V_{2n-1} + \ldots + U_{2n} V_1.$$

All of the above results from some propositions that relate to the canonical forms that we have already recalled.

Green's formula is then written:

(1)
$$\int_{a}^{b} [v L(u) - u M(v)] dx = U_{1} V_{2n} + U_{2} V_{2n-1} + \dots + U_{2n} V_{1}$$

in an infinitude of ways since one can choose the $U_1, ..., U_{2n}$ arbitrarily, on the condition that those forms are nonetheless independent.

It will then be easy to define the adjoint system to the system:

(2)
$$L(u) = 0,$$

 $U_i(u) = 0$ $(i = 1, 2, ..., m),$

in which the $U_i(u)$ have the usual form:

$$U_{i}(u) = \alpha_{i} u(a) + \alpha_{i}' u'(a) + \dots + \alpha_{i}^{(n-1)} u^{(n-1)}(a) + \beta_{i} u(b) + \dots + \beta_{i}^{(n-1)} u^{(n-1)}(b)$$

a, *b* are two points of the interval *A*, *B*. Of course, we suppose that the forms $U_1(u), ..., U_m(u)$ are independent relative to the variables $u(a), ..., u^{(n-1)}(a), u(b), ..., u^{(n-1)}(b)$. For that to be true, it is necessary that we should have $m \le 2n$.

Therefore, let *m* forms be given $U_1(u), \ldots, U_m(u)$. We add 2n - m forms $U_{m+1}(u), \ldots, U_{2n}(u)$ to them such that U_1, U_2, \ldots, U_{2n} are linearly independent. That choice of $U_{m+1}(u), \ldots, U_{2n}(u)$ is

possible in an infinitude of ways. With this choice of $U_1, U_2, ..., U_{2n}$, Green's formula (1) will give 2n other forms $V_1, V_2, ..., V_{2n}$ that are linear in $v(a), v'(a), ..., v^{(n-1)}(a), v(b), v'(b), ..., v^{(n-1)}(b)$. Choose the first 2n - m of those forms and define the differential system:

(3)
$$M(v) = 0, V_i(v) = 0 \quad (i = 1, 2, ..., 2n - m).$$

One calls it the *adjoint system to the system* (2).

The $V_1, V_2, ..., V_{2n-m}$ depend upon the $U_1, U_2, ..., U_{2n}$, so it would seem that the system (3) is not determined completely since a change in the $U_{m+1}, ..., U_{2n}$ would imply a change in the $V_1, ...,$ V_{2n-m} . That is nothing fundamental since if one replaces the $U_{m+1}, U_{m+2}, ..., U_{2n}$ with some other forms $U'_{m+1}, U'_{m+2}, ..., U'_{2n}$ then the $V'_1, ..., V'_{2n}$ that are determined by means of Green's formula will be coupled with the $V_1, ..., V_{2n}$ in a simple way that we have learned how to determine. Upon taking that into account in the canonical form that was adopted for $[P(u,v)]_a^b$, the order of the indices in V is reversed, and one sees that $V'_1, ..., V'_{2n-m}$ are given by forms that are linear in V_1 , $V_2, ..., V_{2n-m}$ such that:

$$V'_i = B_1 V_1 + B_2 V_2 + \ldots + B_{2n-m} V_{2n-m}$$
.

Those forms are independent in $V_1, V_2, ..., V_{2n-m}$ because they are independent in v(a), v'(a), ..., v(b), v'(b), ...

It will then result that the conditions $V_1 = 0, ..., V_{2n-m} = 0$ are equivalent to the conditions $V'_1 = 0, ..., V'_{2n-m} = 0$. The adjoint system is therefore essentially the same (³).

Finally, if one replaces the U_i in the proposed system (2) with some independent linear forms in $U_1(u), ..., U_m(u)$ then the system (2) will not change essentially. An argument that is entirely parallel to the one that was just made will show that the condition on the adjoint system $V_1 = 0$, ..., $V_{2n-m} = 0$ will not change essentially. Hence, the adjoint system is not altered.

Furthermore, the reciprocity between L(u) and M(v) and the symmetry of the right-hand side of Green's formula with respect to the two groups of variables U_i and V_i proves that the adjoint system to (3) is the system (2). There is then a reciprocity between a system and its adjoint.

Let us say a few words about the algebraic analogue.

The differential system corresponds to a homogeneous system:

$$V'_{2n-m+i} = V_{2n-m+i} + A'_{1}{}^{i}V_{1} + \dots + A'_{2n-m}{}^{i}V_{2n-m} .$$

i is an arbitrary integer from the sequence 1, 2, ..., *m*.

^{(&}lt;sup>3</sup>) The forms V_{2n-m+1} , ..., V_{2n} are less interesting than V_1 , ..., V_{2n-m} . Meanwhile, the effect on those forms of a change in U_{m+1} , ..., U_{2n} is simple: One has some new forms that are coupled with the old ones by formulas such as:

Chapter II

(2')
$$\begin{cases} a_{11}\,\xi_1 + \dots + a_{1N}\,\xi_N = 0, \\ \dots \\ a_{M1}\,\xi_1 + \dots + a_{MN}\,\xi_N = 0. \end{cases}$$

One will find that the adjoint system (3) corresponds to the homogeneous system (3') whose matrix of coefficients is simply the matrix of (2') in which one permutes the rows and columns:

(3')
$$\begin{cases} a_{11}\eta_1 + \dots + a_{M1}\eta_M = 0, \\ \dots \\ a_{1M}\eta_1 + \dots + a_{MN}\eta_M = 0. \end{cases}$$

The analogy is more satisfying here than the one that was pointed out for just the adjoint equation.

We shall not pursue that any further since in the ultimate proofs, we will not use those algebraic analogies, so we shall be content to point them out when they might suggest some new facts.

9. Some properties of adjoint systems. – To abbreviate, we shall consider only the very important case in which m = n. We have:

(1)
$$L(u) = 0, U_1(u) = 0, \dots, U_n(u) = 0, \dots, U_n(u$$

and its adjoint:

(2)
$$\begin{array}{c}
M(v) = 0, \\
V_1(v) = 0, \\
\dots \\
V_n(v) = 0.
\end{array}$$

We will have need for the following lemma:

Lemma:

Let $u_1, u_2, ..., u_k$ be linearly-independent of the system (1). As in the definition of the adjoint system, choose forms $U_{n+1}(u), ..., U_{2n}(u)$ such that $U_1, ..., U_n, U_{n+1}, ..., U_{2n}$ are independent. I then say that the k systems of constants:

$$U_{n+1}(u_1), ..., U_{2n}(u_1),$$

 $U_{n+1}(u_2), ..., U_{2n}(u_2),$ $U_{n+1}(u_k), ..., U_{2n}(u_k)$

are linearly independent, i.e., that it is impossible to find constants $c_1, c_2, ..., c_k$ that are not all zero and are such that one has:

$$c_1 U_{n+i}(u_1) + c_2 U_{n+i}(u_2) + \ldots + c_k U_{n+i}(u_k) = 0$$
 for $i = 1, 2, \ldots, n$.

Indeed, if one can find such constants then that would mean that one has:

$$U_{n+i} (c_1 u_1 + c_2 u_2 + \ldots + c_k u_k) = 0,$$

since the U operations are linear.

Now:

$$u=c_1\ u_1+\ldots+c_k\ u_k$$

is the solution to (1), and that solution will verify the 2n conditions:

$$U_1(u) = 0$$
, ..., $U_{2n}(u) = 0$.

The 2*n* forms $U_1, ..., U_{2n}$ are independent, so the equations $U_1 = ... = U_{2n} = 0$ admit just the solution:

$$u(a) = u'(a) = \dots = u^{(n-1)}(a) = u(b) = \dots = u^{(n-1)}(b) = 0$$

Hence, the solution:

$$u=c_1\ u_1+\ldots+c_k\ u_k$$

is identically zero, and since $c_1, ..., c_k$ are not all zero, that will imply that $u_1, ..., u_k$ are not linearly independent. The contradiction proves the lemma.

Having said that, we remark that in terms of algebraic systems, if one imagines a system:

```
a_{11} \xi_1 + \ldots + a_{1n} \xi_n = 0 ,

a_{n1} \xi_1 + \ldots + a_{nn} \xi_n = 0 ,

a_{11} \eta_1 + \ldots + a_{n1} \eta_n = 0 ,

\ldots

a_{1n} \eta_1 + \ldots + a_{nn} \eta_n = 0 ,
```

and its adjoint:

then the matrices in those two systems will be the same, up to a permutation of the rows and columns, but their rank will be the same. Since the number of variables is also the same, those two

Chapter II

systems have the same *index*, i.e., the same number of independent solutions. That suggests the following theorem:

Theorem:

The index of a homogeneous differential system (1) is equal to that of its adjoint (1).

The index is also the number of linearly-independent solutions of the system here.

If k is the index of (1) then it will suffice to prove that the index of (2) is $\ge k$, and the theorem will then result from the reciprocity between (1) and (2).

We appeal to Green's formula:

$$\int_{a}^{b} [v L(u) - u M(v)] dx = U_1 V_{2n} + U_2 V_{2n-1} + \dots + U_{2n} V_1$$

Let *u* be a solution to the system (1), and let $z_1, z_2, ..., z_n$ be a fundamental system of solutions to M(v) = 0. Upon substituting *u* and z_i for *u* and *v* in Green's identity, one will have:

(3)
$$0 = U_{n+1}(u) V_{2n}(z_i) + U_{n+2}(u) V_{n-1}(z_i) + \ldots + U_{2n}(u) V_1(z_i) \qquad (i = 1, 2, \ldots, n).$$

That is a system of *n* homogeneous equations in $U_{n+1}(u)$, ..., $U_{2n}(u)$. It is satisfied by the values $U_{n+1}(u_i)$, ..., $U_{2n}(u_i)$, in which u_i is an arbitrary solution to (1). Let u_1 , ..., u_k be *k* linearly-independent solutions to (1), which has the index *k*.

From the lemma, the *k* systems of constants:

$$U_{n+1}(u_i), \ldots, U_{2n}(u_i)$$
 (*i* = 1, 2, ..., *k*)

are linearly independent.

The system (3) will then have at least k linearly-independent solutions. Its rank is then $\leq n - k$. Now, its rank is that of the matrix:

$$\begin{bmatrix} V_n(z_1) & \cdots & V_1(z_1) \\ \vdots & \ddots & \vdots \\ V_n(z_n) & \cdots & V_1(z_n) \end{bmatrix}.$$

Now, that matrix is what one must construct in order to find the index of the system (2). One then concludes that the system (1) will have at least *k* independent solutions, i.e., that its index is $\geq k$.

If one now considers an inhomogeneous system then the analogy with the algebraic systems will suggest some further results.

Let the system be:

$$a_{11} \xi_1 + \ldots + a_{1n} \xi_n = b_1,$$

.....
 $a_{n1} \xi_1 + \ldots + a_{nn} \xi_n = b_n.$

The necessary and sufficient condition for it to be compatible is expressed simply with the aid of Rouché's characteristic determinants. Without introducing those determinants explicitly, one can also say that the necessary and sufficient condition for compatibility is that any homogeneous relation that couples the a in the same column of the matrix (a_{ik}) must also couple the b, i.e., that any solution to the equations:

$$\eta_1 a_{11} + \eta_2 a_{21} + \ldots + \eta_n a_{n1} = 0 ,$$

$$\eta_1 a_{1n} + \eta_2 a_{2n} + \ldots + \eta_n a_{nn} = 0 ,$$

in which η_1, \ldots, η_n are the unknowns, must verify the relation:

$$b_1 \eta_1 + b_2 \eta_2 + \ldots + b_n \eta_n = 0$$

One sees how naturally one introduces the adjoint system to the homogeneous system that corresponds to the given system.

We shall infer a compatibility condition that is analogous to Green's formula for the case of an inhomogeneous differential system.

If the given system is:

(4)
$$L(u) = r, U_i(u) = \gamma_i \quad (i = 1, 2, ..., n)$$

then the homogeneous adjoint system will likewise be, by definition:

(5)
$$M(v) = 0, V_i(v) = 0 \quad (i = 1, 2, ..., n).$$

Suppose that (4) has a solution u, and let v be an arbitrary solution to the adjoint system. If one substitutes those two functions in Green's identity then it will become:

(6)
$$\int_{a}^{b} r v \, dx = \gamma_{1} \, V_{2n} \, (v) + \ldots + \gamma_{n} \, V_{n+1} \, (v)$$

Hence, any solution of the adjoint system must verify the relation (6). It is a necessary condition for the compatibility of the system (4). One can show that it is also sufficient with the aid of Green's formula, but we shall not dwell upon that point $(^4)$.

^{(&}lt;sup>4</sup>) See my article that was cited at the beginning of this chapter, in which one will also find some results for differential systems in which $m \neq n$ that are analogous to the ones that we pointed out here for the case in which m = n.

10. Quasi-differential equations. Second-order equations. – In all of the preceding, it was supposed that the coefficients in the expression:

$$L(u) = l_n \frac{d^n u}{dx^n} + \dots + l_0 u$$

have continuous derivatives up to an order that is indicated by their index in order for one to be able to construct the expression:

$$M(v) = (-1)^{n} \frac{d^{n}(l_{n} v)}{dx^{n}} + \dots + l_{0} v.$$

Can we not simply confine ourselves to continuous $l_0, l_1, ..., l_n$, even if it means enlarging the definition of the adjoint expression?

For the generalization that we have in mind, it would be simpler to directly consider a quasidifferential expression instead of L(u) and to develop the theory of the adjoint expression for such an expression. The case of a differential expression L(u) in which l_0, \ldots, l_n are simply continuous will be a special case of the preceding theory.

We shall only sketch out that theory for a third-order quasi-differential expression:

$$L(u) = a_3 \frac{d}{dx} \left\{ a_2 \frac{d}{dx} \left[a_1 \frac{d}{dx} (a_0 u) + b_1 u \right] + b_2 u \right\} + l_2 \frac{d}{dx} \left[a_1 \frac{d}{dx} (a_0 u) + b_1 u \right] + l_1 \frac{d}{dx} (a_0 u) + q u,$$

if $a_0 a_1 a_2 a_3 \neq 0$.

We shall call the expression:

$$M(v) = -a_0 \frac{d}{dx} \left\{ a_1 \frac{d}{dx} \left[a_2 \frac{d}{dx} (a_3 v) - l_2 v \right] + l_1 v \right\} + b_1 \frac{d}{dx} \left[a_2 \frac{d}{dx} (a_3 v) - l_2 v \right] - b_2 \frac{d}{dx} (a_3 v) + q v$$

its adjoint.

Upon constructing v L(u) - u M(v), one sees that this expression is equal to $\frac{d}{dx}P(u,v)$, in which P(u, v) is a linear function in:

$$u, \quad \frac{d}{dx}(a_0 u), \quad \frac{d}{dx} \left[a_1 \frac{d}{dx}(a_0 u) + b_1 u \right],$$
$$v, \quad \frac{d}{dx}(a_3 v), \quad \frac{d}{dx} \left[a_2 \frac{d}{dx}(a_3 v) - l_2 v \right].$$

Those six variables are the ones that correspond to the:
in the ordinary differential equations in the existence theorem for quasi-differential equations.

The formulas that one obtains are more complicated, but do not differ essentially from the ones that are obtained in the usual theory of the adjoint. With some small changes, one can likewise formulate a theory of the adjoint system that is parallel to the one that was already presented. One would then see that it is not essential for one to suppose that the $l_0, l_1, ..., l_n$ in L(u) have derivatives up to certain orders.

That raises a question in regard to the adjoint equation. We shall study it for only the secondorder equations. Let:

 $L(u) = l_2 \frac{d^2 u}{dx^2} + l_2 \frac{du}{dx} + l_0 u = r \qquad [l_2 \neq 0 \text{ in } (A, B)].$

The adjoint of L(u) is:

$$M(v) = l_2 \frac{d^2 v}{dx^2} + (2l'_2 - l_1) \frac{dv}{dx} + (l''_2 - l'_1 + l_0) v.$$

Under what conditions is the expression L(u) identical to M(v)? One sees immediately that it is necessary and sufficient that:

 $l'_{2} = l_{1}$.

Therefore, an arbitrary second-order differential expression is not always its own adjoint, but as one sees, one can make it so by a simple calculation upon multiplying it by the factor:

$$\frac{1}{l_2} \exp\left(\int \frac{l_1}{l_2} dx\right).$$

Hence, it is no loss of generality for a second-order equation to suppose that its left-hand side is self-adjoint and write it in the form that Sturm adopted:

$$\frac{d}{dx}\left(K\frac{du}{dx}\right) - Gu = R \; .$$

That form is even more general than the ordinary form because K can be supposed to be simply continuous, but not differentiable, so u'' will not exist, but u and u' will always exist, and that is generally all that we shall need. It is then a very simple case of a quasi-differential equation.

Thus the adjoint of:

$$L(u) = \frac{d}{dx}(Ku') - Gu$$

is:

$$M(v) = \frac{d}{dx}(Kv') - Gv,$$

so

$$v L(u) - u M(v) = \frac{d}{dx} [K(v u' - u v')]$$

is an identity that will be valid even when K has no derivative.

Now take a homogeneous system:

$$L(u) = \frac{d}{dx}(Ku') - Gu = 0,$$

$$U_1 = a_1 u(a) + a_2 u(b) + a_3 u'(a) + a_4 u'(b) = 0,$$

$$U_2 = b_1 u(a) + b_2 u(b) + b_3 u'(a) + b_4 u'(b) = 0.$$

The search for the adjoint system proceeds with no difficulty. Of course, one supposes that U_1 and U_2 are independent, i.e., the rank-2 matrix:

$$\begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \end{pmatrix}$$

Set:

$$d_{ij}=a_i b_j-a_j b_i.$$

There are several cases to distinguish:

1. $d_{12} \neq 0$: One takes U_3 and U_4 arbitrarily, but such that U_1 , U_2 , U_3 , U_4 are independent.

One can obviously take $U_3 = u'(a)$, $U_4 = u'(b)$ because the determinant of the four forms will then be $d_{12} \neq 0$.

One then, in turn, forms:

$$\int_{a}^{b} [v L(u) - u M(v)] dx = [K(v u' - u v')]_{a}^{b}$$

and reduces the form on the right-hand side to:

 $U_1 V_4 + U_2 V_3 + U_3 V_2 + U_4 V_1$.

One will then find that:

$$V_{1} = \frac{1}{d_{12}} \left[-K(b) d_{12} v(b) + K(a) d_{24} v'(a) + K(b) d_{14} v'(b) \right],$$

$$V_{2} = \frac{1}{d_{12}} \left[-K(a) d_{12} v(a) + K(a) d_{23} v'(a) + K(b) d_{13} v'(b) \right] .$$

There are analogous expressions for V_3 and V_4 , although they are simpler. The adjoint system is:

$$M(v) = \frac{d}{dx}(Kv') - Gv = 0,$$

$$V_1(v) = 0,$$

$$V_2(v) = 0.$$

In order for the proposed system to be its own adjoint, it is necessary and sufficient that the conditions $V_1(v) = 0$, $V_2(v) = 0$ are *essentially the same as* $U_1 = 0$, $U_2 = 0$, or in other words, that V_1 and V_2 are linear combinations of U_1 and U_2 . Since v(a) does not enter into V_1 , upon eliminating u(a) from U_1 and U_2 and comparing them to V_1 , one will get a first necessary condition. Upon eliminating u(b) from U_1 and U_2 and comparing them to V_2 , one will get a second necessary condition, and when those two conditions are taken together, that will give necessary and sufficient conditions. Moreover, they will reduce to the single condition:

(1)
$$d_{24} K(a) = d_{13} K(b)$$
.

2. If $d_{12} = 0$ then there will be several cases to consider. However, one will always recover the preceding condition for the proposed system to be its own adjoint.

Some simple examples will illustrate the circumstances that might present themselves in the process of determining the index of a system.

For example, let the equation be $\frac{d^2u}{dx^2} + u = 0$. Its general integral is $u = A \cos x + B \sin x$. If one takes the two independent conditions:

$$u(0) - u(2\pi) = 0, u'(0) - u'(2\pi) = 0$$

then one will have introduced no essential restriction on the solution. The index of the system is 2, like that of the equation.

If one takes:

$$u(0) - u(2\pi) = 0,$$

$$u'(0) + u'(2\pi) = 0$$

then the first one will introduce no restriction, and one will find that the index is 1. The solution is $A \cos x$.

Finally, if one takes:

$$u(0) + u(2\pi) = 0$$

$$u'(0) + u'(2\pi) = 0$$

then the only solution will be u = 0. Thus, the index is zero.

11. Characteristic numbers. – In conclusion, consider the case that will often come up in what follows in which the coefficients of the system:

$$L(u) = 0,$$

 $U_i(u) = 0$ $(i = 1, 2, ..., n)$

depend upon a parameter λ . The first problem that one poses is that of determining the values of λ for which that system is compatible, i.e., there are solutions that are not identically zero. That problem is less precise than that of determining the index.

If $u_1, ..., u_n$ is a fundamental system of solutions to L(u) = 0 then the compatibility condition for the system will be:

(1)
$$\begin{vmatrix} U_1(u_1) & \cdots & U_1(u_n) \\ \vdots & \ddots & \vdots \\ U_n(u_1) & \cdots & U_n(u_n) \end{vmatrix} = 0.$$

The coefficients of U_i and the solutions u_i are functions of λ , so one will then have an equation in λ :

$$F(\lambda) = 0$$

that we call the *characteristic equation of the differential system*. Its roots $\lambda_1, \lambda_2, \ldots$ will be the *characteristic numbers*.

