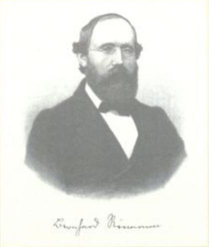
Bernhard Riemann



Collected Papers

Kendrick Press

Translated by
Roger Baker,
Charles Christenson
and Henry Orde

Collected Papers

Bernhard Riemann

Translated from the 1892 edition by Roger Baker, Charles Christenson and Henry Orde

Kendrick Press

Collected Papers
Bernhard Riemann (1826–1866)

Translated by Roger Baker, Charles Christenson and Henry Orde from the papers in 'Bernhard Riemann's Gesammelte Mathematische Werke', 2nd edn., edited by Heinrich Weber, Teubner, Leipzig, 1892.

AMS Subject Classification: 01A75

ISBN numbers:

0-9740427-2-2 (hardback)

0-9740427-3-0 (paperback)

www.kendrickpress.com
520 N. Main St. #341
Heber City, UT 84032, USA

Preface

Bernhard Riemann (1826–1866) was one of the greatest mathematicians of the modern era. He aimed very deliberately at new ways of thinking about existing problems and concepts in mathematics, often with startling success. Looking for a quantitative way to support the assertion that he influenced twentieth century mathematics more than his contemporaries, I turned to the collection 'Development of Mathematics 1900–1950', edited by J-P. Pier. Riemann is mentioned in the index as many times as Gauss, Cauchy, Weierstrass and Dedekind combined.

Over the years, I have often talked to mathematicians and students who would like to read Riemann's papers, but find this difficult or time-consuming in the original German. This is the audience I had in mind in organizing the present volume. A few of the papers have been translated before, but the translations in this book are all new.

Riemann was fortunate enough to attend the lectures (on the least squares method) of the legendary C.F. Gauss (1777–1855) as a young student in Göttingen. He came to idolize Gauss later, but in 1847 he traveled to Berlin for greater stimulus. There he attended the lectures, and took part in the discussions, of C.G.J. Jacobi (1804–1851) and P.G.L. Dirichlet (1805–1859). Both these scholars became important influences on Riemann. He cited Jacobi's great memoir 'Fundamenta nova theoriae functionum ellipticarum' in many places in his own work, and it suggested possibilities to him that went well beyond Jacobi's beautiful calculations. Dirichlet, on the other hand, liked to take an abstract approach to each topic, and Riemann much preferred this way of attacking problems. Dirichlet was also an enthusiastic proponent of mathematical models of physical problems, and Riemann was to carry Dirichlet's approach further in both his writings and his lectures. Returning to Göttingen, he also became fired with enthusiasm for experimental physics and the construction of theoretical explanations for the new phenomena that were being observed. Here his great mentor was Wilhelm Weber (1804–1891).

As a young researcher, Riemann was drawn in many directions. However, his doctoral thesis of 1851 on the foundations of the theory of functions of a complex variable took on great importance in suggesting to him several further lines of research that needed to be carried through. The thesis also became a watershed in the subject, and in 1951 was celebrated with a centennial conference. There were many inspired ideas in the thesis. Riemann surfaces appear there for the first time. Analytic functions are viewed as conformal mappings, and the Riemann mapping theorem is proposed and given a proof, albeit unsatisfactory. Riemann's researches later in the 1850s on hypergeometric functions, Abelian functions and the Riemann zeta function demonstrate his commitment to complex function theory as a tool of exploration. Yet Dieudonné (1985) writes of Riemann's great paper **VI** on Abelian functions as an 'epoch' in algebraic geometry. It took a very long time before all the mathematical ideas in this paper fully bore fruit. Even a brilliant contemporary such as A. Clebsch (1833–1872) felt somewhat defeated by Riemann's memoir, although Clebsch took up the potent idea of genus of an algebraic curve from it. The paper **VII** on the Riemann zeta function, treasured by number theorists as a gem, is the source of the most famous unsolved problem in mathematics, the 'Riemann hypothesis'.

Simply in preparing for his habilitation in 1854, Riemann wrote two works of genius—his habilitation thesis on trigonometric series, and his 'trial lecture' on the foundations of geometry. The methods in the thesis can still be found intact in modern works such as Zygmund's 'Trigonometric Series'. The trial lecture became the inspiration for a new era in differential geometry, and Einstein's theory of general relativity is its (not very indirect) descendant. Yet Riemann never published either the thesis or the trial lecture—this was left to his friend and colleague Richard Dedekind (1831–1916) after Riemann's death. In 1854, these works simply qualified Riemann to be a poorly paid instructor at Göttingen.

It is hard to believe, yet while working on these wonderful ideas, Riemann still had a great deal of time to think about problems of physics. Of the nine papers that he published during his lifetime, four are on problems of mathematical physics. (Riemann's total of publications up to 1866 is brought up to eleven by announcements of his papers on the hypergeometric functions and the propagation of sound waves.) Much of his writing on physics from the 1850s was not submitted for publication.

It is difficult to guess the directions in which Riemann might later have focused his extraordinary abilities. He succeeded Dirichlet as professor at Göttingen in 1859, but his health deteriorated from 1862 onwards. He remained dedicated to scholarship to the last; this touching story is told here in an essay by Dedekind. Among his contemporaries, Dedekind was the one who most fully appreciated Riemann's mathematics.

After Riemann died in Italy in 1866, Dedekind and others oversaw the publication of seven posthumous papers. The early death of Clebsch delayed the appearance of the collected works, but eventually Heinrich Weber (1842–1913) was able to send the first edition to the press in 1876. Weber was ably assisted by Dedekind and by H.A. Schwarz (1848–1921). A dozen papers in the first edition were assembled by poring over Riemann's *Nachlass*, the mass of materials left behind at his death. The 1892 edition, which is the source for the present translations, contained some modest additions and corrections, and its numbering **I–XXXI** of the papers is preserved here for the reader's convenience. However, I omitted three non-mathematical items: **XVIII**, 'The mechanism of the ear', and two fragments on philosophy that can be found on pages 509–525 of Weber (1892).

Footnotes in the text of the papers are mostly Riemann's. However, some footnotes making Riemann's references to the literature more precise were added by Weber. These are indicated by a W.

There is far more in Riemann's work than any one person can comprehend if the influence of the papers is to be properly considered. Nevertheless, I provide notes at the end of the book containing basic information about each paper and suggestions for further reading.

A natural next step would be the book of Laugwitz (1999), which gives a unified account of Riemann as a mathematician and natural philosopher. Riemann also had a very important role as teacher and expositor through the revision and publication of his lectures by Hattendorff, Stahl and others.

Readers who would like to suggest corrections or alternative readings within the papers, or additional remarks for the notes, should send these to baker@math.byu.edu. I will maintain a web page that takes these suggestions into account, at http://www.math.byu.edu/~baker/Riemann/index.htm

During the final stages of the preparation of this book, Henry Orde, one of my fellow translators, died in Kent, England. Henry was born in 1922. He served in the British armed forces throughout the second world war, and then studied mathematics at Cambridge. After working for a firm of merchants in Malaya for several years, he joined the nascent computer industry in England. A heart attack forced him into retirement at age 50, and he was then able to devote his energies to pure mathematics, book collecting and music. Number theorists will recall that he gave an elementary proof of the class number formula for quadratic fields with negative discriminant. See Orde (1978).

Roger Baker

Contents

I.

Foundations for a general theory of functions of a complex variable.

(*Inaugural dissertation, Göttingen 1851; second printing (unchanged), Göttingen 1867.*)

1.

Denote by z a variable that can take successively all possible real values. When there is a unique value of the variable w corresponding to each z, we say that w is a function of z. If w varies continuously when z runs continuously over all values between two given points, we say that the function is continuous in this interval.

This definition clearly enforces no law between individual values of the function. For when the function is specified in a certain interval, the method of continuing it outside the interval remains entirely arbitrary.

The dependence of the quantity w on z can be expressed by a mathematical law, so that definite operations at each value of z yield the corresponding w. The possibility of a single law of dependence for all values of z in a given interval was formerly ascribed only to a certain class of functions (*functiones continuae* in Euler's terminology). More recent researches have shown, however, that there are analytic expressions that represent each continuous function on a given interval. Hence it is one and the same thing to say that w depends on z in some arbitrary given manner; or that w is given by definite operations. Both notions are equivalent, in view of the results mentioned.

This is not the case, however, if z is not restricted to real values, but varies over complex numbers of the form $x + yi$ (where $i = \sqrt{-1}$).

Let

$$x + yi, \quad x + yi + dx + dyi$$

be two values of the quantity z with an infinitely small difference, and let

$$u + vi, \quad u + vi + du + dvi$$

be the corresponding values of w. If the dependence of w on z is defined arbitrarily, the ratio

$$\frac{du + dvi}{dx + dyi}$$

varies, generally speaking, according to the values of dx and dy. Indeed, let $dx + dyi = \epsilon e^{\phi i}$; then

$$\begin{aligned}\frac{du + dvi}{dx + dyi} &= \frac{1}{2}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + \frac{1}{2}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) i \\ &\quad + \frac{1}{2}\left[\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) + \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right) i\right] \frac{dx - dyi}{dx + dyi} \\ &= \frac{1}{2}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + \frac{1}{2}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) i \\ &\quad + \frac{1}{2}\left[\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} + \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right) i\right] e^{-2\phi i}.\end{aligned}$$

In whatever way w is determined from z by a combination of simple operations, the value of the derivative $\frac{dw}{dz}$ will always be independent[1] of the particular value of the differential dz. Obviously we cannot obtain arbitrary dependence of complex w on complex z in this way.

The above characteristic, common to all functions obtained via any operations, will be fundamental to the following investigation, where such a function will be treated independently of its expression. Without proving the general sufficiency and validity of the definition of dependence via operations, we take the following definition as starting point.

A complex variable w is said to be a function of another complex variable z, if w varies with z in such a way that the value of the derivative $\frac{dw}{dz}$ is independent of the value of the differential dz.

2.

Both quantities z and w will be treated as variables that can take every complex value. It is significantly easier to visualize variation over a connected two-dimensional domain, if we link it to a spatial viewpoint.

We represent each value $x + yi$ of the quantity z by a point O of the plane A having rectangular coordinates x, y; and every value $u + vi$ of the quantity w by a point Q of the plane B, having rectangular coordinates u, v. Dependence of w on z is then represented by the dependence of the position of Q on the position of O. When w corresponds to z in a way that varies

[1]This assertion is obviously justified in all the cases where one can obtain from the expression of w in terms of z, using the rules of differentiation, an expression for $\frac{dw}{dz}$ in terms of z. The rigorous general validity of the assertion is left aside for now.

continuously with z, or in other words when u, v are continuous functions of x, y, then to every point of the plane A corresponds a point of the plane B; generally speaking, to every line corresponds a line, and to every connected piece of surface there is a corresponding connected piece of surface. Thus we can think of this dependency of w on z as a mapping of the plane A on the plane B.

3.

We now investigate the properties that this mapping has when w is a function of z, that is, when dw/dz is independent of z.

We denote by o a general point of the plane A in the neighborhood of O, and the image of o in the plane B by q. Further let $x + yi + dx + dyi$ and $u + vi + du + dvi$ be the values of z and w at these points. We may view dx, dy and du, dv as rectangular coordinates of the points o and q with respect to the points O and Q taken as origins. If we write $dx + dyi = \epsilon e^{\phi i}$ and $du + dvi = \eta e^{\psi i}$, then the quantities $\epsilon, \phi, \eta, \psi$ become polar coordinates of these points relative to these origins. Now let o' and o'' be any two points infinitely close to O. For quantities depending on o', o'', we use the above notations with appropriate accents. By hypothesis,

$$\frac{du' + dv'i}{dx' + dy'i} = \frac{du'' + dv''i}{dx'' + dy''i}.$$

Consequently

$$\begin{aligned}\frac{du' + dv'i}{du'' + dv''i} &= \frac{\eta'}{\eta''}\, e^{(\psi' - \psi'')i} = \frac{dx' + dy'i}{dx'' + dy''i} \\ &= \frac{\epsilon'}{\epsilon''}\, e^{(\phi' - \phi'')i}\end{aligned}$$

and so

$$\frac{\eta'}{\eta''} = \frac{\epsilon'}{\epsilon''} \quad , \quad \psi' - \psi'' = \phi' - \phi''.$$

That is, in the triangles $o'Oo''$ and $q'Qq''$ the angles $o'Oo''$, $q'Qq''$ are equal, and the corresponding sides are proportional.

This yields the similarity of two corresponding infinitely small triangles; and consequently, in general, similarity of the smallest parts of the plane A and their images on the plane B. An exception to this result occurs only in the special case where the corresponding variations of the quantities z and w are not in finite ratio. We tacitly excluded this exception from our deduction.[2]

4.

If we write the differential quotient $\frac{du+dvi}{dx+dyi}$ in the form

$$\frac{\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial x}\,i\right)\,dx+\left(\frac{\partial v}{\partial y}-\frac{\partial u}{\partial y}\,i\right)\,dyi}{dx+dyi}$$

it is plain that it will have the same value for any two values of dx and dy, exactly when

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y}\quad,\quad\frac{\partial v}{\partial x}=-\frac{\partial u}{\partial y}.$$

Hence this condition is necessary and sufficient for $w = u + vi$ to be a function of $z = x + yi$. For the individual terms of the function, we deduce the following:

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0,\ \frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}=0.$$

This equation is the basis for the investigation of the properties of the individual terms of such a function. We give the proof of the most important of these properties before undertaking a deeper treatment of the complete function. However, we first establish some points concerning more general matters, in order to smooth the ground of the investigation.

5.

For the following treatment we permit x, y to vary only over a finite region. The position of the point O is no longer considered as being in the plane A, but in a surface T spread out over the plane. We choose this wording since it is inoffensive to speak of one surface lying on another, to leave open the possibility that the position of O can extend more than once over a given part of the plane. However, in such a case we suppose that the portions of

[2]On this subject, see: 'General solution of the problem of mapping the portions of a surface so that the image is similar to the original in the smallest parts', C.F. Gauss. (Response to the 1822 prize question proposed by the Royal Society of Sciences in Copenhagen. *Astronomische Abhandlungen*, edited by Schumacher, vol. III, Altona 1825; Gauss, *Collected works*, vol. IV, p. 189.)

surface lying upon one another do not connect along a line. Thus a folding of the surface, or a splitting of the surface into superimposed parts, does not occur.

Now the number of pieces of surface superimposed in each part of the plane is completely determined, when we give the boundary of the region and its direction (that is, its inner and outer sides). A transit of these pieces can take different forms.

Indeed, if we draw a line through the part of the plane covered by the surface, the number of superimposed surfaces only changes on crossing the boundary. Indeed, the number changes on moving from outside to inside by +1, in the opposite case by −1. Thus the number is determined everywhere. Along the edge of the line, each bordering portion of surface continues in a definite way, as long as the line does not meet the contour. For indeterminacy can only occur at an isolated point, and consequently occurs either at a point of the line itself, or at a finite distance from it. If we confine ourselves to part of the line ℓ in the interior of the surface, and both sides of it on a sufficiently narrow strip of surface, we may speak of definite bordering parts of the surface, whose number is equal on either side. Specifying a definite direction of the line, denote the parts of the surface on the left side by $a_1, \ldots, a_n$ and those on the right by $a'_1, \ldots, a'_n$. Each part a will continue into one of the parts a'. Indeed, this will in general be the same part along the length of the line ℓ, except that for certain positions of ℓ it can change at one point. Suppose that above such a point σ (that is, along the preceding part of ℓ) the surface portions $a_1, \ldots, a_n$ are connected respectively with the portions $a'_1, \ldots, a'_n$, while below σ, $a_{\alpha_1}, a_{\alpha_2}, \ldots, a_{\alpha_n}$ connect respectively with $a'_1, \ldots, a'_n$. Here $\alpha_1, \ldots, \alpha_n$ is a permutation of $1, \ldots, n$. A point of a_1 that moves into a'_1, above σ, will pass to a_{σ_1} if we go back to the left side below σ. If a point moves around σ from left to right, the index of the part of surface in which it lies takes successively the values

$$1, \alpha_1, \alpha_{\alpha_1}, \ldots, \mu, \alpha_\mu, \ldots .$$

In this sequence, as long as the term 1 does not recur, the terms are necessarily distinct. For an arbitrary intermediate term α_μ is necessarily preceded by μ, and in direct succession by all preceding terms up to 1. But when after a certain number of terms, say m, a number evidently less than n, the term 1 reappears, the other terms must then recur in the same order. The point moving around σ comes back after every m circuits into the same part of the surface and is restricted to m superimposed surface parts, that meet above σ at a unique point. We call this point a branch point of order $m-1$ of the surface T. Repeating the procedure, the $n-m$ remaining surface parts, provided they are not isolated, divide into systems of $m_1, m_2, \ldots$ surface parts. In this case, further branch points of orders $m_1-1, m_2-1, \ldots$ are located at σ.

When the position and direction of the boundary of T, and the position of the branch points, are given, then T is either completely determined, or restricted to one of a finite number of distinct forms. The latter occurs, in so far as these determining portions can lie on different parts of the superimposed surfaces.

A variable which in general, that is after excluding isolated lines or points,[3] takes a definite value at each point O of the surface T that varies continuously with position, can obviously be considered as a function of x, y. When we refer to functions of x, y in what follows, the term will always be employed in this fashion.

Before passing to the treatment of such functions, we introduce some clarifications on the connectivity of a surface. We restrict ourselves to surfaces that do not split apart along a line.

6.

We regard two parts of surface as connected, or belonging to a single piece, if from any point of one part to any point of the other, a line can be drawn interior to the surface. We regard two parts as separate, when this procedure is not possible.

The study of the connectivity of a surface is based on its decomposition via transverse cuts, that is lines which cut through the interior from one boundary point simply (no point occurring multiply) to another boundary point. The terminal point can lie in the part of the boundary thereby adjoined, and thus at an earlier point of the transverse cut.

A connected surface that is split apart by each transverse cut, is said to

[3]This restriction does not arise from the definition of a function, but is needed for the application of infinitesimal calculus to it. As an example of a function discontinuous at all points of a surface, take the function whose value is 1 for rational x and y, and otherwise is 2. We cannot apply to the function either differentiation or integration, and so cannot directly use infinitesimal calculus. The arbitrary restriction placed here on the surface T will be justified later (Section 15).

be simply connected; otherwise it is multiply connected.

Theorem 1 A simply connected surface is divided by each transverse cut ab into two simply connected pieces.

Suppose that one of the pieces is not split apart by a transverse cut cd. There are three possible cases: neither of the endpoints c, d is situated on ab; only c is on ab; both are on ab. Rejoining the surface respectively along the whole line ab; along the part cb; or along the part cd of the line, we obtain a connected surface that arises from a transverse cut of A, contrary to hypothesis.

Theorem 2 Suppose that a surface T divides via n_1 transverse cuts[4] q_1 into a system T_1 of m_1 simply connected pieces, and via n_2 transverse cuts q_2 into a system T_2 of m_2 pieces. Then $n_2 - m_2 \leq n_1 - m_1$.

Every line q_2, not completely contained in the system of transverse cuts q_1, yields together with the system of cuts q_1 one or several transverse cuts q_2' of the surfaces T_1. As endpoints of the transverse cuts q_2', we have:

1) The $2n_2$ endpoints of the transverse cuts q_2, except for those whose endpoints meet part of the system of lines q_1.

2) Every intermediate point of a transverse cut q_2, at which this cut meets an intermediate point of a line q_1, except for the case where the point is already on another line of q_1, that is, when an end of a cut q_1 coincides with this point.

Denote now by μ the number of times that lines of both systems join or separate (where in consequence an isolated common point is to be counted twice); by ν_1, the number of times an end part of a line q_1 coincides with an intermediate part of a line q_2; by ν_2, the number of times an end part of a line q_2 coincides with an intermediate part of a line q_1; and finally, by ν_3 the number of times an end portion of a line q_1 coincides with an end part of a line q_2. Then case 1) yields $2n_2 - \nu_2 - \nu_3$ endpoints, case 2) yields $\mu - \nu_1$ endpoints, of the transverse cuts q_2'. Both cases taken together yield all endpoints, each one only once. The number of these transverse cuts is

[4]By a decomposition through several transverse cuts we understand a successive decomposition. That is, the surface obtained by one transverse cut is decomposed further by a new transverse cut.

thus

$$\frac{2n_2 - \nu_2 - \nu_3 + \mu - \nu_1}{2} = n_2 + s.$$

In an entirely analogous way we conclude that the number of transverse cuts q_1' of the surface T_2, formed by the lines q_1, is

$$\frac{2n_1 - \nu_1 - \nu_3 + \mu - \nu_2}{2} = n_1 + s.$$

Now the surface T_1 will obviously be transformed by the $n_2 + s$ transverse cuts q_2' into the same surface obtained from T_2 by the $n_1 + s$ transverse cuts q_1'. But T_1 comprises m_1 simply connected pieces and consequently, by Theorem 1, is decomposed by $n_2 + s$ transverse cuts into $m_1 + n_2 + s$ pieces. Consequently, if $m_2 < m_1 + n_2 - n_1$, the number of pieces of the surface T_2 would be increased by more than $n_1 + s$ by the effect of $n_1 + s$ transverse cuts, which is absurd.

As a consequence of this theorem, if the indeterminate number of transverse cuts is denoted by n, and the number of pieces by m, then $n - m$ is constant for all divisions of the surface into simply connected pieces. For consider any two divisions, by n_1 transverse cuts into m_1 pieces and by n_2 cuts into m_2 pieces. Since the first set of pieces are simply connected, we have

$$n_2 - m_2 \leq n_1 - m_1.$$

Since the second set are simply connected, we have

$$n_1 - m_1 \leq n_2 - m_2;$$

and since both inequalities hold, we have $n_2 - m_2 = n_1 - m_1$.

This number can properly be designated the 'connectivity' of a surface. By definition, it will be diminished by 1 by a transverse cut. It will be unchanged by the effect of a simple cut, starting from a point of the interior and ending either at a point of the boundary, or at a point of a previous cut. It will be increased by 1, by a simple cut in the interior of the surface with two endpoints. For in the first case, the cut becomes a transverse cut, if we make a new transverse cut. In the latter case, two new transverse cuts are needed.

Finally the connectivity of one surface formed from several pieces is obtained by adding the connectivities of these pieces.

In the following, we usually restrict ourselves to surfaces that have only one piece, and for their connectedness we simply speak of simple, double, and so on. Here we understand, by an n-fold connected surface, a surface that can be decomposed by $n-1$ transverse cuts into one that is simply connected.

As to the dependence of the connectedness of the boundary on the connectedness of the surface, it is clear that:

1) the boundary of a simply connected surface necessarily comprises a single closed line.

Suppose the boundary comprised two separate pieces. A transverse cut q that joins a point of one piece a to a point of another piece b, merely divides connected parts of the surface from each other. For one can form a line interior to the surface along a, starting from one side of the transverse cut q and ending on the opposite side. So q does not split the surface, contrary to hypothesis.

2) Each transverse cut either decreases by 1, or increases by 1, the number of pieces of the boundary.

For a transverse cut q there are three cases: Either q joins a point of one piece of the boundary a, to a point of another piece b. In this case, all these lines taken in the sequence a, q, b, q form a single closed piece of the boundary. Or, q joins two points of a single piece of the boundary. In this case the boundary falls into two pieces via the two endpoints, each of which, taken together with the transverse cut, gives a closed piece of the boundary. Or finally, q ends at one of its own preceding points. In this case q can be considered as comprising a closed line o and a line ℓ joining a point of o to a point of a boundary piece a. In this case o on the one hand, and a, ℓ, o, ℓ on the other hand, are closed paths each forming a piece of the boundary.

Instead of two boundary pieces we obtain one, in the first case; and two instead of one, in the last two cases. This yields our theorem.

The number of pieces comprising the boundary of an n-fold connected portion of surface, is thus either n, or less than n by an even number.

From this we obtain a corollary:

If the number of boundary pieces of an n-fold connected surface is n, the surface is split into two separate parts by every simple closed cut in its interior.

For the connectivity is not altered by this cut and the number of boundary pieces is increased by 2. Consequently, if the new surface were connected, it would be n-tuply connected with $n+2$ boundary pieces, which is impossible.

7.

Let X and Y be two functions of x, y continuous at all points of the surface T spread out over A. With the surface integral extended over all elements dT of this surface, we have

$$\int \left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y}\right) dT = -\int (X \cos \xi + Y \cos \eta) ds.$$

Here, at each point of the boundary, ξ denotes the inclination of the interior normal to the x-axis, and η denotes its inclination to the y-axis. On the right side, the integral extends over all elements ds of the boundary line.

In order to transform the integral $\int \frac{\partial X}{\partial x}\, dT$, we divide the part of the plane A covered by T into strips, via a system of lines parallel to the x axis, in such a way that each branch point of T falls on one of these lines. With this assumption, each part of T corresponding to one of these strips is formed of one or more separate trapezoidal pieces. The contribution to $\int \frac{\partial X}{\partial x}\, dT$ of one of these surface strips, cutting off an element dy of the y axis, will obviously be

$$dy \int \frac{\partial X}{\partial x}\, dx.$$

Here the integration is taken over the (one or more) straight lines belonging to the surface T that fall on a normal issuing from some point of this element dy. Now let the lower endpoints of these lines be $O_1, O_2, O_3, \ldots$, (that is, corresponding to the smallest value of x); the upper points $O^{(1)}, O^{(2)}, O^{(3)}, \ldots$, and denote by $X_1, X_2, \ldots, X^{(1)}, X^{(2)}, \ldots$ the values of X at these points. Let $ds_1, ds_2, \ldots, ds^{(1)}, ds^{(2)}, \ldots$ be the corresponding elements cut off from the boundary of the surface strips, and $\xi_1, \xi_2, \ldots, \xi^{(1)}, \xi^{(2)}, \ldots$ the values of ξ at these elements. Then

$$\int \frac{\partial X}{\partial x}\, dx = -X_1 - X_2 - X_3 \cdots + X^{(1)} + X^{(2)} + X^{(3)} \ldots.$$

The angles ξ are evidently acute at the lower endpoints, obtuse at the upper endpoints; consequently

$$dy = \cos \xi_1 ds_1 = \cos \xi_2 ds_2 = \\ \cdots = -\cos \xi^{(1)} ds^{(1)} = -\cos \xi^{(2)} ds^{(2)} \ldots.$$

Substituting these values yields

$$dy \int \frac{\partial X}{\partial x}\, dx = -\sum X \cos\xi\, ds,$$

the summation extending over all boundary elements whose projection onto the y-axis is dy.

By integration over all the elements dy that occur, it is clear that all elements of the surface T and all elements of the boundary will be exhausted. Accordingly we obtain, with the integration taken over the perimeter,

$$\int \frac{\partial X}{\partial x}\, dT = -\int X \cos\xi\, ds.$$

By entirely analogous reasoning, we obtain

$$\int \frac{\partial Y}{\partial y}\, dT = -\int Y \cos\eta\, ds$$

so that

$$\int \left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y}\right) dT = -\int (X\cos\xi + Y\cos\eta) ds,$$

as we wished to prove.

8.

On the boundary line, we denote by s the distance of a general point O_o, from a fixed initial point, in a direction to be fixed below. On the normal at O_o, let p denote the distance of an indeterminate point O from O_o; the distance is taken to be positive for interior points of the normal. The values of x and y at the point O can clearly be considered as functions of s and p. At the points of the boundary line we then have the partial derivatives

$$\frac{\partial x}{\partial p} = \cos\xi,\ \frac{\partial y}{\partial p} = \cos\eta,\ \frac{\partial x}{\partial s} = \pm\cos\eta,\ \frac{\partial y}{\partial s} = \mp\cos\xi.$$

Here we take the upper sign if the direction in which s increases makes the same angle with p as the x-axis with the y-axis. In the opposite case, we take the lower sign. We take the direction of increasing s throughout the boundary in such a way that

$$\frac{\partial x}{\partial s} = \frac{\partial y}{\partial p}, \text{ and so } \frac{\partial y}{\partial s} = -\frac{\partial x}{\partial p}.$$

This does not essentially restrict the generality of our results.

Obviously we can extend this rule to lines interior to T. Here, in determining the signs of dp and ds, if their mutual dependence is fixed as above, we must still indicate how to fix the sign of dp or of ds. For a closed line, we shall indicate one of the two parts into which it cuts the surface, and take the line as its boundary, which determines the sign of dp. For a line that is not closed, we indicate instead the origin of the line, that is, the endpoint where s takes the least value.

Substituting the values obtained for $\cos\xi$ and $\cos\eta$ into the equation proved in the previous section, with the domains of integration as before, we have

$$\int \left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y}\right) dT = -\int \left(X\frac{\partial x}{\partial p} + Y\frac{\partial y}{\partial p}\right) ds$$
$$= \int \left(X\frac{\partial y}{\partial s} - Y\frac{\partial x}{\partial s}\right) ds.$$

9.

We apply the result at the end of the previous section to the case where, throughout the surface,

$$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = 0.$$

This yields the following theorems.

I. Let X and Y be two functions finite and continuous throughout T satisfying

$$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = 0.$$

Then, integrating over the entire boundary of T,

$$\int \left(X\frac{\partial x}{\partial p} + Y\frac{\partial y}{\partial p}\right) ds = 0.$$

Consider an arbitrary surface T_1 spread over A and divide it in arbitrary fashion into two pieces T_2, T_3. The integral

$$\int \left(X\frac{\partial x}{\partial p} + Y\frac{\partial y}{\partial p}\right) ds$$

over the boundary of T_2 can be regarded as the difference between the integrals over the boundaries of T_1 and T_3. For where T_3 borders T_1 the two integrals cancel out; every other element corresponds to an element of the boundary of T_2.

Using this transformation, we deduce from I:

II. The value of the integral

$$\int \left(X \frac{\partial x}{\partial p} + Y \frac{\partial y}{\partial p} \right) ds$$

taken over the boundary of a surface spread over A, remains constant when the surface is increased or diminished in size in a manner that does not cross any part of the surface where the hypotheses of Theorem I are violated.

When the functions X, Y satisfy the above differential equation throughout the surface T, but have discontinuities at isolated lines or points, we can surround each of these lines or points with an arbitrarily small region of surface. Applying Theorem II, we obtain:

III. The integral

$$\int \left(X \frac{\partial x}{\partial p} + Y \frac{\partial y}{\partial p} \right) ds$$

over the boundary of T is equal to the sum of the integrals

$$\int \left(X \frac{\partial x}{\partial p} + Y \frac{\partial y}{\partial p} \right) ds$$

taken around all discontinuities. The contribution to the integral from each discontinuity is constant, however narrow the boundary that encloses it.

The contribution from a point of discontinuity is necessarily zero if, ρ denoting distance of O from the discontinuity, ρX and ρY become infinitely small with ρ. Take such a point, as origin, and an arbitrary direction as initial line, for polar coordinates ρ, ϕ; choose as the boundary curve a circle of radius ρ and center at this point. Then the integral in question is expressed by

$$\int_0^{2\pi} \left(X \frac{\partial x}{\partial p} + Y \frac{\partial y}{\partial p} \right) \rho \, d\phi.$$

This cannot take a nonzero value χ, since whatever the value of χ, we can taken ρ so small that the absolute value of $\left(X \frac{\partial x}{\partial p} + Y \frac{\partial y}{\partial p} \right) \rho$ is smaller than

$\frac{\chi}{2\pi}$ for every ϕ, giving

$$\int_0^{2\pi} \left(X \frac{\partial x}{\partial p} + Y \frac{\partial y}{\partial p} \right) \rho \, d\phi < \chi.$$

IV. Suppose that for a simply connected surface spread over A, we have

$$\int \left(X \frac{\partial x}{\partial p} + Y \frac{\partial y}{\partial p} \right) ds = 0$$

when we integrate around an arbitrary part of the surface; or

$$\int \left(Y \frac{\partial x}{\partial s} - X \frac{\partial y}{\partial s} \right) ds = 0.$$

Then this integral, taken between any two fixed points O_o and O, has the same value for any line from O_o to O.

Any two paths s_1 and s_2 joining O_o to O, taken together, form a closed line s_3. Either this line itself passes through no point more than once, or one can divide it into several simple closed lines as follows. Start from an arbitrary point and describe the contour. Each time we meet a point already traversed, separate off the intermediate part, and consider the following part as the immediate extension of the preceding portion of the curve. Each such line divides the surface into a simply connected piece and a doubly connected piece, and forms the entire boundary of one of these pieces. The integral

$$\int \left(Y \frac{\partial x}{\partial s} - X \frac{\partial y}{\partial s} \right) ds$$

corresponding to the simply connected piece will be 0 by hypothesis. Consequently the same property holds for this integral taken over the whole path s_3, with the quantity s treated as increasing in the same direction throughout. It follows that the integrals along the lines s_1 and s_2 cancel out, if there is no change in this direction; that is, one line passes from O_o to O and the other from O to O_o. If we alter the sense of the latter integral, they become equal.

Consider any surface T in which, generally speaking,

$$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = 0.$$

We remove the discontinuities if necessary, so that for any portion of the remaining pieces of the surface

$$\int \left(Y \frac{\partial x}{\partial s} - X \frac{\partial y}{\partial s} \right) ds = 0,$$

and reduce the remaining surface via transverse cuts to a simply connected surface T^*. For each path in the interior of T^*, from a point O_o to another point O, the above integral has the same value, which for brevity we denote by

$$\int_{O_o}^{O} \left(Y \frac{\partial x}{\partial s} - X \frac{\partial y}{\partial s} \right) ds.$$

The integral, where O_o is fixed and O is arbitrary, is independent for each O of the path joining the points; so it can be treated as a function of x, y. The variation of this function when O moves along an arbitrary line element ds is

$$\left(Y \frac{\partial x}{\partial y} - X \frac{\partial y}{\partial s} \right) ds;$$

this is continuous everywhere in T^* and equal on both sides of a transverse cut of T.

V. The integral

$$Z = \int_{O_o}^{O} \left(Y \frac{\partial x}{\partial s} - X \frac{\partial y}{\partial s} \right) ds$$

thus represents for fixed O_o a function of x, y continuous throughout T^*, but varies by a constant on crossing anywhere along a transverse cut of T from one branch point to another. The function has partial derivatives

$$\frac{\partial Z}{\partial x} = Y, \quad \frac{\partial Z}{\partial y} = -X.$$

The variations on crossing transverse cuts depend on certain quantities, independent of each other, equal in number to the number of transverse cuts. For when one runs over the system of transverse cuts in a retrograde sense—the later parts first—this variation is everywhere determined if its value at the start of each cut is known. However, the latter values are independent of each other.

10.

For the functions so far denoted by X and Y, take

$$u\frac{\partial u'}{\partial x} - u'\frac{\partial u}{\partial x}\ ,\ u\frac{\partial u'}{\partial y} - u'\frac{\partial u}{\partial y}$$

respectively. Then

$$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = u\left(\frac{\partial^2 u'}{\partial x^2} + \frac{\partial^2 u'}{\partial y^2}\right) - u'\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right).$$

Thus if u, u' satisfy the equations

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0\ ,\ \frac{\partial^2 u'}{\partial x^2} + \frac{\partial^2 u'}{\partial y^2} = 0,$$

we have

$$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = 0.$$

We can apply the results of the previous section to

$$\int\left(X\frac{\partial x}{\partial p} + Y\frac{\partial y}{\partial p}\right)ds,$$

which is equal to

$$\int\left(u\frac{\partial u'}{\partial p} - u'\frac{\partial u}{\partial p}\right)ds.$$

Suppose now that u and its first order partial derivatives never have discontinuities along a line. Suppose further that at each point of discontinuity, ρ denoting distance from O to the discontinuity, $\rho\frac{\partial u}{\partial x}$ and $\rho\frac{\partial u}{\partial y}$ become infinitely small with ρ. According to the remark on III of the previous section, the discontinuities of u can be disregarded.

For, given a straight line from one of these discontinuities, take a value R of ρ such that

$$\rho\frac{\partial u}{\partial \rho} = \rho\frac{\partial u}{\partial x}\frac{\partial x}{\partial \rho} + \rho\frac{\partial u}{\partial y}\frac{\partial y}{\partial \rho}$$

remains finite for $\rho < R$. Let $u = U$ for $\rho = R$, let M be the largest absolute value of $\rho\frac{\partial u}{\partial \rho}$ in the interval of the line with $0 < \rho < R$. Then, again disregarding sign,

$$u - U < M(\log\rho - \log R),$$

so that $\rho(u - U)$, and indeed ρu, become infinitely small with ρ. By hypothesis, this holds also for $\rho\,\frac{\partial u}{\partial x}$, $\rho\,\frac{\partial u}{\partial y}$. Consequently, if u' does not become discontinuous, the same holds for

$$\rho\left(u\,\frac{\partial u'}{\partial x} - u'\,\frac{\partial u}{\partial x}\right) \text{ and } \rho\left(u\,\frac{\partial u'}{\partial y} - u'\,\frac{\partial u}{\partial y}\right);$$

so that the case considered in the previous paragraph applies here.

Suppose now that the surface T formed from the positions of O is simple wherever it covers A. Take an arbitrary fixed point O_o of T where $u = u_0$, $x = x_0$, $y = y_0$. The quantity

$$\frac{1}{2}\,\log((x - x_0)^2 + (y - y_0)^2) = \log r,$$

considered as a function of x and y, has the property that

$$\frac{\partial^2 \log r}{\partial x^2} + \frac{\partial^2 \log r}{\partial y^2} = 0$$

and has a discontinuity only for $x = x_0$, $y = y_0$; and so in our case for only one point of the surface T.

Consequently, from Section 9, III with $u' = \log r$, the integral

$$\int \left(u\,\frac{\partial \log r}{\partial p} - \log r\,\frac{\partial u}{\partial p}\right) ds$$

has the same value for the whole boundary of T as it does for an arbitrary circuit around the point O_o. Take for this circuit the circumference of a circle on which r is constant, and denote by ϕ the angle of a radius with endpoint O, measured from an arbitrary point of the circumference in some definite direction. The integral is

$$-\int_0^{2\pi} u\,\frac{\partial \log r}{\partial r}\,r\,d\phi - \log r \int \frac{\partial u}{\partial p}\,ds$$

which, since

$$\int \frac{\partial u}{\partial p}\,ds = 0,$$

reduces to

$$-\int_0^{2\pi} u\,d\phi.$$

This value becomes $-u_0 2\pi$ for infinitely small r if u is continuous at O_o.