We suppose that the coefficients of the system are continuous in (x, λ) and analytic in λ in a certain Weierstrass domain \mathcal{D} . In this case, $F(\lambda)$ will be an analytic function of λ . Its zeroes will then be isolated in that domain (but possibly with some limit points on the boundary of the domain), unless of course $F(\lambda)$ is identically zero.

Any characteristic number λ_i will make the system compatible. The system will then have a certain index $k_i > 0$, whereas for any non-characteristic number, one will have $k_i = 0$. k_i will be called the *index* of λ_i .

A characteristic number λ_i has a certain order of multiplicity m_i if one considers it to be a zero of $F(\lambda) = 0$, and that order of multiplicity will not depend upon the fundamental system that one appeals to in order to define the characteristic equation since a change of that fundamental system will have the effect of multiplying the left-hand side of the characteristic equation by a function of λ with no zero, as one easily confirms. The m_i can differ from k_i , but one will always have:

$$m_i \geq k_i$$
.

(One has seen that $k_i \le n$, moreover.) In effect, let *k* be the index of a characteristic number λ . One can show that this number annuls:

$$F'(\lambda)$$
, ..., $F^{(k-1)}(\lambda)$.

Each of those derivatives is a sum of determinants, each of which includes n - k + 1 unaltered rows from the determinant $F(\lambda)$, while the other k - 1 rows might have been differentiated.

If one develops those determinants by using the Laplace formula, while taking the minors that are composed of those n - k + 1 rows, then one will see that all of those minors are zero since they are determinants of order n - k + 1 that come from the determinant $F(\lambda)$, which has rank n - k, from the definition of the index k. One then concludes that:

$$F'(\lambda) = 0, ..., F^{(k-1)}(\lambda) = 0.$$

Hence, $m \ge k$.

Examples:

1. Let:

$$\frac{d^2u}{dx^2} + \lambda^2 u = 0$$

Its general solution is:

$$u = A \cos \lambda x + B \sin \lambda x .$$

Take the conditions:

$$u(-1) - u(+1) = 0$$
,
 $u'(-1) + u'(+1) = 0$.

For any λ , $A \cos \lambda x$ is even, and its odd derivative will be a solution to the system. Therefore, any value of λ will be characteristic. Indeed, one finds that:

 $F(\lambda) \equiv 0$.

$$\frac{d^2u}{dx^2} + \lambda^2 u = 0,$$

$$u(0) = 0,$$

$$u'(0) = 0$$

then, from the existence theorem, the only solution for any λ will be zero. There will be no characteristic number, and one will effectively find that:

$$F(\lambda) \equiv 1$$
.

3.

$$\frac{d^2u}{dx^2} + \lambda^2 u = 0 ,$$

$$u(0) = 0 ,$$

$$u(\pi) = 0 .$$

The characteristic numbers here are $\lambda = \pm 1, \pm 2, ...$, and one will easily see that their orders of multiplicity and their indices are all equal to 1.

CHAPTER III

REAL SOLUTIONS AND THEIR ZEROES IN THE SIMPLEST CASES (⁵)

12. Solutions for one invariable equation. – In this chapter, we shall study the results that Sturm gave in his first article, and certain extensions of those results. First of all, imagine the homogeneous differential equation:

$$\frac{d}{dx}(Ku') - Gu = 0$$

If one supposes that its coefficients are (real or complex) continuous functions of the variable x in the closed interval $A \le x \le B$ then it will be easy to see that *any solution u, whether real or complex, of that equation cannot have an infinitude of zeroes in* (A, B) *without being identically zero.* That is because those zeroes will admit at least one limit point c in $A \le x \le B$. Since the function u is continuous at c :

$$u\left(c\right)=0$$

One also knows that u'(c) exists. If one constructs it while considering the value of u at c and at a zero of u(x) that tends to c then one can confirm that:

$$u'(c)=0.$$

From our existence theorem, u(x) will be identically zero then.

That extends to homogeneous linear equations of order *n* with coefficients that are continuous in $A \le x \le B$. If the zeroes of a solution (⁶) define an infinite set in (*A*, *B*) then they will have at least one limit point *c*. At that point, one has:

$$u(c) = 0$$
, $u'(c) = 0$.

However, since u(x) is continuous, along with its derivative, there will be at least one zero of u' between two zeroes of u. Thus, u'(x) has an infinitude of zeroes that admit c for a limit point. Hence:

^{(&}lt;sup>5</sup>) STURM, J. de Math. pures et appl. **1** (1836), pp. 106. – PICARD, *Traité d'Analyse*, t. III, Chap. VI. – BÔCHER, Trans. Amer. Math. Soc. **1** (1900), pp. 414; *ibid.*, **2** (1901), pp. 428; *ibid.* **3** (1902), pp. 196. – PICONE, Ann. di Reale Scuola Normale sup. di Pisa **11** (1909), pp. 3.

 $^(^{6})$ We supposed that this solution is real in the proof. It can be applied to a complex solution with a slight modification.

$$u''(c)=0.$$

Step-by-step, the argument will show that:

$$u(c) = u'(c) = \dots = u^{(n-1)}(c) = 0$$

Therefore, u is identically zero.

The property of the function u that it has a finite number of zeroes in (A, B) does not extend to either the derivatives u'(c), u''(c), ... or the solutions to the inhomogeneous equations, as one can see in some examples that are easy to construct.

Nonetheless, there exist certain combinations of u and some of its successive derivatives u', u'', ... that can have only a finite number of zeroes.

For example, one proves (⁷) that the Wronskian of k independent solutions $u_1, ..., u_k$ of homogeneous linear equation of order $n \ge k$ cannot have an infinite number of zeroes in (A, B).

In the case of a second-order equation:

(1)
$$\frac{d}{dx}(Ku') - Gu = 0,$$

consider some functions of the form:

$$\Phi = \varphi_1 u - \varphi_2 K u' = 0 ,$$

in which φ_1 , φ_2 are two given functions of x. Under what conditions will Φ have only a finite number of zeroes? We suppose, to simplify, that φ_1 and φ_2 have continuous derivatives. The same thing will be true for Φ .

One has:

$$\Phi' = (\varphi_1' - \varphi_2 G)u + (\varphi_1 - \varphi_2' K)u',$$

upon taking (1) into account.

If Φ has an infinitude of zeroes in $(A \le x \le B)$ then those zeroes will have at least one limit point *c* in (A, B). Hence:

(2) $\begin{cases} \Phi(c) = 0, \\ \Phi'(c) = 0. \end{cases}$

Now, if *u* is not identically zero then one cannot have both:

$$u(c) = 0$$
, $u'(c) = 0$

at *c*.

The two equations (2) then must have a zero determinant at u(c) and u'(c). That zero determinant will then give the equality:

^{(&}lt;sup>7</sup>) See Bull. Amer. Math. Soc. **8** (1901), pp. 53.

(3)
$$K\left(\varphi_{1}'\varphi_{2}-\varphi_{1}\varphi_{2}'-\frac{\varphi_{1}^{2}}{K}-G\varphi_{1}^{2}\right)=0 \quad \text{for} \quad x=c.$$

K is supposed to be $\neq 0$ in $(A \le x \le B)$ in order that (1) should have continuous coefficients. Hence, one can suppress *K* in (3) and arrive at the following conclusion:

If φ_1 , φ_2 are functions of x with continuous derivatives such that:

(4)
$$\{\varphi_1, \varphi_2\} = \varphi_1' \varphi_2 - \varphi_1 \varphi_2' - \frac{\varphi_1^2}{K} - G \varphi_1^2$$

is 0 in the closed interval (A, B) then one will be assured that Φ has only a finite number of zeroes in (A, B) as long as u is not identically zero in (A, B).

Sturm's first theorem. – Consider a second-order homogeneous linear differential equation with real coefficients for which the real variable x belongs to the interval a, b.

If it admits a real solution u_1 that is annulled at least twice in a, b, and x_1 , x_2 are two consecutive zeroes of u_1 then any other real solution u_2 that is independent of u_2 will be annulled once and only once between x_1 and x_2 .

It obviously suffices to show that u_2 has a zero between x_1 and x_2 because if there were two then u_1 would have a third zero between x_1 and x_2 that would be found between the previous two zeroes of u_2 . Now, u_2 cannot be annulled at x_1 or x_2 because u_2 would then be the product of u_1 with a constant. If u_2 were everywhere non-zero between x_1 and x_2 then u_1 / u_2 would be a continuous function in $x_1 \le x \le x_2$ that is annulled at x_1 and x_2 , so it would have a zero derivative at some point between x_1 and x_2 .

That derivative is:

$$\frac{u_2 \, u_1' - u_1 \, u_2'}{u_2^2} \, .$$

Up to the factor $1/u_2^2$, it is the Wronskian of the two independent solutions u_1 , u_2 , and one knows that it is non-zero in (a, b). Hence, we have a contradiction, and as a result u_2 will have at least one zero *between* x_1 and x_2 .

We have seen that it cannot have more than one. We can summarize that by saying that *the* zeroes of the independent real solutions to a second-order homogeneous linear equation with real coefficients are separate from each other.

We cite the example of the equation:

$$\frac{d^2u}{dx^2} + u = 0 ,$$

for which:

$$u_1 = \cos x , \qquad u_2 = \sin x$$

are two independent solutions.

The zeroes of u_1 are separate from those of u_2 , like those of any solution $A \cos x + B \sin x$ ($B \neq 0$) that is independent of u_1 .

Imagine the imaginary solution $\cos x + i \sin x = e^{ix}$ to that equation. It has no zero in the interval *a*, *b* of the real variable *x*. If *x* varies from *a* to *b* then the point $u = e^{ix}$ in the plane of the complex variable will rotate from e^{ia} to e^{ib} on the circle of radius 1 that is described around the origin as its center. The real part and the imaginary part of e^{ix} , which are solutions, oscillate between -1 and +1 in a known way.

In order to see that this special case is typical, imagine the general case:

$$\frac{d}{dx}(Ku')-Gu=0.$$

If u_1 and u_2 are two arbitrary real solutions then $u = u_1 + i u_2$ will represent the most-general complex solution.

If u_1 , u_2 are two real independent solutions then u will not be the product of a real solution with a complex constant. We call that solution *essentially imaginary*.

Consider the point whose coordinates are u_1 , u_2 . It represents $u = u_1 + i u_2$, so u_1 / u_2 will be the angular coefficient of ou.

Now:

$$\frac{d}{dx}\left(\frac{u_2}{u_1}\right) = \frac{u_2 u_1' - u_1 u_2'}{u_2^2}$$

Since the Wronskian always has the same sign, since it is never zero, the vector ou will always rotate in the same sense around o. (If u_1 and u_2 are proportional then u will oscillate along a line that passes through o. That will be a limiting case of the preceding one where the trajectory of u is flattened into a line that passes through o.)

Take two essentially-imaginary solutions:

$$u = u_1 + i u_2$$
, $v = u_3 + i u_4$

in which u_1 , u_2 , u_3 , u_4 are real solutions, while u_1 and u_2 are linearly independent, as well as u_3 and u_4 .

Hence, one has:

$$u_3 = \alpha u_1 + \beta u_2$$
, $u_4 = \gamma u_1 + \delta u_2$ with $\begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix} \neq 0$,

in which α , β , γ , δ are real, in addition.

In the plane of the complex variable u, one passes from u to v by the preceding linear substitution, which bears the name of an *affinity*. That transformation changes lines into lines and multiplies all areas by the same quantity $(\alpha \delta - \beta \gamma)$. In particular, a line that issues from the origin is changed into a line that issues from another origin o. If the radius vector ou makes a certain number of circuits around o then the radius vector ov will make an equal number of them. More

precisely, if one considers x to be like time, and one looks for the areal velocity with respect to o of the moving body that the solution $u_1 + i u_2$ represents then one will see that this velocity is:

$$\frac{ds}{dx}=u_1u_2'-u_2u_1',$$

which is the Wronskian of the two functions u_1 , u_2 .

When one passes from the solution u to the solution v, that Wronskian will be simply multiplied by $\alpha\delta - \beta\gamma$, which should not be surprising since areas are always multiplied by that number. In particular, if u makes less than a half-turn around o when x describes the interval (A, B) then v will make less than a half-turn, and it will then result that any real solution in (A, B) (that is represented by the projection of the vector ou onto ox or oy) cannot be annulled more than once. The differential equation is then said to be *non-oscillatory* in the interval (A, B). We shall ultimately return to the question of how to recognize whether a given equation is oscillatory or not in a given interval.

With those brief indications about complex solutions, we now return to the real solutions of an equation with real coefficients. One then sees that if one considers $\Phi = \varphi_1 u - \varphi_2 K u'$, in which φ_1 , φ_2 have continuous derivatives such that:

$$\{\varphi_1, \varphi_2\} = \varphi_1' \varphi_2 - \varphi_1 \varphi_2' - \frac{\varphi_1^2}{K} - G \varphi_1^2$$

is non-zero in (A, B), in addition, then that expression for Φ will have only a limited number of zeroes in (A, B).

Take two such functions Φ_1 and Φ_2 such that: The first one is constructed from one solution u_1 , while the second one is constructed from a solution u_2 that is independent of u_1 (u_1 , u_2 , and φ_1 , φ_2 are real):

$$\Phi_1 = \varphi_1 \, u_1 - \varphi_2 \, K \, u_1', \qquad \Phi_2 = \varphi_1 \, u_2 - \varphi_2 \, K \, u_2', \quad \{\varphi_1, \varphi_2\} \neq 0 \, .$$

 Φ_1 and Φ_2 have only a limited number of zeroes in (A, B). Under those conditions, *Sturm's first* theorem on the zeroes of two independent solutions u_1 , u_2 will extend to the zeroes of Φ_1 and Φ_2 . There is one and only one zero of Φ_2 between two consecutive zeroes of Φ_1 .

The proof that was given for u_1 and u_2 applies here with almost no changes.

One finds, in the same way, that if *u* is an essentially-imaginary solution of (1) then the vector that represents the quantity $\varphi_1 u - \varphi_2 K u'$ will always rotate in the same sense when *x* increases from *A* to *B*, provided that $\{\varphi_1, \varphi_2\} \neq 0$.

One can pose a further question regarding those expressions Φ :

Take a real solution of:

$$\frac{d}{dx}(Ku') - Gu = 0$$

that is not identically zero. Take four real functions:

 $\varphi_1\,,\, \varphi_2\,, \qquad \psi_1\,,\, \psi_2\,,$

and construct:

$$\Phi = \varphi_1 u - \varphi_2 K u', \qquad \Psi = \psi_1 u - \psi_2 K u'.$$

Can one deduce any results concerning the zeroes of Φ and Ψ ? First of all, suppose that there are a finite number of those zeroes, and in order for that to be true, suppose that $\{\varphi_1, \varphi_2\}$ and $\{\psi_1, \psi_2\}$ are non-zero in (A, B).

One then notes that Ku' and u, which are particular forms of Φ and Ψ , verify a homogeneous Riccati equation, i.e., an equation of the form:

$$\omega_1 \, \omega_2' - \omega_2 \, \omega_1' = A \, \omega_1^2 + B \, \omega_1 \, \omega_2 + C \, \omega_2^2,$$

so one can think that the same thing is true for Φ and Ψ , and indeed, one will have:

(5)
$$\begin{cases} \Phi(\psi_1 u - \psi_2 K u') - \Psi(\varphi_1 u - \varphi_2 K u') = 0, \\ \text{or rather:} \\ (\psi_1 \Phi - \varphi_1 \Psi) u - (\psi_2 \Phi - \varphi_2 \Psi) K u' = 0 \end{cases}$$

identically.

Upon differentiating that and taking the equation that *u* verifies into account, one will have:

(6)
$$(\psi_1 \Phi' - \varphi_1 \Psi' + \psi_1' \Phi - \varphi_1' \Psi - G\psi_2 \Phi + G\varphi_2 \Psi) u + [\psi_1 \Phi - \varphi_1 \Psi - (\psi_2' \Phi - \varphi_2' \Psi + \psi_2 \Phi' - \varphi_2 \Psi') K] u = 0,$$

so (5) and (6) will be two homogeneous linear equations in u and u'. If u is not identically zero then one will not have both u = 0, u' = 0 at a point. Hence, the determinant of those two equations will be identically zero. That will produce a *homogeneous Riccati equation for* Φ and Ψ :

(7)

$$(\varphi_{1}\psi_{2}-\varphi_{2}\psi_{1})(\Phi'\Psi-\Phi\Psi')+\{\psi_{1}\psi_{2}\}\Phi^{2}$$

$$-[\varphi_{2}\psi_{1}'-\varphi_{2}'\psi_{1}+\varphi_{1}'\psi_{2}+2\frac{\varphi_{1}\psi_{1}}{K}-2G\varphi_{2}\psi_{2}]\Phi\Psi+\{\varphi_{1}\varphi_{2}\}\Psi^{2}=0.$$

One sees that a necessary condition for that equation to be regular in (A, B) is that:

$$\varphi_1 \psi_2 - \varphi_2 \psi_1 \neq 0$$
 in (A, B) .

If that condition is realized then Φ and Ψ cannot be annulled at the same point because $\varphi_1 \psi_2 - \varphi_2 \psi_1$ is the determinant of the two equations in u and Ku', namely, $\Phi = 0$, $\Psi = 0$, so one would have u = 0, u' = 0 at that point, which is absurd.

Before passing to the study of zeroes of Φ and Ψ , we point out that at a point where Φ is annulled, but Ψ is not, we will have, from (7), that:

(I)
$$\Phi' = -\frac{\{\varphi_1 \varphi_2\}}{\varphi_1 \psi_2 - \varphi_2 \psi_1} \Psi,$$

and similarly, at a point where $\Psi = 0$, we will have:

(II)
$$\Psi' = \frac{\{\psi_1\psi_2\}}{\varphi_1\psi_2 - \varphi_2\psi_1}\Phi.$$

Having said that, one can show that: There is one and only one zero of Ψ between two consecutive zeroes of Φ , and conversely.

Indeed, first of all, Φ cannot be annulled more than once in any segment of (A, B) where $\Psi \neq 0$. Indeed, if Φ is annulled at least twice at two consecutive points x_1, x_2 then from (I), Φ will have the same sign since neither Ψ nor $\varphi_1 \psi_2 - \varphi_2 \psi_1$ will be annulled. That is obviously impossible.

Conversely, Ψ cannot be annulled more than once in any segment where Φ remains non-zero.

Now suppose that Φ admits several zeroes in (A, B), and let x_1, x_2 be two such consecutive zeroes. Ψ is annulled at least once between x_1 and x_2 because in the contrary case, the interval x_1 , x_2 will be found in a somewhat-larger interval where Ψ is non-zero, and Φ cannot have two zeroes in that interval. Therefore, Ψ has at least one zero *between* x_1 and x_2 , and from a known argument, it will have only one.

Therefore, there exists a case in which the zeroes of Φ *,* Ψ *in* (*A*, *B*) *are separate:*

$$\varphi_1 \ \psi_2 - \varphi_2 \ \psi_1 \neq 0 , \quad \{\varphi_1 \ \varphi_2\} \neq 0 , \quad \{\psi_1 \ \psi_2\} \neq 0 .$$

That theorem will be meaningless if neither Φ nor Ψ has more than one zero in (A, B). Here is a case in which that situation is produced:

Theorem:

If one combines the conditions that:

$$\varphi_1 \ \psi_2 - \varphi_2 \ \psi_1 \neq 0 , \quad \{\varphi_1 \ \varphi_2\} \neq 0 , \quad \{\psi_1 \ \psi_2\} \neq 0$$

in (A, B) with the condition that $\{\varphi_1 \ \varphi_2\}$ and $\{\psi_1 \ \psi_2\}$ have opposite signs then neither Φ nor Ψ can be annulled more than once in (A, B), and furthermore, if one of those functions is annulled once then the other one cannot be annulled.

In other words, consider the product $\Phi \Psi$. Its zeroes are those of Φ and Ψ . From the theorem that was proved above, two consecutive zeroes of that product are, in one case, a zero of Φ , and in

the other, a zero of Ψ . The present theorem says that under the preceding conditions, *the product* $\Phi \Psi$ has only one zero in (A, B).

Indeed, if it has several of them then consider two consecutive zeroes, one of which belongs to Φ and the other to Ψ . We shall show that the derivative:

$$\frac{d}{dx}(\Phi\Psi) = \Phi'\Psi + \Phi\Psi'$$

has the same sign at those points. That contradiction will prove our theorem. For the first zero, from (I), the expression $\frac{d}{dx}(\Phi \Psi)$ will take the value:

$$\frac{-\left\{\varphi_{1}\,\varphi_{2}\right\}}{\varphi_{1}\,\psi_{2}-\varphi_{2}\,\psi_{1}}\Psi^{2}$$

,

and as for the consecutive zero, from (II), its value is:

$$\frac{\{\psi_1\psi_2\}}{\varphi_1\psi_2-\varphi_2\psi_1}\Phi^2.$$

Those two values have the same sign since:

$$\{\varphi_1 \ \varphi_2\} \times \{\psi_1 \ \psi_2\} < 0$$
.
Q. E. D.

Application. – If one takes $\psi_1 = 1$, $\psi_2 = 0$ then one will have $\Psi = u$, and one will obtain the following result:

If there exist a function φ with continuous derivative such that:

$$\varphi' + \frac{\varphi^2}{K} - G < 0$$

then neither a solution u of the equation:

$$\frac{d}{dx}(Ku') - Gu = 0, \qquad \text{in which} \qquad K > 0,$$

that is not identically zero nor the function $\varphi u - K u'$ can have more than one zero in (A, B).

The equation is therefore non-oscillatory.