Under the above hypotheses on u and T, then, we have, for an arbitrary point O_o interior to the surface, where u is continuous,

$$u_0 = \frac{1}{2\pi} \int \left(\log r \frac{\partial u}{\partial p} - u \frac{\partial \log r}{\partial p} \right) ds$$

with the integral taken over the boundary; and

$$u_0 = \frac{1}{2\pi} \int_0^{2\pi} u \, d\phi$$

with the integral taken over a circle around O_o. The first of these expressions yields the following

Theorem Suppose that, in general, the function u satisfies

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

inside a simple surface T spread over the plane A. Suppose further that

1) the points where this differential equation is violated do not fill a region of the surface,

2) the points in which u, $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$ are discontinuous do not fill any line,

3) for each discontinuity the quantities $\rho \frac{\partial u}{\partial x}$, $\rho \frac{\partial u}{\partial y}$ become infinitely small along with the distance ρ of O from the discontinuity, and

4) a discontinuity of u, that can be removed by modifying the value of u at isolated points, is excluded.

Then u and all its partial derivatives are continuous and finite at all points interior to the surface.

For consider O_o as a variable point. The only quantities that vary in the expression

$$\int \left(\log r \frac{\partial u}{\partial p} - u \frac{\partial \log r}{\partial p} \right) ds$$

are $\log r$, $\frac{\partial \log r}{\partial x}$, $\frac{\partial \log r}{\partial y}$. However, these quantities, along with their partial derivatives, are finite and continuous functions of x_0, y_0 at every point of the boundary while O_o remains in the interior of T; the partial derivatives may be expressed as rational functions of these quantities, containing only powers of r in the denominator. The continuity property persists for the value of our integral and consequently for the function u_0. Now u, by the earlier hypotheses, can only differ in value from our integral at isolated points where u is discontinuous. This possibility is excluded by hypothesis 4) of our theorem.

11.

Under the hypotheses on u and T at the end of the previous paragraph, we have the following theorems:

I. If $u = 0$ and $\frac{\partial u}{\partial p} = 0$ along a line, then $u = 0$ everywhere.

We show first of all that a line λ where $u = 0$ and $\frac{\partial u}{\partial p} = 0$ cannot be the boundary of a region a of the surface, where $u > 0$.

For suppose that this did take place. From a take a piece bounded on one side by λ and on the other side by a circle; the center of the circle O_o being excluded from this piece. This construction is always possible. With the integral taken over the boundary of this piece, and with r, ϕ denoting polar coordinates of O with respect to O_o, we have

$$\int \log r \frac{\partial u}{\partial p} ds - \int u \frac{\partial \log r}{\partial p} ds = 0.$$

By hypothesis it follows that, with the integral taken only over the circle,

$$\int u\, d\phi + \log r \int \frac{\partial u}{\partial p} ds = 0.$$

Since

$$\int \frac{\partial u}{\partial p} ds = 0,$$

we obtain

$$\int u\, d\phi = 0,$$

which is incompatible with the hypothesis that $u > 0$ in the interior of a.

In a similar way we show that the equations $u = 0$ and $\frac{\partial u}{\partial p} = 0$ cannot hold on the boundary line of a region of the surface b, where u is negative.

Suppose now that u and $\frac{\partial u}{\partial p}$ are 0 on a line in the surface T, and there is a region of T where $u \neq 0$. Clearly such a region is bounded either by this line or by a region of the surface where $u = 0$. Consequently, it is always bounded by a line where u and $\frac{\partial u}{\partial p}$ are 0, which leads necessarily to one of the hypotheses rejected above.

II. If u and $\frac{\partial u}{\partial p}$ have given values along a line, then u is determined throughout T.

Let u_1 and u_2 be any two functions that satisfy the above conditions on u. These conditions also apply to the difference $u_1 - u_2$. Suppose that u_1, u_2 are identical, along with their first order partial derivatives with respect to p, on a line, but differ in some region of the surface. Then along this line $u_1 - u_2 = 0$ and $\frac{\partial}{\partial p}(u_1 - u_2) = 0$, while $u_1 - u_2$ is not 0 everywhere, contrary to Theorem I.

III. If u is not a constant in T, then the points in T where u takes a constant value are necessarily lines which divide regions where u is larger from regions where u is smaller.

This theorem is obtained by combining the following:

u cannot have a minimum or maximum at an interior point of T;

u cannot be a constant in just a part of T;

the lines in which $u = a$ cannot bound portions of the surface on both sides where $u - a$ has the same sign.

These are all propositions whose negation, as we readily see, would lead a violation of the equation

$$u_0 = \frac{1}{2\pi} \int_0^{2\pi} u \, d\phi$$

or

$$\int_0^{2\pi} (u - u_0) \, d\phi = 0$$

proved in the previous section. The negation is consequently impossible.

12.

We now turn back to the treatment of a complex variable $w = u + vi$ which, generally speaking (that is, excluding isolated lines and points) has a unique value for each point O of the surface T that varies continuously with the position of O. Moreover we suppose that the equation

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad , \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

hold except at the excluded values. As indicated earlier, we say that w is a function of $z = x + iy$. To simplify the following, we suppose that a function of z has no discontinuity that can be removed by changing its value at an isolated point.

The surface T will at first be supposed to be simply connected, and to be spread out simply over the plane A.

Theorem Suppose that the set of discontinuities of the function w of z do not contain a line. Further suppose that at an arbitrary point O' of the surface, where $z = z'$, $(z - z')w$ becomes infinitely small as O tends to O'. Then w, along with all its derivatives, is finite and continuous at all points interior to the surface.

Let $z - z' = \rho e^{\phi i}$. The hypotheses on the variations of the quantity w have the following consequences for u and v.

1) $\dfrac{\partial u}{\partial x} - \dfrac{\partial v}{\partial y} = 0$

and

2) $\dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x} = 0$

throughout T;

3) the functions u and v are not discontinuous along a line;

4) for every point O', the quantities ρu and ρv become infinitely small with the distance ρ of O from O';

5) neither u nor v has discontinuities that can be removed by changing the value at an isolated point.

As a consequence of the hypotheses 2), 3), 4), for any portion of the surface T we obtain

$$\int \left(u \frac{\partial x}{\partial s} - v \frac{\partial y}{\partial s} \right) ds = 0,$$

the integral being taken over the boundary, from III, Section 9. Further the integral

$$\int_{O_o}^{O} \left(u \frac{\partial x}{\partial s} - v \frac{\partial y}{\partial s} \right) ds$$

(by IV, Section 9) has the same value along every line from O_o to O, and for fixed O_o forms a continuous function U of x, y except for isolated points. Recalling 5), the partial derivatives satisfy $\frac{\partial U}{\partial x} = u$ and $\frac{\partial U}{\partial y} = -v$ at every point. By substituting these values for u and v, the hypotheses 1), 3), 4) become the conditions of the theorem at the end of Section 10. Accordingly the function U is finite and continuous, along with all its partial derivatives, at every point of T. Thus the same holds for the complex function $w = \frac{\partial U}{\partial x} - \frac{\partial U}{\partial y} i$, and its derivatives with respect to z.

13.

We now investigate what transpires when, retaining the other hypotheses of Section 12, we suppose that, for a specific point O' interior to the surface,

$$(z - z')w = \rho e^{i\phi} w$$

does not become infinitely small when O tends to O'. In this case w becomes infinitely large when O tends to O'. We suppose that, if the quantity w is not of order of magnitude $\frac{1}{\rho}$, that is, their quotient does not have a finite bound, then at least the order of the two quantities is in finite ratio. That is, there is a power of ρ whose product with w either becomes infinitely small with ρ, or remains finite. If μ is the exponent of such a power and n is the smallest integer with $n > \mu$, then the quantity $(z - z')^n w = \rho^n e^{n\phi i} w$ becomes infinitely small with ρ. Now $(z - z')^{n-1} w$ is a function of z (since $\frac{d}{dz}((z - z')^{n-1} w)$ is independent of dz). This function satisfies, in this region of the surface, the hypotheses of Section 12, and consequently is finite and continuous at the point O'. Denote its value O' by a_{n-1}. Then $(z - z')^{n-1} w - a_{n-1}$ is a function that is continuous at O' and vanishes there, and consequently becomes infinitely small with ρ. By Section 12, $(z - z')^{n-2} w - \frac{a_{n-1}}{z - z'}$ is continuous at the point O'. Continuing this procedure, it is obvious that on subtracting an expression of the form

$$\frac{a_1}{z - z'} + \frac{a_2}{(z - z')^2} + \cdots + \frac{a_{n-1}}{(z - z')^{n-1}},$$

w becomes a function that remains finite and continuous at O'.

Accordingly, if we vary the hypotheses of Section 12, by permitting w to become infinitely large as O tends to a point O' in the interior of T, then the order of this infinite quantity, if finite, is necessarily an integer. (A quantity inversely proportional to the distance from O to O' is considered to have order 1.) If this integer is m, then on attaching to it a function containing $2m$ arbitrary constants, w becomes a function continuous at O'.

Remark. We consider a function to contain one arbitrary constant, when the possible determinations of the constant comprise a continuous one-dimensional domain.

14.

The restrictions on the surface T of Sections 12 and 13 are not essential for the validity of the results obtained. Obviously one can surround each point interior to an arbitrary surface with a piece of the surface which has the properties assumed there. The only exception is the case where this point is a branch point of the surface.

To investigate this case, consider the surface T, or an arbitrary piece of it, containing a branch point O' of order $n-1$, where $z = z' = x' + y'i$, and mapped by the function $\zeta = (z - z')^{1/n}$ onto another plane Λ. That is, we represent the value of the function $\zeta = \xi + \eta i$ at the point O as a point Θ of Λ with rectangular coordinates ξ, η, and treat Θ as the image of O. The image of the region T obtained in this way is a connected surface spread over Λ, without a branch point at the image Θ' of O', as we now show.

To fix ideas, draw a circle of radius R around the point O in the plane A and draw a diameter parallel to the x-axis, so that $z - z'$ is real on this diameter. The piece of the surface T, around the branch point, cut off by this circle, is divided on both sides of this diameter into n separate pieces of semicircular form, provided R is sufficiently small. The surface portions on the side of the diameter where $y - y'$ is positive will be denoted by $a_1, \ldots, a_n$; those on the opposite side by $a'_1, \ldots, a'_n$. We suppose further that for negative values of $z - z'$, $a_1, \ldots, a_n$ are connected respectively to $a'_1, \ldots, a'_n$, while for positive values they are connected respectively to $a'_n, a'_1, \ldots, a'_{n-1}$. Then a point O', encircling the branch point in the appropriate direction, runs in succession over the surfaces $a_1, a'_1, a_2, a'_2, \ldots, a_n, a'_n$ and from a'_n back to a_1; this hypothesis is obviously permissible. We introduce polar coordinates on both planes, writing $z - z' = \rho e^{\phi i}$, $\zeta = \sigma e^{\psi i}$, and choose as the image of the surface portion a_1 the value $(z - z')^{1/n} = \rho^{1/n} e^{\frac{\phi}{n} i}$. Here we suppose that $0 \leq \phi \leq \pi$. Then $\sigma \leq R^{1/n}$, $0 \leq \psi \leq \pi/n$ for all points of a_1; its image in Λ lies in a sector from $\psi = 0$ to $\psi = \frac{\pi}{n}$ of a circle around Θ' of radius $R^{1/n}$. Indeed, to each point of a_1 corresponds a point of this sector, varying continuously along with it, and vice versa. Thus the image of a_1 is a connected surface spread out simply over this sector. Analogously, the images of $a'_1, a_2, \ldots, a'_n$ are respectively sectors from $\psi = \frac{\pi}{n}$ to $\psi = \frac{2\pi}{n}$, from $\psi = \frac{2\pi}{n}$ to $\psi = \frac{3\pi}{n}, \ldots$, from $\psi = \frac{2n-1}{n}\pi$ to $\psi = 2\pi$. Here ϕ is successively to run from π to 2π, from 2π to $3\pi, \ldots$ from $(2n-1)\pi$ to $2n\pi$ for the points of these surfaces, which is possible in exactly one way.

These sectors follow one another in the same manner as the surfaces a and a', in such a way that coincident points correspond to coincident points. The sectors can thus be joined together to give the connected image of a portion of T surrounding O', and this image is obviously a surface spread out simply over the plane Λ.

A variable that has a definite value for each point O likewise has a definite value for each point Θ and conversely, since each O corresponds to only one Θ and each Θ to only one O. If the variable is a function of z, then it is a function of ζ. For if $\frac{dw}{dz}$ is independent of dz, then $\frac{dw}{d\zeta}$ is independent of $d\zeta$, and conversely. We conclude that the results of Section 12 and 13 are applicable to all functions w of z at branch points O', if they are treated as functions of $(z - z')^{1/n}$. This yields the following result:

If a function w of z becomes infinite as O tends to a branch point O' of order $n - 1$, then this infinite quantity is necessarily of the same order as a power of the distance from O to O' whose exponent is a multiple of $\frac{1}{n}$. If this exponent is $-\frac{m}{n}$, then on attaching an expression of the form

$$\frac{a_1}{(z-z')^{1/n}} + \frac{a_2}{(z-z')^{2/n}} + \cdots + \frac{a_m}{(z-z')^{m/n}},$$

where $a_1, a_2, \ldots, a_m$ are arbitrary complex numbers, the function becomes continuous at O'.

This theorem has the corollary that the function w is continuous at O' if $(z - z')^{1/n} w$ becomes infinitely small as O tends to O'.

15.

Now consider a function of z with a definite value for each point of an arbitrary surface T spread over A, not everywhere constant. If we represent the value $w = u + iv$ at O geometrically as a point Q of the plane B with rectangular coordinates u, v, we have the following consequence.

I. the totality of points Q can be considered as forming a surface S, in which each point corresponds to one point O in T that varies continuously with Q.

To prove this, obviously we only need to show that the position of the point Q always changes along with that of O (and, generally speaking, changes continuously). This is contained in the following results:

A function $w = u + vi$ of z cannot be constant along a line, unless w is everywhere constant.

Proof. If w takes the constant value $a + bi$ along a line, then $u - a$ and $\frac{\partial(u-a)}{\partial p} = -\frac{\partial v}{\partial s}$ vanish not only on this line but everywhere, along with

$$\frac{\partial^2(u-a)}{\partial x^2} + \frac{\partial^2(u-a)}{\partial y^2}.$$

By Section 11, I, $u - a = 0$. Since

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad , \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x},$$

we also have $v - b = 0$ everywhere, contrary to hypothesis.

II. By the hypothesis made in I, two parts of S cannot be connected together unless the corresponding parts of T are connected together. Conversely, whenever a connection occurs in T and w is continuous, a corresponding connection holds in S.

Assuming this, the boundary of S corresponds in part to the boundary of T and in part to the discontinuities. The interior of S, however, excluding isolated points, is spread over B smoothly. That is, the surface never splits into superimposed surfaces or folds back on itself.

The first could only happen, since T is connected in the corresponding fashion, if T splits in this way–contrary to hypothesis. The latter case is treated next.

We show firstly, that a point Q' where $\frac{dw}{dz}$ is finite, cannot lie in a fold of the surface S.

To see this, we surround the point O' corresponding to Q', with a piece of the surface T of any shape and unspecified size. According to Section 3, this size can be taken so small that the shape of the corresponding piece on S differs as little as we wish from that of the piece on T. Consequently its boundary cuts out a portion of B that surrounds Q'. However, this is impossible if Q' lies on a fold of the surface S.

By I, $\frac{dw}{dz}$, as a function of z, can only vanish at isolated points. Since w is continuous at the points of T considered here, $\frac{dw}{dz}$ can only become infinite at the branch points of the surface. This yields the desired result.

III. Consequently, S is a surface obeying the hypotheses that were imposed on T in Section 5. In this surface, the variable quantity z has one definite value for each point Q, which varies continuously with the position of Q in such a way that $\frac{dz}{dw}$ is independent of the direction of variation. Accordingly z, in the sense of the term specified earlier, is a continuous function of the complex quantity w over the domain represented by S.

From this, we further obtain:

Let O' and Q' be corresponding interior points of the surfaces T and S, at which $z = z'$, $w = w'$. If neither of them is a branch point, $\frac{w-w'}{z-z'}$ tends to a finite limit as O tends to O', and the image of an infinitely small portion is similar to that portion. However, if Q' is a branch point of order $n-1$ and O' is a branch point of order $m-1$, then $\frac{(w-w')^{1/n}}{(z-z')^{1/m}}$ has a finite limit as O tends to O'. For the adjacent regions we have a form of mapping, readily obtained from Section 14.

16.

Theorem Let α and β be arbitrary functions of x, y for which the integral

$$\int \left[\left(\frac{\partial \alpha}{\partial x} - \frac{\partial \beta}{\partial y} \right)^2 + \left(\frac{\partial \alpha}{\partial y} + \frac{\partial \beta}{\partial x} \right)^2 \right] dT$$

taken over the arbitrary surface T spread over A, is finite. If we vary α by functions vanishing on the boundary that are continuous, or only discontinuous at isolated points, the integral takes a minimum value for one of these functions. If we exclude discontinuities that can be removed by changing the function at isolated points, the minimum is obtained for only one function.

Let λ be an unspecified function vanishing on the boundary, either continuous or possessing only discontinuities at isolated points, for which the integral

$$L = \int \left(\left(\frac{\partial \lambda}{\partial x} \right)^2 + \left(\frac{\partial \lambda}{\partial y} \right)^2 \right) dT,$$

taken over the whole surface, is finite. Denote by ω an unspecified member of the set of functions $\alpha + \lambda$, and finally denote by Ω the integral

$$\int \left[\left(\frac{\partial \omega}{\partial x} - \frac{\partial \beta}{\partial y} \right)^2 + \left(\frac{\partial \omega}{\partial y} + \frac{\partial \beta}{\partial x} \right)^2 \right] dT$$

over the whole surface. The family of functions λ forms a closed connected domain, in that any function passes continuously into any other, while no limit of the functions is discontinuous along a line unless L simultaneously becomes infinite (Section 17). For each λ, writing $\omega = \alpha + \lambda$, the integral Ω has a finite value, that tends to infinity with L and varies continuously with the form of λ, but can never be less than 0. Consequently Ω has a minimum value for at least one form of the function ω.

To prove the second part of our theorem, denote by u one of the functions ω for which Ω attains its minimum value. Let h be an unspecified constant function, so that $u + h\lambda$ satisfies the required conditions for the function ω. For $\omega = u + h\lambda$, we write the value of Ω as

$$\begin{aligned} &\int \left[\left(\frac{\partial u}{\partial x} - \frac{\partial \beta}{\partial y} \right)^2 + \left(\frac{\partial u}{\partial y} + \frac{\partial \beta}{\partial x} \right)^2 \right] dT \\ &+ 2h \int \left[\left(\frac{\partial u}{\partial x} - \frac{\partial \beta}{\partial y} \right) \frac{\partial \lambda}{\partial x} + \left(\frac{\partial u}{\partial y} + \frac{\partial \beta}{\partial x} \right) \frac{\partial \lambda}{\partial y} \right] dT \\ &+ h^2 \int \left(\left(\frac{\partial \lambda}{\partial x} \right)^2 + \left(\frac{\partial \lambda}{\partial y} \right)^2 \right) dT = M + 2Nh + Lh^2, \end{aligned}$$

say. This must be at least M for every λ (by definition of the minimum), provided h is sufficiently small. Now it follows that $N = 0$ for each λ. Otherwise,

$$2Nh + Lh^2 = Lh^2 \left(1 + \frac{2N}{Lh} \right)$$

would be negative if h is opposite in sign to N and of absolute value $< \frac{2N}{L}$. The value of Ω for $\omega = u + \lambda$, a form which obviously contains all possible values of ω, is thus $M + L$. Since L is essentially positive, Ω can attain for no form of the function ω a smaller value than it attains for $\omega = u$.

Now suppose that for another u' the functions ω yield a minimum value M' of Ω. The same deductions give $M' \leq M$ and $M \leq M'$, consequently $M = M'$. Write u in the form $u + \lambda'$; then we obtain for M' the expression

$M + L'$, where L' denotes the value of L for $\lambda = \lambda'$. The equation $M = M'$ yields $L' = 0$. This is only possible if, throughout the surface,

$$\frac{\partial \lambda'}{\partial x} = 0, \quad \frac{\partial \lambda'}{\partial y} = 0.$$

Wherever λ' is continuous, it is necessarily a constant. Since λ' vanishes at the boundary and is not discontinuous along a line, it can only be nonzero at isolated points. Thus two functions ω for which Ω attains a minimum, differ only at isolated points. If all discontinuities of u that can be removed by changing its value at isolated points are discarded, the function is completely determined.

17.

We now need to supply the proof of the fact that λ, if L is to remain finite, cannot tend to a function γ discontinuous along a line. That is, if λ is supposed to coincide with γ outside a portion of surface T' that surrounds the line of discontinuity, and T' is taken sufficiently small, then L exceeds an arbitrary given number C.

We assign s and p their usual meanings with respect to the line of discontinuity. For indeterminate s, denote by χ the curvature, with a convex curvature on the side of positive p taken to be positive. Let p_1 denote the value of p on the boundary of T' on the side where $p > 0$, and p_2 the value on the side where $p < 0$. Denote the corresponding values of γ by γ_1 and γ_2. Consider any portion of this line with continuous curvature. The part of T' between the normals at its endpoints, taken only up to the center of curvature, contributes to L the quantity

$$\int ds \int_{p_2}^{p_1} dp(1 - \chi p) \left[\left(\frac{\partial \lambda}{\partial p}\right)^2 + \left(\frac{\partial \lambda}{\partial s}\right)^2 \frac{1}{(1 - \chi p)^2} \right].$$

The smallest value of the expression

$$\int_{p_2}^{p_1} \left(\frac{\partial \lambda}{\partial p}\right)^2 (1 - \chi p) dp$$

for fixed boundary values γ_1 and γ_2 of λ, by known procedures, is

$$\frac{(\gamma_1 - \gamma_2)^2 \chi}{\log(1 - \chi p_2) - \log(1 - \chi p_1)}.$$

Consequently the above contribution is necessarily

$$> \int \frac{(\gamma_1 - \gamma_2)^2 \chi ds}{\log(1 - \chi p_2) - \log(1 - \chi p_1)}$$

however we choose λ in the interior of T'.

The function γ will be continuous for $p = 0$ if the greatest value of $(\gamma_1 - \gamma_2)^2$, for $\pi_1 > p_1 > 0$ and $\pi_2 < p_2 < 0$, becomes infinitely small with $\pi_1 - \pi_2$. Consequently, for each value of s we can choose a finite quantity m such that, however small $\pi_1 - \pi_2$ is taken, there are values of p_1 and p_2 with $\pi_1 > p_1 \geq 0$, $\pi_2 > p_2 \geq 0$ (the equality signs mutually exclusive) for which

$$(\gamma_1 - \gamma_2)^2 > m.$$

Consider an arbitrary form of T' consistent with the earlier restrictions. Let P_1 and P_2 be the values of p_1 and p_2 for this form, and denote by a the value of the integral

$$\int \frac{m \chi ds}{\log(1 - \chi P_2) - \log(1 - \chi P_1)}$$

over the part of the discontinuity line in question. Obviously we can obtain

$$\int \frac{(\gamma_1 - \gamma_2)^2 \chi ds}{\log(1 - \chi p_2) - \log(1 - \chi p_1)} > C$$

by choosing p_1 and p_2 for each value of s so that the inequalities

$$p_1 < \frac{1 - (1 - \chi P_1)^{a/C}}{\chi} \quad , \quad p_2 > \frac{1 - (1 - \chi P_2)^{a/C}}{\chi}$$

and

$$(\gamma_1 - \gamma_2)^2 > m$$

are satisfied. It follows that, however λ is defined in the interior of T', the contribution to L from the piece of T in question exceeds C. Thus L itself exceeds C, as we wished to prove.

18.

By Section 16, for the function u fixed there and for any function λ, we have $N = 0$. Here

$$N = \int \left[\left(\frac{\partial u}{\partial x} - \frac{\partial \beta}{\partial y} \right) \frac{\partial \lambda}{\partial x} + \left(\frac{\partial u}{\partial y} + \frac{\partial \beta}{\partial x} \right) \frac{\partial \lambda}{\partial y} \right] dT;$$

the integral extends over T. We now draw further consequences from this equation.

Cut from T a piece T' surrounding the discontinuities of u, β, λ. The contribution to N from the remaining part T'' of T is

$$-\int \lambda \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) dT - \int \left(\frac{\partial u}{\partial p} + \frac{\partial \beta}{\partial s} \right) \lambda \, ds,$$

by Sections 7 and 8 with $X = \left(\frac{\partial u}{\partial x} - \frac{\partial \beta}{\partial y} \right) \lambda$ and $Y = \left(\frac{\partial u}{\partial y} + \frac{\partial \beta}{\partial x} \right) \lambda$.

By the boundary condition imposed on λ, the contribution to

$$\int \left(\frac{\partial u}{\partial p} + \frac{\partial \beta}{\partial s} \right) \lambda \, ds,$$

from the part of the boundary of T'', common with that of T, is 0. So one can regard N as the combination of

$$-\int \lambda \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) dT$$

taken over T'', and

$$\int \left[\left(\frac{\partial u}{\partial x} - \frac{\partial \beta}{\partial y} \right) \frac{\partial \lambda}{\partial x} + \left(\frac{\partial u}{\partial y} + \frac{\partial \beta}{\partial x} \right) \frac{\partial \lambda}{\partial y} \right] dT + \int \left(\frac{\partial u}{\partial p} + \frac{\partial \beta}{\partial s} \right) \lambda \, ds,$$

where the integrations are respectively taken over T' and its boundary.

Clearly, if $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$ differs from 0 in some part of T, N would take a nonzero value provided that λ, as is permissible, takes the value 0 in T', and $\lambda \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right)$ has the same sign throughout T''. However if $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$ throughout T, then the contribution to N from T'' is 0 for every λ. Now the condition $N = 0$ yields the result that the contribution from the discontinuities must be 0.

Now the functions $X = \frac{\partial u}{\partial x} - \frac{\partial \beta}{\partial y}$, $Y = \frac{\partial u}{\partial y} + \frac{\partial \beta}{\partial x}$, do not merely, generally speaking, satisfy

$$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = 0,$$

but also

$$\int \left(X \frac{\partial x}{\partial p} + Y \frac{\partial y}{\partial p} \right) ds = 0,$$

the integral being taken over any part of the boundary of T, at least when this integral has a definite value.

If T is multiply connected, we reduce T via transverse cuts (by V of Section 9) to a simply connected surface T^*, and we see that the integral

$$-\int_{O_o}^{O} \left(\frac{\partial u}{\partial p} + \frac{\partial \beta}{\partial s} \right) ds$$

has the same value for every line from O_o to O in the interior of T^*. Taking O_o fixed, the integral is a function of x, y that is continuous throughout T^* and has the same variation along either side of a transverse cut. Attaching to β this function ν yields a function $v = \beta + \nu$ for which

$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}, \quad \frac{\partial v}{\partial y} = \frac{\partial u}{\partial x}.$$

We obtain the following

Theorem Suppose that a complex function $\alpha + \beta i$ of x, y is given throughout a connected surface T that reduces by transverse cuts to a simply connected surface T^*, and

$$\int \left[\left(\frac{\partial \alpha}{\partial x} - \frac{\partial \beta}{\partial y} \right)^2 + \left(\frac{\partial \alpha}{\partial y} + \frac{\partial \beta}{\partial x} \right)^2 \right] dT,$$

taken over the whole surface, is finite. Then there is one and only one choice of function $\mu + \nu i$ of x, y with the following properties, that can be adjoined to $\alpha + \beta i$ to give a function of z:

1) μ vanishes on the boundary, or at least only differs from 0 in isolated points; ν is given arbitrarily at one point;

2) variations of μ in T and of ν in T^* are discontinuous only at isolated points, and the discontinuities are restricted by the finiteness of the integrals

$$\int\left[\left(\frac{\partial\mu}{\partial x}\right)^2+\left(\frac{\partial\mu}{\partial y}\right)^2\right]dT \quad , \quad \int\left[\left(\frac{\partial\nu}{\partial x}\right)^2+\left(\frac{\partial\nu}{\partial y}\right)^2\right]dT$$

over the surface. Further, ν has the same variations on either side of the transverse cuts.

These conditions suffice to determine μ and ν. For μ, from which ν is determined up to an additive constant, always yields a minimum of the integral Ω; since if $u = \alpha + \mu$, then clearly $N = 0$ for every λ. By Section 16, this property is possessed by only one function.

19.

The principles that underlie the result at the end of the previous section open the way for the investigation of specific functions of a complex variable, independently of expressions for the functions.

For orientation in this field, we consider the extent of the conditions needed to determine such a function in a given domain.

We restrict ourselves initially to a particular case. When the surface spread over A that represents the domain is simply connected, the function $w = u + iv$ of z can be determined from the following conditions:

1. the values of u on the boundary are given, and vary, with infinitely small changes of position, by infinitely small quantities of the same order. Elsewhere, the variation of u is unrestricted[5];

2. the value of v is given arbitrarily at one point;

3. the function is finite and continuous at all points.

The function is determined completely by these conditions. Indeed, this follows from the theorem in the previous section; we define $\alpha + \beta i$, as we may, so that α takes the given value on the boundary and, throughout the surface, $\alpha + \beta i$ varies, with infinitely small changes of position, by infinitely small quantities of the same order.

[5]The variations of this value need only be restricted by not being discontinuous along any part of the boundary. A further restriction is made here simply to avoid unnecessary complications.

Thus, generally speaking, u can be given as an arbitrary function of s on the boundary, and thereby v is determined everywhere. Conversely v can be chosen arbitrarily at the boundary points, and the value of u follows. Consequently, the field of play for the choice of the value of w on the boundary encompasses a one-dimensional manifold for each boundary point, and the complete determination of these values requires one equation for each boundary point. All the equations need not necessarily give one specific term a value, at one specific boundary point. We can also arrange our determining conditions so that at each point of the boundary there is one equation containing both terms, varying in form continuously with the position of that point. Alternatively, dividing the boundary into several parts, we associate to each point of one part, $n-1$ specific points, one in each of the other parts, and set up n equations for these n points, varying in form continuously as the points vary. However, these conditions, whose totality forms a continuous manifold, and which are expressed as equations between arbitrary functions, must be subject to a certain restriction in order to be necessary and sufficient to determine a single function continuous in the interior of the domain. The restriction is given by particular conditions (equations containing arbitrary constants), since the exactitude of our determination clearly does not extend as far as these.

For the case where the domain of the variable z is represented by a multiply connected surface, our treatment needs no essential modification. Applying the theorem of Section 18 gives a function with the above properties, up to the variation across the transverse cuts. These variations can be made equal to 0, provided that the boundary conditions contain as many available constants as there are cuts.

The case where continuity is interrupted along a line can be reduced to the foregoing by regarding this line as a cut of the surface.

Finally, suppose that we relax continuity at certain isolated points, that is, by Section 12, allow points where the function becomes infinite. Leaving unchanged the other hypotheses of the case initially studied, we can give arbitrarily a function of z, whose subtraction makes the function we must determine a continuous function. The original function is completely determined by this. Take the quantity $\alpha + \beta i$ equal to the given function in a circle, as small as we wish, with center at the point of discontinuity, and elsewhere conforming to the previous specification. The integral

$$\int \left(\left(\frac{\partial \alpha}{\partial x} - \frac{\partial \beta}{\partial y} \right)^2 + \left(\frac{\partial \alpha}{\partial y} + \frac{\partial \beta}{\partial x} \right)^2 \right) dT$$

over the interior of the circle is 0, and is finite over the remainder of the surface. One can apply the theorem of the previous section to obtain a function with the required properties. We deduce with the help of the theorem of Section 13, that in general, if the function becomes infinite of order n at an isolated discontinuity, we have $2n$ constants at our disposal.

We represent geometrically (by Section 15) a function w of a complex quantity z varying over the interior of a given two-dimensional domain. To a given surface T, covering A, corresponds an image S, covering B, which after excluding isolated points is similar to T in its smallest parts. The conditions found previously, necessary and sufficient to determine the function, are relative to the values on the contour and at the points of discontinuity. They thus arise (according to Section 15) as conditions on the position of the boundary of S, giving a single equation to be satisfied for each boundary point. If each of these equations concerns just one boundary point the conditions are represented geometrically by a family of curves, one of which gives the location of each boundary point. When two boundary points, one varying continuously with the other, appear in a pair of these equations, it follows that between two parts of the contour there is a dependence which fixes one part once the position of the other part is arbitrarily assigned. In a similar way one obtains, for other forms of these equations, a geometrical interpretation, which we shall not pursue here.

20.

The introduction of complex quantities into mathematics has its origin and immediate purpose in the theory of simple [6] laws of dependence between variables arising from algebraic operations. For if the field of these laws of dependence is extended by permitting the variable quantities to have complex values, a formerly hidden harmony and regularity emerges. The cases where this extension has been made are admittedly of limited scope, and reduce

[6]Here we consider the elementary operations to be addition, subtraction, multiplication, division, integration and differentiation: and *simpler* dependence indicates that fewer elementary operations are required. In fact, all functions used up to now in analysis can be defined via a finite number of these operations.

almost entirely to the laws of dependence between two variable quantities of which one is either an algebraic function of the other[7]; or, a function whose derivative is algebraic. Nearly all writings on the subject not only gave simpler, more effective form to results established without the help of complex quantities, but opened the way to new results, as witnessed by the history of the study of algebraic functions, circle or exponential functions, elliptic and Abelian functions.

We indicate briefly the progress which the present study yields for the theory of the above functions.

Previous methods of treating these functions always based the definition of the function on an expression that yields its value for each value of the argument. Our study shows that, because of the general nature of a function of a complex variable, a part of the determination through a definition of this kind yields the rest. Indeed, we reduce this part of the determination to that which is necessary for complete determination of the function. This essentially simplifies the discussion. For example, to show the equality of two expressions for the same function, one formerly needed to transform one into the other: that is, show that the expressions coincide for every value of the variable. Now it suffices to prove the expressions coincide in a much more restricted domain.

A theory of these functions based on the principles introduced here would fix the form of the function (that is, its value for every value of the argument) independently of a means of determining the function via operations on quantities. We determine the function by appending to the general idea of a function of a complex variable just those features that are necessary for the determination. Only at this stage do we pass to the various representations that the function permits. The common character of a class of functions formed in a similar way by operations on quantities, is then represented in the form of boundary conditions and discontinuity conditions imposed on them. Suppose for example that the domain of a complex variable extends over the plane in either simple or multiple fashion, that the function has only isolated points of discontinuity interior to the domain, and only has infinities of finite order. (For infinite z, the quantity z itself; for finite z', the order of $\frac{1}{z-z'}$, is an infinite quantity of first order.) Then the function is necessarily algebraic, and conversely each algebraic function fulfills these conditions.

We leave aside for the present the development of this theory which, as

[7]That is, an algebraic equation holds that connects the two quantities.

observed above, will shed light on the simple laws of dependence arising from operations on quantities, since we exclude for the present the discussion of expressions for a function.

On the same grounds we do not consider here the utility of our results as the basis of a general theory of the laws of dependence. For this, one would need to show that the concept of a function of a complex variable, taken as fundamental here, coincides fully with that of a dependence[8] that can be expressed in terms of operations on quantities.

21.

In order to clarify our general results, it is useful to give an example of their application.

The application mentioned in the previous section, while its discussion attains the immediate objective, nevertheless is somewhat special. When dependence is regulated by a finite number of operations of elementary type, the function contains only a finite number of parameters. As for the form of a system of mutually independent boundary conditions and discontinuity conditions, sufficient to determine the function, this implies that they cannot include arbitrary conditions at each point of a line. For our present objective it is more appropriate to take an example not of this type, but rather an example where the function of a complex variable depends on an arbitrary function.

For a clear and convenient framework, we present this example in the geometric form used at the end of Section 19. It then appears as a study of the possibility of forming a connected image of a given surface, similar to the surface in the smallest parts, whose form is specified. That is, all points of the boundary should be located on particular curve; moreover, recalling Section 5, the sense of the boundary, along with the branch points, are given. We restrict ourselves to the solution of this problem in the case where every point of one surface corresponds to only one point of the other, and the surfaces are simply connected. For this case, the solution is contained in the following theorem.

Two given simply connected plane surfaces can always be related in such

[8]The dependence expressed here denotes dependence via a finite or infinite number of the four simplest operations, addition, subtraction, multiplication and division. The expression 'operations on quantities' (by contrast to 'operations on numbers') indicates operations in which the rationality of the quantities does not play a role.

a way that each point of one surface corresponds to a point of the other, varying continuously with that point, with the corresponding smallest parts similar. One interior point, and one boundary point, can be assigned arbitrary corresponding points; however, this determines the correspondence for all points.

When two surfaces T and R both correspond to a third surface S in such a way that similarity holds between corresponding smallest parts, this yields a correspondence between T and R that clearly has the same property. The problem of producing a correspondence between two arbitrary surfaces with similarity between the smallest corresponding parts now reduces to mapping an arbitrary surface onto a specific surface, with similarity in the smallest parts. To prove our result, draw a circle K in the plane B with radius 1 and center $w = 0$. We need only show that an arbitrary simply connected surface T spread over A can be mapped in a connected way throughout the circle K, with similarity of the smallest parts; and this in a unique way once the center corresponds to an arbitrary given interior point O_o, while an arbitrary given point of the circumference corresponds to an arbitrary given boundary point O' of the surface T.

We denote the particular values of z, Q for the points O_o, O' by corresponding indices, and describe in T an arbitrary circle Θ with midpoint O_o that does not reach the boundary of T and contains no branch point. We introduce polar coordinates: let $z - z_0 = re^{i\phi}$, then $\log(z - z_0) = \log r + \phi i$. The real part is continuous in the entire circle except at the point O_o, where it becomes infinite. Among the possible values of ϕ, we take the smallest positive values, so that along the radius where $z - z_0$ is real and positive, the imaginary part is 0 on one side, and 2π on the other side, of this radius, while varying continuously elsewhere. Clearly we can replace this radius by an arbitrary line ℓ from the center to the boundary in such a way that $\log(z - z_0)$ has a jump of $2\pi i$ where O crosses this line from the negative side (that is, according to Section 8, the side of negative p) to the positive side; otherwise, $\log(z - z_0)$ varies continuously in the whole circle Θ.

We now take the complex function $\alpha + \beta i$ of x, y equal to $\log(z - z_0)$ in the circle, while, outside the circle, having extended ℓ in any fashion up to the boundary T, the function is specified as follows:

1) on the circumference of Θ, the function is $\log(z - z_0)$; on the boundary of T, it is purely imaginary.

2) on crossing from the negative side of ℓ to the positive side the function varies by $-2\pi i$, while otherwise it varies with an infinitely small change of

location by an infinitely small quantity of the same order;

this is always possible. The integral

$$\int\left(\left(\frac{\partial\alpha}{\partial x}-\frac{\partial\beta}{\partial y}\right)^2+\left(\frac{\partial\alpha}{\partial y}+\frac{\partial\beta}{\partial x}\right)^2\right)dT,$$

taken over Θ, has the value 0; and over the remainder of the surface, a finite value. Consequently, $\alpha+\beta i$ can be transformed, by adjoining a continuous function, defined up to an imaginary constant, that has imaginary values on the boundary, into a function $t=m+ni$ of z. The real part m of this function will be 0 on the boundary, $-\infty$ at O_o, and vary continuously throughout the remainder of T. For each value of a of m between $-\infty$ and 0, the surface T is divided by a line, on which $m=a$, into a part where $m<a$, containing O_0 in its interior, and parts where $m>a$ whose boundary is made up of the boundary of T and of lines where $m=a$. Either the connectivity of T, which is -1, is unchanged by this decomposition, or T is divided into two pieces with connectivity 0 and -1, or into more than two pieces. However the last case is impossible, since in at least one of these pieces m is finite, continuous, and constant on all parts of the boundary. It is then either constant in a portion of the surface; or has a maximum or minimum value at a single point or along a line, which contradicts Section 11, III.

Thus the points where m takes a constant value form simple closed lines. One of these lines bounds a region that surrounds O_0, and m must decrease towards the interior: consequently, with a positive circuit of the boundary (where by Section 8, s increases) the quantity n, while continuous, is increasing. Now the function has a jump[9] of -2π only with a crossing from the negative to the positive side of the line ℓ, and so it must take each value between 0 and 2π, excluding multiples of 2π, exactly once. Write $e^t=w$, then e^m and n are polar coordinates of Q with respect to the center of the circle K. The totality of points Q clearly forms a surface S spread out simply over the whole of K; the point Q_o falls at the center of the circle. However, the point Q' can be placed at an arbitrary given point of the circumference by means of the constant still at our disposal in n. This completes the proof.

[9]The line ℓ joins one point interior to this piece to another; and so, if it cuts the boundary several times, it must pass from inside to outside one more time than from outside to inside. Thus the sum of the jumps of n throughout a positive circuit is always -2π.

In the case where the point O_o is a branch point of order $n-1$, we reach our goal by replacing $\log(z-z_0)$ by $\frac{1}{n}\log(z-z_0)$; the reasoning is analogous, and the treatment is easily completed with the help of Section 14.

22.

The complete extension of the investigation in the previous section to the general case, where a point of one surface corresponds to several points on the other, and simple connectedness is not assumed for the surfaces, is left aside here. Above all this is because, from the geometrical point of view, our entire study would need to be put in a more general form. Our restriction to plane surfaces, smooth except for isolated points, is not essential: rather, the problem of mapping one arbitrary given surface onto another with similarity in the smallest parts, can be treated in a wholly analogous way. We content ourselves with a reference to two of Gauss's works; that cited in Section 3, and *Disquisitiones generales circa superficies curvas*, §13.

Contents.

II.

On the laws of distribution of electric charge in a ponderable body that is neither a perfect conductor nor insulator, but may be viewed as resisting the holding of electric charge with finite power.

(*Official report of the 31st assembly of German scientific researchers and physicians, Göttingen, September 1854.*)

The formation of residual charge in the Leyden jar and the formation of static electricity in other apparatus has been examined using the ingenious instruments for static electricity, mentioned by Prof. Kohlrausch in the previous session of this section. The phenomenon is essentially the following: If we let a Leyden jar stand for a long time while charged, discharge it and than let it stand isolated for a while, then a noticeable charge reappears. This leads to the assumption that in the first discharge only a part of the distinct types of electricity reunites; a part, however, remains in the jar. The first part is called the disposable charge, the second the residual charge. The precision of the measurements made by Prof. Kohlrausch, of the loss of the disposable charge and the return of the residual charge, allows me to test a law which was believed on other grounds, which fills a gap in the previous theory of static electricity.

As is well known, the mathematical investigations of static electricity are in reference to its distribution for complete and totally insulated conductors. We consider a ponderable body as either a perfect conductor or insulator. As a consequence, this theory says that when in equilibrium, the aggregate of the electric charge accumulates only at the boundary surface of the conductors and insulators. Admittedly, however, this is a mere fiction. In nature there are neither bodies which cannot be penetrated by electric charge, nor bodies for which the total electric charge collects on a a mathematical surface. Rather we should assume that a ponderable body resists the absorption or holding of the electric charge with finite strength. We should assume the hypothesis, whose consequences seem to agree with reality, that it does not resist *becoming* electrified or *taking* an electric charge, but resists *being* electrified or *holding* electric charge. The law of this resistance is, depending on the dual or unitary conception, the following. By the dual conception, under which which the electric charge is the excess of the positive electricity over the negative, at each point of the ponderable body, this is caused by the intensity of the density of the electric charge trying to decrease in proportion to the density of the excess, with the same sign as the excess (or, trying to increase with a negative excess). From the unitary point of view, by which the electric charge is the excess of electricity in the body over what is natural for it, we must assume this is caused at each point of the body by the intensity of the electric density trying to decrease in proportion to the density of the excess (or increase, in the case of a negative excess).

Besides these reasons for flow, if no noticeable thermal or magnetic or inductive effects or influences take place and if the ponderable bodies rest against one another, the electromotive force is brought into the calculation in conformity with Coulomb's Law. Under the same circumstances we can assume the independence of the consequential current from the current caused by the proportionality between the electromotive force and the current intensity.

In order to express these rules for flow in a formula, let x, y, z be rectangular coordinates, and at (x, y, z) when the time is t, let the density of the electric charge be ϱ. Let u be the 4πth part of the potential of the total charge using Gauss's definition (by which the potential at a particular point is equal to the integral over the quantity of electric charge at each point divided by its distance from the point). Then the electromotive force attributable to Coulomb's Law, when decomposed into the direction of the three axes, is proportional to

$$-\frac{\partial u}{\partial x}, \quad -\frac{\partial u}{\partial y}, \quad -\frac{\partial u}{\partial z}.$$

That arising from the reaction of the ponderable body is proportional to

$$-\frac{\partial \varrho}{\partial x}, \quad -\frac{\partial \varrho}{\partial y}, \quad -\frac{\partial \varrho}{\partial z}.$$

Hence the components of the electromotive force can be set equal to

$$-\frac{\partial u}{\partial x} - \beta^2 \frac{\partial \varrho}{\partial x}, \quad -\frac{\partial u}{\partial y} - \beta^2 \frac{\partial \varrho}{\partial y}, \quad -\frac{\partial u}{\partial z} - \beta^2 \frac{\partial \varrho}{\partial z},$$

where β^2 depends only on the nature of the ponderable body. These components are now equal to quantities proportional to the components of the current intensity, $\alpha\xi, \alpha\eta, \alpha\zeta$ say. Here one designates by ξ, η, ζ the components of the current intensity and by α a constant that depends on the nature of the ponderable body.

We combine this with the phoronomic equation

$$\frac{\partial \varrho}{\partial t} + \frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial y} + \frac{\partial \zeta}{\partial z} = 0,$$

which is obtained by expressing in two ways the incoming electricity in the time element dt into the space element $dxdydz$, and the equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = -\varrho$$

which follows from the concept of potential. We obtain

$$\alpha \frac{\partial \varrho}{\partial t} + \varrho - \beta^2 \left\{ \frac{\partial^2 \varrho}{\partial x^2} + \frac{\partial^2 \varrho}{\partial y^2} + \frac{\partial^2 \varrho}{\partial z^2} \right\} = 0$$

after first multiplying by α and then replacing ξ, η, ζ by their values.

This gives a partial differential equation for u, which is of the first degree in t and of the fourth degree with respect to the space coordinates. In order to determine u completely, from a particular time onwards, inside the ponderable body, we need, besides this equation, one condition for each point of the ponderable body at the starting time, and for the following times two conditions at each point on the surface.

I will now compare the implications of this law with experience in some particular cases.

For equilibrium (in a system of insulated conductors) we have

$$\frac{\partial u}{\partial x} + \beta^2 \frac{\partial \varrho}{\partial x} = 0, \quad \frac{\partial u}{\partial y} + \beta^2 \frac{\partial \varrho}{\partial y} = 0, \quad \frac{\partial u}{\partial z} + \beta^2 \frac{\partial \varrho}{\partial z} = 0$$

or

$$u + \beta^2 \varrho = \text{Const.},$$

or, since

$$-\varrho = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2},$$

$$u - \beta^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) = \text{Const.}$$

For current equilibrium or the steady state distribution (in the closed loops in a constant circuit) we have

$$\frac{\partial \varrho}{\partial t} = 0$$

or

$$\varrho - \beta^2 \left(\frac{\partial^2 \varrho}{\partial x^2} + \frac{\partial^2 \varrho}{\partial y^2} + \frac{\partial^2 \varrho}{\partial z^2} \right) = 0.$$

Now if the length β is very small compared to the dimensions of the ponderable body, then $u-$Const., in the first case, and ρ, in the second, decreases very rapidly off the surface and is very small in the interior, and these quantities indeed change with respect to distance p from the surface approximately like $e^{-p/\beta}$. This must be the case for a metallic conductor. If we set $\beta = 0$, then the known formula for a perfect conductor is obtained.

When applying this rule to the formation of the residual change in the Leyden jar, I must assume that its dimensions can be considered infinitely large compared to the distance of the condensor plates, since information concerning the dimensions of the apparatus is lacking. I will not bore this distinguished audience by carrying out the calculation, but content myself with stating its result.

The measurements of Professor Kohlrausch show that the disposable charge is closely represented by a parabola, when considered as a function of time. Indeed, the parameter of the parabola, which matches the charge curve most closely, slowly decreases, so that if we denote the initial charge by L_0 and the charge at time t by L_t, then $\frac{L_0 - L_t}{\sqrt{t}}$ is a quantity which gradually decreases with increasing t.

The same results from the calculation, if it was assumed that both α and β^2 are very large for glass, as was to be expected here, and may be regarded as infinitely large, while their quotient remains finite. I have not made a close comparison of the calculation with experimental data, since I lack information about the dimensions of the apparatus and, in general, the methods used, in order to determine the necessary corrections in the calculations on account of the differences from the hypotheses. In particular, the determination of the electrical constant for glass is to be desired. But I regard the law derived here for the distribution of electric charge as being fully confirmed by the measurements of Professor Kohlrausch.

I will speak briefly about the application of this rule in a different setting.

The propagation of an electric current in a metallic conductor is well known, as is its consequence, the steady state current caused by the static electricity arising from a constant, or slowly changing, electromotive force. This process, because of its extremely short duration, and the thermal and magnetic effects, is accessible to experimental research only in its effects. The only experimental results available are the measurements of the speed of propagation in telegraph wires, and Ohm's laws of current equilibrium. A more careful analysis of Ohm's laws lead us to the assumptions made here, and, indeed, I was led to them in that way.

Ohm determined the current distribution for current equilibrium by the following two conditions.

1) In order to actually obtain the current intensity as proportional to the electromotive force, we must introduce forces other than the external electromotive force that are derivatives of a function of position, the voltage.

2) For current equilibrium, at each point of the ponderable body, the same amount of electricity flows in as out.

Ohm believed that the voltage, that function of position of which the inner electromotive forces are the derivatives, was dependent on the static electric charge, in proportion to its density. That assumption, in fact, turns out to explain both conditions. However, Professors Weber[1] and Kirchoff[2] have noted, almost simultaneously, that the electricity would have to be in equilibrium when it filled the ponderable body with uniform density, which is indeed the experience under equilibrium after the electricity is distributed on the surface. The voltage must be a function that is constant in the entire conductor when in equilibrium, and thus rather proportional to the potential of the electric charge. These internal electromotive forces are identical with those that follow from Coulomb's Law.

These views about voltage were also assumed by most investigators. However, the cause of the second condition for current equilibrium, that each part of a ponderable body has a constant electrical charge, has not been investigated.

By the dual concept the quantity of negative electricity must remain constant, as well as the positive. It seems that we can explain that no noticeable excess of one of the electricities forms, by Coulomb's Law, from the attraction of the opposite electricities (at least so long as we do not enter into the proportionalities more precisely). We must still assume a reason that the neutral electricity remains constant at each part of the body, and thus

[1] *Abhandlungen d. k. sächs. Ges. d. W.*, 1852, I, p. 293.
[2] *Poggendorff's Annalen*, vol. 79, p. 506.

assume an influence of the ponderable on it. On the suggestion of Professor Weber, I have tried, for several years, to make this assumption the object of calculations, without obtaining satisfactory results.

By the unitary concept only one mechanism is needed, which tries to keep the quantity of electricity contained in a ponderable body constant. We are led immediately to the assumption above, that every ponderable body tries to possess electricity of a certain density, and resists being filled by either a larger or a smaller quantity. This law of resistance can be assumed to take the form that has been confirmed by experiment for glass.

These considerations lead to the acceptance of Franklin's original view about electrical phenomena. In order to lay a foundation for the deeper understanding of these phenomena, either by themselves or with other phenomena, and as a basis for further development and modification, we must subject them to the demands and hints of experiment.

I hope that the distinguished circle of scientists, before whom I have the honor of developing this topic, will find it worthwhile to subject it to closer scrutiny.

III.

On the theory of Nobili's color rings.

(*Poggendorff's Annalen der Physik und Chemie*, vol. 95, March 28, 1855.)

Nobili's color rings represent a valuable tool for the experimental study of the laws of current flow in a body made conducting by decomposition. The way to produce these rings is the following. We pour a solution of lead oxide in concentrated potassium hydroxide over a plate of platinum, gold plated silver, or German silver and let the current from a strong battery flow, through the tip of a fine platinum wire melted in a glass tube, into the fluid layer and exit through the plate. The anion, lead peroxide according to Beetz, is deposited on the metal plate in a delicate transparent layer which has varying thickness depending on the distance from the cathode, so that the plate after removal of the fluid shows Newton's color rings. The relative thickness of the layer at different distances can be determined from the color-rings. Using Faraday's Law (by which the amount deposited must be everywhere proportional to the amount of electricity that passed through), we derive the distribution of the current leaving the fluid.

The first attempt to determine the current flow by calculation and compare the result obtained with experiment was made by E. Becquerel. He assumed that the length and width of the liquid layer as compared to its thickness could be considered as infinitely large, that the current entered at a point on its upper surface, and that it spread out by Ohm's law in that surface. He now believed that using these assumptions the flow curves could be considered as straight lines (with insignificant error). From that assumption he derived the law that the thickness of the deposited layer is inversely proportional to the distance from the cathode. He corroborated the law experimentally.

On the other hand, Mr. Du-Bois-Reymond has shown in a talk to the Berlin Physical Society, that the assumption of straight flow lines resulted in the thickness of the deposited substance at the endpoint being proportional to the inverse cube of the distance. This caused Mr. Beetz to conduct a series of confirming experiments, described in vol. 71, p. 71 of *Poggendorff's Annalen*, that inspire much confidence.

Precise calculation meanwhile shows that the hypothesis of straight lines for current flow is inadmissible and leads to incorrect results. Indeed, the flow lines, at least for large distances to the exit point (since they lie between two very close parallel lines and have at most one inflection point) in the middle part of their path, bend hardly at all. By no means, however, can we conclude that there is not a significant error by replacing them by straight lines running from the cathode to the exit point. I will first develop by exact calculation the consequences ensuing from the assumptions of Mr. E. Becquerel and Mr. Du-Bois-Reymond, and finally come back to the experiments of Mr. Beetz.

I assume that the cathode is in the fluid layer, which is bounded by two horizontal planes, and is restricted to a point. For a point of the fluid layer, denote its horizontal distance from the cathode by r, the height over the lower boundary surface by z. We denote the elevation of its voltage from voltage at the upper side of this boundary layer by u. Furthermore, let S be the strength of the current, w the resistivity of the fluid, $z = \alpha$ at the cathode, and $z = \beta$ at the upper surface. Now u must be determined as a function of r and z. The flow intensity at the point $(r, 0)$, which by Faraday's law must be proportional to the thickness that we seek of the layer deposited there, is then equal to the value of $\frac{1}{w}\frac{\partial u}{\partial z}$ at this point.

If it is initially assumed that the expanse of the fluid layer as compared to its thickness may be considered as infinitely large, then the conditions for determining u are:

(1) for $-\infty < r < \infty, \quad 0 < z < \beta,$

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} = 0;$$

(2) for $-\infty < r < \infty, \; z = 0, \qquad u = 0;$

(3) for $-\infty < r < \infty, \quad z = \beta, \qquad \frac{\partial u}{\partial z} = 0;$

(4) for $r = \pm\infty, \quad 0 < z < \beta, \quad u$ is finite;

(5) for $r = 0, \qquad z = \alpha,$

$$\left.\begin{aligned} u &= \frac{wS}{4\pi}\frac{1}{\sqrt{r^2 + (z-\alpha)^2}} \\ \text{or} \qquad &= \frac{wS}{2\pi}\frac{1}{\sqrt{r^2 + (z-\alpha)^2}} \end{aligned}\right\} +$$

a continuous function of r and z, depending on whether the cathode lies in the interior or on the upper surface.

These conditions are satisfied by

$$u = \frac{Sw}{4\pi} \sum_{m=-\infty}^{\infty} (-1)^m \left(\frac{1}{\sqrt{r^2 + (z + 2m\beta - \alpha)^2}} - \frac{1}{\sqrt{r^2 + (z + 2m\beta + \alpha)^2}} \right)$$

or, if we take $S = \frac{4\pi}{w}$ for simplicity:

$$u = \sum_{m=-\infty}^{\infty} (-1)^m \left(\frac{1}{\sqrt{r^2 + (z + 2m\beta - \alpha)^2}} - \frac{1}{\sqrt{r^2 + (z + 2m\beta + \alpha)^2}} \right).$$

If we set $u = a_1 \sin \frac{\pi z}{2\beta} + a_2 \sin 2\frac{\pi z}{2\beta} + a_3 \sin 3\frac{\pi z}{2\beta} + \cdots$, then for even n the coefficient a_n is 0, and for odd n,

$$\begin{aligned} \beta a_n &= \int_0^{2\beta} \sin n \frac{\pi t}{2\beta} \sum_{m=-\infty}^{\infty} \\ &\quad (-1)^m \left(\frac{dt}{\sqrt{r^2 + (t + 2m\beta - \alpha)^2}} - \frac{dt}{\sqrt{r^2 + (t + 2m\beta + \alpha)^2}} \right) \\ &= \int_{-\infty}^{\infty} \left(\sin n \frac{\pi}{2\beta} (t + \alpha) - \sin n \frac{\pi}{2\beta} (t - \alpha) \right) \frac{dt}{\sqrt{r^2 + t^2}} \\ &= 2 \sin n \frac{\pi\alpha}{2\beta} \int_{-\infty}^{\infty} \cos n \frac{\pi t}{2\beta} \frac{dt}{\sqrt{r^2 + t^2}} \\ &= 2 \sin n \frac{\pi\alpha}{2\beta} \int_{-\infty}^{\infty} \frac{e^{n \frac{\pi}{2\beta} ti} dt}{\sqrt{r^2 + t^2}}. \end{aligned}$$

In the last integral we can also write $2 \int_{ri}^{\infty i}$ instead of $\int_{-\infty}^{\infty}$. If we substitute t for the variable tri, we obtain

$$a_n = \frac{4 \sin n \frac{\pi}{2\beta} \alpha}{\beta} \int_1^{\infty} \frac{e^{-n \frac{\pi}{2\beta} rt} dt}{\sqrt{t^2 - 1}},$$

thus

$$u = \sum \sin n \frac{\pi}{2\beta} z \frac{4 \sin n \frac{\pi}{2\beta} \alpha}{\beta} \int_1^{\infty} \frac{e^{-n \frac{\pi}{2\beta} rt} dt}{\sqrt{t^2 - 1}},$$

summed over all positive odd values of n.

Assume the fluid is bounded by $r = c$ and indeed, for instance, by a nonconductor. Then we must have $\frac{\partial u}{\partial r} = 0$ when $r = c$. Hence a function u'', satisfying conditions (1)—(4) below, will be added to the function u obtained above, which we now write as u'.

(1) for $-c < r < c, \quad 0 < z < \beta,$

$$\frac{\partial^2 u''}{\partial r^2} + \frac{1}{r}\frac{\partial u''}{\partial r} + \frac{\partial^2 u''}{\partial z^2} = 0;$$

(2) for $-c < r < c, \quad z = 0, \qquad u'' = 0;$

(3) for $-c < r < c, \quad z = \beta, \qquad \frac{\partial u''}{\partial z} = 0;$

(4) for $r = \pm c, \quad 0 < z < \beta, \qquad \frac{\partial u''}{\partial r} = -\frac{\partial u'}{\partial r};$

and u'' is continuous everywhere.

As a consequence of conditions (1) through (3), it must be possible to represent u'' in the form

$$b_1 \sin \frac{\pi}{2\beta} z + b_3 \sin 3 \frac{\pi}{2\beta} z + b_5 \sin 5 \frac{\pi}{2\beta} z + \cdots$$

and indeed it follows from (1) that b_n satisfies the condition

$$\frac{d^2 b_n}{dr^2} + \frac{1}{r}\frac{db_n}{dr} - \frac{n^2\pi^2}{4\beta^2} b_n = 0.$$

A particular solution of this equation is, as already known, $\int_1^\infty \frac{e^{-n\frac{\pi}{2\beta}rt}dt}{\sqrt{t^2-1}}$. Another is obtained if the same integral is taken from -1 to 1. Thus the general solution is

$$b_n = c_n \int_1^\infty \frac{e^{-n\frac{\pi}{2\beta}rt}dt}{\sqrt{t^2-1}} + \gamma_n \int_{-1}^1 \frac{e^{-n\frac{\pi}{2\beta}rt}dt}{\sqrt{1-t^2}}$$

where c_n and γ_n are constants. If we denote

$$\int_1^\infty \frac{e^{-2qt}dt}{\sqrt{t^2-1}} \quad \text{by} \quad f(q), \qquad \int_{-1}^1 \frac{e^{-2qt}dt}{\sqrt{1-t^2}} \quad \text{by} \quad \phi(q),$$

then

$$b_n = c_n f\left(n\,\frac{\pi}{4\beta}\,r\right) + \gamma_n \phi\left(n\,\frac{\pi}{4\beta}\,r\right).$$

The expansion by increasing powers of q gives

$$f(q) = \sum_{m=0}^{\infty} \frac{q^{2m}}{(m!)^2}\,(\Psi(m) - \log q),$$
$$\phi(q) = \pi \sum_{m=0}^{\infty} \frac{q^{2m}}{(m!)^2}.$$

Hence $f(q)$ is infinite for $q = 0$ and, since u'' remains continuous for $r = 0$, c_n must be 0. From (4) we obtain

$$\gamma_n = -\frac{4 \sin n\,\frac{\pi}{2\beta}\,\alpha\, f'\left(n\,\frac{\pi}{4\beta}\,c\right)}{\beta \phi'\left(n\,\frac{\pi}{4\beta} c\right)}.$$

Hence

$$u = \sum_n \left(\sin n\,\frac{\pi}{2\beta}\,z\right) \frac{4}{\beta}\,\sin n\,\frac{\pi\alpha}{2\beta} \left\{ f\left(n\,\frac{\pi}{4\beta}\,r\right) - \phi\left(n\,\frac{\pi}{4\beta}\,r\right) \frac{f'\left(n\,\frac{\pi}{4\beta} c\right)}{\phi'\left(n\,\frac{\pi}{4\beta} c\right)} \right\},$$

where the sum is over the positive odd values of n.

We may use the half -convergent series

$$f(q) = e^{-2q}\sqrt{\frac{\pi}{4q}} \sum_{m<4q+1} (-1)^m \frac{(1\cdot 3\cdots(2m-1))^2}{m!(16q)^m},$$
$$\phi(q) = e^{2q}\sqrt{\frac{\pi}{4q}} \sum_{m<4q+1} \frac{(1\cdot 3\cdots(2m-1))^2}{m!(16q)^m}$$

to calculate $f(q)$ and $\phi(q)$ for large values of q, which only gives the value up to a fraction of the order of e^{-4q}. If this is not sufficiently accurate, then it is probably most expedient to apply the expansion by increasing powers of q.

Hence for sufficiently large values of $\frac{r}{\beta}$, neglecting quantities of the order of magnitude $e^{-3\frac{\pi r}{2\beta}}$, we obtain

$$u = \sin\frac{\pi z}{2\beta}\frac{4\sin\frac{\pi\alpha}{2\beta}}{\beta}\sqrt{\frac{\beta}{r}}\left\{e^{-\frac{\pi r}{2\beta}}\sum\frac{(1\cdot 3\cdot(2m-1))^2}{m!}\left(-\frac{\beta}{4\pi r}\right)^m\right.$$
$$-\sum\frac{(1\cdot 3\cdots(2m-1))^2}{m!}\left(\frac{\beta}{4\pi r}\right)^m e^{\frac{\pi(r-2c)}{2\beta}}\times$$
$$\left.\frac{\sum\frac{(1\cdot 3\cdots(2m-1))^2(2m+1)}{m!(2m-1)}\left(-\frac{\beta}{4\pi c}\right)^m}{\sum\frac{(1\cdot 3\cdots(2m-1))^2(2m+1)}{m!(2m-1)}\left(\frac{\beta}{4\pi c}\right)^m}\right\}$$

and the thickness of the layer is proportional to $\left(\frac{\partial u}{\partial z}\right)_0$, or proportional to

$$\frac{e^{-\frac{\pi r}{2\beta}}}{\sqrt{r}}\sum\frac{(1\cdot 3\cdots(2m-1))^2}{m!}\left(-\frac{\beta}{4\pi r}\right)^m$$
$$-\frac{e^{\frac{\pi(r-2c)}{2\beta}}}{\sqrt{r}}\sum\frac{(1\cdot 3\cdots(2m-1))^2}{m!}\left(\frac{\beta}{4\pi r}\right)^m\times$$
$$\frac{\sum\frac{(1\cdot 3\cdots(2m-1))^2(2m+1)}{m!(2m-1)}\left(-\frac{\beta}{4\pi c}\right)^m}{\sum\frac{(1\cdot 3\cdots(2m-1))^2(2m+1)}{m!(2m-1)}\left(\frac{\beta}{4\pi c}\right)^m}.$$

This result also holds in general if instead of the cathode being a point, we assume it is an arbitrary surface of revolution. Then for values of r between c and those values up to which conditions (1) to (3) remain valid, u can be represented by a series of the form

$$u = \sum K_n \sin n\frac{\pi z}{2\beta}\left\{f\left(n\frac{\pi r}{4\beta}\right) - \phi\left(n\frac{\pi r}{4\beta}\right)\frac{f'\left(n\frac{\pi c}{4\beta}\right)}{\phi'\left(n\frac{\pi c}{4\beta}\right)}\right\}.$$

An exception would occur only if K_1 were 0.

The special hypothesis made by Mr. E. Becquerel and essentially retained by Mr. Du-Bois-Reymond is that the cathode is a point of the upper surface, thus $\alpha = \beta$. In this case, according to the calculation, the thickness of the layer for large values of $\frac{r}{\alpha}$ is inversely proportional neither to the distance from the cathode (according to Mr. Becquerel), nor its cube (according to Mr. Du-Bois-Reymond). Rather, the thickness diminishes with growing r/α, as a power with the exponent r/α, such that $\frac{\alpha \log\left(\frac{\partial u}{\partial z}\right)_0}{r}$ approaches a fixed limit $-\pi/2$, to any degree of accuracy. On the other hand the law of Du-Bois-Reymond is not only approximately true for large values of r/α, but is strictly true if $\beta = \infty$, since then

$$u = \sum_{m=-\infty}^{\infty} (-1)^m \left(\frac{1}{\sqrt{r^2 + (z + 2m\beta - \alpha)^2}} - \frac{1}{\sqrt{r^2 + (z + 2m\beta + \alpha)^2}} \right)$$

reduces to

$$\frac{1}{\sqrt{r^2 + (z-\alpha)^2}} - \frac{1}{\sqrt{r^2 + (z+\alpha)^2}}$$

and consequently

$$\left(\frac{\partial u}{\partial z}\right)_0 \text{ reduces to } \frac{2\alpha}{(r^2 + \alpha^2)^{3/2}}.$$

However the assumption from which the result is derived, namely, that the flow lines may be considered as straight, is by no means confirmed. The equation for the flow lines is

$$\int \left(r \frac{\partial u}{\partial z} dr - r \frac{\partial u}{\partial z} dz \right) = v = \text{const.}$$

Indeed, the constant multiplied by $2\pi/w$, if one takes the integral so that it vanishes for $r = 0$, is equal to the part of the flow within a surface of rotation

$$v = \text{const.}$$

So in our case the flow lines are the lines obtained from the equation

$$v = 2 - \frac{z+\alpha}{\sqrt{r^2 + (z+\alpha)^2}} \pm \frac{z-\alpha}{\sqrt{r^2 + (z-\alpha)^2}} = \text{const.}$$

These lines vary considerably from a straight line for large values of the constant. Mr. Du-Bois-Reymond, indeed, made the assumption that the cathode is on the upper surface, but his later conclusions were not essentially dependent on this assumption. This suggests the conjecture that the experiments of Mr. Beetz, which yielded results not too far from the law of the cube, would not depend on the demand of Mr. Du-Bois-Reymond that the cathode is on the upper surface of the fluid. Rather, Mr. Beetz, for greater convenience, used more fluid, so that in the series for $\left(\frac{\partial u}{\partial z}\right)_0$,

$$\sum_{m=0}^{\infty}(-1)^m\left(\frac{2m\beta+\alpha}{(r^2+(2m\beta+\alpha)^2)^{3/2}}-\frac{2m\beta-\alpha}{(r^2+(2m\beta-\alpha)^2)^{3/2}}\right),$$

the later terms, or indeed their sum, can be neglected as compared to the first. In this case the elegant experiments of Mr. Beetz could actually be viewed as a proof that the current flow almost follows from the assumed laws. However, should this conjecture be false, then from Mr. Beetz's experiments we would draw the conclusion that still other conditions must be considered when calculating the current flow, whose determination would require a new experimental investigation.

IV.

Contributions to the theory of the functions represented by the Gauss series $F(\alpha, \beta, \gamma, x)$.

(*Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, VII, 1857.*)

The Gauss series $F(\alpha, \beta, \gamma, x)$, as a function of the fourth component x, represents this function only when the modulus of x does not exceed unity. In order to study this function in its whole domain, with unrestricted variability of x, two methods have been presented in previous works. One can either proceed from a linear differential equation which it satisfies, or from the representation of the function as a definite integral. Each of these methods has its own advantages. However, up to now, in the comprehensive treatment of Kummer in Crelle's Journal, Vol. XV, and in the yet unpublished researches of Gauss[1], only the first method is used. The reason for this may well be that calculation with definite integrals between complex limits was insufficiently developed, or could not be assumed familiar to a wide readership.

In the present work I have treated these transcendental functions by a new method, which essentially applies to any function that satisfies a linear differential equation with algebraic coefficients. The method yields results almost directly from the definition, that were formerly obtained only after somewhat troublesome calculations. This has been done in the part of the work presented here, mainly in order to give a summary of the possible representations of the functions, in view of their numerous applications in physics and astronomy. It is necessary to make some general preliminary remarks on the treatment of a function whose argument varies without restriction.

For ease of visualization, consider the value of an independent variable $x = y + zi$ as a point of an infinite plane with rectangular coordinates y, z. Suppose the function w is given in one part of the plane. By an easily proved theorem, one can extend the function continuously outside this domain, so that it satisfies the equation $\frac{\partial w}{\partial z} = i\frac{\partial w}{\partial y}$, in only one way. Evidently this extension should not take place merely along lines where a partial differential equation could not apply, but must be made along strips of finite width.

For 'multivalued' functions like those studied here, in other words, functions that can take different values for the same value of x, according to the

[1]Gauss, Collected Works vol. III, 1886, p.207. W.

way in which the extension is carried out, there are certain points of the x-plane around which the function extends into another. For example, a is such a point in the case of $\sqrt{x-a}$, $\log(x-a)$, $(x-a)^{\mu}$, where μ is not an integer.

Consider an arbitrary line drawn from the point a. The value of the function in a neighborhood of a can be chosen so that it is continuous outside the line; however, the function takes different values on opposite sides of the line. Thus the extension of the function across the line yields a function different from the existing one.

For simplicity, the different extensions of a single function defined in the same part of the x-plane will be called branches of this function. A point x around which one branch of the function extends into another will be called a branch point. At a value where no branching occurs, the function will be called single-valued or monodromic.

1.

I denote by

$$P\begin{Bmatrix} a & b & c & \\ \alpha & \beta & \gamma & x \\ \alpha' & \beta' & \gamma' & \end{Bmatrix}$$

a function of x satisfying the following conditions:

1) It is finite and single-valued for all x except a, b, c.

2) Any three branches P', P'', P''' of the function are connected by a linear homogeneous equation with constant coefficients:

$$c'P' + c''P'' + c'''P''' = 0.$$

3) The function can be written in the forms

$$c_\alpha P^{(\alpha)} + c_{\alpha'} P^{(\alpha')}, c_\beta P^{(\beta)} + c_{\beta'} P^{(\beta')}, c_\gamma P^{(\gamma)} + c_{\gamma'} P^{(\gamma')}$$

with constants $c_\alpha, c_{\alpha'}, \ldots, c_{\gamma'}$. Here

$$P^{(\alpha)}(x-a)^{-\alpha}, P^{(\alpha')}(x-a)^{-\alpha'}$$

remains single valued, and does not become zero or infinite, at $x = a$; likewise for $P^{(\beta)}(x-b)^{-\beta}$, $P^{(\beta')}(x-b)^{-\beta'}$ at $x = b$, and $P^{(\gamma)}(x-c)^{-\gamma}$, $P^{(\gamma')}(x-c)^{-\gamma'}$ at

$x = c$. For the six numbers $\alpha, \alpha', \ldots, \gamma'$, suppose that none of the differences $\alpha - \alpha'$, $\beta - \beta'$, $\gamma - \gamma'$ is an integer, and that

$$\alpha + \alpha' + \beta + \beta' + \gamma + \gamma' = 1.$$

We leave undecided for now the question of how diverse are the functions that satisfy these conditions. The answer will emerge in the course of our investigation (Section 4). For convenience, I refer to x as the variable; a, b, c as the first, second and third branch points; and α, α'; β, β'; γ, γ' as the first, second and third exponent pairs of the P-function.

2.

There are some immediate consequences of the definition.

In the function

$$P\begin{Bmatrix} a & b & c & \\ \alpha & \beta & \gamma & x \\ \alpha' & \beta' & \gamma' & \end{Bmatrix}$$

the three columns can be permuted arbitrarily, and α, α' can be interchanged, as can β, β' and γ, γ'. Further,

$$P\begin{Bmatrix} a & b & c & \\ \alpha & \beta & \gamma & x \\ \alpha' & \beta' & \gamma' & \end{Bmatrix} = P\begin{Bmatrix} a' & b' & c' & \\ \alpha & \beta & \gamma & x' \\ \alpha' & \beta' & \gamma' & \end{Bmatrix}.$$

Here x' is a rational expression of first degree in x that takes the values a', b', c' for $x = a, b, c$.

The function

$$P\begin{Bmatrix} 0 & \infty & 1 & \\ \alpha & \beta & \gamma & x \\ \alpha' & \beta' & \gamma' & \end{Bmatrix},$$

to which any P-function with the same $\alpha, \alpha', \ldots, \gamma'$ can be reduced, will be denoted briefly by

$$P\begin{pmatrix} \alpha & \beta & \gamma & \\ \alpha' & \beta' & \gamma' & x \end{pmatrix}.$$

In such a function each pair α, α'; β, β'; γ, γ' can be interchanged, and the three pairs can be permuted arbitrarily, provided we substitute for x a rational expression of first degree that takes the values $0, 1, \infty$ for the values of x corresponding to the first, second and third exponent pairs respectively. In this way we express the function

$$P\begin{pmatrix} \alpha & \beta & \gamma & x \\ \alpha' & \beta' & \gamma' & \end{pmatrix}$$

via P-functions with the variables $x, 1-x, \frac{1}{x}, 1-\frac{1}{x}, \frac{x}{x-1}, \frac{1}{1-x}$, and the same exponents in a different order.

From the definition, we also have

$$P\begin{Bmatrix} a & b & c & \\ \alpha & \beta & \gamma & x \\ \alpha' & \beta' & \gamma' & \end{Bmatrix}\left(\frac{x-a}{x-b}\right)^{\delta} = P\begin{Bmatrix} a & b & c & \\ \alpha+\delta & \beta-\delta & \gamma & x \\ \alpha'+\delta & \beta'-\delta & \gamma' & \end{Bmatrix},$$

and consequently

$$x^{\delta}(1-x)^{\epsilon}\, P\begin{pmatrix} \alpha & \beta & \gamma & x \\ \alpha' & \beta' & \gamma' & \end{pmatrix} = P\begin{pmatrix} \alpha+\delta & \beta-\delta-\epsilon & \gamma+\epsilon & x \\ \alpha'+\delta & \beta'-\delta-\epsilon & \gamma'+\epsilon & \end{pmatrix}.$$

By this transformation, two exponents from different pairs can take arbitrary given values. Since $\alpha+\alpha'+\beta+\beta'+\gamma+\gamma'=1$, one can introduce any set of values for which the differences $\alpha-\alpha'$, $\beta-\beta'$, $\gamma-\gamma'$ are the same as before. Consequently I will write later, for convenience,

$$P(\alpha-\alpha', \beta-\beta', \gamma-\gamma', x)$$

for all the functions that can be given the form

$$x^{\delta}(1-x)^{\epsilon}\, P\begin{pmatrix} \alpha & \beta & \gamma & x \\ \alpha' & \beta' & \gamma' & \end{pmatrix}.$$

3.