Upon being given a particular form to φ , we will have various sufficient conditions for an equation to be non-oscillatory. We cite the condition G > 0, which we will obtain by setting $\varphi = 0$, and which is well known. Later on, we shall rediscover those conditions of non-oscillation.

13. Effect produced on the solutions by a change of coefficients in the equation. – Having thus extended the properties of the zeroes of real solutions in various directions, we shall return to those zeroes themselves in order to study the effect that is produced on them by a change of the functions G or K.

First leave K unchanged. Any reduction in G will then increase the rapidity of the oscillation of the solutions to the equation. The proof will clarify the sense of that statement.

Consider u_1 , which is a solution to:

$$\frac{d}{dx}(Ku_1')-G_1u_1=0,$$

and u_2 , which is a solution to:

$$\frac{d}{dx}(Ku_2') - G_2u_2 = 0$$

 G_2 is supposed to be $\langle G_1$, and neither u_1 nor u_2 is identically zero. Upon combining the preceding two equations, multiplied by u_2 and $-u_1$, resp., one will have:

$$\frac{d}{dx}[K(u_1'u_2-u_1u_2')]=(G_1-G_2)u_1u_2.$$

Hence, upon integration:

(1)
$$[K(u_1'u_2 - u_1u_2')]_{x_1}^{x_2} = \int_{x_1}^{x_2} (G_1 - G_2)u_1u_2 dx$$

That formula, which was given by Sturm, can be considered to be a special case of Green's formula.

Suppose that u_1 has several zeroes in (A, B), and let x_1, x_2 be two of its consecutive zeroes. Say that u_2 oscillates more rapidly than u_1 , i.e., that any solution u_2 to the second equation has at least one zero *between* x_1 *and* x_2 that is distinct from x_1 and x_2 .

In effect, if that were not true then one could suppose that u_1 and u_2 were positive between x_1 and x_2 . Since $G_1 - G_2$ is > 0, the integral in the preceding formula would then be > 0.

Now, $u_1 = 0$ at x_1 and x_2 , while it is > 0 at $x_1 < 0$ at x_2 , and not = 0 since u_1 is not = 0. Thus, the left-hand side of the preceding formula is ≤ 0 . That fact contradicts the preceding result. Thus, u_2 is annulled at least once between x_1 and x_2 .

The term "oscillates more rapidly than u_2 " is easy to explain because if one considers a solution u_1 to the first equation and a solution u_2 to the second one that are both annulled at x then the zero of u_2 that follows x will present itself before the zero that follows u_1 in such a way that one half the oscillation of u_2 is more rapid than that of u_1 .

Now suppose that *K* and *G* change. Consider the two equations:

$$\frac{d}{dx}(Ku_1') - G_1 u_1 = 0, \qquad \frac{d}{dx}(Ku_2') - G_2 u_2 = 0,$$
$$G_1 > G_2, \qquad K_1 > K_2 > 0.$$

It is easy to see that the reduction in K and G produces an oscillation of the solutions that is more rapid. In other words: There will be at least one zero of any solution u_2 of the second equation between two consecutive zeroes of an arbitrary solution u_1 of the first one.

We shall give a proof of that fact that is based upon a formula of Sturm that was modified by Picone.

From the identities:

$$\frac{d}{dx}(K_1u_2u_1') = u_2\frac{d}{dx}(K_1u_1') + K_1u_2'u_1' = G_1u_1u_2 + K_1u_2'u_1'$$

and

$$\frac{d}{dx}(K_2 u_1 u_2') = G_2 u_1 u_2 + K_2 u_2' u_1',$$

one deduces that:

$$\frac{d}{dx}(K_1u_2u_1'-K_2u_1u_2')=(G_1-G_2)u_1u_2+(K_1-K_2)u_1'u_2'.$$

That formula was given by Sturm and immediately generalizes the case that we treated above in which $K_1 = K_2$. However, the argument that we presented for $K_1 = K_2$ is no longer appropriate here due to the presence of the factor $u'_1u'_2$. We then establish the following formula:

$$\frac{d}{dx}\left[\frac{u_1}{u_2}(K_1u_2u_1'-K_2u_1u_2')\right]$$

= $\frac{u_1}{u_2}[(G_1-G_2)u_1u_2+(K_1-K_2)u_1'u_2']+\frac{u_1'u_2-u_1u_2'}{u_2^2}(K_1u_2u_1'-K_2u_1u_2'),$

from which, we infer Picone's identity:

$$\frac{d}{dx}\left[\frac{u_1}{u_2}(K_1u_2u_1'-K_2u_1u_2')\right] = (G_1-G_2)u_1^2 + (K_1-K_2)u_1'^2 + K_2\left[u_1'-u_2'\frac{u_1}{u_2}\right]^2.$$

That formula obviously applies at any point where $u_2 \neq 0$. If one then supposes that u_2 is non-zero between x_1 , x_2 , which are consecutive zeroes of u_1 , as well as at x_1 and x_2 , and that one has integrated the two sides from x_1 to x_2 , then since K_1 and K_2 are > 0 in $x_1 x_2$, as well as $K_1 - K_2$ and

in which:

 $G_1 - G_2$, the right-hand side will give a positive integral, and the left-hand side will give a zero integral, so the contradiction proves the theorem.

The formula will still apply when u_2 is annulled at x_1 or x_2 if is also annulled at one of those points because u_1 / u_2 will then tend to the value u'_1 / u'_2 , which it is well-defined (⁸), since u_2 is not identically zero and u' is not zero. When the left-hand side is integrated between x_1 and x_2 , it will then give a well-defined quantity, along with the third term on the right-hand side, and the integration will once more correctly prove that our theorem is exact.

One can further extend it to the case in which one has:

$$G_2 \leq G_1 , \qquad K_2 \leq K_1 ,$$

but, of course, under the condition that equality is not true at all points in the interval $x_1 x_2$. Nonetheless, there is one exceptional case here that needs to be pointed out.

We appeal to Picone's identity:

$$\left[u_{1}^{2}\left(K_{1}\frac{u_{1}'}{u_{1}}-K_{2}\frac{u_{2}'}{u_{2}}\right)\right]_{x_{1}}^{x_{2}}=\int_{x_{1}}^{x_{2}}(G_{1}-G_{2})u_{1}^{2}dx+\int_{x_{1}}^{x_{2}}(K_{1}-K_{2})u_{1}'^{2}dx+\int_{x_{1}}^{x_{2}}K_{2}\left(\frac{u_{1}'u_{2}-u_{1}u_{2}'}{u_{2}^{2}}\right)^{2}dx$$

Upon supposing that $G_1 \ge G_2$, $K_1 \ge K_2$ in *AB* and that the equality signs are not valid at any point in $x_1 x_2$, as we have said, the right-hand side will generally have a non-zero positive value.

Indeed, if $G_1 > G_2$ in part of the interval then the first integral will certainly be positive. If one has $G_1 \equiv G_2$ in $x_1 x_2$ then it might happen that one has $K_1 - K_2 = 0$ in one part of the interval $x_1 x_2$ and $u'_1 = 0$ in the rest of it, in such a way that the second integral will be zero. However, that can happen only if:

$$\frac{d}{dx}(K_1u') - G_1u = 0$$

admits a constant, but non-zero, solution in that partial integral, which would demand that $G_1 \equiv 0$ in that part of $x_1 x_2$, and indeed, if $G_1 \equiv 0$ in part of $x_1 x_2$ then there will be a solution u_1 that is constant in the part. Obviously, if *K* decreases then that particular constant solution will remain invariable. One can then easily construct an example in which the reduction of *K* does not alter the oscillation of the solutions, which would define an exception to the result that we found above.

In order to discard those exceptional cases, we suppose that the identity $G_1 \equiv G_2 \equiv 0$ is not verified on any subinterval of *AB* (⁹).

With those caveats, the reduction of G and K will indeed produce a more-rapid oscillation of the solutions, as we said above.

⁽⁸⁾ When x tends to x_1 through values that are above it or to x_2 through values that are below it.

^{(&}lt;sup>9</sup>) Some much-less-restricting conditions would suffice here.

Application. – The theorem permits us to compare the solutions of two differential equations whose coefficients have some simple relations between them. In particular, one can deduce some conditions for an equation to be non-oscillatory.

Indeed, suppose that one has the equation:

(1)
$$\frac{d}{dx}(Ku') - Gu = 0.$$

As always, *K* is supposed to be > 0, while *x* is included in *AB*. Let min *K* and min *G* be the minimum of *K* and *G*, resp., in the closed interval *AB*, and consider the equation:

$$\frac{d}{dx}[(\min K)u'] - (\min G)u = 0,$$

or

(2)
$$u'' - \frac{\min G}{\min K}u = 0$$

(2) is, in turn, integrated.

If min G > 0 then one will have the exponential solution $\exp\left(\sqrt{\frac{\min G}{\min K}}x\right)$, which is a real

solution that is always non-zero in AB. Therefore, (1) will have no solution that is annulled twice. The same thing will be true if min G = 0.

Therefore, if *G* is everywhere ≥ 0 then equation (1) will not be oscillatory.

If min G < 0 then the two fundamental solutions of (2) will be:

$$u = \sin\left(\sqrt{-\frac{\min G}{\min K}}\,x\right)$$

and

$$u = \cos\left(\sqrt{-\frac{\min G}{\min K}} x\right).$$

They oscillate: Meanwhile, one can form a solution to (2) that is non-zero in an interval ab in AB on the condition that the interval ab must be less than:

$$\frac{\pi}{\sqrt{-\frac{\min G}{\min K}}}.$$

If one has:

$$-\frac{\min G}{\min K} < \frac{\pi^2}{\left(b-a\right)^2}$$

then equation (1) will not be oscillatory in ab.

One has, analogously, that:

$$-\frac{\min G}{\min K} \ge \frac{\pi^2}{\left(b-a\right)^2}$$

is a sufficient condition for oscillation in ab, which implies that max G < 0.

That sufficient condition can even be extended. If:

$$-\frac{\min G}{\min K} \ge \frac{k^2 \, \pi^2}{\left(b-a\right)^2}$$

then any solution to the equation will have at least k zeroes in the closed interval $a \le x \le b$. Those results are often useful in the applications.

In passing, we point out a special case that was imagined by Sturm and Liouville. It is the case of the equation:

$$\frac{d}{dx}(k\,u')+(\lambda\,g-l)\,u\,=0\,,$$

in which k, g, l are functions of x that are continuous in the interval AB. $k > 0, g > 0, l \ge 0$.

That equation is provided by the study of the motion of heat in a heterogeneous bar or by simple vibrations of heterogeneous strings.

If λ varies then k will remain unvaried, but $G = l - \lambda g$ will vary. If λ increases then G will decrease, and the solutions will oscillate more rapidly.

One cannot suppose that g > 0 and assume that $g \ge 0$. To simplify, we confine ourselves to g > 0. The condition that $l \ge 0$ that is provided by the physical problem is superfluous since it does not affect the variation of $G = l - \lambda g$ in any way.

Another case that was envisioned by various geometers since Sturm, and with various methods, is the one in which k > 0, $l \ge 0$, but *g* changes sign in AB.

One might believe that this case is essentially distinct from the preceding one. Nothing of the sort is true. If one divides by $|\lambda|$ then the equation will be written:

$$\frac{d}{dx}\left(\frac{k}{|\lambda|}u'\right) + \left[g\left(\operatorname{sgn}\lambda\right) - \frac{l}{|\lambda|}\right]u = 0, \quad \operatorname{sgn}\lambda = \begin{cases} +1 & \text{if } \lambda > 0, \\ -1 & \text{if } \lambda < 0. \end{cases}$$

If one sets:

$$K = \frac{k}{|\lambda|}, \qquad G = \frac{l}{|\lambda|} - (\operatorname{sgn} \lambda) g$$

then one sees that the equation has the same type as the ones that we studied in the preceding sections.

1. If $\lambda > 0$ then an increase in λ will generally decrease *G* and also *K*. That is the Sturm case with *K* and *G* decreasing. If $\lambda \equiv 0$ then *G* will not vary, but *K* will certainly decrease.

Therefore, the solutions will oscillate more rapidly.

2. $\lambda < 0$, so if $|\lambda|$ increases then the same thing will result.

We have then found the circumstances that will be produced in Sturm's special case as a simple corollary to some of his results.

14. The comparison theorems. – Consider the following system, which includes two conditions at one point:

$$\frac{d}{dx}(Ku') - Gu = 0,$$
$$u(a) = \alpha,$$
$$u'(a) = \alpha'.$$

With the condition $|\alpha| + |\alpha'| > 0$, we exclude the case in which the solution *u*, which is unique, is identically zero. The fundamental problem that is then posed is the study of *how the zeroes of the solution u vary when K and G change*.

First of all, let us say a few words from a point of view that is slightly different from the one that one takes in such a study. If one considers the system:

$$\frac{d}{dx}(Ku') - Gu = 0,$$

$$\alpha' u(a) - \alpha u'(a) = 0$$

then the supplementary condition expresses the idea that for the value a of x, u and u' will be proportional to α and α' , resp., so the solutions of that second system will be deduced from the solution to the initial system by multiplying them by an arbitrary constant, which will not change its zeroes. That happens in such a way that the study of the zeroes of the solutions to the second system will be identical to the same study that was made of the first one.

Therefore, consider two systems:

(1)
$$\begin{aligned} \frac{d}{dx}(K_1u') - G_1u &= 0, \\ u(a) &= \alpha_1, \\ u'(a) &= \alpha_1', \end{aligned}$$

(2)
$$\begin{aligned} \frac{d}{dx}(K_2 u') - G_2 u &= 0, \\ u(a) &= \alpha_2, \\ u'(a) &= \alpha'_2, \end{aligned}$$

in which we suppose that $K_1 \ge K_2$, $G_1 \ge G_2$. In order for their solutions to not be identically zero, we suppose:

$$|\alpha_1| + |\alpha_1'| \neq 0$$
, $|\alpha_2| + |\alpha_2'| \neq 0$.

Moreover, the equality $K_1 = K_2$, $G_1 = G_2$ is not valid at any point in a subset of the interval *ab*, and the identity $G_1 = G_2 = 0$ must not be true in any subset of that interval.

Finally, we shall make the following hypotheses in the α and α' :

If $\alpha_1 \neq 0$ then α_2 must be $\neq 0$, and in such a way that $\frac{K_1(a)\alpha'_1}{\alpha_1} \geq \frac{K_2(a)\alpha'_2}{\alpha_2}$, i.e., that $\frac{K(a)\alpha'}{\alpha}$

will diminish upon passing from (1) to (2).

If $\alpha_1 = 0$ then we will make no supplementary hypothesis.

Under those hypotheses, Sturm established his two comparison theorems:

First comparison theorem:

If the solution u_1 to (1) has a certain number of zeroes that are distinct from a in the interval ab ($a < x \le b$) then the solution u_2 will have at least as many zeroes in that interval, and if one enumerates the zeroes of u_1 in order of increasing magnitude by $x_1, x_2, x_2, ...,$ and those of u_2 by $x'_1, x'_2, x'_3, ...$ then one will always have:

$$x'_k < x_k$$

for any value of k that corresponds to a zero of u_1 and u_2 .

From a result that was obtained in the preceding paragraph, one knows that u_2 has at least one zero between x_1 and x_2 , between x_2 and x_3 , etc. It remains to be proved that u_2 has at least one zero between a and x_1 . If one has:

$$u_1(a) = \alpha_1 = 0$$

then one will be certain that u_2 has at least one zero between a and x_1 from the results that were just mentioned.

Hence, suppose that $\alpha_1 \neq 0$. If u_2 has no zero between a and x_1 then since $u_2(a) \neq 0$, Picone's formula will say that between those limits:

Chapter III

$$\left[u_1^2 \left(\frac{K_1 u_1'}{u_1} - \frac{K_2 u_2'}{u_2}\right)\right]_a^{x_1} = \int_a^{x_1} (G_1 - G_2) u_1^2 dx + \int_a^{x_1} (K_1 - K_2) u_1'^2 dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1} K_2 \frac{(u_1' u_2 - u_2 u_2')^2}{u_2^2} dx + \int_a^{x_1}$$

At least one of the first two integrals on the right-hand side is > 0 and $\neq 0$.

The bracket on the left-hand side is = 0 for the upper limit since u_1 is annulled, and it is > 0 for the lower limit by virtue of the hypothesis:

$$\frac{K_1(a)\alpha_1'}{\alpha_1} \geq \frac{K_2(a)\alpha_2'}{\alpha_2}.$$

Hence, the left-hand side is ≤ 0 . The right-hand side is certainly > 0. That is a contradiction, and the theorem is then proved.

Second comparison theorem:

Under the same hypotheses, if $u_1(b) \neq 0$, and $u_2(b) \neq 0$ then:

$$\frac{K_1(b)u_1'(b)}{u_1(b)} > \frac{K_2(b)u_2'(b)}{u_2(b)},$$

provided that the solutions u_1 of (1) and u_2 of (2) have the same number of zeroes between a and b. (That restriction is meaningful because u_2 can have more zeroes than u_1 between a and b.)

1. First, suppose that u_1 and u_2 have no zeroes in (*ab*). One can then appeal to Picone's identity between the limits *a* and *b*. The right-hand side is positive. If one does not have:

$$\frac{K_1(b)u_1'(b)}{u_1(b)} > \frac{K_2(b)u_2'(b)}{u_2(b)}$$

then the left-hand side will be negative or zero, from the hypotheses that were made.

2. Suppose that u_1 and u_2 has *n* zeroes between *a* and *b*, and let x_n be the last of those zeroes, which is certainly a zero of u_1 (first comparison theorem). One can apply Picone's formula between x_n and *b*. The right-hand side is > 0. The bracket on the left is zero at the lower limit, so in order to avoid any contradiction, one must have that the inequality to be proved is true.

Here are some consequences of the comparison theorems: Let the system be:

(3)
$$\begin{aligned} \frac{d}{dx}(Ku') - Gu &= 0, \\ u(a) &= \alpha, \\ u'(a) &= \alpha', \end{aligned}$$

in which K, G, α , α' depend upon the parameter λ :

$$|\alpha|+|\alpha'|>0.$$

Suppose that λ varies in an interval $\Lambda_1 \Lambda_2$ (in which one can have $\Lambda_1 = -\infty$ or $\Lambda_2 = +\infty$), and *K* and *G* are continuous functions of *x* and λ . The solution *u* to that system and its derivative *u'* will be continuous functions of *x* and λ .

1. The zeroes of u will also be continuous functions of λ , except perhaps the ones that are found at one extremity of the interval ab or the other. That is basically nothing but the classical theorem on the continuity of implicit functions since u and u' are never both zero. It is true that we are not assured of the existence of the derivative $du / d\lambda$, but one can easily carry out the proof without making use of the derivative (¹⁰).

2. Suppose that λ increases from Λ_1 to Λ_2 , K and G decrease or remain constant for each value of x. In order to avoid the exceptional case that was pointed out above, suppose that K and G are not both independent of λ for any value of λ , even in a subset of the interval ab, and that if G is independent of λ in such an interval then they must not be identically zero. Finally, suppose that α either identically zero or that it does not vanish anywhere in $\Lambda_1 \Lambda_2$, and that $K \alpha' / \alpha$ decreases or remains constant for each value of λ in the latter case.

One will then deduce from the comparison theorems that the zeroes of u(x) must be decreasing and that, on the other hand, $\frac{K(b)u'(b)}{u(b)}$ must be decreasing, while the number of zeroes between

a and b will remain constant.

To fix ideas, in this study, it is good to define some zeroes of functions of λ , K, and G to be outside of the interval ab, which does not change anything in the nature of the solutions in the interval ab. In order to do that, one takes a number b' > b.

For x > b', one chooses K = 1. K will vary linearly between b and b' from K (b) up to the value 1. K is then a function that is always positive in $(a, +\infty)$. It will never increase as λ increases.

Similarly, G is taken to be equal to -1 from b' to $+\infty$ and to vary linearly between b and b' from G(b) up to -1.

Under those conditions, the equation will reduce to:

^{(&}lt;sup>10</sup>) See, for example, Osgood, Funktionentheorie, Chapter II, section 4.

$$\frac{d^2u}{dx^2} + u = 0$$

for x > b', and its fundamental solutions are sin x and cos x.

The solution *u* to the proposed system will then be oscillatory, and it will have an infinitude of zeroes for x > b'. Those zeroes will be continuous functions of λ as λ varies.

One has seen that if $(^{11})$:

$$\frac{-\max G}{\max K} \ge \frac{k^2 \, \pi^2}{\left(b-a\right)^2}$$

for a certain value of λ then since the solution to the system is a solution to the differential equation of that system, it will have at least k zeroes in ab. Hence, if one adds the following hypothesis to the previous ones: when λ varies from Λ_1 to Λ_2 :

$$\lim_{\lambda=\Lambda_2}\frac{-\max G}{\max K}=+\infty\,,$$

then one will see that if λ is taken to be sufficiently close to Λ_2 then the solution will have as many zeroes as one desires in (*ab*).

In what follows, we shall suppose that this condition is satisfied.

15. Sturm's oscillation theorems. – Therefore, vary λ from Λ_1 to Λ_2 .

If λ starts from Λ_1 (or a value that is arbitrarily close to Λ_1 if $\Lambda_1 \Lambda_2$ is an open interval) then the solution u will have a number m of zeroes *between* a and b, while the extremities a, b are excluded. If λ increases up to Λ_2 then that number must increase indefinitely. Thus, for a certain value μ_m of λ , the solution u will acquire a new zero at b, which is a zero that will enter into ab for $\lambda > \mu_m$. A new zero will present itself at b for a value $\mu_{m+1} > \mu_m$ of λ , and so on. We will then have a sequence of values μ_m , μ_{m+1} , ... that has Λ_2 for its limit point. If λ is between μ_m and μ_{m+1} then the solution will have m + 1 zeroes between a and b. It will then have m + 2 of them when λ is between μ_{m+1} and μ_{m+2} , ...

Moreover, when λ varies between μ_i and μ_{i+1} , the quantity $\frac{K(b)u'(b)}{u(b)}$, which is always

decreasing, must necessarily decrease from $+\infty$ to $-\infty$ since u(b) = 0 at μ_i and μ_{i+1} without one also having u'(b) = 0.

Having said that, one can look for the values of λ (viz., characteristic numbers) for which the system:

^{(&}lt;sup>11</sup>) Do not forget that here max G and max K denote the maxima of G and K when x varies from a to b, and λ is fixed. Those maxima obviously depend upon λ .

(1)
$$\begin{aligned} \frac{d}{dx}(Ku') - Gu &= 0, \\ \alpha' u(a) - \alpha u'(a) &= 0, \\ \beta' u(b) + \beta u'(b) &= 0, \end{aligned}$$

is compatible. β and β' are functions of λ such that: either $\beta \equiv 0$ or β does not vanish anywhere, and $K(b)\beta'/\beta$ is a decreasing function of λ .

The answer is provided by the following theorem:

Theorem:

The preceding system has an infinitude of characteristic numbers in the interval $\Lambda_1 \Lambda_2$. The first of those numbers can be between Λ_1 and μ_m , but one can certainly assert that there is exactly one of them in each interval ($\mu_m \mu_{m+1}$), ($\mu_{m+1} \mu_{m+2}$), ...

Indeed, if λ increases in one such interval $\mu_i \ \mu_{i+1}$, while u is the solution (¹²) to the system, in the absence of the second condition, then $\frac{K(b)u'(b)}{u(b)}$ will decrease from $+\infty$ to $-\infty$, and by hypothesis, $K(b)\beta'/\beta$ decreases, so $-K(b)\beta'/\beta$ will increase. It is then clear that one will have equality for a certain value of λ between μ_i and μ_{i+1} , which is unique moreover, i.e., $\beta u'(b)$ $+\beta' u(b) = 0$. That will suffice to prove that in each interval $\mu_i \ \mu_{i+1}$, there is a unique value of λ for which the proposed system is compatible.

So far, one has supposed that $\beta \neq 0$. If one has $\beta \equiv 0$ then the second condition will reduce to $u(b) \equiv 0$, and the μ_m , μ_{m+1} , ... will be precisely the characteristic values.

More generally, we shall call λ_{m+1} the characteristic value between μ_m and μ_{m+1} , λ_{m+2} , the characteristic values between μ_{m+1} and μ_{m+2} , ..., and if there is a characteristic value between Λ_1 and μ_m then we shall call it λ_m . It is clear that we can prove, by an argument that is completely analogous to the one that we just gave, that we cannot have more than one characteristic value in the interval $\Lambda_1 \mu_m$.

We let u_m , u_{m+1} , u_{m+2} , ... denote the solutions to the system for the characteristic values λ_m , λ_{m+1} , ..., resp. (We say *the* solutions by neglecting the arbitrary constant factor that multiplies them.)

Those functions differ by the number of their zeroes *between a and b*.

If u_m exists then it will have exactly *m* zeroes between *a* and *b*, u_{m+1} will have m + 1, ...

That result constitutes one of Sturm's oscillation theorems. One can then assert:

 $^(^{12})$ From the nature of the system, one obviously sees that if a solution exists then one can deduce an infinitude of them by multiplying it by an arbitrary constant.

Theorem:

If one is given an arbitrary whole number k > m then there will exist one and only one value of λ for which the proposed system admits a solution that has exactly k zeroes between a and b.

The theorem above is incomplete in the sense that it does not say that λ_m exists, and it does not tell us how to determine *m*.

One can go further with more precise hypotheses about G and K.

Add the following hypothesis to the ones that were made before:

$$\lim_{\lambda=\Lambda_1}\left(\frac{-\min G}{\min K}\right)=-\infty,$$

in which min *K* is necessarily > 0, by hypothesis. Hence, min *G* will be > 0 in the vicinity of Λ_1 . Hence, *G* will be positive for any value of *x* in *ab* for values of λ that are close to Λ_1 . One knows that the equation will then be non-oscillatory. Thus, for λ close to Λ_1 , no solution to the equation, and *a fortiori*, the system (1), will have more than one zero in *ab*. Therefore, one will have *m* = 0 or 1.

More precisely: In order to see that, we compare the differential equation of the system with the following one:

(2)
$$\frac{d}{dx}[(\min K)u'] - (\min G)u = 0$$

or

$$\frac{d^2u}{dx^2} - su = 0 ,$$

with

$$s=\frac{\min G}{\min K}>0.$$

That equation admits the fundamental solutions:

$$e^{\sqrt{s}(x-a)}$$
 and $e^{-\sqrt{s}(x-a)}$.

Hence, choose a number λ_0 between μ_m and Λ_2 , and look for the solution u_2 to (2) that verifies:

(3)
$$(\min K_0) \alpha'(\lambda_0) u_2(a) - (\min K) \alpha(\lambda_0) u_2'(a) = 0,$$

in which we have denoted the value of *K* for $\lambda = \lambda_0$ by K_0 . One finds that:

$$u_{2} = \frac{(\min K)\alpha(\lambda_{0})\sqrt{s} + (\min K_{0})\alpha'(\lambda_{0})\sqrt{s}}{2\sqrt{s}}e^{\sqrt{s}(x-a)}$$
$$+ \frac{(\min K)\alpha(\lambda_{0})\sqrt{s} - (\min K_{0})\alpha'(\lambda_{0})\sqrt{s}}{2\sqrt{s}}e^{-\sqrt{s}(x-a)}$$

Compare a u_2 that verifies (2) and (3) with a function u that verifies the equation and first conditions in the system (1).

Upon passing from (1) to (2), one has reduced the coefficients of the equation and replaced the first condition in (1) with (3) in such a fashion that $K(a)\alpha'/\alpha$ is diminished. [Indeed, the system (1) is studied here for λ that are found between Λ_1 and μ_m , and since λ_0 is $> \mu_m$, the value of $K(a)\alpha'/\alpha$, which is a decreasing function of λ , is greater between Λ_1 and μ_m than it is for λ_0 .] One can then conclude from Sturm's comparison theorems that $\frac{K(b)u'(b)}{u(b)}$ is greater than $(\min K)u'(b)$

 $\frac{(\min K)u'_2(b)}{u_2(b)}$, which is the value that the preceding expression takes for the system (2), (3).

Now, upon appealing to the value of u_2 , we will easily find that:

$$\lim_{\lambda=\Lambda_1}\frac{(\min K)u_2'(b)}{u_2(b)}=+\infty.$$

Hence, we have, a fortiori:

$$\lim_{\lambda=\Lambda_1}\frac{K(b)u'(b)}{u(b)}=+\infty.$$

We reach the conclusion that under our hypotheses, $\frac{K(b)u'(b)}{u(b)}$ will decrease from $+\infty$ to $-\infty$

as λ varies from Λ_1 to μ_m . We then conclude that the system (1) has one and only one characteristic number λ_m in the interval $\Lambda_1 \mu_m$.

Moreover, if one observes that for very large *s*, the preceding expression u_2 describes the sign of $\alpha(\lambda_0)$ for any value of *x* in (*ab*) or the sign of $\alpha'(\lambda_0)$ if $\alpha(\lambda_0) = 0$ then one can conclude that *m* is equal to zero, and not to 1.

In summary: Imagine the system:

(1)
$$\begin{aligned} \frac{d}{dx}(Ku') - Gu &= 0, \\ \alpha' u(a) - \alpha u'(a) &= 0, \\ \beta' u(b) + \beta u'(b) &= 0, \end{aligned}$$

Chapter III

in which K, G, α , α' , β , β' are functions of λ (and the first two are also functions of x) that satisfy the conditions that were stated before. If the conditions:

$$\lim_{\lambda = \Lambda_2} \left(\frac{-\max G}{\max K} \right) = +\infty,$$
$$\lim_{\lambda = \Lambda_1} \left(\frac{-\max G}{\max K} \right) = -\infty$$

are fulfilled, in addition, then the system (1) will have an infinitude of characteristic numbers between Λ_1 and Λ_2 , namely, λ_0 , λ_1 , ..., which are arranged in order of increasing magnitude. The characteristic functions $u_0, u_1, ...,$ which are solutions to the system (1) for $\lambda = \lambda_0, \lambda_1, ...,$ resp., have a number of zeroes between a and b that is equal to their index exactly.

Let us indicate another class of conditions that leads to analogous conclusions.

Suppose that the interval $\Lambda_1 \Lambda_2$ is closed on the left (i.e., Λ_1 belongs to that interval), and suppose that:

$$(\min G)_{\lambda=\Lambda_1} \ge 0, \qquad (\alpha \, \alpha')_{\lambda=\Lambda_1} \ge 0, \qquad (\beta \, \beta')_{\lambda=\Lambda_1} \ge 0$$

at the point Λ_1 . Hence, one has m = 0, and the characteristic value λ_0 will exist without fail. Indeed, imagine the auxiliary system:

(4)
$$\frac{d}{dx} \left[\left(\min_{\lambda = \Lambda_1} K \right) u' \right] - \left(\min_{\lambda = \Lambda_1} G \right) u = 0,$$
$$\alpha'(\Lambda_1) u(a) - \alpha(\Lambda_1) u'(a) = 0.$$

The notation $\left(\min_{\lambda=\Lambda_1} K\right)$ is easy to explain. If x varies in ab then $K(x, \lambda)$ will have a minimum $(\min K)$ that is a function of λ . It is the value of that function at Λ_1 that we shall denote by $\left(\min_{\lambda=\Lambda_1} K\right)$.

Let $v_2(x)$ be a solution to that system, and let $v_1(x)$ be a solution to the system:

(5)
$$\begin{cases} \frac{d}{dx}(Ku') - Gu = 0, \\ \alpha'(\Lambda_1)u(a) - \alpha(\Lambda_1)u'(a) = 0, \end{cases}$$

in which one has set $\lambda = \Lambda_1$ in *K* and *G*.

The comparison theorem says that: If v_2 has no zero in *ab* then v_1 will not have any, *a fortiori*. If one can show that:

$$\frac{(\min K) v_2'(b)}{v_2(b)} \ge 0 ,$$

moreover, then the comparison theorem will give:

$$\frac{(\min K)v_1'(b)}{v_1(b)} \ge 0$$

One can now integrate (4).

Exclude the case in which $\left(\min_{\lambda=\Lambda_1} G\right) = 0$, for the moment, and set: $s = \frac{(\min G)_{\lambda=\Lambda_1}}{(\min K)_{\lambda=\Lambda_1}}$

(s is > 0 and \neq 0). The solution v_2 is given by:

$$v_2(x) = \alpha(\Lambda_1) \cosh \sqrt{s} (x-a) + \frac{\alpha'(\Lambda_1)}{\sqrt{s}} \sinh \sqrt{s} (x-a) ,$$

as a calculation that was already done before will show.

Obviously, if $\left(\begin{array}{c} \alpha \ \alpha' \\ \lambda = \Lambda_1 \end{array} \right) \ge 0$ then one can suppose that $\alpha \ge 0$, $\alpha' \ge 0$ with no loss of generality

since α and α' enter into only the conditions (4) and (5), so $v_2(x)$ is ≥ 0 . Since α and α' cannot be annulled simultaneously, one concludes that $v_2(x) > 0$ in *ab*. Therefore, if $v_1(x)$ is the solution to the differential equation and the first condition of the system (1) for $\lambda = \Lambda_1$ then it will have no zero in *ab*. Hence, m = 0.

A very simple calculation will show, in turn, that:

$$\frac{(\min K)v_2'(b)}{v_2(b)} > 0 \quad \text{and} \quad \neq 0$$

One then concludes that:

$$\frac{K(b)v_1'(b)}{v_1(b)} > 0$$

If λ then increases from Λ_1 to the first value μ_m that makes the solutions v(x) of the differential equation and the first condition of the system (1) take on a zero at *b* then $\frac{K(b)v'(b)}{v(b)}$ will decrease

from a positive value that is $\neq 0$ down to $-\infty$, and since $-K(b)\beta'/\beta$ increases from a negative value under the same conditions, there will be a characteristic number λ_0 of the system (1) between Λ_1 and μ_m .

In the special case in which $\left(\min_{\lambda=\Lambda_1} G\right) = 0$, *s* is = 0, so the solution v_2 will be a linear function of x - a. One again sees that $v_2(x) > 0$ in *ab*, and that $\frac{(\min K)v'_2(b)}{v_2(b)}$ is positive or zero. The same

argument as before shows that m = 0 and λ_0 exists. However, λ_0 can coincide with Λ_1 in a special case. That can happen only if $\alpha'(\Lambda_1) = 0$, $\beta'(\Lambda_1) = 0$, $G(x, \Lambda_1) = 0$, as one sees by some very simple calculations. Finally, if β is zero then μ_m will be precisely the value λ_m that we are studying.

In summary, by means of the conditions:

$$[\min G]_{\lambda=\Lambda_1} \ge 0, \qquad (\alpha \alpha')_{\lambda=\Lambda_1} \ge 0, \qquad (\beta \beta')_{\lambda=\Lambda_1} \ge 0,$$

there is an infinitude of characteristic numbers λ_0 , λ_1 , ... of the system (1) in ($\Lambda_1 \Lambda_2$) that correspond to characteristic functions u_0 , u_1 , ..., resp., that have a number of zeroes between a and b that is equal to their index, and λ_0 can be equal to Λ_1 only if one has:

$$G(x, \Lambda_1) \equiv 0$$
, $\alpha'(\Lambda_1) = 0$, $\beta'(\Lambda_1) = 0$.

We said that Sturm gave some special cases of the oscillation theorem. In particular, he stated it for the system:

$$\frac{d}{dx}(k u') + (\lambda g - l)u = 0,$$

$$\alpha' u(a) - \alpha u'(a) = 0,$$

$$\beta' u(b) + \beta u'(b) = 0,$$

in which the conditions that are imposed upon α , α' , β , β' are the same as before, while k, g, l are, in addition, functions of x that are independent of λ and satisfy the inequalities k > 0, g > 0 in (ab).

k remains invariable here, while $G = l - \lambda g$ will diminish if λ increases from $\Lambda_1 = -\infty$ to $\Lambda_2 = +\infty$.

The boundary conditions are fulfilled. For example, $-\max G = \min (\lambda g - l)$ will increase indefinitely if λ increases up to $+\infty$. Hence, there is an infinitude of increasing characteristic values $\lambda_0, \lambda_1, \ldots$ for that system that tends to $+\infty$, and the corresponding characteristic values have 0, 1, 2, 3, ... zeroes in (*ab*). Sturm added the restrictions that $l \ge 0$ (¹³), $\alpha \alpha' \ge 0$, $\beta \beta' \ge 0$, at least, for $\lambda = 0$, which were imposed by the physical problem that he addressed. Those conditions, which are superfluous for the oscillation theorem, nonetheless permit one to specify the positions

^{(&}lt;sup>13</sup>) More precisely, Sturm's conditions are that $l \ge 0$, and that α , α' , β , β' are constants ≥ 0 that are independent of λ .

of the characteristic numbers by asserting, for example, that they are all positive. Indeed, if one considers the interval $\Lambda_1 = 0$, $\Lambda_2 = +\infty$, which is closed at Λ_1 , then the second class of conditions that is given for the oscillation theorem will be valid for that interval because G = l is ≥ 0 for $\lambda =$ 0, and its minimum will be ≥ 0 . Hence, $\lambda_0, \lambda_1, \ldots$ are positive.

Let us return to a problem that we spoke of before, which related to the equation:

(6)
$$\frac{d}{dx}(k u') + (\lambda g - l)u = 0$$

in which one always supposes that $k > 0, l \ge 0$, but g changes sign. [The case in which g is constantly < 0 is not essentially distinct from the one in which g > 0, so it would suffice to change λ into $-\lambda$ in order to reduce one to the other. Finally, we leave aside the case in which g keeps the same sign but might be annulled.]

In order to obtain some precise results under the new hypotheses, one must suppose that:

$$\alpha \alpha' \ge 0$$
, $\beta \beta' \ge 0$.

We shall show how one can include that particular case in the preceding results. If one divides the two sides of the proposed equation by $|\lambda| = v$ then it will become:

(7)
$$\frac{d}{dx}\left(\frac{k}{v}u'\right) - \left[\frac{l}{v} - g\left(\operatorname{sgn}\lambda\right)\right]u = 0, \quad \operatorname{sgn}\lambda = \begin{cases} +1 & \text{if } \lambda > 0, \\ -1 & \text{if } \lambda < 0. \end{cases}$$

That is an equation of the usual type with:

$$K = \frac{k}{v}, \qquad G = \frac{l}{v} - g (\operatorname{sgn} \lambda),$$

if v increases, while K and G decrease.

Let us see whether $K(a)\alpha'/\alpha$ and $K(b)\beta'/\beta$ diminish when v increases.

In order for that to be true, it is necessary that $\frac{1}{v} \frac{k(a)\alpha'}{\alpha}$ and $\frac{1}{v} \frac{k(b)\beta'}{\beta}$ must diminish.

If we then impose the condition that $K(a)\alpha'/\alpha$ and $K(b)\beta'/\beta$ will diminish when v increases (and for that to be true, it would suffice that α'/α , β'/β should diminish) then due to the facts that $\alpha \alpha' \ge 0$, $\beta \beta' \ge 0$, we will then see that $K(\alpha) \alpha' / \alpha$ and $K(b) \beta' / \beta$ will also diminish.

Previously, we supposed that:

$$\lim_{\lambda=\Lambda_2}\left(\frac{-\max G}{\max K}\right)=+\infty.$$

Here we shall use the interval $\Lambda_1 = 0$, $\Lambda_2 = +\infty$. If v increases then since the dominant term in G is $-g (\text{sgn } \lambda)$, we will see that G will change sign when v sufficiently large. Hence, max G will be > 0, and we will have, in turn:

$$\lim_{\nu=+\infty}\left(\frac{-\max G}{\max K}\right)=-\infty,$$

since max $K = \max k / v$ is > 0 and goes to zero when v goes to infinity. The result that was found seems to contradict the condition that was recalled above. Meanwhile, it is easy to see that the essential part of the preceding condition for the oscillation theorem is again found to be verified here. Indeed, consider the function – g (sgn λ). One can certainly find an interval a'b' in ab in which it remains constantly negative [because – g (sgn λ) changes sign in ab]. Its maximum in a'b' is then < 0.

We confine ourselves to the interval a'b' for the variable x. The preceding condition is found to be fulfilled: We will indeed have:

$$\lim_{\nu=+\infty} \left(\frac{-\max G}{\max K} \right) = +\infty \,.$$

If we recall that this condition permits us to assert that the solution v_1 to the system:

$$\frac{d}{dx}(Ku') - Gu = 0, \quad \alpha'u(a) - \alpha u'(a) = 0$$

can have (if λ is sufficiently close to Λ_2) a number of zeroes that is arbitrarily large in the interval *ab* then we will see that the same conclusions will apply to the interval *a'b'* here, and therefore to the interval *ab*, *a fortiori*.

The same observation can be made for the second type of indicated conditions:

$$(\min G)_{\lambda=\Lambda_1} \ge 0$$
, $\alpha \alpha' \ge 0$, $\beta \beta' \ge 0$

Granted, v = 0 does not belong to the interval of variation that we have assumed for v since the function K in (7) will go to infinity when v goes to zero.

However, one should recall the initial form (6) for which v = 0 did not create a singularity.

Upon operating as before, we can compare the equation (6), in which v = 0, and the condition $\alpha'(0)u(a) - \alpha(0)u'(a) = 0$, to a system that is analogous to the one that we called (4), for which v_2 will be the solution: We deduce, with no difficulty, that $v_2(x) > 0$, $\frac{(\min K)v'_2(b)}{v_2(b)} \ge 0$, and we pass from that to the solution $v_1(x)$ to (6).

Thus:

1. If $\lambda > 0$ then one will the infinite series of positive characteristic values $\lambda_0, \lambda_1, \ldots$ between 0 and $+\infty$.

2. If $\lambda < 0$ then upon setting $v = -\lambda$ in equation (6), one will find some new characteristic values that are different from the preceding ones since $G = l - \lambda g$ will change value when one changes λ into $-\lambda$.

We can denote the positive characteristic values by λ_0^+ , λ_1^+ , ... (they are increasing and go to $+\infty$) and the negative characteristic values by λ_0^- , λ_1^- , ... (they are decreasing and go to $-\infty$).

 λ_0^+ and λ_0^- are generally > 0 in one case and < 0 in the other, and neither of them can be = 0 except in the exceptional case that was pointed out above where *G* reduces to zero for $\lambda = \Lambda_1$, $\alpha'(\Lambda_1) = 0$, $\beta'(\Lambda_1) = 0$, i.e., for l = 0 here, with $\frac{\alpha'(0) = 0}{\beta'(0) = 0} {14 \choose 2}$.

By definition, we are again dealing with an oscillation theorem in the case in which g changes sign. It differs from the theorem that related to g when its sign was invariable only by the fact that for each integer k, there are two characteristic values λ_k^+ and λ_k^- that are positive and negative, resp., and will give a solution to the system (8) that is provided with k zeroes in the interval *ab*.

16. Study of the characteristic values from standpoint of reality and their order of multiplicity. – In the preceding sections, we studied real characteristic values. We now demand to know if there exist imaginary ones.