Before all else, it is now necessary to study the behavior of the function somewhat more closely. To this end, consider a closed line ℓ passing through the branch points of the function, that divides the complex numbers into two regions. In each region, a given branch of the function is continuous and distinct from the other branches. Along the different portions of the boundary, different relations hold between the branches belonging to the two domains. To represent these conveniently, denote the linear expression

$pt + qu$, $rt + su$, formed from t, u via the system of coefficients $S = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$, by

$$(S)(t, u).$$

By analogy with Gauss's proposed name 'positive lateral unit' for $+i$, the positive side of a line with given direction denotes the side that lies in relation to the line as $+i$ lies with respect to 1 (the left side, in the usual representation of the complex numbers). Thus x makes a positive circuit around the branch point a when it travels around a contour that encloses a region containing a and no other branch point, in the positive direction with respect to the direction from inside to outside the region.

Suppose now that the line ℓ passes through c, b, a in that order. In the region on the positive side of ℓ, let P', P'' be two branches of the function, not in constant ratio. For another branch P''', the coefficient c''' in the hypothesized relation

$$c'P' + c''P'' + c'''P''' = 0$$

cannot vanish. Hence P''' is a linear combination of P', P'' with constant coefficients. Now suppose that P', P'' pass into $(A)(P', P'')$, $(B)(P', P'')$, $(C)(P', P'')$ when x makes a positive circuit around a, b, c respectively. Then the periodicity of the function is fully determined by the coefficients of $(A), (B), (C)$.

There are, however, further relations between these coefficients. Let x run along the negative side of the line ℓ. The functions P', P'' must recover their original values, since the path describes in a negative sense the entire boundary of a region, inside which the functions are single valued. This amounts to moving x from one of the values c, b, a to the next along the positive side, making each time a positive circuit around this value. In this way (P', P'') passes successively into $(C)(P', P'')$, $(C)(B)(P', P'')$, $(C)(B)(A)(P', P'')$. Consequently

$$(C)(B)(A) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1}$$

This equation gives four conditions to be satisfied by the twelve coefficients of A, B, C.

For the discussion of the above conditions I restrict myself, to fix ideas, to the function

$$P\begin{pmatrix} \alpha & \beta & \gamma & \\ \alpha' & \beta' & \gamma' & x \end{pmatrix};$$

and so to the case $a = 0$, $b = \infty$, $c = 1$. This does not essentially restrict the generality of the result. For the line ℓ through 1, ∞, 0, take the real axis. This must run from $-\infty$ to $+\infty$ in order to pass successively through c, b, a. Within the region on the positive side of this line, which comprises the complex numbers with positive imaginary parts, the components $P^\alpha, P^{\alpha'}, P^\beta, P^{\beta'}, P^\gamma, P^{\gamma'}$ of P, characterized above, are single valued. These components are determined up to constant factors depending on the choice of $c_\alpha, c_{\alpha'}, \ldots, c_{\gamma'}$, once P is given. The functions $P^\alpha, P^{\alpha'}$ become $P^\alpha e^{\alpha 2\pi i}$, $P^{\alpha'} e^{\alpha' 2\pi i}$ on a positive circuit of x around 0. Likewise P^β, $P^{\beta'}$ become $P^\beta e^{\beta 2\pi i}$, $P^{\beta'} e^{\beta' 2\pi i}$ and P^γ, $P^{\gamma'}$ become $P^\gamma e^{\gamma 2\pi i}$, $P^{\gamma'} e^{\gamma' 2\pi i}$ on a positive circuit of x around ∞ and 1 respectively. Denote by P' the value taken by P after a positive circuit around 0. If $P = c_\alpha P^\alpha + c_{\alpha'} P^{\alpha'}$, then

$$P' = c_\alpha e^{\alpha 2\pi i} P^\alpha + c_{\alpha'} e^{\alpha' 2\pi i} P^{\alpha'}.$$

This pair of expressions has nonzero determinant, since by hypothesis, $\alpha - \alpha'$ is not an integer. Hence $P^\alpha, P^{\alpha'}$ can be expressed as a linear combination with constant coefficients of P, P'; and thus as a linear combination of P^β, $P^{\beta'}$; or $P^\gamma, P^{\gamma'}$. Now let

$$P^\alpha = \alpha_\beta P^\beta + \alpha_{\beta'} P^{\beta'} = \alpha_\gamma P^\gamma + \alpha_{\gamma'} P^{\gamma'},$$
$$P^{\alpha'} = \alpha'_\beta P^\beta + \alpha'_{\beta'} P^{\beta'} = \alpha'_\gamma P^\gamma + \alpha'_{\gamma'} P^{\gamma'},$$

and write briefly

$$\begin{Bmatrix} \alpha_\beta & \alpha_{\beta'} \\ \alpha'_\beta & \alpha'_{\beta'} \end{Bmatrix} = (b), \quad \begin{Bmatrix} \alpha_\gamma & \alpha_{\gamma'} \\ \alpha'_\gamma & \alpha'_{\gamma'} \end{Bmatrix} = (c).$$

Denote the inverse substitution of $(b), (c)$ respectively by $(b)^{-1}, (c)^{-1}$. We obtain for the functions $(P^\alpha, P^{\alpha'})$ the substitutions

$$(A) = \begin{Bmatrix} e^{\alpha 2\pi i} & 0 \\ 0 & e^{\alpha' 2\pi i} \end{Bmatrix}, \quad (B) = (b) \begin{Bmatrix} e^{\beta 2\pi i} & 0 \\ 0 & e^{\beta' 2\pi i} \end{Bmatrix} (b)^{-1},$$
$$(C) = (c) \begin{Bmatrix} e^{\gamma 2\pi i} & 0 \\ 0 & e^{\gamma' 2\pi i} \end{Bmatrix} (c)^{-1}.$$

From the equation $(C)(B)(A) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, since the determinant of a composite substitution is the product of the determinants of its components, we have

$$
\begin{aligned}
1 &= \mathrm{Det}(A)\mathrm{Det}(B)\mathrm{Det}(C)\\
&= e^{(\alpha+\alpha'+\beta+\beta'+\gamma+\gamma')2\pi i}\mathrm{Det}(b)\mathrm{Det}(b)^{-1}\mathrm{Det}(c)\mathrm{Det}(c)^{-1}.
\end{aligned}
$$

Since $\mathrm{Det}(b)\mathrm{Det}(b)^{-1} = 1$, $\mathrm{Det}(c)\mathrm{Det}(c)^{-1} = 1$, we have

$$
\alpha + \alpha' + \beta + \beta' + \gamma + \gamma' = \text{integer}, \tag{2}
$$

which agrees with the above hypothesis that this sum is 1.

The three remaining relations contained in the equation

$$
(C)(B)(A) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
$$

give three conditions for (b), (c). This can be seen more easily in the following way.

If x passes around 0 and then around ∞ in a negative sense, the combined path yields a positive circuit around 1. The resulting value of P^α is therefore

$$
\begin{aligned}
&\alpha_\gamma e^{\gamma 2\pi i} P^\gamma + \alpha_{\gamma'} e^{\gamma' 2\pi i} P^{\gamma'}\\
&\qquad = \left(\alpha_\beta e^{-\beta 2\pi i} P^\beta + \alpha_{\beta'} e^{-\beta' 2\pi i} P^{\beta'}\right) e^{-\alpha 2\pi i}.
\end{aligned}
$$

Multiply this equation by an arbitrary factor $e^{-\sigma\pi i}$, and the equation

$$
\alpha_\gamma P^\gamma + \alpha_{\gamma'} P^{\gamma'} = \alpha_\beta P^\beta + \alpha_{\beta'} P^{\beta'}
$$

by $e^{\sigma\pi i}$, and subtract. Canceling a common factor, we obtain

$$
\begin{aligned}
&\alpha_\gamma \sin(\sigma - \gamma)\pi e^{\gamma\pi i} P^\gamma + \alpha_{\gamma'} \sin(\sigma - \gamma')\pi e^{\gamma'\pi i} P^{\gamma'}\\
&\quad = \alpha_\beta \sin(\sigma + \alpha + \beta)\pi e^{-(\alpha+\beta)\pi i} P^\beta + \alpha_{\beta'} \sin(\sigma + \alpha + \beta')\pi e^{-(\alpha+\beta')\pi i} P^{\beta'}.
\end{aligned}
$$

Analogously, replacing α by α', we have

$$
\begin{aligned}
&\alpha'_\gamma \sin(\sigma - \gamma)\pi e^{\gamma\pi i} P^\gamma + \alpha'_{\gamma'} \sin(\sigma - \gamma')\pi e^{\gamma'\pi i} P^{\gamma'}\\
&\quad = \alpha'_\beta \sin(\sigma + \alpha' + \beta)\pi e^{-(\alpha'+\beta)\pi i} P^\beta + \alpha'_{\beta'} \sin(\sigma + \alpha' + \beta')\pi e^{-(\alpha'+\beta')\pi i} P^{\beta'},
\end{aligned}
$$

for arbitrary σ. Eliminating one of the functions, for example $P^{\gamma'}$, from both equations by a suitable choice of σ, the resulting equations differ only by a constant factor, since $P^\beta/P^{\beta'}$ is not constant. Hence this elimination of $P^{\gamma'}$ gives

$$(3)\qquad \frac{\alpha_\gamma}{\alpha'_\gamma}=\frac{\alpha_\beta\sin(\alpha+\beta+\gamma')\pi e^{-\alpha\pi i}}{\alpha'_\beta\sin(\alpha'+\beta+\gamma')\pi e^{-\alpha'\pi i}}=\frac{\alpha_{\beta'}\sin(\alpha+\beta'+\gamma')\pi e^{-\alpha\pi i}}{\alpha'_{\beta'}\sin(\alpha'+\beta'+\gamma')\pi e^{-\alpha'\pi i}}.$$

The analogous elimination of P^γ gives

$$(4)\qquad \frac{\alpha_{\gamma'}}{\alpha'_{\gamma'}}=\frac{\alpha_\beta\sin(\alpha+\beta+\gamma)\pi e^{-\alpha\pi i}}{\alpha'_\beta\sin(\alpha'+\beta+\gamma)\pi e^{-\alpha'\pi i}}=\frac{\alpha_{\beta'}\sin(\alpha+\beta'+\gamma)\pi e^{-\alpha\pi i}}{\alpha'_{\beta'}\sin(\alpha'+\beta'+\gamma)\pi e^{-\alpha'\pi i}}.$$

These are the four relations sought. From these we obtain the ratios of the quotients $\frac{\alpha_\beta}{\alpha'_\beta}, \frac{\alpha_{\beta'}}{\alpha'_{\beta'}}, \frac{\alpha_\gamma}{\alpha'_\gamma}, \frac{\alpha_{\gamma'}}{\alpha'_{\gamma'}}$. The equality of the two values of $\frac{\alpha_\beta}{\alpha'_\beta} : \frac{\alpha_{\beta'}}{\alpha'_{\beta'}}$ obtained from the second and fourth relations is readily seen to be a consequence of $\alpha+\alpha'+\beta+\beta'+\gamma+\gamma'=1$, with the help of the identity $\sin s\pi=\sin(1-s)\pi$.

Thus each of the numbers $\frac{\alpha_\beta}{\alpha'_\beta}, \frac{\alpha_{\beta'}}{\alpha'_{\beta'}}, \frac{\alpha_\gamma}{\alpha'_\gamma}, \frac{\alpha_{\gamma'}}{\alpha'_{\gamma'}}$ is determined by any one of them, for example $\frac{\alpha_\beta}{\alpha'_\beta}$. Now the three numbers $\alpha'_{\beta'}$, α'_γ, $\alpha'_{\gamma'}$ are determined by the five numbers $\alpha_\beta, \alpha'_\beta, \alpha_{\beta'}, \alpha_\gamma, \alpha_{\gamma'}$. However, these five numbers depend, when P is given, on the factors that are still arbitrary in $P^\alpha, P^{\alpha'}, P^\beta, P^{\beta'}, P^\gamma, P^{\gamma'}$; or rather, on their ratios. By an appropriate choice of these factors, these five numbers can take any finite values.

4.

The remark just made opens the way to the theorem that, in two P-functions with the same exponents, the corresponding components differ only by a constant factor.

In fact, if P_1 is a function with the same exponents as P, the five numbers $\alpha_\beta, \alpha_{\beta'}, \alpha_\gamma, \alpha_{\gamma'}, \alpha'_\beta$ can be given the same values for both functions. Now $\alpha'_\beta, \alpha'_\gamma, \alpha'_{\gamma'}$ are also the same for both functions. Hence one has simultaneously

$$(P^\alpha, P^{\alpha'}) = (b)(P^\beta, P^{\beta'}) = (c)(P^\gamma, P^{\gamma'}),$$
$$(P_1^\alpha, P_1^{\alpha'}) = (b)(P_1^\beta, P_1^{\beta'}) = (c)(P_1^\gamma, P_1^{\gamma'}).$$

Consequently,

$$\begin{aligned}P^\alpha P_1^{\alpha'} - P^{\alpha'}P_1^\alpha &= \mathrm{Det}(b)(P^\beta P_1^{\beta'} - P^{\beta'}P_1^\beta)\\ &= \mathrm{Det}(c)(P^\gamma P_1^{\gamma'} - P^{\gamma'}P_1^\gamma).\end{aligned}$$

Of these three expressions the first, after multiplication by $x^{-\alpha-\alpha'}$, obviously remains single-valued and finite at $x=0$. So does the second, multiplied by $x^{\beta+\beta'}=x^{-\alpha-\alpha'-\gamma-\gamma'+1}$, at $x=\infty$; and the third, multiplied by $(1-x)^{-\gamma-\gamma'}$, at $x=1$. This holds for all three expressions when x has a value other than $0,1,\infty$. Consequently, the function

$$(P^{\alpha}P_1^{\alpha'}-P^{\alpha'}P_1^{\alpha})x^{-\alpha-\alpha'}(1-x)^{-\gamma-\gamma'}$$

is everywhere continuous and single-valued, and is thus a constant. Moreover, this function is 0 for $x=\infty$, and must therefore be 0 everywhere.

It follows that

$$\begin{aligned}
&\frac{P_1^{\alpha'}}{P^{\alpha'}}=\frac{P_1^{\alpha}}{P^{\alpha}},\\
&\frac{P_1^{\beta}}{P^{\beta}}=\frac{P_1^{\beta'}}{P^{\beta'}}=\frac{\alpha_\beta P_1^{\beta}+\alpha_{\beta'}P_1^{\beta'}}{\alpha_\beta P^{\beta}+\alpha_{\beta'}P^{\beta'}}=\frac{P_1^{\alpha}}{P^{\alpha}},\\
&\frac{P_1^{\gamma}}{P^{\gamma}}=\frac{P_1^{\gamma'}}{P^{\gamma'}}=\frac{\alpha_\gamma P_1^{\gamma}+\alpha_{\gamma'}P_1^{\gamma'}}{\alpha_\gamma P^{\gamma}+\alpha_{\gamma'}P^{\gamma'}}=\frac{P_1^{\alpha}}{P^{\alpha}}.
\end{aligned}$$

The function P_1^{α}/P^{α} is accordingly single-valued. Moreover, the function must be finite everywhere–hence, as we wish to prove, a constant – if we can show that P^{α} and $P^{\alpha'}$ cannot vanish together for a value of x distinct from 0, 1 and ∞.

To this end, we observe that

$$\begin{aligned}
P^{\alpha}\frac{dP^{\alpha'}}{dx}-P^{\alpha'}\frac{dP^{\alpha}}{dx}&=\operatorname{Det}(b)\left(P^{\beta}\frac{dP^{\beta'}}{dx}-P^{\beta'}\frac{dP^{\beta}}{dx}\right)\\
&=\operatorname{Det}(c)\left(P^{\gamma}\frac{dP^{\gamma'}}{dx}-P^{\gamma'}\frac{dP^{\gamma}}{dx}\right).
\end{aligned}$$

Consequently this function becomes infinitely small at $0,\infty,1$ of orders $\alpha+\alpha'-1$, $\beta+\beta'+1=2-\alpha-\alpha'-\gamma-\gamma'$, $\gamma+\gamma'-1$ respectively. Elsewhere the function is finite and single valued. Thus

$$\left(P^{\alpha}\frac{dP^{\alpha'}}{dx}-P^{\alpha'}\frac{dP^{\alpha}}{dx}\right)x^{-\alpha-\alpha'+1}(1-x)^{-\gamma-\gamma'+1}$$

is an everywhere finite and single-valued function, and is thus a constant. This constant is necessarily nonzero, for otherwise $\log P^{\alpha}-\log P^{\alpha'}$ is constant and $\alpha = \alpha'$, contrary to hypothesis. However, the constant would obviously be 0 if P^α and $P^{\alpha'}$ vanished simultaneously for an x distinct from $0, 1, \infty$, since $\frac{dP^\alpha}{dx}$, $\frac{dP^{\alpha'}}{dx}$, as derivatives of single-valued continuous functions, cannot become infinite.

Thus P^α, $P^{\alpha'}$ cannot vanish simultaneously for a value of x distinct from $0, 1, \infty$. Now the single-valued functions

$$\frac{P_1^\alpha}{P^\alpha} = \frac{P_1^{\alpha'}}{P^{\alpha'}} = \frac{P_1^\beta}{P^\beta} = \frac{P_1^{\beta'}}{P^{\beta'}} = \frac{P_1^\gamma}{P^\gamma} = \frac{P_1^{\gamma'}}{P^{\gamma'}}$$

are finite everywhere, and indeed constant, as was to be proved.

Suppose that two branches of a given P-function are not in constant ratio. From the above theorem, it follows that every P-function with the same exponents is a linear combination with constant coefficients of these branches. By the properties assumed in Section 1, the latter P-function is completely determined up to two constants linearly contained in it. These constants can easily be found from the values of the function for special values of the variable, most conveniently by taking the variable equal to one of the branch points.

The question of whether a function exists, satisfying these conditions, admittedly remains unanswered. Since this will be resolved later through explicit representations of the function via definite integrals and hypergeometric series, no special investigation is called for.

5.

Besides the possible transformations for all values of the exponents in Section 2, the following transformations follow readily from the definition:

$$(A) \qquad P\begin{Bmatrix} 0 & \infty & 1 & \\ 0 & \beta & \gamma & x \\ \frac{1}{2} & \beta' & \gamma' & \end{Bmatrix} = P\begin{Bmatrix} -1 & \infty & 1 & \\ \gamma & 2\beta & \gamma & \sqrt{x} \\ \gamma' & 2\beta' & \gamma' & \end{Bmatrix},$$

where by the foregoing we must have $\beta + \beta' + \gamma + \gamma' = \frac{1}{2}$;

$$(B) \qquad P\begin{Bmatrix} 0 & \infty & 1 & \\ 0 & 0 & \gamma & x \\ \frac{1}{3} & \frac{1}{3} & \gamma' & \end{Bmatrix} = P\begin{Bmatrix} 1 & \rho & \rho^2 & \\ \gamma & \gamma & \gamma & \sqrt[3]{x} \\ \gamma' & \gamma' & \gamma' & \end{Bmatrix},$$

where $\gamma + \gamma' = \frac{1}{3}$ and ρ is an imaginary cube root of unity.

To get a convenient overview of all the functions that reduce to one another with the help of these transformations, it is useful to introduce, instead of the exponents, their differences. As proposed earlier, we denote by $P(\alpha - \alpha', \beta - \beta', \gamma - \gamma', x)$ all the functions of form

$$x^{\delta}(1-x)^{\epsilon}\, P\begin{pmatrix} \alpha & \beta & \gamma & \\ \alpha' & \beta' & \gamma' & x \end{pmatrix}.$$

Here $\alpha - \alpha', \beta - \beta', \gamma - \gamma'$ may be called the first, second and third exponent differences.

From the formulae in Section 2, it follows that in the function

$$P(\lambda, \mu, \nu, x)$$

the quantities λ, μ, ν may be arbitrarily changed in sign and permuted. The variable then takes one of the six values $x, 1-x, \frac{1}{x}, 1-\frac{1}{x}, \frac{1}{1-x}, \frac{x}{x-1}$. Of the 48 P-functions obtained in this way, any group of 8 obtained by changing signs of λ, μ, ν has the same variable.

Of the transformations A and B given in this section, the first is applicable if one of the exponent differences is $\frac{1}{2}$, or two of them are equal. The second is applicable when two exponent differences are $\frac{1}{3}$, or all three of them are equal. By successive application of these transformations one obtains, each expressible in terms of the other, the functions:

I. $P\left(\mu, \nu, \frac{1}{2}, x_2\right)$, $P(\mu, 2\nu, \mu, x_1)$ and $P(\nu, 2\mu, \nu, x_3)$, where

$$\sqrt{1-x_2} = 1 - 2x_1, \quad \sqrt{1-\frac{1}{x_2}} = 1 - 2x_3,$$

so that

$$x_2 = 4x_1(1-x_1) = \frac{1}{4x_3(1-x_3)}.$$

II. $P(\nu, \nu, \nu, x_3), P\left(\nu, \frac{\nu}{2}, \frac{1}{2}, x_2\right), P\left(\frac{\nu}{2}, 2\nu, \frac{\nu}{2}, x_1\right),$

$$P\left(\frac{1}{3}, \nu, \frac{1}{3}, x_4\right), P\left(\frac{1}{3}, \frac{\nu}{2}, \frac{1}{2}, x_5\right), P\left(\frac{\nu}{2}, \frac{2}{3}, \frac{\nu}{2}, x_6\right),$$

where

$$1-\frac{1}{x_4}=\left(\frac{x_3+\rho}{x_3+\rho^2}\right)^3,$$

hence

$$\frac{1}{x_4}=\frac{3(\rho-\rho^2)x_3(1-x_3)}{(\rho^2+x_3)^3},$$
$$x_4(1-x_4)=\frac{(\rho+x_3)^3(\rho^2+x_3)^3}{27x_3^2(1-x_3)^2}=\frac{(1-x_3(1-x_3))^3}{27x_3^2(1-x_3)^2}.$$

Further, by I,

$$4x_4(1-x_4)=x_5=\frac{1}{4x_6(1-x_6)},$$
$$4x_3(1-x_3)=x_2=\frac{1}{4x_1(1-x_1)}.$$

III. $P\left(\nu,\nu,\frac{1}{2},x_2\right), P(\nu,2\nu,\nu,x_1),$

$$P\left(\frac{1}{4},\nu,\frac{1}{2},x_3\right), P\left(\frac{1}{4},2\nu,\frac{1}{4},x_4\right)$$

where

$$x_3=\frac{1}{4}\left(2-x_2-\frac{1}{x_2}\right)=4x_4(1-x_4),$$
$$x_2=4x_1(1-x_1).$$

All these functions can be further transformed by the general transformations, and in this way their exponent-differences can be permuted arbitrarily and given arbitrary signs.

Apart from the two transcendental functions II and III, if one exponent-difference remains arbitrary, only the function $P\left(\nu,\frac{1}{2},\frac{1}{2}\right)=P(\nu,1,\nu)$ admits a further repetition of transformations A and B. However, this leads to entirely elementary formulae, since

$$P\begin{pmatrix} 0 & 0 & 0 & \\ \nu & -\nu & 1 & x \end{pmatrix} = \text{const. } x^\nu + \text{const.}'$$

In fact, transformation B only applies to $P(\nu, \nu, \nu)$ or $P\left(\frac{1}{3}, \nu, \frac{1}{3}\right)$; that is, only to the transcendental function II. However, transformation A can be repeated more often than in I, when one of the numbers $\mu, \nu, 2\mu, 2\nu$ is $\frac{1}{2}$, or one of the equations $\mu = \nu$, $\mu = 2\nu$, $\nu = 2\mu$ holds. Of these hypotheses, $\mu = 2\nu$ or $\nu = 2\mu$ leads to transcendental functions II; $\mu = \nu$, or $2\mu = \frac{1}{2}$, or $2\nu = \frac{1}{2}$, leads to transcendental functions III. Finally $\mu = \frac{1}{2}$ or $\nu = \frac{1}{2}$ leads to the function $P\left(\nu, \frac{1}{2}, \frac{1}{2}\right)$.

To obtain the number of different expressions given by these transformations for each of the transcendental functions I–III, recall that in the P-functions above, we can admit as variables all roots of their defining equations. Each root belongs to a system of 6 values, which can be introduced as variables in place of one another by means of the general transformation.

However, in Case I, the two values of x_1 and x_3, corresponding to a given x_2, belong to the same set of 6 values. Hence each function I can be represented by P-functions via $6.3 = 18$ different variables.

In Case II, among the values of the variables corresponding to a given value of x_5, the two values of x_6 and x_4, the 6 values of x_3 and the 6 values of x_1, combined in pairs, lead to the same set of 6 values. The three values of x_2 lead to three different systems of 6 values. So x_1 and x_2 each yield three systems, and x_3, x_4, x_5, x_6 each yield one system, of six values. Altogether there are $6.10 = 60$ values through which each function II can be expressed via P-functions.

Finally, in Case III, x_3, the two values of x_2, the two values of x_4, and each pair of the four values of x_1, give a system of 6 values. Hence each of the functions III can be represented via P-functions of $6.5 = 30$ different variables.

In each P function, we can assign arbitrary signs to the exponent differences, without changing the variable, via the general transformation. Since no exponent difference is 0, a given function occurs in 8 distinct ways as a P-function of the same variable. The total number of these expressions is $8.6.3 = 144$ in Case I, $8.6.10 = 480$ in Case II, and $8.6.5 = 240$ in Case III.

6.

If we change all the exponents of a P-function by integer amounts, the quantities

$$\frac{\sin(\alpha+\beta+\gamma')\pi e^{-\alpha\pi i}}{\sin(\alpha'+\beta+\gamma')\pi e^{-\alpha'\pi i}}, \frac{\sin(\alpha+\beta'+\gamma')\pi e^{-\alpha\pi i}}{\sin(\alpha'+\beta'+\gamma')\pi e^{-\alpha'\pi i}},$$

$$\frac{\sin(\alpha+\beta+\gamma)\pi e^{-\alpha\pi i}}{\sin(\alpha'+\beta+\gamma)\pi e^{-\alpha'\pi i}}, \frac{\sin(\alpha+\beta'+\gamma)\pi e^{-\alpha\pi i}}{\sin(\alpha'+\beta'+\gamma)\pi e^{-\alpha'\pi i}}$$

in the equations (3), Section 3, are unchanged.

If in the functions

$$P\begin{pmatrix} \alpha & \beta & \gamma & \\ \alpha' & \beta' & \gamma' & x \end{pmatrix}, \; P_1\begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 & \\ \alpha_1' & \beta_1' & \gamma_1' & x \end{pmatrix},$$

the corresponding exponents α_1 and α, and so on, differ by integers, one can take the eight numbers $(\alpha_\beta)_1, (\alpha'_\beta)_1, (\alpha_{\beta'})_1, \ldots$ equal to the eight numbers $\alpha_\beta, \alpha'_\beta, \alpha_{\beta'}, \ldots$. For, from the equality of any five, the equality of the three others follows.

By the reasoning used in Section 4, we deduce that

$$\begin{aligned} P^\alpha P_1^{\alpha_1'} - P^{\alpha'} P_1^{\alpha_1} &= \mathrm{Det}(b)(P^\beta P_1^{\beta_1'} - P^{\beta'} P_1^{\beta_1}) \\ &= \mathrm{Det}(c)(P^\gamma P_1^{\gamma_1'} - P^{\gamma'} P_1^{\gamma_1}). \end{aligned}$$

Among the quantities $\alpha + \alpha_1'$ and $\alpha_1 + \alpha'$, $\beta + \beta_1'$ and $\beta_1 + \beta'$, $\gamma + \gamma_1'$ and $\gamma_1 + \gamma'$, denote by $\bar\alpha, \bar\beta, \bar\gamma$ the members of the pairs, which are smaller by a positive integer than the others. Then

$$(P^\alpha P_1^{\alpha_1'} - P^{\alpha'} P_1^{\alpha_1}) x^{-\bar\alpha}(1-x)^{-\bar\gamma}$$

is a function of x that is single valued and finite at $x = 0$, $x = 1$ and all other finite values of x. However, this function becomes infinite at $x = \infty$ of order $-\bar\alpha - \bar\gamma - \bar\beta$, so that it is a polynomial F of degree $-\bar\alpha - \bar\beta - \bar\gamma$.

As before, denote the exponent differences $\alpha - \alpha'$, $\beta - \beta'$, $\gamma - \gamma'$ by λ, μ, ν. In regard to these, we see firstly that their sum varies by an even integer, if all exponents vary by an integer. For the sum exceeds the sum of the exponents, which remains equal to 1, by

$$-2(\alpha' + \beta' + \gamma'),$$

a quantity that changes by an even integer. However, the exponent differences can vary by any integers whose sum is even. Denote now $\alpha_1 - \alpha'_1$, $\beta_1 - \beta'_1$, $\gamma_1 - \gamma'_1$ by λ_1, μ_1, ν_1 and write $\Delta\lambda, \Delta\mu, \Delta\nu$ for the absolute values of the differences $\lambda - \lambda_1$, $\mu - \mu_1$, $\nu - \nu_1$. In the pair of numbers $\alpha + \alpha'_1$ and $\alpha' + \alpha_1$, the one that is the positive amount $\Delta\lambda$ smaller than the other is equal to

$$\frac{\alpha + \alpha'_1 + \alpha' + \alpha_1}{2} - \frac{\Delta\lambda}{2}.$$

Hence

$$-\bar{\alpha} = \frac{\Delta\lambda}{2} - \frac{\alpha + \alpha'_1 + \alpha' + \alpha_1}{2},$$

and likewise

$$-\bar{\beta} = \frac{\Delta\mu}{2} - \frac{\beta + \beta'_1 + \beta' + \beta_1}{2},$$
$$-\bar{\gamma} = \frac{\Delta\nu}{2} - \frac{\gamma + \gamma'_1 + \gamma' + \gamma_1}{2}.$$

The degree of the polynomial F, which is equal to the sum of these numbers, is thus

$$\frac{\Delta\lambda + \Delta\mu + \Delta\nu}{2} - 1.$$

7.

Now let

$$P\begin{pmatrix} \alpha & \beta & \gamma & \\ \alpha' & \beta' & \gamma' & x \end{pmatrix}, \quad P_1\begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 & \\ \alpha'_1 & \beta'_1 & \gamma'_1 & x \end{pmatrix}, \quad P_1\begin{pmatrix} \alpha_2 & \beta_2 & \gamma_2 & \\ \alpha'_2 & \beta'_2 & \gamma'_2 & x \end{pmatrix}$$

be three functions in which the corresponding exponents differ by integers. From the above theorem, via the identity

$$P^{\alpha}(P_1^{\alpha_1} P_2^{\alpha'_2} - P_1^{\alpha'_1} P_2^{\alpha_2}) + P_1^{\alpha_1}(P_2^{\alpha_2} P^{\alpha'} - P_2^{\alpha'_2} P^{\alpha})$$
$$+ P_2^{\alpha_2}(P^{\alpha} P_1^{\alpha'_1} - P^{\alpha'} P_1^{\alpha_1}) = 0,$$

we obtain the important result that between their corresponding terms, a linear homogeneous equation holds whose coefficients are polynomials in x. Thus

"For a family of P-functions whose corresponding exponents differ by integers, any one may be expressed as a linear combination of any given pair, with rational functions of x as coefficients."

A particular consequence of the method of proof of this result is that the second derivative of a P-function can be expressed as a linear combination with rational coefficients of the function and its first derivative. Thus the function satisfies a second order linear homogeneous differential equation.

To simplify the derivation, we restrict ourselves to the case $\gamma = 0$: by Section 2, the general case easily reduced to this case. We set $P = y$, $P^\alpha = y'$, $P^{\alpha'} = y''$. Then the functions

$$y' \frac{dy''}{d\log x} - y'' \frac{dy'}{d\log x},$$

$$\frac{d^2y'}{d(\log x)^2}\, y'' - \frac{d^2y''}{d(\log x)^2}\, y',$$

$$\frac{dy'}{d\log x}\, \frac{d^2y''}{d(\log x)^2} - \frac{dy''}{d\log x}\, \frac{d^2y'}{d(\log x)^2},$$

multiplied by $x^{-\alpha-\alpha'}(1-x)^{-\gamma'+2}$, are finite and single valued for finite x, and infinite of order 1 for $x = \infty$. Moreover, the first product vanishes of order 1 at $x = 1$. Hence for

$$y = \text{const.}'y' + \text{const.}''y''$$

we have an equation of form

$$(1-x)\frac{d^2y}{d(\log x)^2} - (A + Bx)\frac{dy}{d\log x} + (A' - B'x)y = 0.$$

Here A, B, A', B' are constants yet to be determined.

By the method of undetermined coefficients, one can expand the solution of this differential equation, up to $x = 1$, in increasing or decreasing powers, as a series

$$\sum a_n x^n.$$

The exponent μ of the first term in the first case, where it is the lowest exponent, is determined by the equation

$$\mu^2 - A\mu + A' = 0.$$

In the second case, where it is the highest exponent, μ is determined by the equation

$$\mu^2 + B\mu + B' = 0.$$

The roots of the first equation must be α and α', and those of the second equation $-\beta$ and $-\beta'$. Consequently

$$A = \alpha + \alpha', A' = \alpha\alpha'$$
$$B = \beta + \beta', B' = \beta\beta',$$

and the function

$$P\begin{pmatrix} \alpha & \beta & 0 & \\ \alpha' & \beta' & \gamma' & x \end{pmatrix} = y$$

satisfies the differential equation

$$(1-x)\frac{d^2y}{d(\log x)^2} - (\alpha + \alpha' + (\beta + \beta')x)\frac{dy}{d\log x} + (\alpha\alpha' - \beta\beta' x)y = 0.$$

Now the coefficients can successively be determined via the recursion formula

$$\frac{a_{n+1}}{a_n} = \frac{(n+\beta)(n+\beta')}{(n+1-\alpha)(n+1-\alpha')}$$

which is satisfied by

$$a_n = \frac{\text{const.}}{\Pi(n-\alpha)\Pi(n-\alpha')\Pi(-n-\beta)\Pi(-n-\beta')}.$$

Thus the series

$$y = \text{const.} \sum \frac{x^n}{\Pi(n-\alpha)\Pi(n-\alpha')\Pi(-n-\beta)\Pi(-n-\beta')}$$

where the exponents commence with α or α' and increase by unity; or when the exponents commence with $-\beta$ or $-\beta'$ and decrease by unity, are solutions of the differential equation. Indeed they are the particular solutions denoted above by P^α, $P^{\alpha'}$, P^β, $P^{\beta'}$ respectively.

Gauss denoted by $F(a, b, c, x)$ a series in which the quotient of term number $n+2$ by the preceding term is $\frac{(n+a)(n+b)}{(n+1)(n+c)}x$, and the first term is 1. Following Gauss, our result in the simplest case $\alpha = 0$ can be expressed as

$$P^\alpha\begin{pmatrix} 0 & \beta & 0 & \\ \alpha' & \beta' & \gamma' & x \end{pmatrix} = \text{const.}\, F(\beta, \beta', 1-\alpha', x)$$

or

$$F(a,b,c,x) = P^{\alpha}\begin{pmatrix} 0 & a & 0 & \\ 1-c & b & c-a-b & x \end{pmatrix}.$$

From this result we easily obtain an expression for the P-function as a definite integral, on replacing the Π-functions by an Euler integral of second type in the general term of the series. We then interchange the order of summation and integration. In this way we find that the integral

$$x^{\alpha}(1-x)^{\gamma}\int s^{-\alpha'-\beta'-\gamma'}(1-s)^{-\alpha'-\beta-\gamma}(1-xs)^{-\alpha-\beta'-\gamma}ds,$$

taken from one of the four values $0, 1, \frac{1}{x}, \infty$ to another of these values, along an arbitrary path, gives a function

$$P\begin{pmatrix} \alpha & \beta & \gamma & \\ \alpha' & \beta' & \gamma' & x \end{pmatrix}.$$

Moreover, by a suitable choice of the limits of integration and the path between them, the integral represents each of the six functions $P^{\alpha}, P^{\beta}, \dots, P^{\gamma'}$. However, one can show directly that this integral has the characteristic properties of such a function. This will be seen in the sequel, where this expression for the P-function as a definite integral will be used to determine the factors that still remain arbitrary in $P^{\alpha}, P^{\alpha'}, \dots$. Here I remark only that to make this expression applicable in general, we need to modify the path of integration when the integrand becomes infinite for one of the values $0, 1, \frac{1}{x}, \infty$, in a manner that precludes integration up to this value.

8.

Recall the equation obtained in Sections 2 and 7,

$$P^{\alpha}\begin{pmatrix} \alpha & \beta & \gamma & \\ \alpha' & \beta' & \gamma' & x \end{pmatrix} = x^{\alpha}(1-x)^{\gamma}P^{\alpha}\begin{pmatrix} 0 & \beta+\alpha+\gamma & 0 & \\ \alpha'-\alpha & \beta'+\alpha+\gamma & \gamma'-\gamma & x \end{pmatrix}$$
$$= \text{Const.}\, x^{\alpha}(1-x)^{\gamma}F(\beta+\alpha+\gamma, \beta'+\alpha+\gamma, \alpha-\alpha'+1, x).$$

We see from this that any function represented via the P-function likewise has an expansion as a hypergeometric series in ascending powers of the variable. From Section 5, there are 8 representations of a function via P-functions with the same variable, obtained by permuting its exponents. Thus, for example, there are 8 representations with the variable x. Among these, however, any two obtained by interchanging the second pair, β and β', yield the same expansion. Hence, we obtain four expansions in increasing powers of x. Two of these, obtained from one another by interchanging γ, γ', represent the function P^{α}; the other two represent $P^{\alpha'}$. These four expansions converge as long as x has modulus < 1, and diverge when the modulus exceeds 1. The four series in decreasing powers of x, representing P^{β} and $P^{\beta'}$, behave in the opposite way.

In the case when the modulus of x is 1, it follows from the Fourier series, that these series cease to converge at $x = 1$, if the function becomes infinite at $x = 1$ or order higher than 1. On the other hand, the series remains convergent at $x = 1$ when the function becomes infinite there of lower order than 1, or remains finite. In this case, then, only half of the 8 expansions in powers of x converge if the real part of $\gamma' - \gamma$ does not lie between -1 and 1. In the contrary case, all these expansions converge.