In order to do that, it is indispensable for us to suppose that the coefficients of the differential system under study are defined for imaginary values of λ .

We shall study only the Sturm systems from that viewpoint:

(1)
$$\frac{d}{dx}(k\,u') + (\lambda\,g - l)u = 0$$

(2)
$$\alpha' u(a) - \alpha u'(a) = 0, \qquad |\alpha| + |\alpha'| > 0$$

(3)
$$\beta' u(b) + \beta u'(b) = 0, \qquad |\beta| + |\beta'| > 0,$$

in which we suppose that k, g, l are functions of only x, and α , α' , β , β' are independent of λ .

Let $v_1(x, \lambda)$ be the solution to (1) that satisfies the conditions $u(a) = \alpha$, $u'(a) = \alpha'$. The two functions $v_1(x, \lambda)$ and $v'_1(x, \lambda)$ are continuous in (x, λ) , and are not only analytic, but entire in λ .

In order for the system to be compatible, it is necessary and sufficient that $v_1(x, \lambda)$ should verify the condition (3). That will give the following equation for determining the characteristic numbers:

^{(&}lt;sup>14</sup>) The complete discussion of this case, which is slightly exceptional, presents no serious difficulty when one uses the method that is employed here, as I pointed out in the Bulletin of the American Mathematical Society, October 1914.

(4)
$$\beta' v_1(b,\lambda) + \beta v_1'(b,\lambda) = 0,$$

whose left-hand side is an entire function of λ . It is indeed the same *characteristic equation* that gave us formula (1) in section 11 when we appeal to the fundamental system $u_1 = v_1$, $u_2 =$ the solution to (1) such that $u_2(a) = \gamma$, $u'_2(a) = \gamma'$, in which γ and γ' are constants, and $\alpha \gamma' - \alpha' \gamma = 1$.

First of all, do there exist imaginary characteristic numbers?

Before we answer that, let us establish an indispensable formula.

Let λ' , λ'' be two distinct characteristic numbers, while u_1 and u_2 are the corresponding characteristic functions [*viz.*, solutions to the system (1), (2), (3)]. From a formula of Sturm [formula (1), § 13]:

$$[k(u'_1u_2 - u_1u'_2)]^b_a + \int^b_a (\lambda' - \lambda'') g u_1 u_2 dx = 0.$$

However, u_1 and u_2 verify the conditions (2) and (3), so what remains is:

$$\int_a^b (\lambda' - \lambda'') g u_1 u_2 dx = 0,$$

and since $\lambda' \neq \lambda''$:

$$\int_a^b g \, u_1 \, u_2 \, dx = 0 \, .$$

That formula is valid for *k*, *g*, *l* that are complex functions of the real variable *x*, and α , α' , β , β' are arbitrary complex constants.

However, suppose that the coefficients α , α' , β , β' are real, as well as the functions k, g, l.

If λ is real then the left-hand side of (4) is obviously a real function. Its roots are then pair-wise conjugate. The imaginary root $\lambda' = \mu + v i$, $v \neq 0$ corresponds to the root $\lambda'' = \mu - v i$, and λ' will be $\neq \lambda''$.

If λ' corresponds to the characteristic function $u_1 = s + i t$, and λ'' corresponds to the conjugate imaginary function $u_2 = s - i t$.

When one applies the preceding relation to those two functions u_1 and u_2 , one will have:

(5)
$$\int_{a}^{b} g(s^{2}+t^{2}) dx = 0,$$

in which *s* and *t* are real functions of *x* that are not both identically zero since u_1 is not that way, so if one supposes that g > 0 or equal to zero in exceptional cases then one will reach a contradiction. Therefore, there are no imaginary characteristic values in this case. That is a result due to Poisson, and all that we did is to reproduce his proof.

We now examine the second case that was mentioned before in the preceding sections in which k > 0, $l \ge 0$, $\alpha \alpha' \ge 0$, $\beta \beta' \ge 0$, and g changes sign. The preceding theorem is still true then: All of the characteristic values are real.

Indeed, let $\lambda'' = \mu + v i$ be an imaginary characteristic, which is assumed to exist. Let $u_1 = s + i t$ be the corresponding function. If one writes out that u_1 verifies equation (1) then, upon separating the real part and the imaginary part on the left-hand side, one will have:

$$\frac{d}{dx}(k\,s') + (\mu\,g - l)\,s - \nu\,g\,t = 0,$$

$$\frac{d}{dx}(k\,t') + \nu\,g\,s + (\mu\,g - l)\,t = 0.$$

One concludes from that, upon multiplying the first one by *s* and the second one by *t*, adding them and integrating:

(6)
$$[k(ss'+tt')]_a^b - \int_a^b k(s'^2+t'^2) dx + \mu \int_a^b g(s^2+t^2) dx - \int_a^b l(s^2+t^2) dx = 0$$

It is clear that the second and fourth term are not positive, and the third one is zero, from (5).

By means of the hypotheses $\alpha \alpha' \ge 0$, $\beta \beta' \ge 0$, one will see that from the condition (2), s and

s' have the same sign at a, as well as t, t'. s and s', as well as t and t', have opposite signs at b. Therefore:

$$[k(ss'+tt')]_a^b$$

is a quantity that is ≤ 0 .

On the other hand:

$$-\int_a^b k(s'^2+t'^2)\,dx$$

is indeed < 0 and is not zero since otherwise one would need to have that s' and t', and as a result u'_1 would have to be constantly zero in ab, and therefore that u_1 would have to be a non-zero constant (in order to not be identically zero). That would demand that:

$$(\mu + \nu i) g - l = 0,$$

i.e., $vg \equiv 0$, and since $v \neq 0$, $g \equiv 0$.

That would be incompatible with the hypothesis that g changes sign. The left-hand side of (6) will therefore indeed be negative, which is the contradiction that proves the stated result.

Two questions are posed in the study of characteristic values:

1. What is the index of the system for one such value? – One immediately sees that in the present case, the index always has the value 1 since otherwise any solution to (1) would have to verify (2).

2. *What is the order of the multiplicity of a root of the characteristic equation?* – First, let us establish a preliminary formula.

If *u* is a characteristic function then one will have:

$$\int_a^b g \, u^2 \, dx \neq 0$$

(except in the exceptional case where g does not change sign, l = 0, $\alpha' = \beta' = 0$. We shall exclude that case; see the note towards the end of section 15). If g > 0 then the proof is immediate.

If g changes sign then let λ be the characteristic number that corresponds to the characteristic function u. Upon multiplying the differential equation that u must satisfy by u and integrating, one will find the formula:

$$\lambda \int_{a}^{b} g \, u^{2} \, dx = - \left[k \, u \, u' \right]_{a}^{b} + \int_{a}^{b} k \, u'^{2} \, dx + \int_{a}^{b} l \, u^{2} \, dx \, .$$

The first term on the right is positive or zero, by virtue of the relations:

$$\alpha \alpha' \ge 0$$
, $\beta \beta' \ge 0$.

Since the second term can be zero only in the exceptional case that we have excluded, we will see that $\lambda \int_{a}^{b} g u^{2} dx$ is positive (not zero). Our inequality is then proved since the characteristic number λ can be = 0 only in the exceptional case.

Now that that result has been established, it is easy to see that any root of the characteristic equation will be simple, except for the exceptional case that was pointed out.

Now imagine the entire function in formula (4):

$$F(\lambda) = \beta' v_1(b,\lambda) + \beta v_1'(b,\lambda).$$

Let λ_1 be any one of its zeroes, so the corresponding characteristic function will be $v_1(x, \lambda_1)$. Now, calculate $F'(\lambda_1)$.

If one combines equation (1), which is verified by $v_1(x, \lambda)$, with the one that is verified by $v_1(x, \lambda_1)$ and eliminates *l* then one will infer that:

$$\{k[v_1(x,\lambda)v_1'(x,\lambda_1)-v_1'(x,\lambda)v_1(x,\lambda_1)]\}_a^b+(\lambda_1-\lambda)\int_a^b g\,v_1(x,\lambda_1)v_1(x,\lambda)\,dx\,=0$$

 $v_1(x, \lambda)$ and $v_1(x, \lambda_1)$ verify the condition (2) that relates to the point *a*, and $v_1(x, \lambda_1)$ verifies the condition (3) that relates to *b*, so that equality will reduce to:

$$\int_a^b g \, v_1(x,\lambda_1) \, v_1(x,\lambda) \, dx = \frac{k(b) \, v_1'(b,\lambda_1)}{\beta'} \frac{\beta' \, v_1(b,\lambda) + \beta \, v_1'(b,\lambda)}{\lambda - \lambda_1}.$$
All of that supposes that $\beta' \neq 0$. [If $\beta' = 0$ then a similar calculation would give the same final result.]

Now:

$$\beta' v_1(b,\lambda) + \beta v_1'(b,\lambda) = F(\lambda) = F(\lambda) - F(\lambda_1),$$

since $F(\lambda_1) = 0$. Therefore, if λ goes to λ_1 then the right-hand side will go to:

$$\frac{k(b)v_1'(b,\lambda_1)}{\beta'}F'(\lambda_1).$$

 $v_1(x, \lambda)$ will go to $v_1(x, \lambda_1)$ uniformly for any x in ab. Hence, the left-hand side will go to:

$$\int_a^b g \left[v_1(x,\lambda_1) \right]^2 dx \,,$$

which is $\neq 0$.

One then concludes that $F'(\lambda_1) \neq 0$. Any root of the characteristic equation will then be a simple root, except in the exceptional case that was pointed out before.

CHAPTER IV

CHARACTERISTIC FUNCTIONS AND THEIR ZEROES IN SOME MORE GENERAL CASES (¹⁵)

17. The reality of characteristic numbers. – In this chapter, we shall treat only two typical problems that go beyond the Sturm problems that were studied in the preceding chapter. Meanwhile, we shall begin with some more general considerations.

Imagine a homogeneous system:

(1)
$$L(u) = l_n \frac{d^n u}{dx^n} + \dots + l_1 \frac{du}{dx} + (\lambda g - l)u = 0,$$
$$U_i(u) = 0 \qquad (i = 1, 2, \dots, n),$$

in which λ is a parameter, and whose functions $g, l, l_1, l_2, ..., l_n$ are independent, as well as the constant coefficients that enter into the $U_i(u)$.

The adjoint system will be written:

(2)
$$M(v) = m_n \frac{d^n v}{dx^n} + \dots + m_1 \frac{dv}{dx} + (\lambda g - m) v = 0,$$
$$V_i(v) = 0 \qquad (i = 1, 2, \dots, n),$$

and one easily sees that $m, m_1, ..., m_n$, and the coefficients of the V_i are independent of λ .

Those two systems have the same index for any λ , so they will have the same characteristic numbers.

Let $u_1, u_2, ...$ be the characteristic functions of the first system, while $v_1, v_2, ...$ are the corresponding characteristic functions of the second one.

For two different characteristic numbers (for example, λ_1 , λ_2), I say that:

$$\int_a^b g \, u_1 \, v_2 \, dx = 0 \; .$$

We shall utilize Green's formula:

^{(&}lt;sup>15</sup>) MASON, Trans. Amer. Math. Soc. 7 (1906), pp. 337. – BIRKHOFF, Trans. Amer. Math. Soc. 10 (1909), pp. 259. – KLEIN, Math. Ann. 18 (1881), pp. 419. – BÔCHER, Bull. Amer. Math. Soc. 4 (1898), pp. 307 and 365; *ibid.*, t. 5, pp. 22. – RICHARDSON, Trans. Amer. Math. Soc. 13 (1912), pp. 22. – Math. Ann. 73 (1912), pp. 289. The main result in section 1 of that article is incorrect.

For higher-order equations, one can consult:

LIOUVILLE, J. de Math. **3** (1838), pp. 561. – DAVIDOGLU, Ann. de l'E. N. S. **17** (1900), pp. 359; *ibid.*, **22** (1905), pp. 539. – BIRKHOFF, Trans. Amer. Math. Soc. **9** (1908), pp. 373. Ann. Math. **12** (1911), pp. 103. – HAUPT, Dissertation, Würzburg, 1911.

$$\int_{d}^{b} [v L(u) - u M(v)] dx = U_1 V_{2n} + \ldots + U_{2n} V_1$$

Let $L_0(u)$ and $M_0(v)$ be what L(u) and M(v), resp., become for $\lambda = 0$. One has:

$$\int_{d}^{b} [v L(u) - u M(v)] dx = \int_{d}^{b} [v L_{0}(u) - u M_{0}(v)] dx.$$

For $u = u_1$, $v = v_2$, the right-hand side of Green's formula is zero [from the boundary conditions in (1) and (2)], and the left-hand side reduces to:

$$\int_{d}^{b} \left[v_0 L_0(u_1) - u_1 M_0(v_2) \right] dx = \int_{d}^{b} (\lambda_2 - \lambda_1) g u_1 v_2 dx$$

One then has:

$$\int_a^b g \, u_1 \, v_2 \, dx = 0 \; .$$

If, in particular, the system is its own adjoint then $v_1, v_2, ...$ will be identical to $u_1, u_2, ...,$ and the formula will reduce to:

$$\int_a^b g \, u_1 \, u_2 \, dx = 0 \; .$$

That is the case for the Sturm systems that were studied in the preceding chapter, and for which the preceding formula was established directly.

One then deduces from this that for a real differential system that is its own adjoint:

1. If g > 0 then all of the characteristic numbers will be real because any characteristic value $\lambda_1 = \mu + v i \ (v \neq 0)$ will correspond to $\lambda_2 = \mu - v i$, which is likewise characteristic, and if $u_1 = s + t i$ then one will have $u_2 = s - t i$, so:

$$\int_a^b g(s^2+t^2)\,dx=0\,,$$

which is impossible if g > 0.

2. If $l \ge 0$ then g will change sign due to certain conditions that are imposed upon the coefficients of the $U_i(u)$, and the characteristic numbers will again be real.

Those restrictions should not be surprising here if one imagines that for the second-order Sturm system, with the conditions:

(3)
$$\begin{cases} \alpha' u(a) - \alpha u'(a) = 0, \\ \beta' u(b) + \beta u'(b) = 0, \end{cases}$$

one must impose the conditions that $\alpha \alpha' \ge 0$, $\beta \beta' \ge 0$ in order to prove the same result.

In order to not complicate the notations, we shall confine ourselves to a second-order system that it is own adjoint. We take it in the form:

(4)
$$\begin{cases} \frac{d}{dx}\left(k\frac{du}{dx}\right) + (\lambda g - l)u = 0, \\ \alpha_{1}u(a) + \alpha_{1}'u'(a) + \beta_{1}u(b) + \beta_{1}'u'(b) = 0, \\ \alpha_{2}u(a) + \alpha_{2}'u'(a) + \beta_{2}u(b) + \beta_{2}'u'(b) = 0. \end{cases}$$

In Chapter II, we saw that this system will be its own adjoint under the condition that:

$$k(a)(\beta_{1} \beta_{2}' - \beta_{2} \beta_{1}') = k(b)(\alpha_{1} \alpha_{2}' - \alpha_{2} \alpha_{1}').$$

One can generally reduce the two conditions in (4) to the form:

(5)
$$\begin{cases} u(a) = \gamma_1 u(b) + \gamma'_1 u'(b), \\ u'(a) = \gamma_2 u(b) + \gamma'_2 u'(b). \end{cases}$$

The only exception is when $\alpha_1 \alpha'_2 - \alpha_2 \alpha'_1$, and as a result, $\beta_1 \beta'_2 - \beta_2 \beta'_1$, are zero.

However, it is then clear that the conditions in (4) can be reduced to the Sturm form (3) by first eliminating the β_1 , β'_1 , β_2 , β'_2 , and then the α_1 , α'_1 , α_2 , α'_2 .

One does not exclude the new case by taking the conditions (4) in the form (5) then. Hence, the condition for the system to be its own adjoint is:

(6)
$$k(b) = k(a)(\gamma_1 \gamma'_2 - \gamma_2 \gamma'_1).$$

In order to prove that the characteristic numbers are real, we deduce that:

$$[k(ss'+tt')]_a^b - \int_a^b k(s'^2+t'^2) dx - \int_a^b l(s^2+t^2) dx = 0,$$

as we did for the Sturm systems in section 16, which is an equality that will imply a contradiction when we have:

$$k(a) s(a) s'(a) - k(b) s(b) s'(b) \ge 0,$$

$$k(a) t(a) t'(a) - k(b) t(b) t'(b) \ge 0.$$

However, s and t satisfy the conditions (5) at a and b that are verified by any characteristic function, so if one is to reach the preceding contradiction then it would suffice that every real function u should verify:

 $k(a)[\gamma_1 u(b) + \gamma'_1 u'(b)] [\gamma_2 u(b) + \gamma'_2 u'(b)] - k(b)u(b)u'(b) \ge 0,$

which reduces to:

$$\gamma_1 \gamma_2 u^2(b) + 2\gamma'_1 \gamma_2 u(b) u'(b)] + \gamma'_1 \gamma'_2 u'^2(b) \ge 0$$

because of formula (6).

By virtue of the inequality $\gamma_1 \gamma'_2 - \gamma_2 \gamma'_1 > 0$, which is a consequence of (6), one easily finds that the condition for that quadratic form in u(b), u'(b) to be defined and positive is that γ_1 , γ_2 , γ'_1 , γ'_2 should have the same sign (some of which can be zero): That is a case whose importance was first pointed out by Mason.

Hence, if those Mason conditions are verified then one will arrive at a contradiction by supposing the existence of imaginary characteristic numbers, which we would like to prove.

A very important case is that of a second-order system for which the conditions are:

$$u(a) = u(b),$$

 $u'(a) = u'(b).$

The condition for it to be its own adjoint is that:

$$k\left(a\right)=k\left(b\right).$$

[If we define k to be a periodic function of period b - a then k (x) will be continuous and positive for any value of x. If we take g and l to be periodic in a similar way, which might introduce a finite number of discontinuities for those two functions in a finite interval of variation for x (but we know that this is no inconvenience), then the conditions that we just wrote out will determine periodic solutions of period b - a for the differential equation. This parenthetic comment shows how one might associate what we shall discuss with the theory of periodic solutions.]

Mason's conditions are fulfilled. Hence, it is not only when g > 0, but also when $l \ge 0$ that g will change sign, so the characteristic numbers will all be real. We will see in the next section that there are always an infinitude of them.

In particular, if one takes:

(7)
$$\begin{cases} \frac{d^2u}{dx^2} + \lambda u = 0, \\ u (0) = u (2\pi), \\ u'(0) = u' (2\pi) \end{cases}$$

then all of the characteristic numbers will be real. In order to have them, one takes two fundamental solutions of the equation that are analytic in λ :

$$y_1 = \frac{1}{\sqrt{\lambda}} \sin x \sqrt{\lambda}$$
, $y_2 = \cos x \sqrt{\lambda}$.

The characteristic numbers are found very easily by a direct method. They are:

$$0, 1^2, 2^2, 3^2, \ldots$$

If one forms the characteristic equation then one will see that, except for zero, all of the characteristic values will have double roots.

One also sees that they have index 2.

That example shows that in the case that we are now considering, neither the multiplicities nor the indices of the characteristic numbers will be necessarily equal to 1.

18. Systems with periodic conditions. – Let us return to a more-detailed study of the system:

(1)
$$\frac{d}{dx}(Ku') - Gu = 0, u(a) - u(b) = 0, u'(a) - u'(b) = 0,$$

which has periodic conditions, and in which we suppose that:

$$K\left(a\right)=K\left(b\right)$$

in order for it to be its own adjoint.

K and G are functions of x and λ that are decreasing in λ . We suppose, in addition, that those two functions satisfy the conditions that we imposed upon them in Chapter III when we applied the Sturm method to the system:

$$\frac{d}{dx}(Ku') - Gu = 0,$$

$$\alpha' u (a) - \alpha u (a) = 0,$$

$$\beta' u'(b) + \beta' u'(b) = 0.$$

Take two principal solutions y_1 , y_2 to the equation:

$$y_1(a, \lambda) = 1$$
, $y_2(a, \lambda) = 0$,
 $y'_1(a, \lambda) = 0$, $y'_2(a, \lambda) = 1$.

Abel's formula gives:

$$y_1 y'_2 - y_2 y'_1 = \frac{K(a)}{K(b)},$$

and in particular, for x = b:

(2)
$$y_1(b) y_2'(b) - y_2(b) y_1'(b) = \frac{K(a)}{K(b)} = 1$$

for any λ .

The characteristic equation here is:

$$\begin{vmatrix} 1-y_1(b,\lambda) & -y_2(b,\lambda) \\ -y_1'(b,\lambda) & 1-y_1'(b,\lambda) \end{vmatrix} = 0.$$

It will reduce to:

$$F(\lambda) = y_1(b, \lambda) + y'_2(b, \lambda) - 2 = 0,$$

by virtue of the identity (2).

Does it have an infinitude of real roots? Sturm's methods permit us to answer that question in the affirmative.

Indeed, consider the auxiliary system:

(3)
$$\frac{d}{dx}(Ku') - Gu = 0, \quad u(a) = 0, \quad u(b) = 0.$$

It is a Sturm system. There are then an infinitude of characteristic numbers μ_0 , μ_1 , μ_2 , ... that correspond to characteristic functions that have 0, 1, 2, ..., resp., zeroes in (a, b).