Thus in general one has 24 different hypergeometric series that represent a P-function, that contain ascending or descending powers of three different variables. For a given value of x half of these, that is 12, converge. In Case I of Section 5, all these numbers are to be multiplied by 3; in Case II, by 10; in Case III, by 5. For numerical calculation, the most convenient choice from these series is usually the one for which the fourth element has least modulus.

Regarding the expressions of a P-function via definite integrals that can be derived from the integrals of Section 7 using the transformation of Section 5, these expressions are all distinct. Hence one obtains in general 48, in Case I, 144, in Case II, 480 and in Case III, 240 definite integrals which represent the same term of a P-function, and therefore have ratios that are independent of x. Among these integrals, groups of 24, which are obtained from one another by an even number of interchanges of exponents, can also be transformed into each other as follows. We employ a substitution of degree 1, chosen so that for three given values among $0, 1, \infty, \frac{1}{x}$ of the integration variable s, the new variable takes the values $0, 1, \infty$. The remaining equations, as far as I have investigated them, require transformations of multiple integrals, if they are to be established by the methods of integral calculus.

V.

Author's announcement of the preceding paper.

(*Göttinger Nachrichten*, 1857, no. 1)

On 6th November, 1856 a mathematical memoir entitled "Contributions to the theory of functions represented by the Gauss series $F(\alpha, \beta, \gamma, x)$" was submitted to the Royal Society, by its assessor, Dr. Riemann.

This memoir deals with a class of functions which are used for solving many of the problems of mathematical physics. The series formed from them perform the same roles in the more difficult problems as are served in the easier ones by the trigonometrical series, now so frequently employed, which proceed in terms of sines and cosines of multiples of a variable.

These applications, particularly in astronomy, appear to have led Gauss—following Euler who had already frequently concerned himself with these functions because of their theoretical interest—to undertake his researches into the series which he denoted by $F(\alpha, \beta, \gamma, x)$. A part of these researches was published in the form of a memoir in the 1812 Journal of the Royal Society.

The series is one in which the quotient of term number $n+2$ by the preceding term is equal to

$$\frac{(n+\alpha)(n+\beta)x}{(n+1)(n+\gamma)}$$

and the first term is 1. The name *hypergeometric series* now usual was earlier proposed by Johann Friedrich Pfaff for a more general type of series in which the above quotient is a rational function of the index; whereas Euler, following Wallis, understood by this name a series in which this quotient is a polynomial of degree 1 in the index.

The unpublished part of Gauss's researches on this series, which were found among his posthumous papers, had meanwhile already been supplemented in 1835 by the work of Kummer in Vol. 15 of Crelle's Journal. This work relates to expressions of the series by similar ones in which the variable x is replaced by an algebraic function thereof. A special case of such a transformation had already been discovered by Euler and had been handled in his treatise on integral calculus as well as several of his papers (in the simplest form in the *N. Acta Acad. Petr.*, XII, p. 58). This relation was later proved in different ways by Pfaff (*Disquis. anal. Helmstadii* 1797), Gudermann (*Crelle's J.*, vol.7, p. 306) and Jacobi. Kummer succeeded in devising a procedure based on Euler's method by means of which all the transformations could be found, but the detailed implementation required such lengthy discussion that he refrained from considering transformations of third degree, confining himself to a complete treatment of first and second degree transformations and their compositions.

In the memoir announced here, the author studies these transcendental functions by a method whose principle was described in his inaugural dissertation (Section 20) and which yields all the earlier results almost without calculation. He hopes soon to be able to submit to the Royal Society some further results found by the same methods.

VI.

The Theory of Abelian Functions.

(*Borchardt's Journal für reine und angewandte Mathematatik*, vol. 54, 1857.)

1.

General assumptions, and methods for the study of functions of unbounded variables.

My aim is to present to the readers of the *Journal für Mathematik* the results of some investigations into various transcendental functions, in particular Abelian functions. To avoid repetition, it is worthwhile to summarize, in a separate section, the general assumptions from which my treatment proceeds.

To represent the independent variable, I always use the now well-known geometrical representation of Gauss, in which a complex number $z = x+yi$ is represented by the point on an infinite plane, whose rectangular coordinates are x and y. I denote the complex numbers and the points representing them by the same letters. I regard as a function of $x + yi$ every complex quantity w which satisfies the equation

$$i\frac{\partial w}{\partial x} = \frac{\partial w}{\partial y},$$

without assuming an expression for w in terms of x and y. It follows from this differential equation, by a known theorem, that the quantity w can be expressed as the sum of a series in increasing integer powers of $z - a$, of the form $\sum_{n=0}^{\infty} a_n(z-a)^n$, as long as w, in the neighborhood of $z = a$, has a single determinate value which varies continuously with z. This representation extends up to a distance, measured by $|z-a|$, at which a point of discontinuity is encountered.

It ensues from the considerations, which constitute the basis of the method of undetermined coefficients, that the coefficients a_n are completely defined once w has been given along a finite segment of a line emanating from a, no matter how small.

By combining these two ideas, one can easily convince oneself of the truth of the following proposition:

A function of $x+yi$ given in one part of the (x, y) plane can be continued beyond the boundaries of this region in a continuous fashion in only one way.

Let us now imagine that the function to be studied is defined, not by any analytical expressions or equations containing z, but rather through its value being given in a limited region of the z-plane and that w is extended continuously while satisfying the partial differential equation:

$$i\frac{\partial w}{\partial x} = \frac{\partial w}{\partial y}.$$

This extension is, thanks to the above-mentioned theorems, completely determinate, on the assumption that the prolongation is carried out along a strip of finite width rather than a mere line (where a partial differential equation could not apply). Depending on the nature of the function concerned, either the function will always assume the same value for a given value of z, irrespective of the path along which the prolongation is effected, or else it will not. In the former case I shall call the function *single-valued*; it is then a function which, for every value of z, has a well-defined value and is never discontinuous along a line. In the latter case the function may be said to be *multi-valued*, and here to understand its behaviour it is necessary above all to focus attention on certain points of the z-plane around which one function changes into another. An example of such a point is the point a for the function $\log(z-a)$. Suppose an arbitrary line be drawn from this point a, then in the neighborhood of a the value of this function can always be chosen so that it varies continuously except on the line itself. On either side of the line, however, the function assumes different values, its value on the negative[1] side being $2\pi i$ greater than its value on the positive side. The prolongation of the function across the line, for example from the negative side, then obviously results in a function different from the initial one, and in this case one whose value is everywhere $2\pi i$ greater.

To simplify the description of these relationships, the different prolongations of a given function in a given region of the z-plane will be called *branches* of the original function and a point around which one branch continues into

[1]Following the example of Gauss, who has proposed the name of positive lateral unit for $+i$, I shall designate a sideways direction as positive, in relation to a given forward direction, if it bears the same relation to it as $+i$ does to 1.

another a *branch-point* of the function. Where no branching occurs, the function is said to be *single-valued* or *monodromic*.

A branch of a function of several independent variables, $z, s, t, \ldots$ is said to be single-valued in the neighborhood of a system of values $z = a, s = b$, $t = c, \ldots$, if for all the combinations of values up to a finite distance from these (or in other words if $|z-a|, |s-b|, |t-c|, \ldots$ are all less than a definite finite number), the branch concerned has a well-defined value which varies continuously with the variables. A branch point or point around which one branch continues into another is defined, in the case of a function of several variables, by a set of particular values of the independent variables.

By one of the theorems quoted above, this property of single-valuedness in a function is equivalent to that of its developability in a series of ascending or descending integral powers of the incremental changes in value of the variables. However, it seems inappropriate to express properties independent of the mode of representation, by criteria based on a particular expression for the function.

In many investigations, notably in the study of algebraic and Abelian functions, it is advantageous to represent the branching of a multi-valued function geometrically in the following way. Over the (x, y) plane, we spread another surface like an infinitely thin membrane covering only that region of the plane in which the function has been defined. When the domain of existence of the function is extended, the surface is likewise extended. In a region of the plane where there are two or more different prolongations of the function, the surface will have two or more layers; it will be composed of superimposed sheets, one sheet for each branch. Around a branch point one sheet of the surface continues into the next, and in the neighborhood of the branch point the surface may be considered as a helicoidal surface whose axis goes through the point perpendicular to the (x, y) plane and whose pitch is infinitely small. If the function resumes its original value after completing a number of turns around the branch point (as for example in the case of $(z-a)^{m/n}$, with m and n relatively prime natural numbers, after n circuits of z around a), then admittedly we must assume that the uppermost sheet of the surface passes into the lowest sheet through the others.

The multi-valued function has, for every point on such a surface that represents its branching, one well-defined value, and can therefore be regarded as a completely determined function of position on this surface.

2.

Theorems of analysis situs for the theory of the integral of a complete differential with two terms.

In the study of functions arising from the integration of total differentials, a few theorems belonging to *analysis situs* are almost indispensable. The name analysis situs was given by Leibnitz to a branch of knowledge—perhaps with not quite the same meaning as here—concerned with that part of the theory of continuous quantities not based on their existence independently of position and on measurement of one quantity against another, but on positional and situational relationships that are independent of relative size.

I reserve for the future a treatment of this subject that avoids measurement entirely. Here I restrict myself to an exposition in a geometrical guise of theorems necessary for the integration of a complete differential with two terms.

Let T be a given surface spread simply or multiply over the (x, y) plane[2], and let X, Y be continuous functions of position on this surface, such that $X dx + Y dy$ is a complete differential. Thus

$$\frac{\partial X}{\partial y} - \frac{\partial Y}{\partial x} = 0.$$

It is well-known that the contour integral

$$\int (X dx + Y dy),$$

is zero when taken in a positive or negative sense around part of T. That is, the integral is taken over the entire boundary in the same direction (positive or negative) in relation to the outward normal (see the footnote on p.). For this integral is equal, in the former case, to the integral

$$\int \left(\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y} \right) dT$$

over this part of T. In the latter case, the sign changes. The integral

$$\int (X dx + Y dy)$$

[2]See above, p. .

therefore has the same value when evaluated over two different paths leading from one fixed point to another, if these two paths together form the complete boundary of a piece of the surface T. Accordingly, if every closed curve interior to T forms the complete boundary of a piece of T, then the integral from a fixed initial point to one and the same end-point always has the same value and represents a function of the position of the endpoint, everywhere continuous in T, independent of the path of integration. This gives rise to a classification of surfaces into simply connected surfaces, in which every closed curve forms the complete boundary of a part of the surface (for example a disc), and multiply connected surfaces for which this is not true (for example the annular surface bounded by two concentric circles). A multiply connected surface can always be cut up into simply connected pieces (see the illustrated examples at the end of this section). As this operation is very useful in the study of integrals of algebraic functions, we shall give a brief summary of the relevant theorems, which hold for arbitrary surfaces in space.

Suppose that two systems of curves a and b in a surface F, taken together, form the complete boundary of a part of this surface. Then any other system of curves which, together with a, forms the complete boundary of a part of F has the property that, together with b, it forms the complete boundary of a part of the surface which is composed of the two former portions of surface joined together along a (their sum or difference, depending on whether they lie along the opposite or the same sides of a). The two systems of curves therefore both play the same role in providing a complete boundary for a part of F, and for this purpose either system can be substituted for the other.

If n closed curves $a_1, a_2, \ldots, a_n$ can be drawn in a surface F which neither separately nor in combination form the complete boundary of this surface, but which taken in combination with any other closed curve constitute the complete boundary of a part of F, then the surface F is said to be $n+1$ times connected.

This character of the surface does not depend on the choice of the system of curves $a_1, a_2, \ldots, a_n$. For any n other closed curves $b_1, b_2, \ldots, b_n$ which are insufficient to form the complete boundary of a part of this surface, when taken together with any other closed curve, will form the complete boundary of a part of F.

Indeed, since b_1 together with the lines a form the complete boundary of a part of F, one of the curves a can be replaced by b_1 and the remaining $n-1$ left unchanged. These new n lines together with another, say b_2, suffice to form the complete boundary of a part of F. We can replace the $n-1$ curves a, and b_1, by b_1, b_2 and $n-2$ of the remaining curves a. When, as is here supposed, the curves b do not form the complete boundary of a part of F, then clearly the process can be repeated until all the curves a have been replaced by the curves b.

An $n+1$ times connected surface F can be converted into an n times connected surface F' by a transverse cut, that is a line starting from a point on the boundary of the surface, ending on another point of the boundary, and lying interior to the surface. The new edges of the pieces of surface which result from the operation of cutting count as a part of the boundary during any further dissection. Accordingly, a transverse cut can never cross a point of the surface more than once, though it may end in one of its earlier points.

As the lines $a_1, a_2, \ldots, a_n$ do not suffice to form the complete boundary of a part of F, it necessarily follows that if we visualize F as being cut up into pieces by these lines, the portion of the surface F lying on the left hand side of a_n, likewise the portion on the right, must have boundary lines different from those of a and hence belonging to the boundary of F. We can therefore draw a line in either of these parts of the surface from a point on the curve a_n to the boundary of F, which does not intersect any of the curves a. These two lines q' and q'' together form a transverse cut q in the surface F which fulfills our objective.

In fact the lines $a_1, a_2, \ldots, a_{n-1}$ are closed curves lying within the interior of F', the surface resulting from the operation of making the transverse cut. Taken together, they are insufficient to form the complete boundary of a part of F, or indeed of F'. Every other closed curve ℓ interior to F', however, forms in conjunction with these $n-1$ lines the complete boundary of a part of F'. For the line ℓ together with the complex consisting of the lines $a_1, a_2, \ldots, a_n$ forms the complete boundary of a part f of F. Now it can be shown that the line a_n cannot be a part of this boundary because, if it were, q' or q'', depending on whether f were on the left or right of a_n, would cross the interior of f to reach a point of the boundary of F, which is a point outside f. Thus it cuts the boundary of f, contrary to the hypothesis that ℓ and the complex a of lines (with the exception of the point of intersection of a_n and q) always remain interior to F'.

Accordingly, the surface F' into which F is converted by the transverse cut q is, as the theorem requires, n times connected.

It will now be proved that the surface F is converted into an n times connected surface F' by any transverse cut p which does not cut F into two separate pieces. If the two portions of surface on either side of the transverse cut are still connected together, then we can draw a line b interior to the surface which starts from a point on the transverse cut and returns to this starting point from the other side of the cut. This line b is a closed curve interior to F which cannot be the complete boundary of either of the two pieces of surface into which it divides F. For the transverse cut leads from b, on either side, to a boundary point. We can therefore replace one of the curves a by the curve b and each of the remaining $n-1$ curves a by a curve interior to F' and, if necessary, by the curve b. We can thus show by the reasoning used earlier that F' is n times connected.

Hence an $n+1$ times connected surface is transformed into an n times connected surface by any transverse cut which does not cut the surface into separate pieces.

The surface obtained by making a transverse cut can be further cut up by making a new transverse cut. By repeating the operation n times, an $n+1$ times connected surface can be transformed, by n successive transverse cuts that do not cut off a piece, into a simply connected surface.

To apply this treatment to a surface which has no boundary, in other words a closed surface, it must first be transformed into one with a boundary by excluding an arbitrary point, so that the first cross-cut is a closed curve which begins and ends in this point. The outer surface of a torus, for example, which is a triply connected surface, can be transformed into a simply connected surface by a closed curve drawn on the surface and a transverse cut.

We now apply this decomposition of multiply connected surfaces into simply connected surfaces to the integration of the complete differentials of the form $Xdx + Ydy$, discussed at the start of the section. Suppose that the surface T, spread over the (x, y)-plane, throughout which X, Y are everywhere continuous functions of position satisfying the equation

$$\frac{\partial X}{\partial y} - \frac{\partial Y}{\partial x} = 0,$$

is n times connected. Then it can be transformed into a simply connected surface T' by making n transverse cuts. The value of the line-integral of $Xdx + Ydy$ calculated from any initial fixed point, along a curve interior to T', then depends only on the position of the end-point and may be regarded as a function of its coordinates. If we substitute x, y for these coordinates, we obtain a function

$$z = \int (X dx + Y dy)$$

of x and y which has a definite value for every point of T', and is everywhere continuous within T', but in general will change in value by a constant across a line leading from one nodal point of the system of cuts to another. The variations across the transverse cuts depend on independent variables equal in number to the number of cuts. For if we run through the system of cuts backwards—the later parts first—the variations are everywhere well-defined once the value at the beginning of each transverse cut is given; the latter values are however independent of each other.

In order to elucidate what is meant by the n times connected surfaces defined above (p. ,), the following illustrations give examples of simply, doubly and triply connected surfaces.

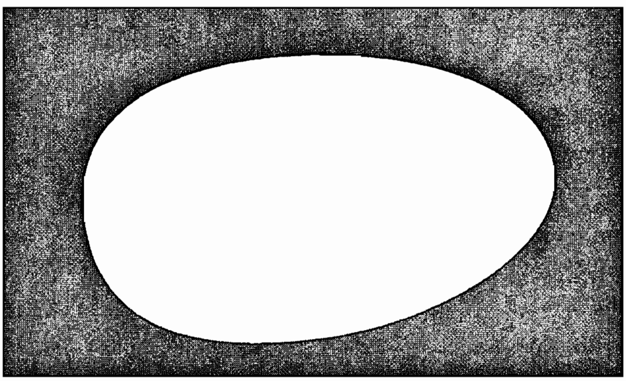


Figure 1: a simply connected surface

The surface in Figure 1 is separated into two pieces by any transverse cut, and any closed curve in the surface forms the complete boundary of a part of the surface.

The surface in Figure 2 is transformed into a simply connected surface by any transverse cut q which does not cut off a piece. Any closed curve, with the curve a adjoined, forms the complete boundary of a part of the surface.

In Figure 3, every closed curve, with the adjunction of the curves a_1 and a_2, forms the complete boundary of a part of this surface. Every transverse cut which does not cut off a piece of the surface transforms the surface into a doubly connected surface, and two such transverse cuts q_1 and q_2 transform it into a simply connected surface.

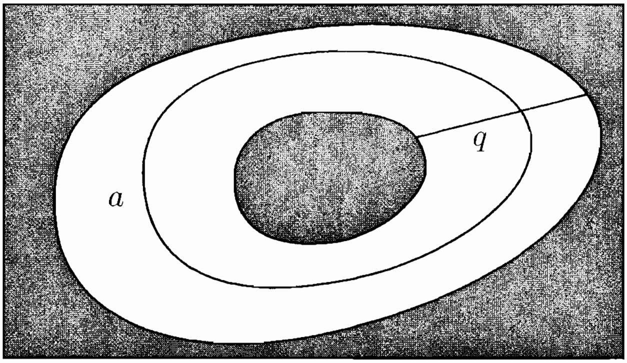


Figure 2: a doubly connected surface

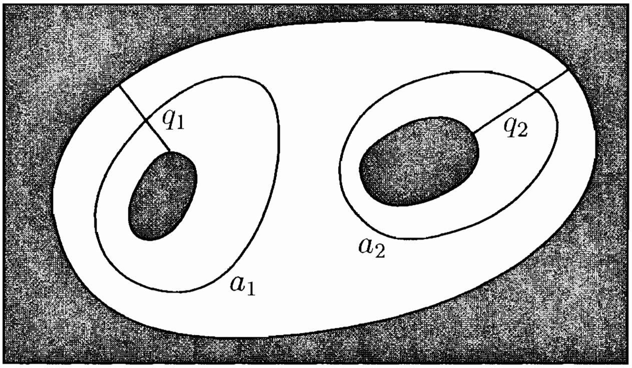


Figure 3: a triply connected surface

In the regions $\alpha, \beta, \gamma, \delta$ of the plane the surface in Figure 4 has two sheets. The sheet containing the curve a_1 is to be regarded as passing underneath the other sheet, as indicated by the dotted lines.

3.

Determination of a function of a complex variable by boundary conditions and discontinuity conditions.

If, in a plane in which the rectangular coordinates of a point are x, y, the value of a function of $x + yi$ is given for the points of a finite line, then the function can be extended continuously beyond this line in only one way and consequently the function is determined throughout its domain (see above,

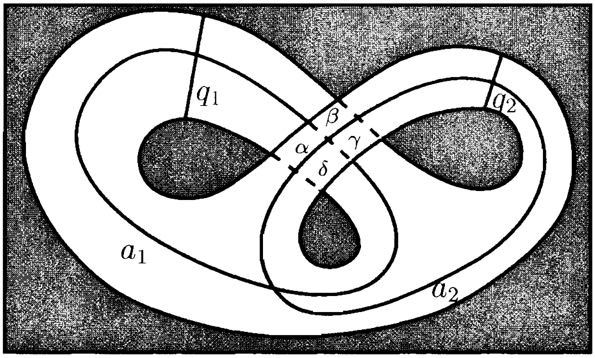


Figure 4: a triply connected surface

p.). It cannot, even on this line, be assigned arbitrary values if it is to be capable of continuation on both sides of the line into the adjacent parts of the surface. For its values on any segment of this line, no matter how small, define the function on the rest of the line. Thus, in this method of determining a function, the conditions which serve to determine it are not independent.

A fundamental requirement in studying a transcendental function is, above all, a set of independent conditions sufficient to define the function. In many cases, notably integrals of algebraic functions and their inverse functions, there is a principle which Dirichlet employed for this purpose. Probably inspired by a similar idea of Gauss, he applied this principle to a function of three variables satisfying Laplace's partial differential equation in his lectures over several years on forces obeying an inverse square law.

For our application to the theory of transcendental functions, one particular case is needed, to which the principle in its simplest form is inapplicable. In Dirichlet's context this case can be neglected as one of lesser importance. This is the case where the function, at certain places in the domain where it must be determined, has prescribed discontinuities. This means that at every such place, the function must satisfy the condition that it becomes discontinuous in the same manner as a given discontinuous function, differing from it by a function continuous there. I shall state the principle in the form required for the application envisaged. Perhaps I may refer, in regard to some related researches, to the account in my doctoral dissertation (*Foundations for a general theory of functions of a complex variable*, Göttingen 1851).

Let us suppose that an arbitrarily bounded surface T is spread over the (x, y) plane simply or multiply, and let two real functions of x, y, denoted by α and β, be given uniquely at each point of T. Denote by $\Omega(\alpha)$ the integral

$$\int \left\{ \left(\frac{\partial \alpha}{\partial x} - \frac{\partial \beta}{\partial y} \right)^2 + \left(\frac{\partial \alpha}{\partial y} + \frac{\partial \beta}{\partial x} \right)^2 \right\} dT$$

extended over T, where the functions α and β can have arbitrary discontinuities provided that the integral does not become infinite. Now $\Omega(\alpha - \lambda)$ remains finite if λ is everywhere continuous with finite partial derivatives. If this continuous function λ is subjected to the condition that it differ from a discontinuous function γ only in an infinitely small part of the surface T, then $\Omega(\alpha - \lambda)$ tends to infinity if γ is discontinuous along a line, or discontinuous at a point in such a way that

$$\int \left\{ \left(\frac{\partial \gamma}{\partial x} \right)^2 + \left(\frac{\partial \gamma}{\partial y} \right)^2 \right\} dT$$

becomes infinite (my inaugural dissertation, Section 17). However, $\Omega(\alpha - \lambda)$ remains finite if γ is discontinuous only at isolated points and so that

$$\int \left\{ \left(\frac{\partial \gamma}{\partial x} \right)^2 + \left(\frac{\partial \gamma}{\partial y} \right)^2 \right\} dT,$$

taken over T, remains finite, for example, when in the neighborhood of a point, at a distance r from it, $\gamma = (-\log r)^\epsilon, 0 < \epsilon < 1/2$. For brevity, we say that a function permitted for λ, not affecting the finiteness of $\Omega(\alpha - \lambda)$, is discontinuous of the first kind. A function without this property is discontinuous of the second kind. Consider now the expression $\Omega(\alpha - \mu)$, where μ is a function which vanishes at the boundary and is continuous, or discontinuous of the first kind. This integral is always finite but, by its nature, can never be negative; and it must therefore at least once, say for $\alpha - \mu = u$, assume a minimum value. Hence Ω, for every function $\alpha - \mu$ differing infinitely little from u, must be greater than $\Omega(u)$.

If therefore σ denotes an arbitrary function of position in the surface T, continuous or discontinuous of the first kind, which vanishes everywhere on the boundary, and h denotes a quantity independent of x, y, then $\Omega(u + h\sigma)$ is greater than $\Omega(u)$ for all small enough positive and negative values of

h. Consequently the coefficient of h, in the expansion of this expression in powers of h, is 0. If this coefficient is 0, then

$$\Omega(u+h\sigma)=\Omega(u)+h^2\int\left\{\left(\frac{\partial\sigma}{\partial x}\right)^2+\left(\frac{\partial\sigma}{\partial y}\right)^2\right\}dT,$$

and hence Ω is always a minimum. This minimum occurs only for one function u; because if there were another minimum for $u+\sigma$, then $\Omega(u+\sigma)$ could not exceed $\Omega(u)$, otherwise

$$\Omega(u+h\sigma)<\Omega(u+\sigma)$$

with an $h<1$. If, however, $\Omega(u+\sigma)=\Omega(u)$ then σ must be a constant and, since it vanishes on the boundary, must be zero. Thus the integral Ω has a minimum only for one function u. As for the variation of the first order, the term in $\Omega(u+h\sigma)$ proportional to h, we have

$$2h\int dT\left\{\left(\frac{\partial u}{\partial x}-\frac{\partial\beta}{\partial y}\right)\frac{\partial\sigma}{\partial x}+\left(\frac{\partial u}{\partial y}+\frac{\partial\beta}{\partial x}\right)\frac{\partial\sigma}{\partial y}\right\}=0.$$

It follows from this equation that the integral

$$\int\left(\left(\frac{\partial\beta}{\partial x}+\frac{\partial u}{\partial y}\right)dx+\left(\frac{\partial\beta}{\partial y}-\frac{\partial u}{\partial x}\right)dy\right),$$

taken over the whole boundary of a part of T, vanishes. If we now convert T, if it is multiply connected, into a simply connected surface T' (in the above manner), then by integrating from a fixed point to the point (x,y) along a path interior to T', we obtain a function of x,y,

$$\nu=\int\left\{\left(\frac{\partial\beta}{\partial x}+\frac{\partial u}{\partial y}\right)dx+\left(\frac{\partial\beta}{\partial y}-\frac{\partial u}{\partial x}\right)dy\right\}+\text{const.}$$

Now ν is either continuous in T' or discontinuous of the first kind, and changes in value across a transverse cut by finite amounts which are constant between the nodes of the network of cuts. The function $v=\beta-\nu$ now satisfies the equations

$$\frac{\partial v}{\partial x}=-\frac{\partial u}{\partial y},\ \frac{\partial v}{\partial y}=\frac{\partial u}{\partial x}.$$

Consequently $u + vi$ is a solution of the differential equation

$$\frac{\partial}{\partial y}(u + iv) - i\,\frac{\partial}{\partial x}(u + iv) = 0$$

and is a function of $x + yi$.

We thus obtain the following theorem, stated in Section 18 of my dissertation:

Let T be a surface which is transformed into a simply connected surface T' by transverse cuts, and let $\alpha + \beta i$ be a given complex function of x, y on T such that the integral

$$\int \left\{ \left(\frac{\partial \alpha}{\partial x} - \frac{\partial \beta}{\partial y} \right)^2 + \left(\frac{\partial \alpha}{\partial y} + \frac{\partial \beta}{\partial x} \right)^2 \right\} dT$$

extended over the whole surface has a finite value. Then this function can be transformed, in only one way, into a function of $x + yi$, by subtraction of a function $\mu + \nu i$ of x, y having the following properties:

1. $\mu = 0$ on the boundary except possibly at isolated points; ν is given arbitrarily at a single point;

2. The variations of μ in T and of ν in T' are discontinuous only in isolated points and such that

$$\int \left\{ \left(\frac{\partial \mu}{\partial x} \right)^2 + \left(\frac{\partial \mu}{\partial y} \right)^2 \right\} dT$$

and

$$\int \left\{ \left(\frac{\partial \nu}{\partial x} \right)^2 + \left(\frac{\partial \nu}{\partial y} \right)^2 \right\} dT$$

taken over T, remain finite. The variations of ν are the same on each edge of the transverse cuts.

When the function $\alpha + \beta i$, at those points where its derivatives become infinite, is discontinuous in the same manner as a given discontinuous function of $x + yi$, and has no discontinuity which can be removed by altering its value at isolated points, then $\Omega(\alpha)$ remains finite and $\mu + \nu i$ is everywhere continuous in T'. This is because a function of $x + yi$ can never have certain kinds of discontinuity (for example, discontinuities of the first kind) at such points (my dissertation, Section 12) and so the difference between two such functions must be continuous, provided that it is not discontinuous of the second kind.

Thus, by the theorem which has just been proved, a function of $x + yi$ can always be found which, ignoring the discontinuities of the imaginary part across the transverse cuts, has prescribed discontinuities interior to T, and whose real part has arbitrarily prescribed values along the boundary; always assuming that at any points where the derivatives become infinite the prescribed discontinuity must be the same as that of a given function of $x + yi$. The conditions at the boundary could, as is easily seen, be put into various other forms without essentially affecting the conclusions which have been drawn.

4.

The theory of Abelian functions.

In the work which follows, I have treated the Abelian functions by a method whose principles were laid down in my inaugural dissertation, and in a somewhat altered form described in the three sections above. To provide an overall view, I begin by briefly summarizing the material.

The first part contains the theory of a system of equivalently branching algebraic functions and their integrals in so far as this can be developed without the theory of θ-series. In Sections 1 to 5, we consider the determination of these functions by means of their branching type and discontinuities. In Sections 6 to 10, we study their representations as ratios of functions of two variables connected by an algebraic equation. In Sections 11 to 13, we discuss the transformation of such expressions by rational substitutions. This study leads to the concept of a class of algebraic equations whose members can be transformed into one another by rational substitutions, which may be important in other researches. The transformation of an equation of this kind into an equation of its class of lowest possible degree (Section 13) may likewise be useful in other circumstances. Lastly, in Sections 14 to 16, in preparation for Part 2, we deal with the application of Abel's addition theorem for an arbitrary system of everywhere finite integrals of equivalently branching algebraic functions to the integration of a system of differential equations.

In the second part, we study a system of everywhere finite integrals of equivalently branching, $2p+1$ times connected algebraic functions. We express the inverse functions of Jacobi in p variables by means of p-fold infinite

θ-series of the form:

$$\theta(v_1, v_2, \ldots, v_p) = \sum_{-\infty}^{\infty} \cdots \sum_{-\infty}^{\infty} e^{\sum_1^p \sum_1^p a_\mu a_{\mu'} m_\mu m_{\mu'} + 2\sum_1^p v_\mu m_\mu}.$$

Here the summation in the exponent is taken over μ and μ'; the outer summation is taken over $m_1, \ldots, m_p$.

It turns out that for the general solution of this problem, when $p > 3$, a special class of θ-series suffices, in which there are $\frac{1}{2}(p-2)(p-3)$ relations between the $\frac{1}{2}p(p+1)$ numbers a, so that only $3p-3$ of them are arbitrary. This part of the memoir constitutes at the same time a theory of this special class of θ-functions. We exclude general theta functions, but they can be treated by an entirely similar method.

Jacobi's inversion problem, of which a solution is given here, has already been solved in several different ways for hyperelliptic integrals through the persistent efforts of Weierstrass, which have been crowned with such success. A survey of his work has appeared in Crelle's Journal (vol. 47, p. 289). However, until now the only parts of that investigation which have been worked out fully are those mentioned in Sections 1 and 2 and the first half of Section 3 relating to elliptic functions (vol. 52, p. 285 of Crelle's Journal). The extent to which there is agreement between the later stages of that work and mine presented here, not only in the results but in the methods used to derive them, will to a large extent emerge only when the promised exposition appears.

The present work, with the exception of the last two sections, 26 and 27, whose subject-matter could be touched upon only briefly in my lectures, is based on a part of my lectures in Göttingen from Michaelmas 1855 to Michaelmas 1856. As regards the discovery of particular results, those in Sections 1 to 5, 9 and 12, and the necessary preliminary theorems which I had to elaborate for my lectures in the manner explained in this memoir, were found in autumn 1851 and the beginning of 1852 in the course of researches into the conformal representation of multiply connected surfaces. I was, however, diverted from these investigations by another matter. It was not until Easter 1855 that I resumed my research, and during the Easter and Michaelmas vacations progressed as far as Section 21; the remaining sections were added by Michaelmas 1856. Some complementary results were added in various places while writing up the work.

Part 1.

1.

If s is the root of an irreducible equation of degree n, whose coefficients are polynomials of degree m in z, then to every value of z correspond n different values of s which vary continuously with z everywhere except when they become infinite. If, therefore, the branching of this algebraic function is represented by an unbounded surface T (as on p.) spread over the z-plane, this surface will have n sheets in every part of the plane, and s will be a single-valued function of position on this surface. An unbounded surface can be regarded either as one whose boundary is infinitely far away or as a closed surface. We shall look on T as closed, so that there will be one point on each of the n sheets corresponding to the value $z = \infty$, except when ∞ is a branch point.

Every rational function of s and z is obviously likewise a single-valued function of position on T and is of the same branching type as the function s. As we shall see later, the converse is also true.

Integration of such a function yields a function whose different prolongations for the same part of T differ only by constants, because their derivatives at the same point of the surface all have the same value.

Such a system of equivalently branching algebraic functions and their integrals constitutes the first object of our study. Instead of proceeding from the above expression for these functions, we define them via their discontinuities, using Dirichlet's principle (p.).

2.

For simplicity, we say that a function is *infinitely small of first order, for a point on T*, if its logarithm is increased by $2\pi i$ by a positive circuit of the boundary of a piece of surface surrounding the point, in which the function remains finite and nonzero. Thus, for a point around which the surface winds μ times, at which z has the finite value a, $(z-a)^{1/\mu}$, and therefore $(dz)^{1/\mu}$, is infinitely small of first order. If $z = \infty$, then $(1/z)^{1/\mu}$ is infinitely small of first order. The case where a function becomes infinitely great or infinitely small of the νth order at a point of T can be treated as though the function became infinitely small or infinitely great of first order at ν coincident (or infinitely close) points; we shall occasionally do this later.

The precise manner, in which the functions considered here become discontinuous, can now be expressed as follows. If one of the functions becomes infinite at a point on T, and if r denotes an arbitrary function which becomes infinitely small of the first order at this point, then the function can always be transformed into one which is continuous at this point by subtracting a finite expression of the form

$$A \log r + Br^{-1} + Cr^{-2} + \dots.$$

This ensues from known theorems on the expansion of functions in power series due to Cauchy, which can also be proved via Fourier series.

3.

Consider an unbounded, connected, everywhere n-sheeted surface T spread over the z-plane. As above we regard T as closed. Let T be cut up into a simply connected surface T'. The boundary of a simply connected surface consists of a single piece, and a closed surface breaks up into an even number of boundary pieces after making an odd number of cuts, and an odd number after an even number of cuts. Hence an even number of cuts is needed. Let the number of transverse cuts required be $2p$. To simplify what follows, we shall suppose that after the first cut, each subsequent cut is made from a point on an earlier cut to the adjacent point on its opposite edge. Now consider a quantity which varies continuously along the boundary of T', with the same variations on both sides of each cut in the system. Then the difference between its values on the two sides, at a point of a transverse cut, is constant along that cut.

We now write $z = x + yi$ and consider a function $\alpha + \beta i$ of x, y in T as follows:

In the neighborhood of the points $\epsilon_1, \epsilon_2, \dots$ specify an arbitrary function r_ν of z that is infinitely small of first order at ϵ_ν, and take the function at ϵ_ν to be a finite sum

$$A_\nu \log r_\nu + B_\nu r_\nu^{-1} + C_\nu r_\nu^{-2} + \dots = \phi_\nu(r_\nu).$$

Here $A_\nu, B_\nu, C_\nu, \dots$ are arbitrary constants. We then draw to an arbitrary point, from each of the points ϵ_ν for which $A_\nu \neq 0$, non-intersecting lines interior to T', the line from ϵ_ν being denoted by ℓ_ν. Lastly we define the function in the remaining part of T so that, except on the transverse cuts and the lines ℓ, it is continuous, while on the positive (left-hand) side of the line ℓ_ν its value is greater by $-2\pi i A_\nu$ than its value on the opposite side of the line, whereas on the positive side of the νth transverse cut, the value is greater than that on the opposite side by the given constant $h^{(\nu)}$. Further, the integral

$$\int \left\{ \left(\frac{\partial \alpha}{\partial x} - \frac{\partial \beta}{\partial y} \right)^2 + \left(\frac{\partial \alpha}{\partial y} + \frac{\partial \beta}{\partial x} \right)^2 \right\} dT,$$

taken over T, has a finite value. It is easily seen that this is always possible provided that the sum of the constants A is zero. This condition is also necessary, because only then is it possible for the function to resume its original value after a circuit of the system of lines ℓ.

We call the constants $h^{(1)}, h^{(2)}, \dots h^{(2p)}$, by which the values of this function on the positive edges of the transverse cuts exceed the values on the negative edges, the *moduli of periodicity* of the function.

By Dirichlet's principle the function $\alpha + \beta i$ can now be transformed into a function ω of $x + yi$, determined apart from an additive constant, by subtracting a similar function of x, y continuous in T' whose moduli of periodicity are purely imaginary. The function ω has the same discontinuities as $\alpha + \beta i$ in the interior of T', and the real parts of the moduli of periodicity of the two functions coincide. Thus ω can be assigned the functions ϕ_ν, and the real parts of the moduli of periodicity, arbitrarily. These conditions determine ω within an additive constant, and the same holds for the imaginary parts of its moduli of periodicity.

It will be seen that this function ω includes, as special cases, the functions indicated in Section 1.

4.

Functions ω that are everywhere finite. (Integrals of the first kind).