 $F(\lambda)$ is not zero for those values μ_i , in general, Indeed, for those values, any solution of $\frac{d}{dx}(Ku') - Gu = 0$ that is zero at *a* will also be zero at *b*, so:

and as a result, (2) will give:

$$y_1(b, \mu_i) y'_2(b, \mu_i) = 1$$
,

 $y_2(b,\mu_i)=0,$

and one will have:

$$F(\mu_{i}) = \left[\sqrt{y_{1}(b,\mu_{i})} - \sqrt{y_{2}'(b,\mu_{i})}\right]^{2}.$$

Hence, $F(\mu_i)$ will be ≥ 0 or ≤ 0 according to whether $y_1(b, \mu_i)$ and $y'_2(b, \mu_i)$ are ≥ 0 or < 0, resp.

Let us confine ourselves to y'_2 (*b*, μ_i).

If $y'_{2}(b, \mu_{i}) < 0$ then:

$$F(\mu_i) \leq 0$$
,

so $y'_{2}(b, \mu_{i})$ is not = 0 because $y_{2}(b, \mu_{i}) = 0$.

If $y'_2(b, \mu_i) > 0$ then:

$$F(\mu_i) \geq 0$$

Those inequalities give the sign of $F(\lambda)$ at the points $\mu_0, \mu_1, \mu_2, \dots$



Figure 1.

Indeed, μ_0 corresponds to $y_2(x, \mu_0)$, which is zero at a, b and > 0 between a and b since $y'_2(a) = 1$.

Therefore:

For μ_1 (see Fig. 1):

 $y'_{2}(b,\mu_{1}) > 0$, $F(\mu_{1}) \ge 0$.

 $F(\mu_0) \leq 0$.

One can continue in that way, and one will see that the equation $F(\lambda) = 0$ will have an infinitude of roots that are separated by the numbers:

 μ_0, μ_1, \ldots

One can be even more precise. Any solution of (1) that is not identically zero must have an even number of zeroes in the interval $a \le x \le b$, which is closed at a and open at b, due to the facts that u(a) = u(b) and u'(a) = u'(b).

It will then result that μ_0 , μ_1 , μ_2 , ... cannot be roots of $F(\lambda) = 0$. Indeed, $y_2(x, \mu_0)$ has just one zero in $(a \le x < b)$, which is *a*. From Sturm's first theorem, which was proved in section **12**, all of the other solutions to the differential equation will also have exactly one zero in the interval. Therefore, they cannot verify the periodicity conditions at *a* and *b*, and the same argument will be true for μ_2 , μ_4 , ... Thus:

$$F(\mu_0) < 0$$
, $F(\mu_2) < 0$, ...

Later on, we will go on to consider Sturm's second auxiliary system:

(4)
$$\frac{d}{dx}(Ku') - Gu = 0, \\ u'(a) = 0, \\ u'(b) = 0.$$

There is an infinitude of characteristic numbers v_0 , v_1 , v_2 , ... that correspond to characteristic functions that have 0, 1, 2, ... zeroes in (a, b).

One reasons with the v_i as one did with the μ_i and one will find that $F(\lambda)$ has a sign at those points that one can fix and which is given by the first model below (Fig. 2).

One can relate v_i and μ_i to each other.

_	\mathcal{V}_0	v_1	<i>V</i> ₂
$F(\lambda)$	≥ 0	< 0	≥ 0

	I	$v_1 \mu_0$	$v_2 \mu_1$	$V_3 \mu_2$
	\mathcal{V}_0			
$F(\lambda)$	≥ 0	≤ 0	≥ 0	< 0

Figure 2.

First of all, $\nu_0 < \mu_0$, because for $\lambda = \nu_0$, the differential equation will have a solution that is everywhere ≤ 0 in (a, b), whereas for $\lambda = \mu_0$, it will have one that is zero at a and b, and consequently, for every value $\lambda \geq 0$, any solution will have at least one zero in (a, b).

One likewise sees that $\nu_1 < \mu_1$, because for $\lambda \ge \mu_1$, every solution will have at least two zeroes in (a, b), whereas for $\lambda = \nu_1$, there will be a solution that has only one zero in the interval.

One can continue in that way and show that $v_i < \mu_i$.

However, v_i can be $> \mu_{i-1}$ or $< \mu_{i-1}$. We then draw small segments along the λ -axis that join the points v_1, μ_0 ; v_2, μ_1 ; ... Those segments will never overlap with each other. They can reduce to points. That is the case for the system (7) in the preceding section.

One then sees that the system (1) admits some characteristic numbers that are infinite in number and are distributed:

- 1. Between v_0 and the first segment.
- 2. Between the consecutive segments.

Those characteristic numbers correspond to characteristic functions that have numbers of zeroes in $(a \le x \le b)$ that are easy to determine.

We shall now make those results more precise.

The conclusions that we will arrive at as a result will be true in any case. However, in order to simplify their presentation, we shall confine ourselves to supposing that K(x) is independent of λ , and $G(x, \lambda)$ is analytic in λ . Hence, the left-hand side $F(\lambda)$ of the characteristic equation will be a function that is analytic between Λ_1 and Λ_2 , so we can speak of the order of multiplicity of its roots. Furthermore, we have already supposed that G is a decreasing function of λ , and here we shall suppose that $\partial G / \partial \lambda < 0$, while excluding the case in which $\partial G / \partial \lambda$ can become zero.

Those conditions still leave enough freedom for the results that we will establish to apply to the important systems in which $G = l - \lambda g$, g > 0.

The characteristic equation is, as one knows:

$$F(\lambda) = y_1(b, \lambda) + y'_2(b, \lambda) - 2 = 0.$$

Consider the sign of $F'(\lambda)$ for various roots of $F(\lambda) = 0$:

$$F' = \frac{\partial}{\partial \lambda} [y_1(b,\lambda)] + \frac{\partial}{\partial \lambda} [y_2'(b,\lambda)] \, .$$

In order to calculate the derivatives above, we remark that if we consider, more generally, the solution u to the differential equation for which $u(a) = \alpha$, $u'(a) = \alpha'$, in which α and α' are two arbitrary constants, then $\partial u / \partial \lambda$ will verify the equation:

$$\frac{d}{dx}\left[K\left(\frac{\partial u}{\partial \lambda}\right)'\right] - G \frac{\partial u}{\partial \lambda} = \frac{\partial G}{\partial \lambda}u$$

It is an inhomogeneous linear equation in $\partial u / \partial \lambda$, and the equation with no right-hand side is the proposed equation that admits the solutions $y_1(x, \lambda)$, $y_2(x, \lambda)$. The method of variation of constants

gives $\frac{\partial u}{\partial \lambda}$ and $\left(\frac{\partial u}{\partial \lambda}\right)'$, while taking into account the fact that:

$$\left[\frac{\partial u}{\partial \lambda}\right]_{a} = 0,$$
$$\left[\left(\frac{\partial u}{\partial \lambda}\right)'\right]_{a} = 0,$$

since the values of u and u' at a will always be α and α' , resp., for any λ , and one will have:

$$\frac{\partial u}{\partial \lambda} = \int_{a}^{x} \frac{\partial G(\xi,\lambda)}{\partial \lambda} u(\xi,\lambda) \frac{y_{1}(\xi,\lambda) y_{2}(x,\lambda) - y_{2}(\xi,\lambda) y_{1}(x,\lambda)}{K(a)} d\xi,$$

$$\left(\frac{\partial u}{\partial \lambda}\right)' = \int_{a}^{x} \frac{\partial G(\xi,\lambda)}{\partial \lambda} u(\xi,\lambda) \frac{y_{1}(\xi,\lambda) y_{2}'(x,\lambda) - y_{2}(\xi,\lambda) y_{1}'(x,\lambda)}{K(a)} d\xi.$$

Hence, for x = b and $u = y_1$ or y_2 , one will have:

$$\frac{\partial y_1(b,\lambda)}{\partial \lambda} = \int_a^b \frac{\partial G(\xi,\lambda)}{\partial \lambda} \frac{y_1(\xi,\lambda)}{K(a)} [y_1(\xi,\lambda) y_2(x,\lambda) - y_2(\xi,\lambda) y_1(x,\lambda)] d\xi,$$
$$\frac{\partial}{\partial \lambda} [y_2'(b,\lambda)] = \int_a^b \frac{\partial G(\xi,\lambda)}{\partial \lambda} \frac{y_2(\xi,\lambda)}{K(a)} [y_1(\xi,\lambda) y_2'(x,\lambda) - y_2(\xi,\lambda) y_1'(x,\lambda)] d\xi,$$

$$F'(\lambda) = \int_{a}^{b} \frac{\partial G(\xi,\lambda)}{\partial \lambda} \{ y_{2}(b,\lambda)y_{1}^{2}(\xi,\lambda) + [y_{2}'(b,\lambda) - y_{1}(b,\lambda)] y_{1}(\xi,\lambda) y_{2}(\xi,\lambda) - y_{1}'(b,\lambda) y_{2}^{2}(\xi,\lambda) \} d\xi.$$

In order to determine the sign of the quadratic form in y_1 (ξ , λ), y_2 (ξ , λ) that is between brackets, imagine its discriminant. From Abel's formula, one reduces it to:

$$\left[\frac{y_2'(b,\lambda)+y_1(b,\lambda)}{2}\right]^2-1.$$

We shall study the sign of $F'(\lambda)$ for those values of λ that annul $F(\lambda)$. Now, for those values:

$$y_1(b,\lambda) + y'_2(b,\lambda) = 2$$

so the preceding determinant will be zero. The form is the square of:

so:

$$\left[\sqrt{y_2(b)} y_1(\xi) \pm \sqrt{-y_1'(b)} y_b(\xi)\right],\,$$

so it will be positive or negative according to whether $y_2(b)$ and $-y'_1(b)$ are positive or negative, resp. [It is implicit that the values of λ that enter into $y_2(b)$, $y'_1(b)$ are the roots of $F(\lambda) = 0$.]

Let us first overlook the case in which that expression is identically zero, and since it is a solution to the differential equation, that amounts to saying that one does not have both $y_2(b) = 0$ and $y'_1(b) = 0$. We shall return to that case later.

The quadratic form will not be identically zero upon integrating it since K(a) > 0, $\frac{\partial G(\xi, \lambda)}{\partial \lambda} < 0$. $F'(\lambda)$ will then be found to have the opposite sign to the form.

Therefore, $F(\lambda)$ cannot have more than one zero in any interval of variation for λ in which $y_2(b) \neq 0$ because the form, and as a result $F'(\lambda)$, will keep the same sign at any point of that interval where $F(\lambda) = 0$. The same remark will be true for an interval in which $y'_1(b)$ remains $\neq 0$.

We then remember that the μ_i are the roots of the equation:

and the v_i are the roots of:

 $y_1'(b,\lambda) = 0$.

 $y_2(b,\lambda)=0,$

Neither y'_1 nor y_2 are zero between two consecutive segments ($v_i \ \mu_{i-1}$) (Fig. 2), so $F(\lambda)$ can have only one zero between two such segments. Moreover, the distribution of the signs of $F(\lambda)$ at the extremities proves that $F(\lambda)$ is annulled at least once between two consecutive segments, so $F(\lambda)$ will admit exactly one zero between two consecutive segments ($v_i \mu_{i-1}$). Furthermore, there are no zeroes along the segments themselves since there is only one in (μ_0, μ_1) , for example, (see Fig. 2), and it is between μ_0 and ν_2 . (It is easy to infer the same conclusion for all cases in the figure.)

For the same reason, there is one and only zero to $F(\lambda)$ between v_0 and the segment (v_1, μ_0) . There are none between Λ_1 and ν_0 .

Those results fix the positions of the characteristic numbers $\lambda_0, \lambda_1, \dots$ (Fig. 3).





Let us move on to the characteristic functions u_0, u_1, \ldots

One knows that for $\lambda = \mu_0$, the solution to the auxiliary system (3) is annulled at only *a* and *b*. Thus, for $\lambda \leq \mu_0$, any solution to the equation can have only one zero in (a, b). Now, any solution to (1) has an even number of zeroes. Hence, u_0 will not be annulled in (a, b).

Some entirely-analogous arguments will prove that u_1 and u_2 have two zeroes, u_3 and u_4 have four, etc., which we summarize in the table:

Characteristic functions....
$$u_0$$
 u_1 u_2 u_3 u_4 ...
Number of zeroes in ab 0 2 2 4 4 ...

Those results constitute an oscillation theorem for the present case.

Once more, the only exceptional cases that we have to consider are the ones for which some of the values $\mu_1 \nu_2$, $\mu_1 \nu_2$, ... coincide [since we know the sign of $F(\lambda) \neq 0$ precisely for $(\mu_0 \nu_1)$, $(\mu_2 \nu_3), \ldots$, and for which the λ coincide with those values. That is exactly the case that we overlooked in the preceding discussion, in which $y_2(b) = y'_1(b) = 0$ for a root of the equation $F(\lambda) = 0.$

That exceptional case essentially presents itself in the particular example that was treated before [see (7), § 16]. One must then study $F''(\lambda)$ for those values of λ for which $F(\lambda) = F'(\lambda)$ = 0.

Now, a calculation that is analogous to the one that was made for $F'(\lambda)$ will prove that for those values:

$$F''(\lambda) = 2 \int_{a}^{b} \frac{\frac{\partial G(x)}{\partial \lambda}}{K(a)} \left[y_{1}(x) \frac{\partial y_{2}(x)}{\partial \lambda} - y_{2}(x) \frac{\partial y_{1}(x)}{\partial \lambda} \right] dx,$$

so upon replacing $\frac{\partial y_2}{\partial \lambda}$ and $\frac{\partial y_1}{\partial \lambda}$ with their values that were found before:

$$F''(\lambda) = -2\int_{a}^{b}\int_{a}^{x} \frac{\partial G(x)}{\partial \lambda} \frac{\partial G(\xi)}{\partial \lambda} \left[y_{1}(x) y_{2}(\xi) - y_{2}(x) y_{1}(\xi) \right]^{2} d\xi dx,$$

and since the bracket that appears in the double integral is not identically zero (because y_1 and y_2 are linearly independent), one will have:

$$F''(\lambda) < 0$$
.

Those multiple roots of $F(\lambda)$ are exactly double roots then, and F will have a maximum at all of those points. It will then follow that the double roots will replace two simple roots that would otherwise be found on two sides of the segment that reduces to that point (¹⁶).

Finally, one very easily finds that the index of the system (1) for those values is 2 since the periodic problem has two linearly-independent solutions, y_1 and y_2 . The double root then corresponds to two characteristic functions that have the numbers of zeroes that are indicated by the oscillation theorem.

Our results generalize in various directions. We can confine ourselves to supposing that G is decreasing without supposing that G possesses a derivative with respect to λ . The methods that we must employ are more delicate, but the results will be the same. We can also consider non-periodic conditions for systems that are always adjoint to themselves. However, the oscillation theorems will be less precise then. One can give the number of zeroes of characteristic functions in (a, b) only up to one unit. Finally, there are extensions to equations of order higher than two, which are very incomplete up to now.

^{(&}lt;sup>16</sup>) If one has, for example, that μ_1 and ν_2 coincide, and λ is a characteristic value that coincides with μ_1 and ν_2 then one can demand to know if there are any other characteristic numbers λ_1 and λ_2 between $(\mu_1 \nu_2)$ and the preceding segment $(\mu_0 \nu_1)$ that are different from the λ that is in $(\mu_1 \nu_2)$, as well as ones between $(\mu_1 \nu_2)$ and $(\mu_2 \nu_3)$. Now, that would be impossible since $F(\lambda)$ is a maximum for the double root with the values λ_1 and λ_2 that are supposed to exist, so $F'(\lambda)$ must have the signs – and +.

However, since $F(\lambda) < 0$ for $(\mu_0 \nu_1)$ and for $(\mu_1 \nu_2)$, one will see that its derivative $F'(\lambda)$ at λ_1 can only be positive. The same argument applies to λ_2 . The contradiction shows that λ_1 and λ_2 do not exist since they will coincide with μ_1 and ν_2 , resp.

19. Second extension of the Sturm problem. Klein's oscillation theorem. – In Chapter III, we studied systems of the form:

$$\frac{d}{dx}(Ku') - Gu = 0,$$

$$\alpha' u(a) - \alpha u'(a) = 0,$$

$$\beta' u(b) + \beta u'(b) = 0,$$

in which K contains only the variable x, while G depends upon x and a parameter, and α , β , α' , β' are constants that are independent of λ .

In what follows, instead of one interval (a, b) of variation for x, we shall imagine an arbitrary number of such intervals:

$$(a_0 b_0);$$
 $(a_1 b_1);$...; $(a_n b_n)$

that follow each other along the *x*-axis in order of increasing indices and have no point in common pairwise (not even the extremities). In addition, we shall consider n + 1 parameters $\lambda_0, \lambda_1, ..., \lambda_n$ that the function *G* depends upon, such that each parameter varies in an interval that it is associated with. For each segment $a_i b_i$, we consider conditions such as:

$$\begin{aligned} &\alpha'_0 u(a_0) - \alpha_0 u'(a_0) = 0 , \\ &\beta'_0 u(b_0) + \beta_0 u'(b_0) = 0 , \\ &\alpha'_1 u(a_1) - \alpha_1 u'(a_1) = 0 , \\ &\beta'_1 u(b_1) + \beta_1 u'(b_1) = 0 , \end{aligned}$$

The problem that we pose is the following one:

Determine $\lambda_0, \lambda_1, ..., \lambda_n$ in such a way that the differential equation:

$$\frac{d}{dx}(Ku') - Gu = 0$$

will admit n + 1 solutions u_0, u_1, \ldots, u_n such that:

u_0	verifies the conditions that relate to			$(a_0 b_0)$
u_1	"	"	"	$(a_1 b_1)$
 u _n				$(a_n b_n)$

For n = 0, it is clear that one has the Sturm problem that was studied in Chapter III.

Klein was led to that problem by studying the work of Lamé on the distribution of heat in an ellipsoid. The only case that we shall consider, along with Klein, is the one in which *G* has the particular form:

(1)
$$G = l(x) - g(x) [\lambda_0 + \lambda_1 x + \lambda_2 x^2 + \dots + \lambda_n x^n]$$

 $(\lambda_0, \lambda_1, ..., \lambda_n \text{ vary from } -\infty \text{ to } +\infty).$

We suppose, to simplify, that g(x) > 0 in the interval (a_i, b_i) . In reality, it suffices to suppose that g is not annulled in any of those intervals, but its sign can be different in two different intervals. K, l, and g are supposed to be continuous in each of the closed intervals (a_i, b_i) . We suppose nothing about the other intervals beyond those ones (1^7) .

Of course, one can consider only one interval for x and n + 1 differential systems by changing the independent variable, but the stated result would not be as simple as it was in the form that that was given to the problem above. Under those conditions, one has the oscillation theorem that is due to Klein:

There exists an infinitude of real systems $(\lambda_0, \lambda_1, ..., \lambda_n)$ for which the desired functions $u_0, u_1, ..., u_n$, resp., will exist without being identically zero.

Those systems of characteristic numbers are distinguished from each other by the number of zeroes that the characteristic functions $u_0, u_1, ..., u_n$ possess in $(a_0 \ b_0), ..., (a_n \ b_n)$, respectively. If one is given n + 1 numbers in advance that are positive or zero:

$$m_0, m_1, \ldots, m_n$$

then one can find one and only one system $\lambda_0, \lambda_1, ..., \lambda_n$ for which the functions $u_0, u_1, ..., u_n$ have $m_0, m_1, ..., m_n$ zeroes, respectively, in each of the intervals $(a_0 \ b_0), (a_1 \ b_1), ..., (a_n \ b_n)$, resp.

For n = 0, one will simply have the Sturm theorem that was established in Chapter III. We shall then proceed by recurrence.

Suppose that the theorem is true up to index n - 1. Prove that it is true for the index n. One can write:

(2)
$$G(x) = [l(x) - \lambda_n x^n g(x)] - g(x)[\lambda_0 + \lambda_1 x + \dots + \lambda_{n-1} x^{n-1}].$$

Give an arbitrary, but fixed, value to λ_n . There will then exist one and only one system of values of $\lambda_0, \lambda_1, ..., \lambda_{n-1}$ for which the characteristic functions have:

 m_0 m_1 ... m_{n-1} zeroes in $(a_0 b_0)$, $(a_1 b_1)$, ... $(a_{n-1} b_{n-1})$,

respectively.

^{(&}lt;sup>17</sup>) And the coefficients in the Lamé equation essentially have singularities between the intervals $(a_i b_i)$.

It remains to be seen whether λ_n can be chosen in such a way that u_n will exist that verify the (n + 1) conditions.

The preceding $\lambda_0, \lambda_1, ..., \lambda_{n-1}$ are functions of λ_n since they will be determined when λ_n is known. Upon expressing them as functions of λ_n in G, G will become a function of x and $\lambda_n : G(x, \lambda_n)$.

We shall show that $G(x, \lambda_n)$ satisfies the conditions that are required by Sturm's oscillation theorem.

In order to do that, consider the difference $G(x, \lambda_n) - G(x, \lambda'_n)$. It is necessarily annulled for a value of x in each interval $(a_0 b_0), (a_1 b_1), ..., (a_{n-1} b_{n-1})$ because if that difference is not annulled in $(a_0 b_0)$, for example, then it would have a constant sign; for example, one would have $G(x, \lambda_n) > G(x, \lambda'_n)$. However, for the value λ'_n of the parameter, the solution u_0 to the equation that verifies the boundary conditions that relate to $a_0 b_0$ will oscillate more rapidly than the solution for the value λ_n (Sturm's comparison theorem), and that would contradict the fact that u_0 always has a number of zeroes that is equal to m_0 for any λ_n .

There is therefore at least:

one point	x_{n-1}	in	$(a_{n-1} b_{n-1})$
one point	<i>x</i> ₁	in	$(a_1 b_1),$
one point	x_0	in	$(a_0 b_0),$

such that the preceding difference is zero.