We now consider the simplest of these functions, beginning with those which are always finite and therefore continuous throughout the interior of T'. Let $w_1, w_2, \dots, w_p$ be such functions, then so is

$$w = \alpha_1 w_1 + \alpha_2 w_2 + \dots + \alpha_p w_p + \text{const.},$$

where $\alpha_1, \alpha_2, \dots, \alpha_p$ are arbitrary constants. Let $k_1^{(\nu)}, k_2^{(\nu)}, \dots, k_p^{(\nu)}$ be the moduli of periodicity of the functions $w_1, w_2, \dots, w_p$ for the νth transverse cut. The modulus of periodicity of w for this transverse cut is

$$\alpha_1 k_1^{(\nu)} + \alpha_2 k_2^{(\nu)} \cdots + \alpha_p k_p^{(\nu)} = k^{(\nu)}.$$

If the variables α are expressed in the form $\gamma + \delta i$, the real parts of the $2p$ numbers $k^{(1)}, k^{(2)}, \dots, k^{(2p)}$ are linear functions of $\gamma_1, \dots, \gamma_p$, $\delta_1, \dots, \delta_p$. Now if $w_1, w_2, \dots w_p$ are connected by no linear equation with constant coefficients, the determinant of these linear expressions cannot vanish. For otherwise the ratios of the α could be given values such that the moduli of periodicity of the real parts of the function w were all zero, and consequently the real part of w and therefore also w itself would have to be constant, by Dirichlet's principle. The $2p$ numbers γ and δ can therefore be determined in such a way that the real parts of the moduli of periodicity take given values. Consequently w can represent any function ω which always remains finite, provided that $w_1, w_2, \dots, w_p$ do not satisfy any linear equation with constant coefficients. These functions, however, can always be chosen so that they satisfy this condition because, as long as $\mu < p$, there are always linear equations between the moduli of periodicity of the real part of

$$\alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_\mu w_\mu + \text{const.}$$

Thus $w_{\mu+1}$ is not a function expressible in this form, if the moduli of periodicity of the real part of the function are determined (as is always possible in view of the foregoing) in such a way that these linear equations are not satisfied.

Functions ω, which become infinite of first order at a single point of the surface T. (Integrals of the second kind.)

Suppose that ω becomes infinite at only one point ϵ of the surface T, and that for this point all the coefficients in ϕ other than B vanish. Such a function is then defined up to an additive constant by the number B and the real parts of its moduli of periodicity. If we denote any such function by $t^0(\epsilon)$, then the constants $\beta, \alpha_1, \alpha_2, \dots, \alpha_p$ in the expression

$$t(\epsilon) = \beta t^0(\epsilon) + \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_p w_p + \text{const.}$$

can always be determined so that B and the real parts of the moduli of periodicity are given specified values. Thus every such function can be expressed in the above form.

Functions ω which become logarithmically infinite at two points of the surface T. (Integrals of the third kind.)

Thirdly, consider the case where the function ω has only logarithmic infinities. Since the sum of the numbers A must be zero, there must be at least two such points on the surface T, say ϵ_1 and ϵ_2, and we must have $A_2 = -A_1$. Denote by $\tilde{\omega}^0(\epsilon_1, \epsilon_2)$ any one of the functions with this property for which $A_2 = 1$. By similar reasoning to that above, all other such functions are of the form

$$\tilde{\omega}(\epsilon_1, \epsilon_2) = \tilde{\omega}^0(\epsilon_1, \epsilon_2) + \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_p w_p + \text{const.}$$

In the following remarks, we assume for simplicity that the points ϵ are neither branch points nor at infinity. We can then set $r_\nu = z - z_\nu$, where z_ν denotes the value of z at the point ϵ_ν. If we now differentiate $\tilde{\omega}(\epsilon_1, \epsilon_2)$ with respect to z_1 in such a way that the real parts of the moduli of periodicity (or p of the moduli themselves) and the value of $\tilde{\omega}(\epsilon_1, \epsilon_2)$ remain constant for any point of the surface T, we obtain a function $t(\epsilon_1)$ which becomes discontinuous at ϵ_1 in the same way as $\frac{1}{z-z_1}$. Conversely if $t(\epsilon_1)$ is such a function,

$$\int_{z_2}^{z_3} t(\epsilon_1) dz_1,$$

taken over an arbitrary line in T from ϵ_2 to ϵ_3, is a function $\tilde{\omega}(\epsilon_2, \epsilon_3)$. Similarly, by differentiating $t(\epsilon_1)$ n times with respect to z_1 we obtain a function ω which is discontinuous at the point ϵ_1 in the same way as $n!(z - z_1)^{-n-1}$ but is elsewhere finite.

For the positions of the point ϵ which we excluded, these theorems require a slight modification.

Obviously it is always possible to construct a linear expression with constant coefficients in functions w, and functions $\tilde{\omega}$ and their derivatives with respect to the discontinuity values, which have in the interior of T' arbitrary given discontinuities of the same type as those of ω, and whose moduli of periodicity have real parts with arbitrary given values. Thus every given function ω can be represented by such an expression.

5.

The general expression for a function ω which becomes infinite of first order at m points $\epsilon_1, \ldots, \epsilon_m$ of the surface T is, as shown above,

$$s = \beta_1 t_1 + \beta_2 t_2 + \cdots + \beta_m t_m + \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_p w_p + \text{const.},$$

where t_ν is an arbitrary function $t(\epsilon_\nu)$ and the numbers α and β are constants. If ρ of the m points ϵ coincide at the point η of the surface T, the functions

t corresponding to these ρ points must be replaced by a function $t(\eta)$ and its $\rho - 1$ first derivatives with respect to the discontinuity value (Section 2).

The $2p$ moduli of periodicity of this function s are homogeneous linear functions of the $p+m$ numbers α and β. If $m \geq p+1$, then among these $p+m$ numbers α, β there are $2p$ which can be expressed as linear forms in the remainder in such a way that all the moduli of periodicity become zero. The function then contains $m-p+1$ arbitrary constants of which it is a linear homogeneous function, and it may be regarded as a linear form in $m-p$ functions, each of which has only $p+1$ infinities of first order.

If $m = p+1$, the ratios of the $2p+1$ numbers α and β are completely determined for every position of the $p+1$ points ϵ. However, for particular positions of these points, some of the β may be zero. Suppose there are $m-\mu$ such points; the function will be infinite of first order at only μ points. The positions of these μ points must therefore be such that among the $2p$ equations which hold between the other $p+\mu$ numbers β and α, there are $p+1-\mu$ equations which are identical consequences of the others, and consequently only $2\mu-p-1$ points can be chosen arbitrarily. Moreover, the function still contains two arbitrary constants.

Let us now determine s so that μ is as small as possible. If s becomes infinite of first order μ times, then this will also be the case for every rational function of first degree in s. In solving this problem, then, one of the μ points can be chosen arbitrarily. The position of the other points must then be determined in such a way that $p+1-\mu$ of the equations between the α and β are identical consequences of the others. This implies, in the absence of certain equations relating the branch points of the surface T, that

$$p+1-\mu \leq \mu - 1, \quad \text{or} \quad \mu \geq \frac{1}{2}p+1.$$

The number of arbitrary constants in a function s, which has m infinities on the surface T of first order and is continuous everywhere else, is $2m-p+1$ in all cases.

Such a function is the root of an equation of degree n, whose coefficients are polynomials of degree m in z.

If $s_1, s_2, \ldots, s_n$ are the n values of the function s for given z, and if σ denotes a variable, then $(\sigma - s_1)(\sigma - s_2)\ldots(\sigma - s_n)$ is a single-valued function of z which becomes infinite only at those points of the z-plane which coincide with a point ϵ, with an order of infinity equal to the number of points ϵ that coincide there. In fact, to every point ϵ that falls there, that is not a branch point, only one factor of this product is infinite of order higher than 1. However, if ϵ is a point around which the surface T winds μ times, there are μ infinite factors of order higher than $1/\mu$. If we now denote the values of z for those points ϵ at which z is not infinite, by $\zeta_1, \zeta_2, \ldots, \zeta_\nu$, and $(z-\zeta_1)(z-\zeta_2)\ldots(z-\zeta_\nu)$ by a_0, then $a_0(\sigma - s_1)\ldots(\sigma - s_n)$ is a single-valued function of z which is finite for all finite values of z and is infinite of order m for $z = \infty$; and thus is a polynomial of degree m in z. It is equally a polynomial of degree n in σ which vanishes for $\sigma = s$. Let us denote it by F and from now on use the notation $F(\overset{n}{\sigma}, \overset{m}{z})$ for a polynomial function of degree n in σ and m in z. Thus s is a root of the equation

$$F(\overset{n}{s}, \overset{m}{z}) = 0.$$

The function F is a power of an irreducible polynomial, that is, one which cannot be expressed as the product of two polynomials in σ and z. This is because every rational polynomial factor of $F(\sigma, z)$, since it has to vanish for some of the roots $s_1, s_2, \ldots, s_n$, must, when $\sigma = s$, be a function of z which vanishes in some part of the surface T. As T is connected, the function must vanish everywhere. Two irreducible factors of $F(\sigma, z)$ could only vanish together for a finite number of pairs of values if one factor were a constant multiple of the other. It follows that F must be a power of an irreducible polynomial.

If the exponent ν of this power exceeds 1, then the branching type of the function s is not represented by the surface T but by a surface τ, everywhere of n/ν sheets, spread over the z-plane and itself covered ν times by the surface T. Although we could regard s as a function which branches in the same way as T, it would not be true to say that T branches in the same way as s.

Another function which, like s, is discontinuous only at certain points of T, is the function $d\omega/dz$. For this function has the same value at the contiguous points on each edge of the transverse cuts and the lines ℓ, because the differences between the corresponding values of ω are constant along these curves. Hence the function can become infinite only where ω does so or at branch points of the surface. It is continuous elsewhere, since the derivative of a single-valued finite function is necessarily single-valued and finite.

All the functions ω are therefore algebraic functions of z that branch like T, or integrals of such functions. This system of functions is determined by the given surface T and depends only on the position of the branch points.

6.

Let us now suppose that the irreducible equation

$$F(\overset{n}{s}, \overset{m}{z}) = 0$$

has been given and that we have to determine the branching of the function s or the surface T representing it. If, for a value β of z, there are μ branches of the function which connect so that after μ circuits of the variable z around this branch point β, the branch first changes back into itself, then the μ branches of the function can easily be proved (by Cauchy's theorem or using Fourier series) to be represented by a power series in ascending rational powers of $z - \beta$ with exponents which have least common denominator μ, and the converse is true.

A point of the surface T at which only two branches are connected, so that one branch continues into the other and vice versa around this point, is called a *simple branch point.*

A point of the surface around which it winds $\mu + 1$ times can then be regarded as the equivalent of μ coincident (or infinitely near) simple branch points.

To show this, suppose that $s_1, s_2, \ldots, s_{\mu+1}$ are single-valued branches of the function s in a piece of the z-plane surrounding one such point and suppose that $a_1, a_2, \ldots, a_\mu$ are simple branch points following one another on a positive circuit of the boundary of this piece. A circuit around a_1 has the effect of interchanging s_1 and s_2, one around a_2 interchanges s_1 and $s_3, \ldots$, and one around a_μ interchanges s_1 and $s_{\mu+1}$. A circuit enclosing all these branch points, but no others, transforms

$$s_1, s_2, \ldots, s_\mu, s_{\mu+1}$$

into

$$s_2, s_3, \ldots, s_{\mu+1}, s_1$$

and when all the simple branch points coincide, a branch point of order μ ensues.

The properties of the functions ω depend essentially on the connectivity of the surface T. To determine this, we first need to count the simple branch points of the function s.

At a branch point, the branches of the function which connect there all assume the same value, and the equation

$$F(s) = a_0 s^n + a_1 s^{n-1} + \cdots + a_n = 0$$

therefore has two or more equal roots. This can only happen if

$$F'(s) = a_0 n s^{n-1} + a_1(n-1)s^{n-2} + \cdots + a_{n-1} = 0$$

or equivalently, if the single-valued function of z,

$$F'(s_1)F'(s_2)\ldots F'(s_n),$$

vanishes. This latter function becomes infinite for finite values of z only when $s = \infty$, and thus when $a_0 = 0$, and to remain finite needs to be multiplied by a_0^{n-2}. It is then a single-valued function of z which has a finite value for all finite z, which becomes infinite of order $2m(n-1)$ when $z = \infty$ and is therefore a polynomial of degree $2m(n-1)$. The values of z for which $F(s)$ and $F'(s)$ simultaneously vanish are therefore the roots of the equation of degree $2m(n-1)$,

$$Q(z) = a_0^{n-2} \prod_i F'(s_i) = 0,$$

or since $F'(s_i) = a_0 \prod_{i' \neq i} (s_i - s_{i'})$,

$$Q(z) = a_0^{2(n-1)} \prod_{i \neq i'} (s_i - s_{i'}) = 0,$$

the equation which is the result of eliminating s from the equations

$$F'(s) = 0 \text{ and } F(s) = 0.$$

If $F(s,z) = 0$ for $s = \alpha$, $z = \beta$, then

$$F(s,z) = \frac{\partial F}{\partial s}(s-\alpha) + \frac{\partial F}{\partial z}(z-\beta)$$
$$+ \frac{1}{2}\left\{\frac{\partial^2 F}{\partial s^2}(s-\alpha)^2 + 2\frac{\partial^2 F}{\partial s \partial z}(s-\alpha)(z-\beta) + \frac{\partial^2 F}{\partial z^2}(z-\beta)^2\right\} + \cdots,$$

$$F'(s) = \frac{\partial F}{\partial s} + \frac{\partial^2 F}{\partial s^2}(s-\alpha) + \frac{\partial^2 F}{\partial s \partial z}(z-\beta) + \cdots$$

Hence if for $(s = \alpha, z = \beta)$, $\frac{\partial F}{\partial s} = 0$ and $\frac{\partial F}{\partial z}$, $\frac{\partial^2 F}{\partial s^2}$ do not vanish, it must follow that $s - \alpha$ tends to zero like $(z-\beta)^{1/2}$, and we have a simple branch point.

In the product $\prod\limits_i F'(s_i)$, two factors become infinitely small of the order of magnitude of $(z-\beta)^{1/2}$, which shows that $Q(z)$ has $z-\beta$ as a factor. In the case where $\frac{\partial F}{\partial z}$ and $\frac{\partial^2 F}{\partial s^2}$ never vanish when F and $\frac{\partial F}{\partial s}$ are both zero, each linear factor of $Q(z)$ corresponds to one simple branch point and the number of such points is accordingly $2m(n-1)$.

The position of the branch points depends on the coefficients of the powers of z in the functions a and varies continuously with them.

If these coefficients take values such that two simple branch points belonging to the same pair of branches coincide, the two branch points cancel and $F(s)$ has two equal roots, without giving rise to a branch point. If each of s_1, s_2 continues into the other, the effect of making a circuit around a piece of the z-plane containing the two points is that s_1 changes into s_1 and s_2 into s_2, so that the two branches are single-valued functions when they coincide. Their derivative $\frac{ds}{dz}$ is likewise single-valued and finite and hence $\frac{\partial F}{\partial z} = -\frac{ds}{dz}\frac{\partial F}{\partial s} = 0$.

If $F = \frac{\partial F}{\partial s} = \frac{\partial F}{\partial z} = 0$ for $s = \alpha, z = \beta$, the next three terms in the expansion of $F(s,z)$ give two values for the ratio $(s-\alpha)/(z-\beta) = \frac{ds}{dz}$ when $s = \alpha$, $z = \beta$. If these values are distinct and finite, the two branches of the function s to which they belong cannot connect and therefore are not branches of one another. Thus $\frac{\partial F}{\partial s}$ is infinitely small of order $z-\beta$ for each of the two branches, and consequently $Q(z)$ has the factor $(z-\beta)^2$. So only two simple branch points coincide.

In order to decide, in every case where, for $z=\beta$, the equation $F(s)=0$ has more than one root equal to α, how many simple branch points coincide for $s=\alpha$, $z=\beta$ and how many of these cancel each other out, we must express these roots in the form of a power-series in ascending powers of $z-\beta$ by Lagrange's method[3], taking the expansion far enough to ensure that the individual expansions all become different, and thereby determining the true number of distinct branches. We also need to find the order of vanishing of $F'(s)$ for each of these roots, in order to determine the number of the corresponding linear factors of $Q(z)$ or, in other words, the number of coincident simple branch points for which $s=\alpha$, $z=\beta$.

If we denote by ρ the number of times the surface T winds around the point (s,z), then $F'(s)$ will be infinitely small of first order as often as simple

[3]Lagrange, Nouvelle méthode pour résoudre les équations littérales par le moyen des séries. Mémoires de l'Académie de Berlin XXIV, 1780, Oeuvres de Lagrange, Tome III p. 5. W.

branch points coincide at the points z; $dz^{1-1/\rho}$ will be infinitely small of first order as often as there are truly existing branch points, and consequently $F'(z)dz^{1/\rho-1}$ as often as there are pairs of simple branch points which cancel.

It follows therefore that if w denotes the number of actually existing branches, and $2r$ the number of those which cancel each other,

$$w + 2r = 2(n-1)m.$$

If we suppose that the branch points coincide only in pairs which cancel each other out, then for r pairs of values $s = \gamma_\rho$, $z = \delta_\rho$,

$$F = \frac{\partial F}{\partial s} = \frac{\partial F}{\partial z} = 0$$

and

$$\frac{\partial^2 F}{\partial z^2}\frac{\partial^2 F}{\partial s^2} - \left(\frac{\partial^2 F}{\partial s \partial z}\right)^2 \neq 0,$$

while for w pairs of values of s and z,

$$F = 0, \ \frac{\partial F}{\partial s} = 0, \ \frac{\partial F}{\partial z} \neq 0, \ \frac{\partial^2 F}{\partial s^2} \neq 0.$$

We shall confine ourselves for the most part to the treatment of this case because the results can easily be extended to the others as limiting cases. We can do this the more readily since as we have based the theory of these functions on principles which are independent of the form in which the functions are expressed and which admit no exceptions.

7.

In a simply connected surface spread over a finite region of the z-plane, there is a relationship between the number of simple branch points and the number of rotations of direction made by the boundary line of the surface; the latter exceeding the former by one. From this can be deduced a relation, for a multiply connected surface, between these numbers and the number of transverse cuts needed to transform it into a simply connected surface. This relation, which does not depend on metrical considerations, and belongs to *analysis situs*, can be derived in the following way for the surface T.

By Dirichlet's principle, the function $\log \zeta$ of z can be defined as a single-valued function of z in the simply connected surface T' in such a manner that ζ , at an arbitrary interior point T', is infinitely small of first order and $\log\zeta$, along an arbitrary simple line joining this point to the boundary, has a value $-2\pi i$ greater on the positive edge of the line than the negative edge, and elsewhere varies continuously; while its value at all boundary points of T' is purely imaginary. The function ζ defined in this way assumes every complex value once whose modulus is less than 1. The totality of its values can therefore be represented by a surface spread simply over a circle in the ζ-plane. To each point inside the circle corresponds a point of T' and conversely. Thus for an arbitrary point of the surface where $z = z'$, $\zeta = \zeta'$, the function $\zeta - \zeta'$ is infinitely small of first order. Consequently for every finite z', when the surface T' turns upon itself $\mu + 1$ times,

$$(\mu + 1)\,\frac{(z - z')}{(\zeta - \zeta')^{\mu+1}} = \frac{dz}{d\zeta(\zeta - \zeta')^{\mu}}$$

remains finite. However, for infinite z',

$$(\mu + 1)\,\frac{z^{-1}}{(\zeta - \zeta')^{\mu+1}} = \frac{-dz}{z^2 d\zeta(\zeta - \zeta')^{\mu}}$$

remains finite. The integral $\int d\log\frac{dz}{d\zeta}$, taken around the circle in a positive direction, is equal to the sum of the integrals around the points where $\frac{dz}{d\zeta}$ is infinite or zero and its value is therefore $2\pi i(w - 2n)$.

If s denotes distance on the boundary of T' from one and the same fixed point to a variable point on the boundary, and σ the corresponding arc of the circumference of the circle, then

$$\log\frac{dz}{d\zeta} = \log\frac{dz}{ds} + \log\frac{ds}{d\sigma} - \log\frac{d\zeta}{d\sigma}$$

and, integrating over the whole boundary,

$$\int d\log\frac{dz}{ds} = (2p - 1)2\pi i, \int d\log\frac{ds}{d\sigma} = 0, -\int d\log\frac{d\zeta}{d\sigma} = -2\pi i,$$

and therefore

$$\int d\log\frac{dz}{d\zeta} = (2p - 2)2\pi i.$$

This proves that

$$w - 2n = 2(p - 1).$$

Since

$$w = 2((n-1)m - r),$$

we obtain

$$p = (n-1)(m-1) - r.$$

8.

The general expression for a function s' of z branching in the same way as T, which becomes infinite of first order at m' given points of T and remains continuous elsewhere contains, as shown above, $m'-p+1$ arbitrary constants and is a linear function of these (Section 5). If, therefore, it can be shown, as we now intend, that rational functions of s and z can be constructed which are infinite of first order for m' arbitrarily given pairs of values of s and z satisfying the equation $F = 0$, and which are also linear functions of $m'-p+1$ arbitrary constants, then every function s' can be represented in this form.

The quotient of two polynomials $\chi(s, z)$ and $\psi(s, z)$ can only take arbitrary finite values for $s = \infty$ and $z = \infty$ when both are of the same degree. The expression for s' is thus of the form $\frac{\psi(\overset{\nu}{s},\overset{\mu}{z})}{\chi(\overset{\nu}{s},\overset{\mu}{z})}$: moreover, $\nu \geq n-1$, $\mu \geq m-1$. If two unconnected branches of the function s become equal so that at two distinct points of the surface T we have $z = \gamma$ and $s = \delta$, then, generally speaking, s' will have different values at these two points. Thus in order to have $\psi - s'\chi$ identically zero, it is necessary that for two different values of s', $\psi(\gamma, \delta) - s'\chi(\gamma, \delta) = 0$ and consequently $\chi(\gamma, \delta) = 0$ and $\psi(\gamma, \delta) = 0$. The functions χ and ψ must therefore vanish for the r pairs of values $s = \gamma_\rho$, $z = \delta_\rho$ (p.).[4]

The function χ vanishes for a value of z for which the following single-valued function $K(z)$ (finite-valued for all finite z) vanishes:

$$K(z) = a_0^\nu \chi(s_1)\chi(s_2)\ldots\chi(s_n).$$

This function K is infinite of order $m\nu + n\mu$ for infinite z and is therefore a polynomial of degree $m\mu + n\nu$. Since two factors of the product $\prod\limits_i \chi(s_i)$

[4]Here, as already mentioned, we consider only the case in which the branch points of the function s coincide only in pairs canceling each other. In general, the functions χ and ψ, at a point of T where, as envisaged in §6, there are branches which cancel each other when T winds ρ times around the point in question, become infinitely small in the same way as $F'(s)dz^{1/\rho-1}$, so that the first terms in the expansion of the function to be represented, in powers of $(\Delta z)^{1/\rho}$, can take arbitrary values.

become infinitely small of first order for the pairs (γ, δ), it follows that $K(z)$ becomes infinitely small of second order, and hence χ is also infinitely small of the first order for

$$i = m\nu + n\mu - 2r$$

pairs of values of s and z, or points of T.

If $\nu > n-1$, $\mu > m-1$, the value of the function χ remains unchanged if we take

$$\chi(\overset{\nu}{s}, \overset{\mu}{z}) + \rho(\overset{\nu-n}{s}, \overset{\mu-m}{z})F(\overset{n}{s}, \overset{m}{z})$$

in place of $\chi(\overset{\nu}{s}, \overset{\mu}{z})$, where ρ is arbitrary. Consequently, among the coefficients of this expression, there are

$$(\nu - n + 1)(\mu - m + 1)$$

which can be chosen arbitrarily. Now if among the remaining

$$(\mu + 1)(\nu + 1) - (\nu - n + 1)(\mu - m + 1)$$

constants, r of them are determined as linear functions of the others in such a way that χ vanishes for the r pairs of values (γ, δ), the function χ still has

$$\begin{aligned} \epsilon &= (\mu + 1)(\nu + 1) - (\nu - n + 1)(\mu - m + 1) - r \\ &= n\mu + m\nu - (n-1)(m-1) - r + 1 \end{aligned}$$

arbitrary constants. We therefore have

$$i - \epsilon = (n-1)(m-1) - r - 1 = p - 1.$$

If we now choose μ and ν so that $\epsilon > m'$, we can then determine χ in such a way that, for any m' pairs of given values, it becomes infinitely small of first order. Thus when $m' > p$, one can fix ψ so that $\frac{\psi}{\chi}$ remains finite for all other values. In fact ψ is likewise a homogeneous linear function of ϵ arbitrary constants, and therefore when $\epsilon - i + m' > 1$ it is possible to determine $i - m'$ of them as linear combinations of the others so that ψ also vanishes for the $i - m'$ pairs of values of s and z for which χ becomes infinitely small of first order. The function ψ thus contains $\epsilon - i + m' = m' - p + 1$ arbitrary constants, and therefore $\frac{\psi}{\chi}$ can represent every function s'.

9.

As the functions $\frac{d\omega}{dz}$ are algebraic functions of z which branch in the same manner as the function s (Section 5), they can, by the theorem which has just been proved, be expressed rationally in terms of s and z, and all the functions ω can be expressed as integrals of rational functions of s and z.

If w is an everywhere finite function ω, its derivative $\frac{dw}{dz}$ is infinite of first order at each simple branch point of the surface T, because dw and $(dz)^{1/2}$ are both infinitely small of first order at these points. Everywhere else, dw/dz remains continuous, and is infinitely small of the second order for $z = \infty$. Conversely, the integral of a function which exhibits this behaviour is everywhere finite.

In order to express this function $\frac{dw}{dz}$ as a quotient of two polynomials in s and z we must (by Section 8) take as denominator a function which vanishes at the branch points and for the r pairs of values (γ, δ). The simplest way to satisfy this requirement is to take a function which vanishes only for these values. Now

$$\frac{\partial F}{\partial s} = a_0 n s^{n-1} + a_1(n-1)s^{n-2} + \cdots + a_{n-1}$$

is such a function.

This function becomes infinite of order $n-2$ when s is infinite (since a_0 is then infinitely small of first order) and infinite of mth order for an infinite z. Thus, to ensure that $\frac{dw}{dz}$ should be finite at all finite points other than branch points, and infinitely small of second order for infinite z, the numerator has to be a polynomial $\phi(\overset{n-2}{s}, \overset{m-2}{z})$ which vanishes for the r pairs of values (γ, δ) (p.). Thus

$$w = \int \frac{\phi(\overset{n-2}{s}, \overset{m-2}{z})}{\frac{\partial F}{\partial s}}\, dz = -\int \frac{\phi(\overset{n-2}{s}, \overset{m-2}{z})}{\frac{\partial F}{\partial z}}\, ds$$

where $\phi = 0$ for $s = \gamma_\rho$, $z = \delta_\rho$, $\rho = 1, 2, \ldots, r$.

The function ϕ contains $(n-1)(m-1)$ constant coefficients, and if r of them are determined as linear functions of the others so that $\phi = 0$ for the r pairs of values $s = \gamma$, $z = \delta$, there still remain $(m-1)(n-1) - r$, or p constants which can be chosen arbitrarily, and ϕ takes the form

$$\alpha_1\phi_1 + \alpha_2\phi_2 + \cdots + \alpha_p\phi_p,$$

in which $\phi_1, \phi_2, \ldots, \phi_p$ are particular functions ϕ, none of which is a linear function of the others, and $\alpha_1, \alpha_2, \ldots, \alpha_p$ are arbitrary constants. A more general expression for w follows, already obtained in a different way, namely

$$\alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_p w_p + \text{const.}$$

The functions ω which are not everywhere finite, and so the integrals of the second and third kinds, can be expressed as rational functions of s and z by using the same principles, but we shall not dwell on this now, because the general rules of the preceding paragraphs do not call for any more detailed explanations. Moreover, definite forms for such integrals arise in the theory of θ-functions.

10.

The function ϕ, as well as for the r pairs of values (γ, δ), is also infinitely small of first order for $m(n-2)+n(m-2)-2r$, or $2(p-1)$, pairs of values of s and z satisfying the equation $F=0$. Now if

$$\phi^{(1)} = \alpha_1^{(1)}\phi_1 + \alpha_2^{(1)}\phi_2 + \cdots + \alpha_p^{(1)}\phi_p$$

and

$$\phi^{(2)} = \alpha_1^{(2)}\phi_1 + \alpha_2^{(2)}\phi_2 + \cdots + \alpha_p^{(2)}\phi_p$$

are any two functions ϕ, the numerator of the expression $\phi^{(2)}/\phi^{(1)}$ can be determined so that it is equal to zero for $p-1$ arbitrarily given pairs of values of s and z satisfying the equation $F=0$; and then the denominator so that it vanishes for $p-2$ of the other pairs of values for which $\phi^{(1)}=0$. The expression so constructed is a linear function of two arbitrary constants and is consequently a general expression of a function which can become infinite of first order at only p points of the surface T. A function which becomes infinite at fewer than p points constitutes a special case of this function. Thus every function which is infinite of the first order at fewer than $p+1$ points of the surface T can be expressed in the form $\phi^{(2)}/\phi^{(1)}$ or in the form $\frac{dw^{(2)}}{dw^{(1)}}$, where $w^{(1)}$ and $w^{(2)}$ are two everywhere finite integrals of rational functions of s and z.

11.

A function z_1 of z that branches like T, and which becomes infinite of first order at n_1 points of this surface is, by the foregoing (p.), the root of an equation of the form

$$G(\overset{n}{z_1}, \overset{n_1}{z}) = 0$$

and therefore takes every value at n_1 points of the surface T. If therefore we imagine each point of T to be mapped onto by a point on a plane representing geometrically the value of z_1 at this point, the totality of these points forms a surface T_1, covering the the z_1-plane n_1 times. The image is known to be similar to T in its smallest parts. To each point in the one surface corresponds one and only one point in the other surface. Thus the functions ω, that is to say the integrals of functions of z that branch like T, transform into functions which have a unique value throughout the surface T_1 and which have the same discontinuities as ω in the corresponding points of T, when z is replaced by z_1 as an independent variable. These functions are therefore integrals of functions of z_1 branching like T_1.

If we denote by s_1 any other function of z branching like T, which for m_1 points of T—and consequently also for m_1 points of T_1—becomes infinite of first order, then (Section 5) an equation of the form

$$F_1(\overset{n_1}{s_1}, \overset{m_1}{z_1}) = 0$$

holds, in which F_1 is a power of an irreducible polynomial in s_1 and z_1. When this power is the first, all functions of z_1 which branch like T_1, and hence all rational functions of s and z, can be expressed rationally in terms of s_1 and z_1 (Section 8).

The equation

$$F(\overset{n}{s}, \overset{m}{z}) = 0$$

can therefore be transformed into

$$F(\overset{n_1}{s_1}, \overset{m_1}{z_1}) = 0$$

by a rational substitution, and conversely.

The domains (s, z) and (s_1, z_1) have the same connectivity because to each point of one corresponds one point of the other. Thus if r_1 denotes the number of cases in which s_1 and z_1 both assume the same value for two different points of the domain (s_1, z_1) and consequently F_1, $\frac{\partial F_1}{\partial s_1}$, $\frac{\partial F_1}{\partial z_1}$ vanish while

$$\frac{\partial^2 F_1}{\partial s_1^2}\,\frac{\partial^2 F_1}{\partial z_1^2} - \left(\frac{\partial^2 F_1}{\partial s_1 \partial z_1}\right)^2$$

does not, then it follows that

$$(n_1 - 1)(m_1 - 1) - r_1 = p = (n - 1)(m - 1) - r.$$

12.

We shall now consider *all irreducible algebraic equations between two complex variables, which can be transformed into one another by rational transformations, as belonging to the same class.* Thus $F(s,z) = 0$ and $F_1(s_1, z_1) = 0$ belong to the same class, if rational functions of s_1 and z_1 can be found which, when substituted for s and z respectively, transform the equation $F(s,z) = 0$ into the equation $F_1(s_1, z_1) = 0$; while equally s_1 and z_1 are rational functions of s and z.

The rational functions of s and z regarded as functions of any one of them, say ζ, constitute a system of similarly branching algebraic functions. In this way, every equation clearly gives rise to a class of systems of similarly branching algebraic functions, which by introducing one of them as an independent variable, can be rationally transformed into each other. Moreover, all the equations of one class lead to the same class of systems of algebraic functions, and conversely (Section 11) each class of such systems leads to one class of equations.

If the (s, z) domain is $2p+1$ times connected and the function ζ becomes infinite of first order at μ points of this domain, the number of branch points of equivalently branching functions of ζ, which can be formed by the other rational functions of s and z, is $2(\mu - p + 1)$, and the number of arbitrary constants in the function ζ is therefore $2\mu - p + 1$ (Section 5). These constants can always be chosen so that $2\mu - p + 1$ branch points take arbitrarily assigned values, when these branch points are mutually independent functions of the constants. This can be done in only a finite number of different ways because the conditions are algebraic. In each class of of similarly branching functions with connectivity $2p+1$, there is consequently only a finite number of μ-valued functions such that $2\mu + p - 1$ branch points have prescribed values. If, on the other hand, the $2(\mu + p - 1)$ branch points of a surface with connectivity $2p + 1$, covering the whole ζ-plane μ times, are arbitrarily prescribed, then (Sections 3–5) there is always a system of algebraic functions of ζ branching like the surface. The remaining $3p - 3$ branch points in these systems of similarly branching μ-valued functions can therefore be assigned any given values; and thus a class of systems of similarly branching functions with connectivity $2p + 1$, and the corresponding class of algebraic equations, depends on $3p - 3$ continuous variables, which we shall call the moduli of the class.

This determination of the number of moduli of a class of algebraic functions with connectivity $2p+1$ is, however, valid only with the proviso that there are $2\mu - p + 1$ branch points which are independent functions of the arbitrary constants in the function ζ. This hypothesis implies that $p > 1$, and in this case the number of moduli is $3p - 3$. When $p = 1$, the number of moduli is 1. A straightforward determination of this number is made difficult by the precise way in which the arbitrary constants enter into ζ. Accordingly in order to determine the number of moduli, we introduce into a system of similarly branching functions with connectivity $2p+1$, as an independent variable, not one of these functions, but rather an everywhere finite integral of such a function.

The values which the function w of z takes in the surface T' are represented geometrically by a surface which we shall call S, which covers simply or multiply a finite portion of the w-plane and is similar to T' in the smallest parts. Since w, on the positive edge of the νth transverse cut, is greater by a constant $k^{(\nu)}$ than on the negative edge, the boundary of S consists of pairs of parallel curves each of which is the image of the same part of the network of cuts forming the boundary of T'. The difference in location of corresponding points in the parallel portions of the boundary of S corresponding to the νth transverse cut is expressed by the complex number $k^{(\nu)}$. The number of simple branch points of the surface S is $2p-2$, because dw becomes infinitely small of second order in $2p-2$ points of T. The rational functions of s and z are thus functions of w, which, for every point of S at which they do not become infinite, have a unique value which varies continuously with s, z and is the same at corresponding points of parallel boundary portions. They therefore constitute a system of similarly branching $2p$-fold periodic functions of w. It can now be shown (in the same way as in Sections 3 to 5) that if the $2p-2$ branch points and the $2p$ differences of location of parallel boundary portions of S are assigned arbitrarily, a system of functions exists which branches similarly to the surface S, and which assumes the same value in corresponding points of parallel boundary portions and is therefore $2p$-fold periodic. Further, this system of functions, regarded as a function of one of them, constitutes a system of similarly branching algebraic functions with connectivity $2p+1$ and consequently defines a class of algebraic functions with connectivity $2p+1$. In fact, by Dirichlet's principle, a function of w can be defined on the surface S, to within an additive constant, by the following conditions: in the interior of S its arbitrarily prescribed discontinuities are of the same form as those of ω in T'; and on the corresponding points of the parallel boundary portions, it is assigned constant values whose real part is given. We can conclude from this, in the same way as in Section 5, the existence of functions which are discontinuous only at isolated points of S and have the same value at corresponding points on parallel boundary portions. If such a function z becomes infinite of first order at n points of S and is not discontinuous elsewhere, it takes every complex value at n points of S. For if a is an arbitrary constant, then $\int d\log(z-a) = 0$ around the boundary of S, because the contributions to the integral from the parallel boundary portions cancel. Hence $z - a$ is infinitely small as often as it is infinitely large of first order. The values assumed by z are consequently represented by a surface covering the z-plane n times everywhere. The other functions of w which are similarly branching and periodic constitute a system of algebraic functions of z, branching in the same way as the surface, with connectivity $2p+1$, as we wished to prove.

For any given class of algebraic functions with connectivity $2p+1$, we introduce the quantity

$$w = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_p w_p + c$$

as an independent variable. We can determine the numbers α so that p of the $2p$ moduli of periodicity have given values, and then choose c, when $p > 1$, so that one of the $2p-2$ branch points of the periodic functions of w has a given value. This completely determines w. Hence there are $3p-3$ remaining quantities, on which the form of branching and periodicity of these functions of w depends. Since, to arbitrary values of these $3p-3$ quantities, corresponds a class of algebraic functions with connectivity $2p+1$, such a class depends on $3p-3$ independent variables.

When $p = 1$ there are no branch points, and in the expression

$$w = \alpha_1 w_1 + c$$

the number α_1 can be determined so that one of the moduli of periodicity has a given value, and the value of the other modulus of periodicity is thereby determined. Accordingly the number of moduli for a class is 1.

13.

In accordance with the principles of transformation developed in Section 11, it is clear that in order to transform any given equation $F(s,z) = 0$ into an equation

$$F_1(\overset{n_1}{s_1}, \overset{m_1}{z_1}) = 0,$$

by a rational substitution, of the same class and of lowest possible degree, we must first determine for z_1 an expression $r(s,z)$ rational in s and z such that n_1 shall be as small as possible. We then likewise determine for s_1 another expression $r'(s,z)$ such that m_1 shall be the smallest possible, while the values of s_1 corresponding to an arbitrary value of z_1 do not split up into groups equal to one another. This ensures that $F_1(\overset{n_1}{s_1}, \overset{m_1}{z_1})$ is not a higher power than the first of an irreducible polynomial.

If the domain of values (s,z) has connectivity $2p+1$, the smallest value which can be taken by n_1 is, generally speaking, $\geq \frac{p}{2}+1$ (Section 5) and the number of cases in which s_1 and z_1 can both take the same values for two different points in the domain is

$$(n_1-1)(m_1-1)-p.$$

In a class of algebraic equations between two variables, the equations of lowest degree are therefore, in the absence of any special relations between the moduli, as follows:

$$\begin{array}{lll} \text{for} \quad p=1, & F(\overset{2}{s}, \overset{2}{z})=0, & r=0, \\ \qquad\ \ \, p=2, & F(\overset{2}{s}, \overset{3}{z})=0, & r=0. \\ \text{For} \quad p>2, & & \\ \qquad\ \ \, p=2\mu-3, & F(\overset{\mu}{s}, \overset{\mu}{z})=0, & r=(\mu-2)^2, \\ \qquad\ \ \, p=2\mu-2, & F(\overset{\mu}{s}, \overset{\mu}{z})=0, & r=(\mu-1)(\mu-3). \end{array}$$

Of the coefficients of the powers of s and z in the polynomials F, r of them must be determined as linear homogeneous functions of the others in such a way that $\frac{\partial F}{\partial s}$ and $\frac{\partial F}{\partial z}$ simultaneously vanish for r pairs satisfying $F=0$. The rational functions of s and z, regarded as functions of one of them, will then represent all systems of algebraic functions with connectivity $2p+1$.