Let λ'_0 , λ'_1 , ..., λ'_{n-1} denote the values of the first *n* parameters that correspond to λ'_0 . One will have:

(3)
$$G(x, \lambda_n) - G(x, \lambda'_n) = g(x)[(\lambda'_0 - \lambda_0) + (\lambda'_1 - \lambda_1)x + \dots + (\lambda'_n - \lambda_n)x^n]$$

Since g(x) was supposed to be $\neq 0$ in all of $(a_i b_i)$, the polynomial in brackets will admits x_0 , x_1, \ldots, x_n for its roots. Hence:

(4)
$$\lambda_0' - \lambda_0 + (\lambda_1' - \lambda_1)x + \dots + (\lambda_n' - \lambda_n)x^n = (\lambda_n' - \lambda_n)(x - x_0)(x - x_1)\cdots(x - x_{n-1})$$

One deduces two consequences from that:

1. The continuity of $G(x, \lambda_n)$ with respect to λ_n : That is because $x_0, x_1, ..., x_{n-1}$ remain in the intervals $(a_0 \ b_0), (a_1 \ b_1), ..., (a_{n-1} \ b_{n-1})$, so they are finite. Thus, the difference $G(x, \lambda_n) - G(x, \lambda'_n)$ will go to zero with $\lambda'_n - \lambda_n$. Moreover, it is obvious from the formula that gives that difference that $G(x, \lambda'_n)$ will go to $G(x, \lambda_n)$ uniformly for any x in a_n , b_n when λ'_n goes to $\lambda_n \cdot G(x, \lambda_n)$ will then be continuous with respect to the two independent variables x, λ_n .

2. If λ_n increases then $G(x, \lambda_n)$ will decrease for all x in $(a_n b_n)$. That is because one sees from (3) and (4) that $G(x, \lambda_n) - G(x, \lambda'_n)$ has the same sign as $\lambda'_n - \lambda_n$, so $G(x, \lambda_n)$ will be a decreasing function, and since the product $(x - x_0)(x - x_1) \dots (x - x_{n-1})$ can never be zero when x is in $(a_n b_n)$, the intervals $(a_0 b_0)$, ..., $(a_{n-1} b_{n-1})$ will have no point in common with $(a_n b_n)$ (¹⁸), and one will also see from (3) and (4) that:

$$\lim_{\lambda_n=+\infty} G(x,\lambda_n) = -\infty,$$
$$\lim_{\lambda_n=-\infty} G(x,\lambda_n) = +\infty.$$

The consequences 1 and 2 suffice to show that the conditions for the validity of Sturm's oscillation theorem are fulfilled here. For each number m_n , there will then exist one and only one real value of λ_n for which the proposed equation will admit an (n + 1)th solution u_n that satisfies the boundary conditions relative to $(a_n b_n)$ and has exactly m_n zeroes in that interval.

Klein's theorem is thus proved for any value of *n*.

Let us point out, in passing, a physical application of the theorem:

In order to interpret Sturm's oscillation theorem in a particular case, one can say that if a heterogeneous cord is given then one can make it execute simple vibrations for which the cord presents a number of nodes that is fixed in advance.

One can interpret Klein's oscillation theorem by saying that if a homogeneous membrane is given that is bounded by two arcs of ellipses, along with two homofocal hyperbolic axes, then one can make it vibrate in such a way that it will present a number m_0 of nodal homofocal ellipses, and a number m_1 of nodal homofocal hyperbolas, when m_0 and m_1 were established in advance.

In the foregoing, we found an infinitude of real systems of characteristic values. Do there exist imaginary characteristic values for the λ , i.e., systems of imaginary values λ for which the equation will admit n + 1 solutions that satisfy the boundary conditions that related to the n + 1 intervals a_i, b_i , respectively? We shall see that there are none.

To simplify the writing, we confine ourselves to the typical case of three parameters.

Let $(\lambda'_0, \lambda'_1, \lambda'_2), (\lambda''_0, \lambda''_1, \lambda''_2)$ be two systems of characteristic values that correspond to the characteristic functions:

$$u_1 v_1 w_1, u_2 v_2 w_2$$

respectively. One has:

$$\frac{d}{dx}(Ku_1') + [g(\lambda_0' + \lambda_1'x + \lambda_2'x^2) - l]u_1 = 0,$$

$$\frac{d}{dx}(Ku_2') + [g(\lambda_0'' + \lambda_1''x + \lambda_2''x^2) - l]u_2 = 0.$$

^{(&}lt;sup>18</sup>) This theorem will still be true when the intervals a_i , b_i touch each other or when g(x) vanishes without changing sign at some isolated points. However, one must then appeal to a form of Sturm's oscillation theorem that is a bit more refined than the one that we obtained.

If one multiplies the two equations by u_2 and u_1 , respectively, subtracts corresponding sides, and integrates from a_0 to b_0 then one will have:

$$\int_{a_0}^{b_0} g\left[\lambda_0' - \lambda_0'' + (\lambda_1' - \lambda_1'')x + (\lambda_2' - \lambda_2'')x^2\right] u_1 u_2 dx = 0$$

due to the conditions that u_1 and u_2 verify at a_0 and b_0 . Similarly:

$$\int_{a_1}^{b_1} g\left[\lambda_0' - \lambda_0'' + (\lambda_1' - \lambda_1'') x + (\lambda_2' - \lambda_2'') x^2\right] v_1 v_2 dx = 0,$$

$$\int_{a_2}^{b_2} g\left[\lambda_0' - \lambda_0'' + (\lambda_1' - \lambda_1'') x + (\lambda_2' - \lambda_2'') x^2\right] w_1 w_2 dx = 0.$$

If the two systems $(\lambda'_0, \lambda'_1, \lambda'_2), (\lambda''_0, \lambda''_1, \lambda''_2)$ are distinct, with the three differences $\lambda'_i - \lambda''_i$ not all being zero, then the determinant of the preceding linear equations in $\lambda'_i - \lambda''_i$ will be zero. That determinant will reduce to the simple formula:

$$\int_{a_0}^{b_0} \int_{a_1}^{b_1} \int_{a_2}^{b_2} g(x_0) g(x_1) g(x_2) \begin{vmatrix} 1 & x_0 & x_0^2 \\ 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \end{vmatrix} u_1(x_0) u_2(x_0) v_1(x_0) v_2(x_0) w_1(x_0) w_2(x_0) dx_0 dx_1 dx_2 = 0.$$

That formula generalizes the formula:

$$\int_a^b g \, u_1 \, u_2 \, dx = 0$$

that is due to Poisson and was encountered in section **16**. It shows that the characteristic values are all real for g > 0 because if λ'_0 , λ'_1 , λ'_2 are imaginary and characteristic then they will correspond to λ''_0 , λ''_1 , λ''_2 that are characteristic and imaginary and conjugate to the preceding values. The functions u_1 , v_1 , w_1 will be conjugate to u_2 , v_2 , w_2 , resp. The differential element of the triple integral above will always be > 0, and not $\equiv 0$, since the intervals $(a_0 b_0)$, $(a_1 b_1)$, $(a_2 b_2)$ for x_0 , x_1 , x_2 , resp., will have no point in common. One will then have a contradiction to the supposition that imaginary characteristic values exist.

Klein's theorem then gives all of the characteristic values.

The preceding remark suggests an extension of Klein's theorem. We shall present it briefly for three intervals $(a_0 b_0)$, $(a_1 b_1)$, $(a_2 b_2)$, but it is general in scope.

Take a function *G* of the form:

$$G = l - (\lambda_0 g_0 + \lambda_1 g_1 + \lambda_2 g_2).$$

Can one determine the λ in such a way that Sturm's three systems of conditions in relation to the three intervals are satisfied?

First of all, one can demand to know whether there exist imaginary characteristic values for λ . By an argument that is very analogous to the preceding one, one will be led to consider the expression:

$$\int_{a_0}^{b_0} \int_{a_1}^{b_1} \int_{a_2}^{b_2} \begin{vmatrix} g_0(x_0) & g_1(x_0) & g_2(x_0) \\ g_0(x_1) & g_1(x_1) & g_2(x_1) \\ g_0(x_2) & g_1(x_2) & g_2(x_2) \end{vmatrix} u_1(x_0)u_2(x_0)v_1(x_0)v_2(x_0)w_1(x_0)w_2(x_0)dx_0dx_1dx_2 = 0,$$

and one can complete the proof as before if the determinant that enters into the triple integral does not change sign. One will then be certain that all characteristic values are real.

The study of that question for real values of λ was made recently: Richardson extended Klein's oscillation theorem to that case by some methods that are different from ours and in which we suppose essentially that the difference under study $G(x, \lambda_n) - G(x, \lambda'_n)$ is a polynomial in x, up to a factor g(x).

CHAPTER V

GREEN'S FUNCTIONS AND THEIR APPLICATIONS (¹⁹)

20. Existence and fundamental properties of Green functions. – As one knows, the Green function that solves the Dirichlet problem is defined by saying that: It is a harmonic function in the domain considered that is annulled on its boundary and becomes infinite at a point A in the domain like $1/r^{n-2}$ if $n \ge 3$ (n is the dimension of the space, and r is the distance from a variable point M in that domain to the point A), and like $\log 1/r$ when n = 2.

One can seek to give an analogous definition for a one-dimensional space, and one will be led to imagine the equation:

$$\frac{d^2u}{dx^2}=0$$

and a solution to that equation that is annulled at a and b and becomes infinite at a point ξ in (a, b). That would be impossible since no such solution exists. One then seeks those of the properties of the Green function that are capable of being extended to ordinary linear differential equations.

From that standpoint, we shall appeal to the analogy between linear systems of algebraic equations and differential systems.

Imagine the system:

(1)
$$a_{i1} u_1 + \ldots + a_{in} u_n = b_i$$
 $(i = 1, 2, \ldots, n)$

whose determinant $|a_{ik}| \neq 0$. One can demand to have a formula for solving it that makes the role of the b_i explicit. In order to present it simply, one can consider the *n* systems that are obtained by successively replacing $b_1, b_2, ..., b_n$ with 1 in (1), while all of the other coefficients b_i are replaced with zero, and one will have a system such as:

BÔCHER, Ann. Math. 13 (1911), pp. 71.

PICARD, Traité d'Analyse, t. III, Chapter 6.

^{(&}lt;sup>19</sup>) BIRKHOFF, Trans. Amer. Math. Soc. 9 (1908), pp. 377.

BOUNITZKY, J. de. Math. pure et appl. (6) 5 (1909), pp. 65.

For the relationship between differential systems and integral equations (but only for the case in which the system is its own adjoint), *see*:

HILBERT, Gött. Nachr. (1904), Zweite Mitteilung.

For the method of successive approximations in some particular cases, see:

STEKLOFF, Ann. Fac. Sci. Toulouse 3 (101), pp. 281.

KNESER, Math. Ann. 58 (1903), 109-116.

(2)
$$\begin{cases} a_{11} u_1 + a_{12} u_2 + \dots + a_{1n} u_n = 1, \\ a_{21} u_1 + a_{22} u_2 + \dots + a_{2n} u_n = 0, \\ \dots \\ a_{n1} u_1 + a_{n2} u_2 + \dots + a_{nn} u_n = 0, \end{cases}$$

whose solutions we shall call u'_1 , u'_2 , ..., u'_n .

We will then have *n* systems that do not depend upon $b_1, ..., b_n$:

$$u'_1, u'_2, \cdots u'_n, u''_1, u''_2, \cdots u''_n, \vdots \vdots \vdots \ddots \vdots u_1^{(n)}, u_2^{(n)}, \cdots u_n^{(n)}.$$

We will see, in turn, that the solution to the system (1) is:

(3)
$$\begin{cases} u_1 = b_1 u'_1 + b_2 u''_1 + \dots + b_n u_1^{(n)}, \\ u_2 = b_1 u'_2 + b_2 u''_2 + \dots + b_n u_2^{(n)}, \\ \dots \\ u_n = b_1 u'_n + b_2 u''_n + \dots + b_n u_n^{(n)}. \end{cases}$$

An expression that is analogous to the solution of the Poisson equation:

$$\Delta\left(u\right)=r\left(x,y\right)$$

and vanishes on the boundary can be provided with the aid of the two-dimensional Green function $G(x, y; \xi, \eta)$, which gives the solution as an explicit function of r(x, y); it is:

$$u(x, y) = \iint r(\xi, \eta) G(x, y; \xi, \eta) d\xi d\eta.$$

One seeks to define the Green function for differential equations with boundary conditions in such a way that it will permit one to express the solutions of inhomogeneous differential systems in such a way that it will exhibit the right-hand sides of those systems.

Take the system:

(4)
$$\begin{array}{c} L(u) = r(x), \\ U_i(u) = 0 \qquad (i = 1, 2, ..., n). \end{array}$$

L(u) is a linear expression of order n.

We suppose that the system:

Chapter V

(5)
$$L(u) = 0, \\ U_i(u) = 0$$

is incompatible. (4) will then have exactly one solution. We shall specify it as a function of r(x).

Since the system (5) has no solution that is not $\equiv 0$, take a point ξ in *a*, *b* and try to determine a function *u* that verifies the conditions on the system (5) that has continuous derivatives u', u'', ..., $u^{(n-2)}$ in *a*, *b*, and one derivative $u^{(n-1)}$ that is continuous, except at ξ , where its discontinuity will be:

$$u^{(n-1)}(\xi+0)-u^{(n-1)}(\xi-0)=\frac{1}{l_n(\xi)},$$

in which $l_n(\xi)$ is the coefficient of $\frac{d^n u}{dx^n}$ in L(u), and finally satisfies L(u) = 0 at every point in (a, b)

b), except at the point ξ . I say that this solution is unique, and I call it $G(x, \xi)$.

Indeed, let $y_1, ..., y_n$ be fundamental system of integrals of L(u) = 0. We shall try to determine $c_1, c_2, ..., c_n$ in such a way that:

$$u_1(x) = c_1 y_1 + \ldots + c_n y_n$$

represents $G(x, \xi)$ in terms of (a, ξ) and d_1, d_2, \dots, d_n in such a way that:

$$u_2(x) = d_1 y_1 + \ldots + d_n y_n$$

represents $G(x, \xi)$ in terms of (ξ, b) .

1. In order for $G(x, \xi)$ to have the desired discontinuity at ξ , it is necessary that:

If one takes $z_i = d_i - c_i$, for the time being, then one will have the following equations for determining the z_i :

(6)
$$\begin{cases} z_1 y_1 \quad (\xi) + \cdots + z_n y_n \quad (\xi) = 0, \\ z_1 y'_1 \quad (\xi) + \cdots + z_n y'_n \quad (\xi) = 0, \\ \vdots \qquad \cdots \vdots \qquad \vdots \\ z_1 y_1^{(n-1)} (\xi) + \cdots + z_n y_n^{(n-1)} (\xi) = \frac{1}{l_n(\xi)}. \end{cases}$$

The determinant of those equations is the Wronskian of y_1 (ξ), ..., y_n (ξ). It is non-zero, by hypothesis, so the z_i are then determined uniquely.

2. Let us write out what it means to say that the boundary conditions (5) are verified by G. Let:

$$U_{i}(u) = [\alpha_{i}u(a) + \dots + \alpha_{i}^{(n-1)}u^{(n-1)}(a)] + [\beta_{i}u(b) + \dots + \beta_{i}^{(n-1)}u^{(n-1)}(b)]$$

= $A_{i}(u) + B_{i}(u)$.

Now:

$$A_i(G) = A_i(u_1),$$

 $B_i(G) = B_i(u_2).$

Therefore, the equations are:

$$A_i(u_1) + B_i(u_2) = 0$$
 (*i* = 1, 2, ..., *n*).

Upon making it explicit that:

$$A_i(u_1) = c_1 A_i(y_1) + c_2 A_i(y_2) + \dots + c_n A_i(y_n),$$

$$B_i(u_2) = d_1 B_i(y_1) + d_2 B_i(y_2) + \dots + d_n B_i(y_n),$$

and remarking that:

 $c_i=d_i-z_i,$

(7)
$$d_1 U_i(y_1) + \dots + d_n U_i(y_n) = z_1 A_i(y_1) + z_2 A_i(y_2) + \dots + z_n A_i(y_n) (i = 1, 2, \dots, n).$$

Those equations determine the d_i when one knows the z_i since the system (5) is incompatible, so the determinant:

$$\begin{vmatrix} U_1(y_1) & \cdots & U_1(y_n) \\ U_2(y_1) & \cdots & U_1(y_n) \\ \vdots & \cdots & \vdots \\ U_n(y_1) & \cdots & U_n(y_n) \end{vmatrix} \neq 0.$$

The c_i are also determined then.

We must then determine a function $G(x, \xi)$ that satisfies our desiderata. It is unique.

From the formulas that give the z_i , and then the d_i and the c_i , it is obvious that those quantities are continuous in x, and as a result, $G(x, \xi)$ will be continuous in (x, ξ) .

The same thing will be true of the derivatives:

$$\frac{\partial G}{\partial x}$$
, ..., $\frac{\partial^{n-2}G}{\partial x^{n-2}}$

We shall now show that the solution to the semi-homogeneous system (4) is very simple with the aid of $G(x, \xi)$ and that the unique solution u(x) to that system is given by the formula:

(8)
$$u(x) = \int_a^b r(\xi) G(x,\xi) d\xi,$$

just as it is for the Poisson problem. Indeed, due to the continuity that was just mentioned, we will have:

$$u'(x) = \int_{a}^{b} r(\xi) \frac{\partial}{\partial x} G(x,\xi) d\xi,$$

...,
$$u^{(n-2)}(x) = \int_{a}^{b} r(\xi) \frac{\partial^{n-2}}{\partial x^{n-2}} G(x,\xi) d\xi,$$

for the (n - 2) first derivatives of the function *u* that was defined by (8). Since $G^{(n-1)}(x,\xi)$ is discontinuous for $x = \xi$, we write:

$$u^{(n-2)}(x) = \int_a^x r(\xi) \frac{\partial^{n-2}}{\partial x^{n-2}} G(x,\xi) d\xi + \int_x^b r(\xi) \frac{\partial^{n-2}}{\partial x^{n-2}} G(x,\xi) d\xi.$$

Each of those integrals can be differentiated, and we will have:

$$u^{(n-2)}(x) = \int_{a}^{x} r(\xi) \frac{\partial^{n-2}}{\partial x^{n-2}} G(x,\xi) d\xi + \int_{x}^{b} r(\xi) \frac{\partial^{n-2}}{\partial x^{n-2}} G(x,\xi) d\xi$$
$$+ \begin{cases} \left[r(\xi) \frac{\partial^{n-2}}{\partial x^{n-2}} G(x,\xi) \right] & (\text{for } \xi = x - 0), \\ - \left[r(\xi) \frac{\partial^{n-2}}{\partial x^{n-2}} G(x,\xi) \right] & (\text{for } \xi = x + 0). \end{cases}$$

Since the derivative $\frac{\partial^{(n-2)}}{\partial x^{n-2}}G(x,\xi)$ is continuous for $x = \xi$, one will further have:

$$u^{(n-1)}(x) = \int_a^b r(\xi) \frac{\partial^{n-1}}{\partial x^{n-1}} G(x,\xi) d\xi$$

The same calculation is performed with $u^{(n)}(x)$. However, the correction term (which was zero for $u^{(n-1)}$) will have the value:

$$r(\xi) \left[\frac{\partial^{n-1} G(x,\xi)}{\partial x^{n-1}} \right]_{x=\xi+0} - r(\xi) \left[\frac{\partial^{n-1} G(x,\xi)}{\partial x^{n-1}} \right]_{x=\xi-0} = \frac{r(x)}{l_n(x)}$$

here. Hence:

$$u^{(n)}(x) = \int_a^b r(\xi) \frac{\partial^n}{\partial x^n} G(x,\xi) d\xi + \frac{r(x)}{l_n(x)}$$

It results from this calculation that the expression:

$$u(x) = \int_a^b r(\xi) G(x,\xi) d\xi$$

verifies the equation:

$$L(u) = r$$
.

If one defines $U_i(u)$ then one will see that:

$$U_i(u) = \int_a^b r(\xi) U_i(G) d\xi$$

and since $U_i(G) = 0$, one will have:

$$U_{i}\left(u\right)=0$$

The expression $\int_{a}^{b} r(\xi) G(x,\xi) d\xi$ then represents the unique solution to the system (4).

In the preceding, we just regarded G as no longer a function of x, but a function of ξ . We now demand to know, in general, what is the nature of G when it is regarded as a function ξ ? We find the following remarkable result here:

When regarded as function of ξ , $G(x, \xi)$ is the Green function of the adjoint system to (5), in which ξ is, of course, the independent variable of that system, and ξ is the singular point of the Green function.

In order to prove that theorem, one lets $H(x, \xi)$ denote the Green function of the adjoint system, in which x is the independent variables, as in (5). Consider two arbitrary points, ξ_1 , ξ_2 , of a, b. In order to be more specific, we suppose that $\xi_1 < \xi_2$. We then apply Green's formula by setting:

$$u = G(x, \xi_1),$$

$$v = H(x, \xi_2),$$

and upon first taking a, $\xi_1 - \varepsilon$ to be the limit of integration $\xi_1 + \varepsilon$, $\xi_2 - \varepsilon$, and finally $\xi_2 + \varepsilon$, b. The integrals in those three formulas reduce to zero, since L(G) = 0, M(H) = 0. If one adds the three formulas then one will get the result:

$$[P(G,H)]_{n}^{\xi_{1}-\varepsilon} + [P(G,H)]_{\xi_{1}-\varepsilon}^{\xi_{2}-\varepsilon} + [P(G,H)]_{\xi_{1}+\varepsilon}^{b} = 0$$

Due to the boundary conditions that are satisfied by G and H, one easily sees that the two terms on the left-hand side that refer to the points a and b will vanish. Finally, upon recalling the explicit expression for P(u, v) that was cited in Chapter II, one will see that most of the other terms will vanish in the limit $\varepsilon = 0$, and what will remain is:

$$H(\xi_1, \xi_2) = G(\xi_2, \xi_1).$$

The proof proceeds in the same way when $\xi_1 > \xi_2$, or when $\xi_1 = \xi_2$. One has then established the identity $H(x, \xi) = G(\xi, x)$, which proves our theorem.