14.

I shall now, following Jacobi[5], (Crelle's Journal, vol.9, No. 32, Section 8) use Abel's addition theorem for the integration of a system of differential equations, confining myself to what will be required later in this memoir.

If, in an everywhere finite integral w of a rational function of s and z, we introduce as an independent variable a rational function ζ of s and z_1 which

[5]Jacobi, *Gesammelte Werke*, vol. II, p. 15. W.

becomes infinite of first order for m pairs of values of s and z, then $\frac{dw}{dz}$ is an m-valued function of ζ. If we denote the m values of w for the same value of ζ by $w^{(1)}, w^{(2)}, \ldots, w^{(m)}$, then

$$\frac{dw^{(1)}}{d\zeta} + \frac{dw^{(2)}}{d\zeta} + \cdots + \frac{dw^{(m)}}{d\zeta}$$

is a single-valued function of ζ whose integral is everywhere finite. Consequently,

$$\int d(w^{(1)} + w^{(2)} + \cdots + w^{(m)})$$

is likewise everywhere finite and single-valued and therefore constant. Similarly, if $\omega^{(1)}, \omega^{(2)}, \ldots, \omega^{(m)}$ are the values of an arbitrary integral ω of a rational function of s and z, corresponding to the same value of ζ, then

$$\int d(\omega^{(1)} + \omega^{(2)} + \cdots + \omega^{(m)})$$

is, apart from an additive constant, a well-defined function whose value is determined by the discontinuities of ω. It is the sum of a rational function of ζ and a linear combination (with constant coefficients) of logarithms of rational functions of ζ.

With the aid of this theorem, we shall now show that the following set of p simultaneous differential equations between the $p+1$ pairs of values of s and z, $(s_1, z_1), (s_2, z_2), \ldots, (s_{p+1}, z_{p+1})$, for which $F(s,z) = 0$:

$$\frac{\phi_\pi(s_1, z_1)}{\frac{\partial F(s_1,z_1)}{\partial s_1}}\, dz_1 + \frac{\phi_\pi(s_2, z_2)}{\frac{\partial F(s_2,z_2)}{\partial s_2}}\, dz_2 + \cdots + \frac{\phi_\pi(s_{p+1}, z_{p+1})}{\frac{\partial F(s_{p+1},z_{p+1})}{\partial s_{p+1}}}\, dz_{p+1} = 0$$

for $\pi = 1, \ldots, p$, have a general (complete) solution.

Through these differential equations any p of the pairs of values (s_μ, z_μ) are well-defined functions of the remaining one pair if, when a value of the latter is given, the values of the others follow. If therefore the $p+1$ pairs are determined as functions of a single variable ζ, in such a way that, for the value 0 of this single variable, the pairs are assigned arbitrary initial values (s_1^0, z_1^0), $(s_2^0, z_2^0), \ldots, (s_{p+1}^0, z_{p+1}^0)$, and by the requirement that they satisfy the differential equations, this provides the complete general solution of the differential equations. Now the quantity $1/\zeta$ can always be determined as a single-valued and therefore rational function of (s, z) so that it becomes infinite only for some or all of the $p+1$ pairs of values (s_μ^0, z_μ^0), and then only infinite of the first order, because in the expression

$$\sum_{\mu=1}^{p+1} \beta_\mu t(s_\mu^0, z_\mu^0) + \sum_{\mu=1}^{p} \alpha_\mu w_\mu + \text{ const.},$$

the ratios of the α and β can always be determined so that all the moduli of periodicity are 0. If no β is 0, the $p+1$ branches of the $(p+1)$-valued equivalently branching functions s and z of ζ, $(s_1, z_1), (s_2, z_2), \ldots, (s_{p+1}, z_{p+1})$, that take the values $(s_1^0, z_1^0), (s_2^0, z_2^0), \ldots, (s_{p+1}^0, z_{p+1}^0)$ for $\zeta = 0$, satisfy the required differential equations. If, however, some of the β (say, the last $p+1-m$) vanish, the differential equations are satisfied by the m branches of the m-valued functions s and z of ζ, $(s_1, z_1), (s_2, z_2), \ldots, (s_m, z_m)$, which assume the values

$$(s_1^0, z_1^0), (s_2^0, z_2^0), \ldots, (s_m^0, z_m^0)$$

when $\zeta = 0$, and by constants whose values are the initial values of s_{m+1}, z_{m+1}; $\ldots; s_{p+1}, z_{p+1}$: that is, $s_{m+1}^0, \ldots, z_{p+1}^0$. In the latter case, of the p linear homogeneous equations

$$\sum_{\mu=1}^{m} \frac{\phi_\pi(s_\mu, z_\mu)}{\frac{\partial F(s_\mu, z_\mu)}{\partial s_\mu}} dz_\mu = 0$$

with $\pi = 1, \ldots, p$, which hold between the differentials $dz_\mu / \frac{\partial F(s_\mu, z_\mu)}{\partial s_\mu}$, $p+1-m$ of them are consequences of the others. It follows therefore that $p+1-m$ conditions must hold between the functions $(s_1, z_1), (s_2, z_2), \ldots, (s_m, z_m)$ and hence also between their initial values $(s_1^0, z_1^0), \ldots, (s_m^0, z_m^0)$ for this case to occur. Accordingly only $2m-p-1$ of the constants can have arbitrary values, as established in Section 5.

15.

We now suppose that the integral

$$\int \frac{\phi_\pi(s, z)}{\frac{\partial F(s,z)}{\partial s}} dz + \text{ const.},$$

over the interior of T', is equal to w_π and that the modulus of periodicity of w_π for the νth transverse cut is $k_\pi^{(\nu)}$, so that the functions $w_1, w_2, \ldots, w_p$ of the pairs of values (s, z) simultaneously increase by the amounts $k_1^{(\nu)}, k_2^{(\nu)}, \ldots, k_p^{(\nu)}$

when the point (s, z) crosses from the negative to the positive edge of the νth transverse cut. For brevity, we shall say that a system of p numbers $(b_1, b_2, \ldots, b_p)$ is *congruent* to another such system $(a_1, a_2, \ldots, a_p)$ *with respect to* $2p$ *systems of associated moduli* when one can be derived from the other by simultaneously changing the values of their components by multiples of the associated moduli. Thus if the πth component of the νth system is $k_\pi^{(\nu)}$, then we write

$$(b_1, b_2, \ldots, b_p) \equiv (a_1, a_2, \ldots, a_p)$$

if

$$b_\pi = a_\pi + \sum_{\nu=1}^{2p} m_\nu k_\pi^{(\nu)}$$

for $\pi = 1, 2, \ldots, p$, where $m_1, m_2, \ldots, m_{2p}$ are integers.

Now p arbitrary numbers $a_1, a_2, \ldots, a_p$ can always be uniquely expressed in the form

$$a_\pi = \sum_{\nu=1}^{2p} \xi_\nu k_\pi^{(\nu)}$$

in such a way that the $2p$ numbers ξ are real, and all congruent systems, and only these, are obtained when the numbers ξ are changed by integral amounts. Hence, if in these expressions each number ξ is allowed to increase continuously by 1 from a given value, excluding one of the boundary values, the above expressions run precisely once over representatives of congruent systems.

This point having been established, it follows on integrating the above differential equations or from the p equations

$$\sum_{\mu=1}^{p+1} dw_\pi^{(\mu)} = 0 \quad (\pi = 1, \ldots, p)$$

that

$$\left(\sum w_1^{(\mu)}, \sum w_2^{(\mu)}, \ldots, \sum w_p^{(\mu)}\right) \equiv (c_1, c_2, \ldots, c_p)$$

where $c_1, c_2, \ldots, c_p$ are constants depending on the values (s^0, z^0).

16.

Expressing ζ as the quotient of two polynomials in s, z, say χ/ψ, the pairs $(s_1, z_1), (s_2, z_2), \ldots, (s_m, z_m)$ are the common roots of the equations $F = 0$

and $\frac{\chi}{\psi} = \zeta$. Since the polynomial

$$\chi - \zeta\psi = f(s, z)$$

vanishes for all pairs of values for which χ and ψ simultaneously vanish, whatever the value of ζ, the pairs $(s_1, z_1), \ldots, (s_m, z_m)$ can also be defined as the common roots of the equation $F = 0$ and of an equation $f(s, z) = 0$ whose coefficients vary in such a manner that all the other common roots remain constant. If $m < p + 1$, ζ can be expressed in the form $\frac{\phi^{(1)}}{\phi^{(2)}}$ (Section 10) and f in the form

$$\phi^{(1)} - \zeta\phi^{(2)} = \phi^{(3)}.$$

The most general pairs of functions $(s_1, z_1), (s_2, z_2), \ldots, (s_p, z_p)$ to satisfy the p equations

$$\sum_{\mu=1}^{p} dw_\pi^{(\mu)} = 0 \quad \text{for } \pi = 1, 2, \ldots, p$$

are therefore those which are formed from p common roots of those equations $F = 0$ and $\phi = 0$ which vary in such a way that the remaining other common roots remain constant. From this fact may easily be deduced the following proposition, needed later. The problem of determining $p - 1$ of the $2p - 2$ pairs of values $(s_1, z_1), \ldots, (s_{2p-2}, z_{2p-2})$ as functions of the other $p - 1$ pairs so that the p equations

$$\sum_{\mu=1}^{2p-2} dw_\pi^{(\mu)} = 0 \quad \text{for } \pi = 1, \ldots, p$$

are all satisfied, can be solved in full generality by taking, for the $2p - 2$ quantities, common roots of the equations $F = 0$ and $\phi = 0$ that differ from the r roots $s = \gamma_\rho$, $z = \delta_\rho$ (Section 6). Alternatively we can take the $2p - 2$ pairs of values for which dw is infinitely small of second order. Hence this problem has only one solution. Such pairs of values may be said to be *associated through the equation* $\phi = 0$. By virtue of the equations

$$\sum_{\mu=1}^{2p-2} dw_\pi^{(\mu)} = 0,$$

the p-tuple

$$\left(\sum_1^{2p-2} w_1^{(\mu)}, \sum_1^{2p-2} w_2^{(\mu)}, \ldots, \sum_1^{2p-2} w_p^{(\mu)}\right)$$

the sum being taken over such pairs, is congruent to a system of constants $(c_1, c_2, \ldots, c_p)$, where c_π depends only on the additive constant in the function w_π, that is, on the initial value of the integral which expresses this function.

Part 2.

17.

For further investigation of the integrals of algebraic functions with connectivity $2p+1$, a certain p-fold infinite θ-series is very useful. This is a series in which the logarithm of the general term is a quadratic polynomial in its indices. Suppose that, in this polynomial, the coefficient of the square m_μ^2 is $a_{\mu,\mu}$, the coefficient of the product $m_\mu m_{\mu'}$ is $a_{\mu,\mu'} = a_{\mu',\mu}$, that of $2m_\mu$ is v_μ, and the constant term is 0. We denote by $\theta(v_1, v_2, \ldots, v_p)$ the sum of this series over all positive or negative integer values of the m; the sum $\theta(v_1, \ldots, v_p)$ is regarded as a function of the p variables v. Thus
(1)

$$\theta(v_1, v_2, \ldots, v_p) = \sum_{m_1=-\infty}^{\infty} \cdots \sum_{m_p=-\infty}^{\infty} \exp\left(\sum_{\mu,\mu'=1}^{p} a_{\mu,\mu'} m_\mu m_{\mu'} + 2\sum_{\mu=1}^{p} v_\mu m_\mu \right).$$

For this series to converge, the real part of $\sum_{\mu,\mu'=1}^{p} a_{\mu,\mu'} m_\mu m_{\mu'}$ must be essentially negative. In other words, as a sum of positive or negative squares of independent linear forms in the variables m, it is a sum of p negative squares.

The function θ has the property that systems of simultaneous variations in the p arguments $v_1, v_2, \ldots, v_p$ exist for which $\log\theta$ changes only by a linear combination of the v. Indeed, there are $2p$ such systems, independent in the sense that none is a linear combination of the others. For, leaving out arguments which suffer no change, we have, for $\mu = 1, 2, \ldots, p$,

$$(2) \qquad \theta = \theta(v_\mu + \pi i)$$

and

$$(3) \qquad \theta = e^{2v_\mu + a_{\mu,\mu}} \theta(v_1 + a_{1,\mu}, v_2 + a_{2,\mu}, \ldots, v_p + a_{p,\mu})$$

because when the index m_μ in the θ-series is changed to $m_\mu + 1$, the sum of the series is unaffected, and its value becomes the expression on the right side.

The function θ is defined, apart from a constant factor, by these two last relations and the property of being always finite. For by the finiteness property and (2), it is a single-valued function of $e^{2v_1}, e^{2v_2}, \ldots, e^{2v_p}$ that is finite for finite v, and can therefore be developed in a p-fold infinite series of the form

$$\sum_{m_1=-\infty}^{\infty} \cdots \sum_{m_p=-\infty}^{\infty} A_{m_1,m_2,\ldots,m_p} \exp\left(2\sum_1^p v_\mu m_\mu\right)$$

with constant coefficients A. The relations (3) imply, however, that

$$A_{m_1,\ldots,m_\nu+1,\ldots,m_p} = A_{m_1,\ldots,m_\nu,\ldots,m_p} \exp\left(2\sum_1^p a_{\mu,\nu} m_\mu + a_{\nu,\nu}\right)$$

and consequently

$$A_{m_1,\ldots,m_p} = \text{ const. } \exp\left(\sum_{\mu,\mu'=1}^{p} a_{\mu,\mu'} m_\mu m_{\mu'}\right)$$

as we wished to prove.

These properties of the function can thus be used to define it. The systems of simultaneous variations of the v which cause $\log\theta$ to change by a linear function of these quantities will be called *systems of associated moduli of periodicity of the independent variables* for this θ-function.

18.

I now substitute for the p variables $v_1, \ldots, v_p$ a set $u_1, u_2, \ldots, u_p$ of everywhere finite integrals of rational functions of z and s. Here s is an algebraic function of z with connectivity $2p+1$. For the associated moduli of periodicity of the variables v, I assign the associated moduli of periodicity of these integrals (that is, associated relative to the same transverse cuts). In this way $\log\theta$ becomes a function of the variable z alone, which changes its value by amounts that are linear functions of the u whenever s and z recover their previous values after a continuous variation of z.

It will first be shown that a substitution of this kind is possible for every function s with connectivity $2p+1$. To this end the surface T must be decomposed by closed cuts $a_1, \ldots, a_p$, $b_1, \ldots, b_p$ fulfilling the following conditions: When $u_1, u_2, \ldots, u_p$ are chosen so that the modulus of periodicity of u_μ at the cut a_μ is equal to πi and is equal to zero at the other cuts a, and if the modulus of periodicity of u_μ at the cut b_μ is denoted by $a_{\mu,\nu}$, we must have

$$a_{\mu,\nu} = a_{\nu,\mu}$$

and the real part of $\sum_{\mu,\mu'} a_{\mu,\mu'} m_\mu m_{\mu'}$ must be negative for all real (integer) values of $m_1, \ldots, m_p$.

19.

The dissection of the surface T will be accomplished, not as hitherto by means of closed transverse cuts, but as follows.

We begin by making a cut a_1 that returns to its starting-point but does not cut the surface into two pieces; we then make a transverse cut b_1 from a point on the positive edge of a_1 to the corresponding point on the negative edge; this yields the boundary as a single piece. If the cut surface is not already simply connected, a third transverse cut (not splitting the surface) can be made starting from an arbitrary point on this boundary and going to another point on this boundary (an earlier point of this transverse cut). The latter cut is done in such a way that it consists of a closed line a_2, followed by a continuation part c_1 which links it to the preceding system of cuts. The next transverse cut b_2 is now made from a point on the positive edge of a_2 along a line leading to the same point on the negative edge, so that again we have the boundary consisting of one piece. Further cutting up of the surface, if necessary, can then again be done by two cuts a_3 and b_3 with the same initial point and endpoint, and a line c_2 linking them to the system of lines a_2 and b_2. This procedure can be followed until the surface becomes simply connected, and one obtains a network of cuts consisting of p pairs of lines with the same initial point and endpoint, a_1 and b_1, a_2 and $b_2, \ldots, a_p$ and b_p, and $p-1$ lines $c_1, c_2, \ldots, c_{p-1}$, linking each pair with the succeeding one. The line c_ν can go from a point of b_ν to a point of $a_{\nu+1}$. The network of cuts can be regarded as follows: the $(2\nu-1)$-th transverse cut consists of $c_{\nu-1}$ and the line a_ν starting from the endpoint of $c_{\nu-1}$ and returning to this point. The 2ν-th transverse cut consists of the line b_ν, starting from the positive edge of a_ν and returning to the same point on the negative edge. The boundary of the surface is made up of one single piece after an even number of cuts, and of two pieces after an odd number.

An everywhere finite integral w of a rational function of s and z has the same value on both sides of a line c. For the entire existing boundary consists of one piece. In an integration along the boundary starting on one side of c and ending on the other, the integral $\int dw$ is taken over each element of the earlier cuts twice, in opposite directions. Such a function is accordingly continuous throughout T outside the lines a and b. The surface resulting from cutting up the surface T with these lines may be denoted by T''.

20.

Now let $w_1, w_2, \ldots, w_p$ be p such functions, independent of each other, and let the modulus of periodicity of w_μ at the transverse cut a_ν be $A_\mu^{(\nu)}$ and at the transverse cut b_ν be $B_\mu^{(\nu)}$. The integral $\int w_\mu dw_{\mu'}$ taken around the boundary of the surface T'' in a positive direction, is 0, since the integrand is everywhere finite. In the integration, each of the lines a and b is run through twice, once in a positive direction and once in a negative direction. Throughout the integration, where these lines serve as the boundary of the domain traversed positively, we denote by w_μ^+ the value of w_μ on the positive side of the path and by w_μ^- the value on the negative side. Thus the integral is equal to the sum of all the integrals $\int (w_\mu^+ - w_\mu^-) dw_{\mu'}$ taken over the lines a and b. The lines b lead from the positive to the negative sides of the lines a, and so the lines a lead from the negative to the positive sides of the lines b. The integral along the line a_ν is therefore

$$\int A_\mu^{(\nu)} dw_{\mu'} = A_\mu^{(\nu)} \int dw_{\mu'} = A_\mu^{(\nu)} B_{\mu'}^{(\nu)},$$

and the integral along the line b_ν is

$$\int B_\mu^{(\nu)} dw_{\mu'} = -B_\mu^{(\nu)} A_{\mu'}^{(\nu)}.$$

Hence the integral $\int w_\mu dw_{\mu'}$ around the boundary of T'' in a positive direction is

$$\sum_\nu \left(A_\mu^{(\nu)} B_{\mu'}^{(\nu)} - B_\mu^{(\nu)} A_{\mu'}^{(\nu)} \right)$$

and this sum is consequently zero. This equation is valid for every pair of the functions $w_1, w_2, \ldots, w_p$ and thus yields $p(p-1)/2$ relations between the moduli of periodicity.

If we take for the functions w the corresponding functions u, or choose them so that $A_\mu^{(\nu)}$ is zero whenever $\nu \neq \mu$, and $A_\nu^{(\nu)} = \pi i$, then the relations become $B_{\mu'}^{(\mu)} \pi i - B_\mu^{(\mu')} \pi i = 0$ or $a_{\mu,\mu'} = a_{\mu',\mu}$.

21.

It still remains to show that the numbers a possess the second property which we found above to be necessary.

We set $w = \mu + \nu i$ and suppose that the modulus of periodicity of this function at the cut a_ν is $A^{(\nu)} = \alpha_\nu + \gamma_\nu i$ and at the cut b_ν is $B^{(\nu)} = \beta_\nu + \delta_\nu i$. The integral

$$\int \left\{ \left(\frac{\partial \mu}{\partial x}\right)^2 + \left(\frac{\partial \mu}{\partial y}\right)^2 \right\} dT$$

or

$$\int \left(\frac{\partial \mu}{\partial x}\frac{\partial \nu}{\partial y} - \frac{\partial \mu}{\partial y}\frac{\partial \nu}{\partial x} \right) dT,^{6}$$

taken over the surface T'', is equal to the contour integral $\int \mu d\nu$ taken over the boundary of T'' in a positive direction, and therefore is equal to the sum of the integrals $\int (\mu^+ - \mu^-) d\nu$ taken over the lines a and b. The integral over the line a_ν is $\alpha_\nu \int d\nu = \alpha_\nu \delta_\nu$ and the integral along b_ν is $\beta_\nu \int d\nu = -\beta_\nu \gamma_\nu$. Thus

$$\int \left(\left(\frac{\partial \mu}{\partial x}\right)^2 + \left(\frac{\partial \mu}{\partial y}\right)^2 \right) dT = \sum_{\nu=1}^{p} (\alpha_\nu \delta_\nu - \beta_\nu \gamma_\nu).$$

Hence the last sum is always positive.

The required property of the numbers a can now be deduced by setting $w = u_1 m_1 + u_2 m_2 + \cdots + u_p m_p$. For we then have $A^{(\nu)} = m_\nu \pi i$, $B^{(\nu)} = \sum_\nu a_{\mu,\nu} m_\mu$. Consequently α_ν is always zero and

$$\int \left(\left(\frac{\partial \mu}{\partial x}\right)^2 + \left(\frac{\partial \mu}{\partial y}\right)^2 \right) dT = -\sum \beta_\nu \gamma_\nu = -\pi \sum m_\nu \beta_\nu,$$

which is equal to the real part of $-\pi \sum_{\mu,\nu} a_{\mu,\nu} m_\mu m_\nu$, so that the latter is positive for all real values of the m.

[6]This integral represents the area of the surface in the w-plane occupied by the totality of values assumed by w in T''.

22.

If, in the θ-series (1) of Section 17, we take the number $a_{\mu,\mu'}$ to be the modulus of periodicity of the function u_μ at the cut $b_{\mu'}$ and, with $e_1, e_2, \ldots, e_p$ denoting arbitrary constants, take v_μ to be $u_\mu - e_\mu$, we obtain the well-defined single-valued function of z throughout T,

$$\theta(u_1 - e_1, u_2 - e_2, \ldots, u_p - e_p),$$

which is finite and continuous everywhere except on the lines b. On the positive side of the line b_ν the function is $\exp(-2(u_\nu - e_\nu))$ times greater than on the negative side, if one regards the functions u as having on the line b the arithmetic mean of the values on the two edges. The number of points of T', or pairs of values of s and z, for which the function becomes infinitely small of first order, can be found by considering the contour integral $\int d\log\theta$ taken in a positive direction around the boundary of T'. For this integral is equal to the number of such points multiplied by $2\pi i$. On the other hand, this integral is also equal to the sum of the integrals

$$\int (d\log\theta^+ - d\log\theta^-)$$

taken over all the lines a, b, c. The integrals over the lines a and c are all 0, while the integral along b_ν is $-2\int du_\nu = 2\pi i$ and thus the sum is $2\pi i p$. It follows that the function θ becomes infinitely small of first order at p points of the surface T', which we may denote by $\eta_1, \eta_2, \ldots, \eta_p$.

The function $\log\theta$ increases by $2\pi i$ when the point (s, z) makes a circuit in a positive direction around one of these points and by $-2\pi i$ when a circuit is made in a positive direction around the pair of cuts a_ν and b_ν. In order to define the function $\log\theta$ uniquely throughout the domain, we make a cut in the interior of the domain from each of the points η; the cut ℓ_ν from η_ν going to a_ν and b_ν, indeed to their common starting- and ending-points. We take the function continuous throughout the surface T^* that we obtain. The function then has a value on the positive side of the lines ℓ_ν, a_ν, b_ν exceeding that on the negative side by the amounts $-2\pi i$, $g_\nu 2\pi i$, $-2(u_\nu - e_\nu) - h_\nu 2\pi i$, respectively, where g_ν and h_ν are integers.

The positions of the points η and the values of the numbers g and h depend on the constants e, and the dependence can be determined more closely in the following way. The integral $\int \log\theta\, du_\mu$, taken around the boundary of T^* in the positive sense, vanishes because $\log\theta$ remains continuous throughout T^*. This integral is, however, equal to the sums of the integrals $\int(\log\theta^+ - \log\theta^-)du_\mu$ taken over all the cuts ℓ, a, b, and c, and has the value, when the value of u_μ at the point η_ν is denoted by $\alpha_\mu^{(\nu)}$,

$$2\pi i\left(\sum_\nu \alpha_\mu^{(\nu)} + h_\mu \pi i + \sum_\nu g_\nu a_{\nu,\mu} - e_\mu + k_\mu\right)$$

in which k_μ depends neither on the position of the points η nor on the e, g, h. This expression must therefore be 0.

The number k_μ depends on the choice of the function u_μ, which is defined only to within an additive constant by the condition that the modulus of periodicity must be πi at the cut a_μ and 0 at the other cuts a. If we choose another function for u_μ greater by a constant c_μ and at the same time increase e_μ by c_μ, the function θ and consequently the points η and numbers g and h would be unchanged. The value of the new function u_μ at the point η_ν is, however, $\alpha_\mu^{(\nu)} + c_\mu$. Hence k_μ becomes $k_\mu - (p-1)c_\mu$, which vanishes if we take $c_\mu = \frac{k_\mu}{p-1}$.

We can therefore, as we shall do in what follows, determine the additive constants in the functions u, or the initial values in the expressions of these functions as integrals, so that by substituting $u_\mu - \sum \alpha_\mu^{(\nu)}$ for v_μ in $\log\theta(v_1, \ldots, v_p)$, a function is obtained which becomes logarithmically infinite at the points η, and, extended in a continuous fashion throughout T^*, has values on the positive edges of the lines ℓ_ν, a_ν, b_ν exceeding the corresponding values on the negative edges by $-2\pi i, 0, -2\left(u_\nu - \sum_1^p \alpha_\nu^{(\mu)}\right)$ respectively. Later we shall present a method of finding these initial values that is easier than using the above integral expression for k_μ.

23.

If we set

$$(u_1, u_2, \ldots, u_p) \equiv (\alpha_1^{(p)}, \alpha_2^{(p)}, \ldots, \alpha_p^{(p)})$$

with respect to the $2p$ systems of moduli of the functions u (Section 15), so that

$$(v_1, v_2, \ldots, v_p) \equiv \left(-\sum_1^{p-1} \alpha_1^{(\nu)}, -\sum_1^{p-1} \alpha_2^{(\nu)}, \ldots, -\sum_1^{p-1} \alpha_p^{(\nu)}\right),$$

then $\theta = 0$. Conversely if $\theta = 0$ for $v_\mu = r_\mu$, then $(r_1, r_2, \ldots, r_p)$ is congruent to a system of numbers of the form

$$\left(-\sum_1^{p-1} \alpha_1^{(\nu)}, -\sum_1^{p-1} \alpha_2^{(\nu)}, \ldots, -\sum_1^{p-1} \alpha_p^{(\nu)}\right).$$

For if we set $v_\mu = u_\mu - \alpha_\mu^{(p)} + r_\mu$, with η_p being chosen arbitrarily, the function θ is infinitely small of first order at η_p, and at $p-1$ other points. Denote these by $\eta_1, \eta_2, \ldots, \eta_{p-1}$, then we have

$$\left(-\sum_1^{p-1} \alpha_1^{(\nu)}, -\sum_1^{p-1} \alpha_2^{(\nu)}, \ldots, -\sum_1^{p-1} \alpha_p^{(\nu)}\right) \equiv (r_1, r_2, \ldots, r_p).^7$$

The function θ remains unchanged if all its arguments v are replaced by their negatives. For the sum of the series is not affected when the signs of the indices m are replaced by their opposites, since the $-m_\nu$ run through the same set of values as m_ν, and $\theta(v_1, v_2, \ldots, v_p)$ becomes $\theta(-v_1, -v_2, \ldots, -v_p)$.

If the points $\eta_1, \eta_2, \ldots, \eta_{p-1}$ are now chosen arbitrarily, then

$$\theta\left(-\sum_1^{p-1} \alpha_1^{(\nu)}, \ldots, -\sum_1^{p-1} \alpha_p^{(\nu)}\right) = 0$$

and as this function is even (as just mentioned) it follows that

$$\theta\left(\sum_1^{p-1} \alpha_1^{(\nu)}, \ldots, \sum_1^{p-1} \alpha_p^{(\nu)}\right) = 0.$$

The $p-1$ points $\eta_p, \eta_{p+1}, \ldots, \eta_{2p-2}$ can therefore be determined so that

$$\left(\sum_1^{p-1} \alpha_1^{(\nu)}, \ldots, \sum_1^{p-1} \alpha_p^{(\nu)}\right) \equiv \left(-\sum_p^{2p-2} \alpha_1^{(\nu)}, \ldots, -\sum_p^{2p-2} \alpha_p^{(\nu)}\right)$$

and therefore

$$\left(\sum_1^{2p-2} \alpha_1^{(\nu)}, \ldots, \sum_1^{2p-2} \alpha_p^{(\nu)}\right) \equiv (0, \ldots, 0).$$

[7] See, in this connection, paper **XI**. W.

The position of the last $p-1$ points depends on that of the first $p-1$ in such a manner that for arbitrary continuous variations, $\sum_1^{2p-2} d\alpha_\pi^{(\nu)} = 0$ for $\pi = 1, \ldots, p$. Therefore (Section 16) the points η are $2p-2$ points for which one dw is infinitely small of second order; in other words, denoting by (σ_ν, ζ_ν) the value of the pair (s, z) at the point η_ν, the pairs of values $(\sigma_1, \zeta_1), (\sigma_2, \zeta_2), \ldots, (\sigma_{2p-2}, \zeta_{2p-2})$ are associated through the equation $\phi = 0$ (Section 16).

With this choice of the initial values for the integrals u, we thus have

$$\left(\sum_1^{2p-2} u_1^{(\nu)}, \ldots, \sum_1^{2p-2} u_p^{(\nu)} \right) \equiv (0, \ldots, 0)$$

where the summation is over all the common roots $(\gamma_\rho, \delta_\rho)$ (Section 6) of the equation $F = 0$ and the equation $c_1\phi_1 + c_2\phi_2 + \cdots + c_p\phi_p = 0$; the constants c are arbitrary.

If $\epsilon_1, \epsilon_2, \ldots, \epsilon_m$ are points for which a rational function ξ of s and z, that becomes infinite of first order m times, has the same value, and if $u_\pi^{(\mu)}, s_\mu, z_\mu$ denote the values of u_π, s, z at ϵ_μ, then (Section 15)

$$\left(\sum_1^m u_1^{(\mu)}, \sum_1^m u_2^{(\mu)}, \ldots, \sum_1^m u_p^{(\mu)} \right)$$

is congruent to a constant, that is, a system $(b_1, b_2, \ldots, b_p)$ independent of ξ. It is then possible, for an arbitrarily given position of one point ϵ, to determine the position of the others so that

$$\left(\sum_1^m u_1^{(\mu)}, \ldots, \sum_1^m u_p^{(\mu)} \right) \equiv (b_1, \ldots, b_p).$$

It is therefore possible, when $m = p$, to bring $(u_1 - b_1, \ldots, u_p - b_p)$, and when $m < p$, to bring

$$\left(u_1 - \sum_1^{p-m} \alpha_1^{(\nu)} - b_1, \ldots, u_p - \sum_1^{p-m} \alpha_p^{(\nu)} - b_p \right)$$

for every position of the point (s, z) and the $p - m$ points η, into the form

$$\left(-\sum_1^{p-1} \alpha_1^{(\nu)}, \ldots, -\sum_1^{p-1} \alpha_p^{(\nu)} \right)$$

by allowing one of the points ϵ to coincide with (s, z). It follows that

$$\theta\left(u_1 - \sum_1^{p-m} \alpha_1^{(\nu)} - b_1, \ldots, u_p - \sum_1^{p-m} \alpha_p^{(\nu)} - b_p\right) = 0$$

for all values of the pair (s, z) and the $p - m$ pairs (σ_ν, ζ_ν).

24.

It follows from the investigations of Section 22 as a corollary, that a given system $(e_1, \ldots, e_p)$ is always congruent to one and only one system of the form

$$\left(\sum_1^p \alpha_1^{(\nu)}, \ldots, \sum_1^p \alpha_p^{(\nu)}\right)$$

provided that the function $\theta(u_1 - e_1, \ldots, u_p - e_p)$ does not vanish identically. For the points η must then be the p points for which this function becomes zero. If, on the other hand, $\theta(u_1^{(p)} - e_1, \ldots, u_p^{(p)} - e_p)$ vanishes for every value of (s_p, z_p), then by Section 23 one can set

$$(u_1^{(p)} - e_1, \ldots, u_p^{(p)} - e_p) \equiv \left(-\sum_1^{p-1} u_1^{(\nu)}, \ldots, -\sum_1^{p-1} u_p^{(\nu)}\right)$$

and therefore for every pair (s_p, z_p) the pairs $(s_1, z_1), \ldots, (s_{p-1}, z_{p-1})$ can be determined so that

$$\left(\sum_1^p u_1^{(\nu)}, \ldots, \sum_1^p u_p^{(\nu)}\right) \equiv (e_1, \ldots, e_p).$$

Consequently, when (s_p, z_p) varies continuously,

$$\sum_1^p du_\pi^{(\nu)} = 0 \quad \text{for } \pi = 1, 2, \ldots, p.$$

The p pairs (s_ν, z_ν) are therefore p roots different from the pairs $(\gamma_\rho, \delta_\rho)$, of an equation $\phi = 0$ whose coefficients vary in such a way that the remaining $p - 2$ roots stay constant. If we denote the values of u_π for these $p - 2$ pairs of values of s and z by $u_\pi^{(p+1)}, u_\pi^{(p+2)}, \ldots, u_\pi^{(2p-2)}$, then

$$\left(\sum_1^{2p-2} u_1^{(\nu)}, \ldots, \sum_1^{2p-2} u_p^{(\nu)}\right) \equiv (0, \ldots, 0)$$

and consequently

$$(e_1,\ldots,e_p)\equiv\left(-\sum_{p+1}^{2p-2}u_1^{(\nu)},\ldots,-\sum_{p+1}^{2p-2}u_p^{(\nu)}\right).$$

Conversely, if this congruence holds,

$$\theta(u_1^{(p)}-e_1,\ldots,u_p^{(p)}-e_p)=\theta\left(\sum_{p}^{2p-2}u_1^{(\nu)},\ldots,\sum_{p}^{2p-2}u_p^{(\nu)}\right)=0.$$

An arbitrarily given system $(e_1,\ldots,e_p)$ *is therefore congruent to only one system of the form*

$$\left(\sum_1^p\alpha_1^{(\nu)},\ldots,\sum_1^p\alpha_p^{(\nu)}\right)$$

if it is not congruent to any system of the form

$$\left(-\sum_1^{p-2}\alpha_1^{(\nu)},\ldots,-\sum_1^{p-2}\alpha_p^{(\nu)}\right)$$

and is otherwise congruent to infinitely many.

Since

$$\theta\left(u_1-\sum_1^p\alpha_1^{(\mu)},\ldots,u_p-\sum_1^p\alpha_p^{(\mu)}\right)=\theta\left(\sum_1^p\alpha_1^{(\mu)}-u_1,\ldots,\sum_1^p\alpha_p^{(\mu)}-u_p\right),$$

θ is an exactly similar function of each of the p pairs (σ_μ,ζ_μ) as it is of the pair (s,z). This function of (σ_μ,ζ_μ) vanishes for the pair of values (s,z) and for the other $p-1$ pairs (σ,ζ) associated through the equation $\phi=0$. For, if we denote the value of u_π at these points by $\beta_\pi^{(1)},\beta_\pi^{(2)},\ldots,\beta_\pi^{(p-1)}$, then

$$\left(\sum_1^p\alpha_1^{(\mu)},\ldots,\sum_1^p\alpha_p^{(\mu)}\right)\equiv\left(\alpha_1^{(\mu)}-\sum_1^{p-1}\beta_1^{(\nu)},\ldots,\alpha_p^{(\mu)}-\sum_1^{p-1}\beta_p^{(\nu)}\right)$$

and consequently $\theta=0$, when η_μ coincides with one of these points or with the point (s,z).

25.

The properties of the function θ developed above yield the expression of $\log\theta$ in terms of integrals of algebraic functions of $(s,z),(\sigma_1,\zeta_1),\ldots,(\sigma_p,\zeta_p)$.

The quantity

$$\log\theta\left(u_1^{(2)}-\sum_1^p\alpha_1^{(\mu)},\ldots\right)-\log\theta\left(u_1^{(1)}-\sum_1^p\alpha_1^{(\mu)},\ldots\right)$$

regarded as a function of (σ_μ,ζ_μ), is a function of the position of the point η_μ, which becomes discontinuous at the point ϵ_1, in the same way as $-\log(\zeta_\mu-z_1)$ and at the point ϵ_2, in the same way as $\log(\zeta_\mu-z_2)$. On the positive edge of a line joining ϵ_1 to ϵ_2, this function is greater by $2\pi i$, and on the positive edge of the line b_ν by $2(u_\nu^{(1)}-u_\nu^{(2)})$, than on the negative edge. Outside the lines b and the line joining ϵ_1 to ϵ_2, the function is continuous.