One sees that a differential system that it is own adjoint will admit a Green function $G(x, \xi)$ that is symmetric with respect to its two variables:

$$G(x, \xi) = G(\xi, x).$$

Conversely, if the Green function of a system is symmetric then the system will coincide with its adjoint $(^{20})$. The symmetry of the Green function characterizes the systems that are identical to their own adjoints.

To conclude this section, we point out that the Green function of the system (5) likewise provides the immediate solution to the system:

(9)
$$L(u) = r(x), U_i(u) = \gamma_i \qquad (i = 1, 2, ..., n),$$

which has a unique solution, since (5) is incompatible. In order to make that solution explicit as a function of *r* and the γ_i , we shall be guided by what we said about algebraic equations. Take the *n* systems:

$$\begin{split} L(u) &= 0, \quad L(u) = 0, \quad L(u) = 0, \\ U_1(u) &= 1, \quad U_1(u) = 0, \quad U_1(u) = 0, \\ U_2(u) &= 0, \quad U_2(u) = 1, \quad \vdots \\ \vdots & \vdots & U_{n-1}(u) = 0, \\ U_n(u) &= 0, \quad U_n(u) = 0, \quad U_n(u) = 1, \end{split}$$

which are independent of r and the γ_i . Each of them has a unique solution. Let $G_1(x), \ldots, G_n(x)$, respectively, denote those solutions. An extremely simple calculation will then show that the solution to the system (9) is:

 $[\]binom{20}{1}$ That is a corollary to the more general theorem that two homogeneous and incompatible differential systems are identical if their Green functions are identical.

$$u(x) = \int_a^b G(x,\xi) r(\xi) d\xi + \gamma_1 G_1(x) + \dots + \gamma_n G_n(x) ,$$

whose analogy with the solution to a system of algebraic equations is obvious.

In the preceding considerations, we have supposed that a, b are the limits of the interval in which x varies. We can drop that restriction.

Indeed, take two arbitrary points *a*, *b* in the interval (*A*, *B*) of variation for *x* and suppose that the conditions $U_i(u)$ on the systems (4), (5), or (9) relate to those points *a*, *b*. We then try to define the Green function outside of (*a*, *b*) (²¹). In order to do that, it is good to return to the case in which *a*, *b* are the extremities in order to see how we can modify the definition of the Green function.

Let $G_a(x, \xi)$ and $-G_b(x, \xi)$ denote the two functions $u_1(x)$ and $u_2(x)$, resp., that have served to define $G(x, \xi)$. One has:

(10)
$$u(x) = \int_{a}^{x} r(\xi) G_{a}(x,\xi) d\xi + \int_{b}^{x} r(\xi) G_{b}(x,\xi) d\xi$$

as the solution to the system (4) in (a, b).

When one puts the solution into that form, it is easy to pass to the case in which *a*, *b* are no longer the extremities by abandoning the idea of a Green function in order to adopt the idea of a pair of Green functions.

One can take $G_a(x, \xi)$ and $G_b(x, \xi)$ to be two functions that verify the equation in (A, B) (²²):

$$L\left(u\right) =0$$

when they are regarded as functions of *x* and satisfy the conditions:

(11)
$$\begin{cases} G_{a}(\xi,\xi) + G_{b}(\xi,\xi) = 0, \\ \frac{\partial}{\partial x} [G_{a}(x,\xi) + G_{b}(x,\xi)]_{x=\xi} = 0, \\ \vdots \\ \frac{\partial^{n-2}}{\partial x^{n-2}} [G_{a}(x,\xi) + G_{b}(x,\xi)]_{x=\xi} = 0, \\ \frac{\partial^{n-1}}{\partial x^{n-1}} [G_{a}(x,\xi) + G_{b}(x,\xi)]_{x=\xi} = \frac{1}{l_{n}(\xi)}, \end{cases}$$

and

(12)
$$A_i(G_b) = B_i(G_a)$$
 $(i = 1, 2, ..., n),$

in addition.

^{(&}lt;sup>21</sup>) It is, above all, when a, b are two points that are interior to the *domain of the complex variable x* that this extension becomes interesting. Of course, the coefficients of L(u) = r are supposed to be analytic.

 $^(^{22})$ Or in the domain of the complex variable x.

One immediately sees that two such functions exist and are well-determined. The calculations that are made for the case in which (a, b) coincides with (A, B) are once more valid, so formula (10) is valid entirely. Finally, one sees the method that one would have to follow in order to extend the Green function to the case in which the conditions $U_i(u)$ relate to more than two points of (A, B).

21. The relationship between the theory of differential systems and that of integral equations. – Consider a linear differential system. Here, we shall suppose that we can put it into the form:

(1)
$$L(u) = g(x)u + r, U_i(u) = \gamma_i \qquad (i = 1, 2, ..., n),$$

such that the system:

(2)
$$\begin{aligned} L(u) &= 0, \\ U_i(u) &= 0 \end{aligned}$$

is incompatible. (Later on, we shall say a few words about the possibility of carrying out such an operation.)

One can then find an integral equation that is equivalent to the system (1), i.e., it has the same solutions as that system. Indeed, consider the Green function of the system (2), as well as the solutions $G_1(x), \ldots, G_n(x)$ of the *n* systems that are obtained by replacing one of the conditions $U_i(u) = 0$ in (2) with $U_i(u) = 1$ without changing the other. Let u_1 be an arbitrary solution of (1), and imagine that the system:

(3)
$$L(u) = g(x)u_1 + r,$$
$$U_i(u) = \gamma_i.$$

It has a unique solution because (2) is incompatible, and it is u_1 . One will then have:

$$u_{1}(x) = \gamma_{1} G_{1}(x) + \ldots + \gamma_{n} G_{n}(x) + \int_{a}^{b} [g(\xi)u_{1}(\xi) + r(\xi)]G(x,\xi)d\xi,$$

as one saw in section 20.

If one regards that equation as the one that determines $u_1(x)$ then when one sets:

$$f(x) = \gamma_1 G_1(x) + \gamma_2 G_2(x) + \ldots + \gamma_n G_n(x) + \int_a^b r(\xi) G(x,\xi) d\xi,$$

and

$$K(x, \xi) = g(\xi) G(x, \xi),$$

it will be an integral equation that is written:

(4)
$$u(x) = f(x) + \int_a^b K(x,\xi) u(\xi) d\xi$$

Equation (4) then admits all solutions to the differential system (1).

Let u_1 be any solution of equation (4) then. Now imagine the system (3), in which the righthand side u_1 is the preceding solution. That system has a unique solution u_2 that is given by:

$$u_{2}(x) = \gamma_{1} G_{1}(x) + \ldots + \gamma_{n} G_{n}(x) + \int_{a}^{b} [g(\xi)u_{1}(\xi) + r(\xi)]G(x,\xi) d\xi$$

Now, u_1 verifies the equation:

$$u_{1}(x) = \gamma_{1} G_{1}(x) + \ldots + \gamma_{n} G_{n}(x) + \int_{a}^{b} [g(\xi)u_{1}(\xi) + r(\xi)]G(x,\xi) d\xi,$$

as one will see upon replacing f and K with their values in (4). Thus, $u_2(x) = u_1(x)$. It will then follow that u_2 is a solution to not only (3), but also (1). The system (1) will then admit any solution to equation (4).

Any differential system (1) is equivalent to an integral equation (4).

If $r \equiv 0$, $\gamma_i = 0$ then the system (1) is homogeneous. The same thing is true for equation (4) because $f \equiv 0$. Conversely, if $f \equiv 0$, u = 0 verifies (4) and also (1). Hence, (1) is homogeneous.

Recall that the two equations:

(5)
$$u(x) = \int_{a}^{b} K(x,\xi) u(\xi) d\xi$$

and

(6)
$$v(x) = \int_a^b K(\xi, x) v(\xi) d\xi,$$

which we call *adjoint equations*, are closely related to each other. We shall show the parallelism that this situation exhibits with regard to adjoint differential systems.

Suppose that equation (5) is equivalent to the system:

(7)
$$L(u) = g u, U_i(u) = 0 \qquad (i = 1, 2, ..., n),$$

whose adjoint system is:

(8)
$$M(v) = g v, V_i(v) = 0 \quad (i = 1, 2, ..., n).$$

Since the Green function of L(u) = 0, $U_i(u) = 0$ is $G(x, \xi)$, one sees that the Green function of M(v) = 0, $V_i(v) = 0$ will be $G(\xi, x)$.

The integral equations that are equivalent to the systems (7) and (8) will then be:

$$u(x) = \int_{a}^{b} [g(\xi)G(x,\xi)]u(\xi)d\xi,$$
$$v(x) = \int_{a}^{b} [g(\xi)G(\xi,x)]v(\xi)d\xi,$$

although their relationship to the preceding is not that they are two adjoint equations. Nonetheless, if one writes:

w(x) = g(x) v(x)

$$w(x) = \int_a^b g(x)G(\xi, x)w(\xi)d\xi,$$

which is indeed the adjoint of the integral equation that is equivalent to (7).

Therefore, there is a parallelism between two adjoint integral equations and two adjoint differential systems.

The link that we just pointed out permits us to prove some of the results that we gave in the second chapter simply by appealing to the theory of integral equations. For example, according to Fredholm, the number of linearly-independent solutions to an integral equation is equal to that of its adjoint, and some linearly-independent functions v correspond to functions w that are also that way, and conversely, one sees that the index of a differential system is equal to that of its adjoint.

Nonetheless, the study of differential systems leads to some very specialized kernels $K(x, \xi)$ for equivalent integral equations, and the properties of some differential systems often imply those specialized forms for their kernels in such a way that few of the results of the theory of differential systems are provided by the theory of integral equations of the general type. The direct method is generally preferable for the advanced study of differential systems.

The question of whether it is possible to write a differential system in the form:

$$L(u) = g(x)u + r,$$

 $U_i(u) = \gamma_i$ (*i* = 1, 2, ..., *n*),

such that the system:

$$L(u) = 0,$$
$$U_i(u) = 0$$

is incompatible demands a proof that is rather long, and we shall not give it $(^{23})$. We shall say only that not only can one always choose g(x) in such a fashion as to obtain the preceding form, but one can choose it to be real and positive in $a \le x \le b$, and that can be done in an infinitude of ways.

^{(&}lt;sup>23</sup>) I gave that proof in the Bulletin of the American Mathematical Society for October, 1914, pp. 1.

If one supposes that g(x) > 0 then one can some more precise statements about the integral equation.

One can suppose that one has reduced g to unity by dividing both sides of the differential equation by g(x) [which is always > 0 in (a, b)]. The Green function is obviously altered by that transformation, but in what follows, one will have the advantage of taking the kernel $K(x, \xi)$ to be equal to $G(x, \xi)$.

Without supposing that g has been reduced to 1, one can remark that if the system:

$$L(u) = g u$$
, $U_i(u) = 0$

is its own adjoint then the same thing will be true of:

$$L(u) = 0$$
, $U_i(u) = 0$.

In this case, one has seen that:

$$G(x, \xi) = G(\xi, x).$$

The integral equation to which one is led:

$$u(x) = f(x) + \int_a^b K(x,\xi)u(\xi) d\xi$$

does not have a symmetric kernel because:

$$K(x, \xi) = G(x, \xi) g(\xi),$$

and g is not equal to unity. However, if one imagines:

$$u_0 = u\sqrt{g(x)}$$

then one will see that u_0 verifies the equation with a symmetric kernel:

$$u_0(x) = f(x)\sqrt{g(x)} + \int_a^b \left[\sqrt{g(x)}G(x,\xi)\sqrt{g(\xi)}\right]u(\xi) d\xi.$$

By a small transformation, one can then bring back the theory of systems that are adjoint to themselves to that of equations with symmetric kernels.

Finally, if one has some conditions that relate to a single point a, instead of having $U_i(u)$ that bear upon two points a and b, as in the preceding, then one will know that the system:

$$L(u) = 0$$
, $U_i(u) = 0$,

in which the U_i are independent, reduces to the system:

$$L (u) = 0 ,$$

 $u (a) = 0 ,$
 $u'(a) = 0 ,$
....,
 $u^{(n-1)}(a) = 0 ,$

which is always incompatible.

The Green function $G(x, \xi)$ of that system will obviously be zero from a to ξ . Its $(n - 1)^{\text{th}}$ derivative will have a discontinuity at ξ . $G(x, \xi)$ will no longer be zero at (ξ, b) .

In summary:

$$G(x, \xi) \equiv 0$$
 for $x < \xi$.

In all integrals in which $G(x, \xi)$ is a factor, the part that relates to the interval (a, ξ) will be zero. In particular, the integral equation that is equivalent to a system in which conditions are given at just one point will be a Volterra equation:

$$u(x) = f(x) + \int_a^x K(x,\xi)u(\xi) d\xi$$

22. The method of successive approximations for differential systems. – We conclude by giving an application of Green's functions to the solution of differential systems by successive approximations in a form that is much more general than the one that was considered in the first chapter. We suppose that the differential system is given in the form:

(1)
$$L'(u) = L''(u) + r, U'_i(u) = U''_i(u) + \gamma_i \quad (i = 1, 2, ..., n).$$

L' and L'' are linear and homogeneous differential expressions here, the first of which has order n, while the second one has order less than n, and their coefficients are continuous functions of x.

We suppose, in addition, that the coefficient of $\frac{d^n u}{dx^n}$ in L' has no zeroes in (a, b). The U' and U" are linear forms in $u(a), ..., u^{n-1}(a), u(b), u'(b), ..., u^{n-1}(b)$.

Along with the system (1), consider the homogeneous system:

(2)
$$L'(u) = 0, U'_{i}(u) = 0 \quad (i = 1, 2, ..., n),$$

which we assume to be incompatible. Let $G(x, \xi)$ be the Green system of the system (2), and let $G_1(x), G_2(x), ..., G_n(x)$ be the supplementary functions that are constructed as in section **20**.

Start with an arbitrary function $u_0(x)$ and define the functions $u_1, u_2, ...$ that satisfy the equations:

$$L'(u_m) = L''(u_{m-1}) + r,$$

$$U'_i(u_m) = U''_i(u_{m-1}) + \gamma_i \qquad (i = 1, 2, ..., n).$$

Those functions are determined uniquely since the system (2) is incompatible. If one writes:

$$u_1 = v_1$$
, $u_2 - u_1 = v_2$, $u_3 - u_2 = v_3$, ...

then one will see that for $m \ge 2$, one will have:

(3)
$$v_m = U_1''(v_{m-1})G_1(x) + \dots + U_n''(v_{m-1})G_n(x) + \int_a^b L''[v_{m-1}(\xi)]G(x,\xi)d\xi$$

One then proves, in a manner that is entirely analogous to the one in section 1, that a solution u to the system (1) is given by the series:

(4)
$$v_1 + v_2 + v_3 + \dots,$$

and the derivatives u', u'', ..., $u^{(n-1)}$ are given by the series:

(5)
$$v_1^{(k)} + v_2^{(k)} + \cdots$$
 $(k = 1, 2, ..., n-1),$

provided that all of those series (4), (5) converge uniformly in (a, b). That convergence will not always be true, even in the case where the system (1) has a unique solution. In order to treat that question, let A denote the largest maximum of the moduli of the functions:

$$G(x, \xi), \quad \frac{\partial G}{\partial x}, \quad \dots, \quad \frac{\partial^{n-1}G}{\partial x^{n-1}},$$

$$G_k(x), \quad G'_k(x), \quad \dots, \quad G_k^{(n-1)}(x) \qquad (k=1, 2, \dots, n).$$

Let Σ be the sum of the moduli of the coefficients of all of the expressions U''_i , and let F(x) be the sum of the moduli of the coefficients of L''. Finally, let *B* denote the constant:

$$B=\Sigma+\int_{a}^{b}F(x)\,dx\,.$$

One easily sees then that all of the series (4), (5) are absolutely and uniformly convergent if B < 1/A. Since A depends upon only L' and the U'_i , while B depends upon only L" and the U''_i , the method of successive approximations will certainly converge if the moduli of the coefficients of L" and the U''_i are very small when L' and the U'_i are given.

Therefore, introduce a parameter λ and consider the system:

(6)
$$L'(u) = \lambda [L''(u) + r''] + r', U'_i(u) = \lambda [U''_i(u) + \gamma''_i] + \gamma'_i \quad (i = 1, 2, ..., n),$$

in which r' + r'' = r, $\gamma'_i + \gamma''_i = \gamma_i$. The system (6) reduces to the system (1) when $\lambda = 1$, and on the other hand, from what we have said, the method of successive approximations will certainly converge when it is applied to the system (6) when the modulus of λ very small.

From (3), the series that are analogous to (4) and (5) that one obtains by integrating the system (6) will be series of powers of λ provided that $u_i = v_i$ does not depend upon λ , which will be true when the functions u_0 is subject to the conditions:

$$L''(u_0) + r'' = 0,$$

$$U''_i(u_0) + \gamma''_i = 0 \qquad (i = 1, 2, ..., n).$$

We shall suppose in what follows that those equations are satisfied. The power series that one obtains by integrating the system (6) by the method of successive approximations will certainly converge then if $|\lambda|$ is very small. The question is then whether it will converge when $\lambda = 1$.

First imagine a more general system that makes (6) only a special case:

(7)
$$L(u) = r, U_i(u) = \gamma_i \quad (i = 1, 2, ..., n),$$

in which the coefficients of L(u) and the function r are analytic in λ in a certain Weierstrass domain and continuous in (x, λ) , while the γ_i and the coefficients of the U_i are analytic in λ . We shall not assume that all values of λ in the Weierstrass domain are characteristic numbers.

Under those conditions, the solution to (7) will be an analytic function in λ and continuous in (x, λ) , except for the characteristic values of λ . In order to prove that proposition, it suffices to consider the function:

in which u_0 is a solution of the equation L(u) = r that is analytic in λ and continuous in (x, λ) , and y_1, \ldots, y_n are functions that are analytic in λ and continuous in (x, λ) that define a fundamental system of the equation L(u) = 0. When λ is not a characteristic number, the function (8), which is analytic in λ and continuous in (x, λ) , will satisfy the system (7). Since the system (7) has a unique solution in this case, the proof is complete.

It then follows that the power series that one obtains by applying the method of successive approximation to the system (6) will converge in all of the disc that is described around the point $\lambda = 0$ as its center and contains no point that corresponds to a characteristic number. That result can also be stated in another form:

The method of successive approximations, as we applied it to the system (1), will converge if the auxiliary system (6) has no characteristic number whose modulus is less than or equal to 1.

As an example, take the case in which all of the coefficients in the U''_i are zero, along with all of the coefficients of u(b), u'(b), ..., $u^{n-1}(b)$ in the U'_i . The system (6) will then possess one and only one solution for all values of λ . Therefore, the method of successive approximations will converge in this case when it is applied to the system (1).

Let now return to the system (7) and its solution (8). Let λ_1 be a characteristic number of (7), i.e., a value of λ for which the denominator of (8) vanishes. In general, that point will be a pole of the function (8), which is a pole whose order cannot be higher than the *multiplicity* of the characteristic λ_1 . If the function (8) does not have a pole at the point λ_1 then it will again be a solution to the system (7) for that value of λ , and we will have before us the exceptional case that was considered at the end of section **9** in which the inhomogeneous system (7) has a solution, and even for a characteristic number.

The method of successive approximations, as we applied it to the system (1), will not converge if the auxiliary system (6) has a characteristic number whose modulus is smaller than 1 for which the system (6) has no solution.

Now consider the case of a characteristic number whose multiplicity and index have the same value *k*. It is a very important special case that will always present itself, for example, when the characteristic numbers have multiplicity 1 (*see*, § **11**). If the system (7) has a solution for $\lambda = \lambda_1$ then the rank of the system:

(9)
$$\begin{pmatrix} U_1(u_0) - \gamma_1 & U_1(y_1) & U_1(y_2) & \cdots & U_1(y_n) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ U_n(u_0) - \gamma_n & U_n(y_1) & U_n(y_2) & \cdots & U_n(y_n) \end{pmatrix}$$

will be equal to n - k for $\lambda = \lambda_1$, which will also be the rank of the determinant in the denominator of (8) (*see*, § 5). It will then follow that the determinants of order *n* of the system (9) have zeroes whose order is greater than or equal *k* or $\lambda = \lambda_1$, as one will see upon taking the successive derivatives of those determinants with respect to λ . Since the denominator in (8) has a zero of order *k* at the point $\lambda = \lambda_1$, by hypothesis, one sees that *the function* (8) *has no pole at that point* in the present case.

Upon applying that result to the system (6), we will get the following theorem:

The method of successive approximations, as we applied it to the system (1), will even converge when the auxiliary system (6) has characteristic numbers whose modulus is ≤ 1 provided that for each of those characteristic numbers:

- 1. The system (6) has a solution, and
- 2. The index is equal to the multiplicity.

That is a generalization of a very special result that was given by Liouville in 1840 (J. de math. pures appl., t. V, pp. 356).