Let us now denote by $\tilde{\omega}^{(\mu)}(\epsilon_1,\epsilon_2)$ any function of (σ_μ,ζ_μ) which—except on the lines b—is discontinuous in a similar fashion and whose values on the opposing edges of these lines differ by the same constants. This function differs from the above function by an amount independent of (σ_μ,ζ_μ) (Section 3). Consequently, the above function differs from $\sum_1^p\tilde{\omega}^{(\mu)}(\epsilon_1,\epsilon_2)$ by an amount independent of all the quantities (σ,ζ), and therefore depending only on (s_1,z_1) and (s_2,z_2). Now $\tilde{\omega}^{(\mu)}(\epsilon_1,\epsilon_2)$ represents the value of a function $\tilde{\omega}(\epsilon_1,\epsilon_2)$ of Section 4 for $(s,z)=(\sigma_\mu,\zeta_\mu)$, whose moduli of periodicity on the cuts a are 0. If a constant c is added to this function, the sum $\sum_1^p\tilde{\omega}^{(\mu)}(\epsilon_1,\epsilon_2)$ is increased by pc. The additive constant in the function $\tilde{\omega}(\epsilon_1,\epsilon_2)$, or the initial value in the integral of the third kind representing this function, may therefore be taken below so that

$$\log\theta^{(2)}-\log\theta^{(1)}=\sum_1^p\tilde{\omega}^{(\mu)}(\epsilon_1,\epsilon_2).$$

Since θ depends on each pair of values (σ,ζ) in the same way as on (s,z), the variation undergone by $\log\theta$ when any of the pairs $(s,z),(\sigma_1,\zeta_1),\ldots,(\sigma_p,\zeta_p)$ undergoes a finite variation while the others remain unchanged, can be expressed as a sum of functions $\tilde{\omega}$.

It clearly follows that, by making successive changes in the individual pairs of values $(s,z),(\sigma_1,\zeta_1),\ldots,(\sigma_p,\zeta_p)$, we can always express $\log\theta$ as a sum of functions $\tilde{\omega}$ and

$$\log \theta(0, 0, \ldots, 0)$$

or the value of $\log \theta$ for some arbitrary system of values. The determination of $\log \theta(0, 0, \ldots, 0)$ as a function of the $3p-3$ moduli of the system of rational functions of s and z (Section 12) requires considerations analogous to those used by Jacobi for the determination of $\Theta(0)$ in his work on elliptic functions. We can arrive at the desired result using the equations

$$4\,\frac{\partial \theta}{\partial a_{\mu,\mu}} = \frac{\partial^2 \theta}{\partial v_\mu^2} \quad \text{and} \quad 2\,\frac{\partial \theta}{\partial a_{\mu,\mu'}} = \frac{\partial^2 \theta}{\partial v_\mu \partial v_{\mu'}} \quad (\mu \neq \mu')$$

to express the partial derivatives of $\log \theta$ with respect to the quantities a in the expression

$$d \log \theta = \sum \frac{\partial \log \theta}{\partial a_{\mu,\mu'}}\, da_{\mu,\mu'}$$

via integrals of algebraic functions. The execution of this calculation appears, however, to need a more extensive theory of those functions satisfying a linear differential equation with algebraic coefficients. I intend to provide this shortly, using the the principles applied here.

If (s_2, z_2) differs infinitely little from (s_1, z_1), then $\tilde{\omega}(\epsilon_1, \epsilon_2)$ becomes $dz_1 t(\epsilon_1)$, where $t(\epsilon_1)$ is an integral of the second kind of a rational function of s and z, which is discontinuous at ϵ_1 in the same way as $\frac{1}{z-z_1}$ and has moduli of periodicity 0 at the cuts a. It follows that the modulus of periodicity of such an integral at the cut b_ν is equal to $2\,\frac{du_\nu^{(1)}}{dz_1}$, and the constant of integration can be determined so that the sum of the values of $t(\epsilon_1)$ for the p pairs $(\sigma_1, \zeta_1), \ldots, (\sigma_p, \zeta_p)$ is equal to $\frac{\partial \log \theta^{(1)}}{\partial z_1}$. It then follows that $\frac{\partial \log \theta^{(1)}}{\partial \zeta_\mu}$ is equal to the sum of the values of $t(\eta_\mu)$ for the $p-1$ pairs different from (σ_μ, ζ_μ) associated through the equation $\phi = 0$ and the pair (s, z). We obtain for

$$\frac{\partial \log \theta^{(1)}}{\partial z_1}\, dz_1 + \sum_1^p \frac{\partial \log \theta^{(1)}}{\partial \zeta_\mu}\, d\zeta_\mu = d \log \theta^{(1)},$$

an expression that Weierstrass has given for the case where s is only a two-valued function of z (Crelle's Journal, vol. 47, p. 300, formula 35).

The properties of $\tilde{\omega}(\epsilon_1, \epsilon_2)$ and $t(\epsilon_1)$ as functions of (s_1, z_1) and (s_2, z_2) can be deduced from the equations

$$\tilde{\omega}(\epsilon_1, \epsilon_2) = \frac{1}{p}\left(\log \theta(u_1^{(2)} - pu_1, \ldots) - \log \theta(u_1^{(1)} - pu_1, \ldots)\right)$$

and

$$t(\epsilon_1) = \frac{1}{p}\,\frac{\partial \log\theta(u_1^{(1)} - pu_1, \ldots)}{\partial z_1}$$

which are special cases of the preceding expressions for $\log\theta^{(2)} - \log\theta^{(1)}$ and $\frac{\partial \log\theta^{(1)}}{\partial z_1}$.

26.

We now consider the problem of expressing an algebraic function of z as the quotient of two functions, each a product of the same number of functions $\theta(u_1 - \epsilon_1, \ldots)$ and powers of the quantities e^u.

Any such expression undergoes multiplication by constants whenever (s, z) crosses a transverse cut, and these constants must be roots of unity if they are to depend algebraically on z and consequently assume only a finite number of different values for the same z when z undergoes continuous variation. If all these factors are μth roots of unity, the μth power of the expression sought will be a single-valued and therefore rational function of s and z.

Conversely, it can easily be shown that every algebraic function r of z that can be prolonged continuously throughout the whole surface T', is single-valued and undergoes multiplication by a constant factor whenever a transverse cut is crossed, can be expressed in a multiplicity of different ways as the quotient of two products of θ-functions and powers of the e^u. Let us denote by β_μ a value of u_μ for $r = \infty$ and by γ_μ a value of u_μ for $r = 0$, and let us draw from each point at which r becomes infinite of first order a line interior to T', joining it to a point at which r becomes infinitely small of first order. The function $\log r$ is taken to be continuous in T', except on these lines. Accordingly, if $\log r$ is greater by an amount $g_\nu 2\pi i$ on the positive edge of the line b_ν and by an amount $-h_\nu 2\pi i$ on the positive edge of the line a_ν, than on the corresponding negative edges, consideration of the contour integral $\int \log r du_\mu$ yields:

$$\sum \gamma_\mu - \sum \beta_\mu = g_\mu \pi i + \sum_\nu h_\nu a_{\mu,\nu} \qquad (\mu = 1, \ldots, p)$$

where g_ν and h_ν must, in view of the foregoing remarks, be rational numbers, and in the left side of the equation the summation is over points where r is infinitely small or infinitely large of first order. Any points where the order exceeds 1 are regarded as consisting of several such points (Section 2). If all but p of these points are given, then these p points can always be chosen, generally speaking in only one way, so that the $2p$ factors $e^{g_\nu 2\pi i}$ and $e^{-h_\nu 2\pi i}$ have given values (Sections 15 and 24).

Suppose that in the expression

$$\frac{P}{Q}e^{-2\Sigma h_\nu u_\nu}$$

in which P and Q are products of an equal number of functions

$$\theta\left(u_1 - \sum \alpha_1^{(\pi)}, \dots\right)$$

with the same (s, z) and different (σ, ζ), we substitute the pairs of values of s and z, for which r becomes infinite, for (σ, ζ) in the θ-functions of the denominator, and the pairs for which r vanishes for the pairs (σ, ζ) in the θ-functions of the numerator, and take the remaining pairs (σ, ζ) to be the same in the denominator and numerator. Then the logarithm of the resulting expression has the same discontinuities as $\log r$ in the interior of T', and in crossing the lines a and b its value, like that of $\log r$, changes by purely imaginary numbers constant along these lines. It thus differs from $\log r$, by Dirichlet's principle, only by a constant. The expression itself differs from r by a constant factor. It goes without saying that the substitution is admissible only when none of the θ-functions is identically zero for each value of z. This would happen (Section 23) if all the pairs of values for which a single-valued function of (s, z) vanishes were substituted for the pairs of values of (σ, ζ) in one and the same θ-function.

27.

A single-valued or rational function of (s, z) thus cannot be represented as a quotient of two θ-functions multiplied by powers of the e^u. All functions r, however, that have more than one value for the same pair of values of s and z, and become infinite of first order for p or fewer value-pairs, can be represented in this form, and comprise all the algebraic functions of z that can be represented in this way. Apart from a constant factor, each r is obtained once and once only if in

$$\frac{\theta\left(v_1 - g_1\pi i - \sum\limits_\nu h_\nu a_{1,\nu}, \dots\right)}{\theta(v_1, \dots, v_p)} e^{-2\sum\limits_\nu v_\nu h_\nu}$$

we assign rational fractions less than 1 to h_ν and g_ν and set v_ν equal to $u_\nu - \sum_1^p \alpha_\nu^{(\mu)}$.

The resulting expression is an algebraic function of each of the quantities ζ, and the principles developed in the previous section fully suffice to express it as an algebraic function of $z, \zeta_1, \ldots \zeta_p$.

This can be seen as follows. As a function of (s, z), when continuously extended throughout the surface T', it has one definite value everywhere. It becomes infinite of first order for the pairs $(\sigma_1, \zeta_1), \ldots, (\sigma_p, \zeta_p)$ and, in crossing from the positive to the negative side of the lines a_ν and b_ν, acquires factors $e^{h_\nu 2\pi i}$ and $e^{-g_\nu 2\pi i}$ respectively. Every function of (s, z) satisfying these specifications differs from it only by a factor independent of (s, z). Regarded as a function of (σ_μ, ζ_μ), it has a well-defined value when extended throughout the whole surface T', becomes infinite of first order at the point (s, z) and for the other $p - 1$ pairs $(\sigma_1^{(\mu)}, \zeta_1^{(\mu)}), \ldots, (\sigma_{p-1}^{(\mu)}, \zeta_{p-1}^{(\mu)})$ associated through the equation $\phi = 0$, and across a_ν it acquires a factor $e^{-h_\nu 2\pi i}$, across b_ν a factor $e^{g_\nu 2\pi i}$; and every other function of (σ_μ, ζ_μ) satisfying these specifications differs from it only by a factor independent of (σ_μ, ζ_μ). If therefore we define an algebraic function of $z, \zeta_1, \ldots \zeta_p$ in the form

$$f((s, z); (\sigma_1, \zeta_1), \ldots, (\sigma_p, \zeta_p))$$

possessing these properties as a function of each pair, it differs from the above function only by a factor independent $z, \zeta_1, \ldots \zeta_p$. We can write the above function as Af, where A is this factor. To determine the factor, let the pairs in f other than (σ_μ, ζ_μ) be denoted by $(\sigma_1^{(\mu)}, \zeta_1^{(\mu)}), \ldots, (\sigma_{p-1}^{(\mu)}, \zeta_{p-1}^{(\mu)})$ so that f now becomes

$$g((\sigma_\mu, \zeta_\mu); (s, z), (\sigma_1^{(\mu)}, \zeta_1^{(\mu)}), \ldots, (\sigma_{p-1}^{(\mu)}, \zeta_{p-1}^{(\mu)})).$$

Obviously we get the inverse value of the function to be represented, and thus an expression equal to $\frac{1}{Af}$, if we substitute in Ag the pair (s, z) for (σ_μ, ζ_μ) and substitute, for the pairs (s, z), $(\sigma_1^{(\mu)}, \zeta_1^{(\mu)}), \ldots, (\sigma_{p-1}^{(\mu)}, \zeta_{p-1}^{(\mu)})$, the pairs (s, z) for which the function to be represented is 0, and thus $f = 0$. This yields A^2, and thus A up to sign. The sign can be found by direct treatment of the θ-series in the expression to be represented.

VII.

On the number of primes less than a given magnitude.

(*Monatsberichte der Berliner Akademie*, November 1859.)

I think that I can best express my thanks for the honor which the Academy has conferred in admitting me as a correspondent, by making prompt use of the privilege now afforded me to report on an investigation into the frequency of prime numbers. This is a subject to which Gauss and Dirichlet have devoted much effort, and may therefore be considered to be not unworthy for such a communication.

I take as my starting point in this investigation the remark made by Euler, that

$$\prod \frac{1}{1-\frac{1}{p^s}} = \sum \frac{1}{n^s},$$

where p runs through all prime numbers, and n through all natural numbers. I shall denote by $\zeta(s)$ the function of the complex variable s, which is represented by each of these two expressions when they converge. Both expressions converge only if the real part of s exceeds 1; however, it is easy to find an expression for the function which is always valid.

The equation

$$\int_0^\infty e^{-nx} x^{s-1} dx = \frac{\Pi(s-1)}{n^s}$$

gives immediately

$$\Pi(s-1)\zeta(s) = \int_0^\infty \frac{x^{s-1} dx}{e^x - 1}.$$

If we now consider the contour integral

$$\int \frac{(-x)^{s-1} dx}{e^x - 1}$$

taken from $+\infty$ to $+\infty$ in a positive sense over a closed path, which includes in its interior the point 0 but no other point of discontinuity of the integrand, then it is easily seen to be equal to

$$(e^{-\pi s i} - e^{\pi s i}) \int_0^\infty \frac{x^{s-1} dx}{e^x - 1}$$

provided that, in the multi-valued function $(-x)^{s-1} = e^{(s-1)\log(-x)}$, the logarithm is determined in such a way that it is real when x is a negative real number.

It follows that

$$2\sin\pi s\ \Pi(s-1)\zeta(s) = i\int_{\infty}^{\infty}\frac{(-x)^{s-1}dx}{e^x-1}$$

when the integral is interpreted as above.

This equation now gives the value of $\zeta(s)$ for every complex s and shows that it is a single-valued function whose value is finite for every finite s with the exception of 1. It also shows that $\zeta(s)$ vanishes when s is an even negative integer.

If the real part of s is negative, the integral can also be evaluated by being taken over a path which, instead of surrounding in a positive direction the domain described earlier, surrounds in a negative direction all other complex numbers; because the integral is then infinitely small for all s of infinitely large modulus. In the interior of this domain, the integrand is discontinuous only when x is equal to an integral multiple of $\pm 2\pi i$, and the integral is therefore equal to the sum of the integrals taken in a negative sense around each of these points. The value of the integral around the point $n2\pi i$ is $(-n2\pi i)^{s-1}(-2\pi i)$. Hence

$$2\sin\pi s\ \Pi(s-1)\zeta(s) = (2\pi)^s\sum n^{s-1}((-i)^{s-1}+i^{s-1}),$$

which gives a relation between $\zeta(s)$ and $\zeta(1-s)$. This can also be expressed, by known properties of the function Π, as follows:

$$\Pi\left(\frac{s}{2}-1\right)\pi^{-\frac{s}{2}}\zeta(s)$$

remains unchanged when s is replaced by $1-s$.

This property of the function led me to introduce the integral $\Pi\left(\frac{s}{2}-1\right)$ instead of $\Pi(s-1)$ as a multiplier of the general term of the series $\sum\frac{1}{n^s}$ and thus to obtain a very convenient expression for the function $\zeta(s)$. In fact,

$$\frac{1}{n^s}\Pi\left(\frac{s}{2}-1\right)\pi^{-\frac{s}{2}} = \int_0^{\infty} e^{-n^2\pi x}x^{\frac{s}{2}-1}dx.$$

Therefore, if we set

$$\sum_1^{\infty} e^{-n^2\pi x} = \psi(x),$$

then

$$\Pi\left(\frac{s}{2}-1\right)\pi^{-\frac{s}{2}}\zeta(s)=\int_0^\infty \psi(x)x^{\frac{s}{2}-1}dx.$$

Since

$$2\psi(x)+1=x^{-1/2}\left(2\psi\left(\frac{1}{x}\right)+1\right)$$

(Jacobi, Fund., p. 184)[1] we have

$$\begin{aligned}\Pi\left(\frac{s}{2}-1\right)\pi^{-\frac{s}{2}}\zeta(s)&=\int_1^\infty \psi(x)x^{(s/2)-1}dx+\int_0^1\psi\left(\frac{1}{x}\right)x^{(s-3)/2}dx\\&\quad+\frac{1}{2}\int_0^1(x^{(s-3)/2}-x^{(s/2)-1})dx\\&=\frac{1}{s(s-1)}+\int_1^\infty\psi(x)(x^{(s/2)-1}+x^{-(1+s)/2})dx.\end{aligned}$$

I now write $s=\frac{1}{2}+ti$ and

$$\Pi\left(\frac{s}{2}\right)(s-1)\pi^{-\frac{s}{2}}\zeta(s)=\xi(t)$$

so that

$$\xi(t)=\frac{1}{2}-\left(t^2+\frac{1}{4}\right)\int_1^\infty\psi(x)x^{-\frac{3}{4}}\cos\left(\frac{1}{2}t\log x\right)dx$$

and also

$$\xi(t)=4\int_1^\infty\frac{d(x^{3/2}\psi'(x))}{dx}x^{-1/4}\cos\left(\frac{1}{2}t\log x\right)dx.$$

This function is finite for all finite values of t and can be expanded in a very rapidly convergent series in powers of t^2. Since $\log\zeta(s)=-\sum\log(1-p^{-s})$ remains finite for any value of s whose real part exceeds 1, and since the same is true of the logarithms of the other factors of $\xi(t)$, it is clear that $\xi(t)$ can vanish only if the imaginary part of t lies between $\frac{1}{2}i$ and $-\frac{1}{2}i$. The number of roots of the equation $\xi(t)=0$, whose real parts lie between 0 and T, is about

$$\frac{T}{2\pi}\log\frac{T}{2\pi}-\frac{T}{2\pi}$$

because the integral $\int d\log\xi(t)$ taken in a positive sense around a contour which includes in its interior all the values of t whose imaginary parts lie between $\frac{1}{2}i$ and $-\frac{1}{2}i$ and whose real parts lie between 0 and T (ignoring a small

[1]Jacobi's collected works vol. I, p. 235. W.

fractional term whose order of magnitude is $\frac{1}{T}$) has the value $\left(T \log \frac{T}{2\pi} - T\right) i$; this integral is however equal to $2\pi i$ times the number of roots of the equation $\xi(t) = 0$ in this region. Now we find in fact that there are about this number of real roots in this domain, and it is very likely that all the roots are real. A rigorous proof of this would certainly be desirable; however after a few brief and fruitless attempts to find one, I have put this on one side for the time being, because it did not seem to be essential to the immediate object of my investigation.

If we denote by α an arbitrary root of the equation $\xi(\alpha) = 0$, then $\log \xi(t)$ can be expressed by

$$\sum \log\left(1 - \frac{t^2}{\alpha^2}\right) + \log \xi(0).$$

Since the density of the roots increases with t only as fast as $\log \frac{t}{2\pi}$, this expression converges and is of order $t \log t$ as t tends to infinity. The expression thus differs from $\log \xi(t)$ by an amount which is a continuous function of t^2 and which remains finite and continuous for all finite t and tends to zero for infinite t after division by t^2. The difference is therefore a constant whose value can be found by setting $t = 0$.

With the aid of these results, the number of primes less than x can now be determined.

Let $F(x)$ denote this number when x is not a prime number, but this number plus $\frac{1}{2}$ when x is a prime, so that whenever $F(x)$ has a jump in value,

$$F(x) = \frac{F(x+0) + F(x-0)}{2}.$$

In the series

$$\log \zeta(s) = -\sum \log(1 - p^{-s}) = \sum p^{-s} + \frac{1}{2} \sum p^{-2s} + \frac{1}{3} \sum p^{-3s} + \cdots$$

we now replace

$$p^{-s} \text{ by } s \int_p^\infty x^{-s-1} dx,\ p^{-2s} \text{ by } s \int_{p^2}^\infty x^{-s-1} dx, \ldots$$

We obtain

$$\frac{\log \zeta(s)}{s} = \int_1^\infty f(x) x^{-s-1} dx$$

where $f(x)$ denotes

$$F(x) + \frac{1}{2}F(x^{\frac{1}{2}}) + \frac{1}{3}F(x^{\frac{1}{3}}) + \cdots$$

This equation is valid for every complex value $a + bi$ of s, when $a > 1$. However, if the equation

$$g(s) = \int_0^\infty h(x)x^{-s} d\log x$$

holds in this domain, then the function h can be expressed in terms of the function g using Fourier's theorem. If $h(x)$ is real and

$$g(a + bi) = g_1(b) + ig_2(b),$$

the equation splits into the two following equations:

$$g_1(b) = \int_0^\infty h(x)x^{-a}\cos(b\log x)d\log x,$$

$$ig_2(b) = -i\int_0^\infty h(x)x^{-a}\sin(b\log x)d\log x.$$

If we now multiply these two equations by

$$(\cos(b\log y) + i\sin(b\log y))db$$

and integrate from $-\infty$ to $+\infty$, the right-hand side of each equation becomes $\pi h(y)y^{-a}$, by Fourier's theorem. Accordingly, after adding the two equations and multiplying by iy^a, we obtain

$$2\pi i h(y) = \int_{a-\infty i}^{a+\infty i} g(s)y^s ds,$$

where the integration is carried out so that the real part of s remains constant.

The integral represents, for every value of y at which the function $h(y)$ jumps in value, the mean of the values on either side of the jump. The function $f(x)$ defined as above possesses this same property, and therefore the equation

$$f(y) = \frac{1}{2\pi i}\int_{a-\infty i}^{a+\infty i} \frac{\log\zeta(s)}{s}\, y^s\, ds$$

holds in full generality.

The expression found earlier for $\log\zeta$,

$$\frac{s}{2}\log\pi - \log(s-1) - \log\Pi\left(\frac{s}{2}\right) + \sum_{\alpha}\log\left(1+\frac{\left(s-\frac{1}{2}\right)^2}{\alpha^2}\right) + \log\xi(0),$$

could now be substituted in this equation. However, the integrals of the individual terms of the resulting expression do not converge when the limits of integration are infinite. It is therefore expedient to begin by transforming the equation by partial integration into

$$f(x) = -\frac{1}{2\pi i}\frac{1}{\log x}\int_{a-\infty i}^{a+\infty i}\frac{d\log\frac{\zeta(s)}{s}}{ds}x^s\,ds.$$

Since

$$-\log\Pi\left(\frac{s}{2}\right) = \lim_{m\to\infty}\left(\sum_{n=1}^{m}\log\left(1+\frac{s}{2n}\right) - \frac{s}{2}\log m\right),$$

and therefore

$$-\frac{d\frac{1}{s}\log\Pi\left(\frac{s}{2}\right)}{ds} = \sum_{1}^{\infty}\frac{d\frac{1}{s}\log\left(1+\frac{s}{2n}\right)}{ds},$$

all the terms of the expression for $f(x)$ with the exception of the term

$$\frac{1}{2\pi i}\frac{1}{\log x}\int_{a-\infty i}^{a+\infty i}\frac{1}{s^2}\log\xi(0)x^s ds = \log\xi(0)$$

assume the form

$$\pm\frac{1}{2\pi i}\int_{a-\infty i}^{a+\infty i}\frac{d\left(\frac{1}{s}\log\left(1-\frac{s}{\beta}\right)\right)}{ds}x^s ds.$$

Now

$$\frac{d\left(\frac{1}{s}\log\left(1-\frac{s}{\beta}\right)\right)}{d\beta} = \frac{1}{(\beta-s)\beta}$$

and, when the real part of s exceeds that of β,

$$-\frac{1}{2\pi i}\int_{a-\infty i}^{a+\infty i}\frac{x^s}{(\beta-s)\beta} = \frac{x^\beta}{\beta} = \int_{\infty}^{x}x^{\beta-1}dx,$$

or

$$= \int_0^x x^{\beta-1} dx,$$

depending on whether the real part of β is negative or positive. Accordingly,

$$\begin{aligned}
\frac{1}{2\pi i}\,\frac{1}{\log x}\int_{a-\infty i}^{a+\infty i} \frac{d\left(\frac{1}{s}\log\left(1-\frac{s}{\beta}\right)\right)}{ds}\, x^s ds
&= -\frac{1}{2\pi i}\int_{a-\infty i}^{a+\infty i} \frac{1}{s}\log\left(1-\frac{s}{\beta}\right) x^s ds \\
&= \int_\infty^x \frac{x^{\beta-1}}{\log x}\, dx + \text{ const. (in the first case)}
\end{aligned}$$

and

$$= \int_0^x \frac{x^{\beta-1}}{\log x}\, dx + \text{ const. (in the second case).}$$

The constant of integration can be determined in the first case by allowing the real part of β to become negative infinite. In the second case, the value of the integral from 0 to x takes on two values depending on whether the path of integration lies above or below the real axis, the values differing by $2\pi i$. In the former case, the integral will become infinitely small when the coefficient of i in β is infinite and positive. In the latter case, the integral is infinitely small when this coefficient is infinite and negative. This shows how the expression $\log\left(1-\frac{s}{\beta}\right)$ on the left side is to be determined so that the constant of integration disappears.

By inserting these values in the expression for $f(x)$, we obtain

$$\begin{aligned}
f(x) = Li(x) - \sum_\alpha \left(Li(x^{\frac{1}{2}+\alpha i}) + Li(x^{\frac{1}{2}-\alpha i})\right) \\
+ \int_x^\infty \frac{1}{x^2-1}\,\frac{dx}{x\log x} + \log\xi(0),
\end{aligned}$$

where the summation $\sum_\alpha$ is over all positive roots (or more precisely all complex roots having a positive real part) of the equation $\xi(\alpha) = 0$, arranged in order of increasing moduli. It can be shown, by a more detailed discussion of the function ξ, that with this ordering the series

$$\sum_{\alpha}\left(Li(x^{\frac{1}{2}+\alpha i})+Li(x^{\frac{1}{2}-\alpha i})\right)\log x$$

converges to the limit of the integral

$$\frac{1}{2\pi i}\int_{a-bi}^{a+bi}\frac{d\,\frac{1}{s}\sum\log\left(1+\frac{\left(s-\frac{1}{2}\right)^2}{\alpha^2}\right)}{ds}\,x^s ds$$

when b tends to infinity. If the order were to be changed, however, the series could converge to an arbitrary real value.

The function $F(x)$ can be found from $f(x)$ by inverting the relation

$$f(x)=\sum\frac{1}{n}\,F(x^{1/n}),$$

which yields

$$F(x)=\sum(-1)^{\mu}\frac{1}{m}\,f(x^{\frac{1}{m}}),$$

in which m runs through all the natural numbers not divisible by any square other than 1, and μ denotes the number of prime factors of m.

If, in the sum $\sum_{\alpha}$, we restrict the summation to a finite number of terms, then the derivative of the expression for $f(x)$ (neglecting a term which decreases very rapidly with increasing x) becomes

$$\frac{1}{\log x}-2\sum_{\alpha}\frac{\cos(\alpha\log x)x^{-\frac{1}{2}}}{\log x},$$

which gives an approximate expression for the density of the primes of magnitude $\leq x$, plus half the density of squares of primes, plus one-third of the density of cubes of primes, and so on.

The well-known approximation $F(x)=Li(x)$ is therefore correct only to within an order of magnitude $x^{1/2}$ and gives rather too large a value. For the non-periodic terms in the expression for $F(x)$ are, excluding those which remain bounded as x increases without limit,

$$\begin{aligned}Li(x)-\frac{1}{2}\,Li(x^{1/2})-\frac{1}{3}\,Li(x^{1/3})-\frac{1}{5}\,Li(x^{1/5})\\+\frac{1}{6}\,Li(x^{1/6})-\frac{1}{7}\,Li(x^{1/7})+\cdots\end{aligned}$$

In fact the comparison between the number of primes less than x and $Li(x)$, undertaken by Gauss and Goldschmidt and taken up to $x = 3,000,000$, has revealed that the number of primes is already less than $Li(x)$ after the first hundred thousand and that the difference, with many fluctuations, gradually increases with x. The increases and decreases of density in the prime numbers due to the periodic terms had already been observed in the counts, but it had escaped notice that it is regulated by a certain law. If a future count is undertaken, it would be interesting to follow up the influence of the individual periodic terms in the expression for the density of the primes. The function $f(x)$ should exhibit a more regular behaviour than $F(x)$, and indeed already substantially coincides on average with $Li(x) + \log \xi(0)$ in the first hundred.

VIII.

On the propagation of planar air waves of finite amplitude.

(*Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, vol. 8, 1860.*)

The differential equations that determine the motion of gases have long been set down. However, their integration has essentially been worked out only for the case when the pressure differences can be treated as infinitely small fractions of the total pressure. Until very recently, one had to be content to take into account only the first power of these fractions. A short while ago, Helmholtz took account of second order terms in the calculation and thereby explained the actual origin of combination tones. However, in the case where the initial motion is entirely in one direction, and velocity and pressure are constant in every plane perpendicular to this direction, the exact differential equations can be completely solved. The previous treatment is entirely sufficient to explain the experimental phenomena already observed. Nevertheless, with the great progress which Helmholtz also made very recently in the experimental treatment of acoustical questions, the results of the present more exact calculation should perhaps give some reference points for experimental research in the not too distant future. This may justify their communication, quite apart from the theoretical interest of the treatment of nonlinear partial differential equations.

Boyle's law may be assumed to give the dependence of the pressure on the density, when the temperature differences due to variations in pressure are compensated so quickly that the temperature of the gas can be considered constant. Probably, however, the thermal exchange is entirely negligible. Hence we must take, as a foundation for this dependence, the law that the pressure of the gas increases in proportion to the density, when heat is neither taken up nor released.

According to the laws of Boyle and Gay-Lussac, denoting by v the volume of a unit mass, p the pressure and T the temperature measured from $-273°$C,

$$\log p + \log v = \log T + \text{ const.}$$

Here we treat T as a function of p and v. We denote the specific heat for constant pressure by c, and the specific heat for constant volume by c', in both cases for unit mass. Then this unit mass, when p and v vary by dp and dv, takes up the quantity of heat

$$c\,\frac{\partial T}{\partial v}\,dv + c'\,\frac{\partial T}{\partial p}\,dp.$$

Since $\frac{\partial \log T}{\partial \log v} = \frac{\partial \log T}{\partial \log p} = 1$, this quantity is

$$T(c\,d\log v + c' d\log p).$$

Thus if no heat absorption takes place, $d\log p = -\frac{c}{c'}\,d\log v$. If we assume, with Poisson, that the ratio $\frac{c}{c'} = k$ of the two specific heats is independent of temperature and pressure, we have

$$\log p = -k\log v + \text{const.}$$

According to recent experiments of Regnault, Joule and W. Thomson, these laws are probably very close to validity for oxygen, nitrogen, hydrogen and their mixtures under all pressures and temperatures treated.

A very close fit has been determined for by Regnault for these gases to the laws of Boyle and Gay-Lussac, and the independence of the specific heat c from temperature and pressure.

For atmospheric air, Regnault obtained:

$c = 0.2377$	between $-30°$ C	and $+\ 10°$C,
$c = 0.2379$	between $+10°$ C	and $+\ 100°$C,
$c = 0.2376$	between $+100°$ C	and $+\ 215°$C.

Likewise, atmospheric pressures between 1 and 10 atmospheres yielded no noticeable difference in specific heat.

From experiments of Regnault and Joule, the Mayer hypothesis adopted by Clausius appears to be very nearly correct for these gases: a gas expanding at constant temperature takes up only the amount of heat required to perform the external work. If the volume of the gas varies by dv while the temperature remains constant, then $d\log p = -d\log v$, the quantity of heat absorbed is $T(c - c')d\log v$, and the work done is pdv. Hence the hypothesis gives, denoting by A the mechanical equivalent of heat,

$$AT(c - c')d\log v = pdv$$

or

$$c - c' = \frac{pv}{AT},$$

which is independent of pressure and temperature.

Accordingly $k = \frac{c}{c'}$ is independent of pressure and temperature. If we take $c = 0.237733$, $A = 424.55$ kilogram meters, according to Joule, and for temperature 0°C, or $T = \frac{100°\mathrm{C}}{0.3655}$, $pv = 7990.267$ according to Regnault, we find that $k = 1.4101$. The velocity of sound in dry air at 0°C amounts to

$$\sqrt{7990.267 \times 9.8088k}$$

meters per second, and with this value of k is found to be 332.440. The two most complete series of experiments of Moll and van Beek, respectively give 332.528 and 331.867 separately, 322.271 taken together; the experiment of Martins and A. Bravais gives 332.37, according to their own calculations.

1.

To begin with, we do not need to make a definite hypothesis on the dependence of pressure on density. Hence we assume that the pressure is $\varphi(\rho)$ for density ρ, and leave the function φ undetermined for the present.

We now introduce rectangular coordinates x, y, z, taking the x-axis in the direction of motion. Denote by ρ the density, p the pressure and u the velocity at coordinate x and time t. Let ω be an element of the plane having coordinate x.

The volume of a right cylinder standing on the element ω, of height dx, is then $\omega\, dx$, and the mass it contains is $\omega\rho\, dx$. The variation of the mass during time element t, or the quantity $\omega\, \frac{\partial\rho}{\partial t}\, dt\, dx$, is determined by the mass flowing into the element, which is found to be $-\omega\, \frac{\partial}{\partial x}\,(\rho u)\, dx\, dt$. Its acceleration is $\frac{\partial u}{\partial t} + u\, \frac{\partial u}{\partial x}$, and the force that drives it in the direction of the positive axis is

$$-\frac{\partial p}{\partial x}\,\omega\, dx = -\varphi'(\rho)\,\frac{\partial\rho}{\partial x}\,\omega\, dx$$

where $\varphi'(\rho)$ denotes the derivative of $\varphi(\rho)$. Hence we have the two differential equations for ρ and u:

$$\frac{\partial\rho}{\partial t} = -\frac{\partial(\rho u)}{\partial x} \quad , \quad \rho\left(\frac{\partial u}{\partial t} + u\,\frac{\partial u}{\partial x}\right) = -\varphi'(\rho)\,\frac{\partial\rho}{\partial x},$$

or

$$\frac{\partial u}{\partial t} + u\,\frac{\partial u}{\partial x} = -\varphi'(\rho)\,\frac{\partial}{\partial x}\,\log\rho$$

and

$$\frac{\partial}{\partial t}\log\rho + u\frac{\partial}{\partial x}\log\rho = -\frac{\partial u}{\partial x}.$$

Multiplying the second equation by $\pm\sqrt{\varphi'(\rho)}$ and adding the first, and writing for brevity

$$(1)\qquad \int\sqrt{\varphi'(\rho)}\,d\,\log\rho = f(\rho),$$

$$(2)\qquad f(\rho)+u = 2r,\quad f(\rho)-u = 2s,$$

these equations take the simpler form

$$(3)\qquad \frac{\partial r}{\partial t} = -\left(u+\sqrt{\varphi'(\rho)}\right)\frac{\partial r}{\partial x}\quad,\quad \frac{\partial s}{\partial t} = -\left(u-\sqrt{\varphi'(\rho)}\right)\frac{\partial s}{\partial x},$$

where u and ρ are defined as functions of r and s via the equation (2). It follows that

$$(4)\qquad dr = \frac{\partial r}{\partial x}\left(dx - \left(u+\sqrt{\varphi'(\rho)}\right)dt\right)$$

$$(5)\qquad ds = \frac{\partial s}{\partial x}\left(dx - \left(u-\sqrt{\varphi'(\rho)}\right)dt\right).$$

Under the hypothesis, always found in physical reality, that $\varphi'(\rho)$ is positive, these equations signify that r remains constant if x and t vary in such a way that $dx = \left(u+\sqrt{\varphi'(\rho)}\right)dt$, and that s remains constant if x and t vary so that $dx = \left(u-\sqrt{\varphi'(\rho)}\right)dt$.

Thus a particular value of r, or of $f(\rho)+u$, moves towards larger values of x with velocity $\sqrt{\varphi'(\rho)}+u$, while a particular value of s, or of $f(\rho)-u$, moves towards smaller values of x with velocity $\sqrt{\varphi'(\rho)}-u$.

A definite value of r will gradually meet with each value of s lying ahead of r, and the velocity of its progress will depend at a given moment on the value of s with which it meets.

2.

The analysis at once offers the means of answering the question of where and when a value r' of r encounters a value s' of s lying ahead of r'; that is, we determine x and t as functions of r and s. Indeed, if we introduce r and s as independent variables in equations (3) of the previous section, these equations become linear differential equations in x and t. Thus they can be solved by known methods. In order to effect the reduction of the differential equation to one linear equation, it is convenient to write equations (4) and (5) of the previous section in the form:

$$(1)\qquad dr = \frac{\partial r}{\partial x}\left\{ d\left(x - \left(u + \sqrt{\varphi'(\rho)}\right)t\right) + \left[dr\left(\frac{d\log\sqrt{\varphi'(\rho)}}{d\log\rho} + 1\right) + ds\left(\frac{d\log\sqrt{\varphi'(\rho)}}{d\log\rho} - 1\right)\right] t \right\},$$

$$(2)\qquad ds = \frac{\partial s}{\partial x}\left\{ d\left(x - \left(u - \sqrt{\varphi'(\rho)}\right)t\right) - \left[ds\left(\frac{d\log\sqrt{\varphi'(\rho)}}{d\log\rho} + 1\right) + dr\left(\frac{d\log\sqrt{\varphi'(\rho)}}{d\log\rho} - 1\right)\right] t \right\}.$$

Considering s and r as independent variables, we obtain for x and t the two linear differential equations:

$$\frac{\partial\left(x - \left(u + \sqrt{\varphi'(\rho)}\right)t\right)}{\partial s} = -t\left(\frac{d\log\sqrt{\varphi'(\rho)}}{d\log\rho} - 1\right),$$

$$\frac{\partial\left(x - \left(u - \sqrt{\varphi'(\rho)}\right)t\right)}{\partial r} = t\left(\frac{d\log\sqrt{\varphi'(\rho)}}{d\log\rho} - 1\right).$$

Consequently,

$$(3)\qquad \left(x - \left(u + \sqrt{\varphi'(\rho)}\right)t\right)dr - \left(x - \left(u - \sqrt{\varphi'(\rho)}\right)t\right)ds$$

is a complete differential, whose integral w satisfies the equation

$$\frac{\partial^2 w}{\partial r\partial s} = -t\left(\frac{d\log\sqrt{\varphi'(\rho)}}{d\log\rho} - 1\right) = m\left(\frac{\partial w}{\partial r} + \frac{\partial w}{\partial s}\right).$$

Here $m = \frac{1}{2\sqrt{\varphi'(\rho)}}\left(\frac{d\log\sqrt{\varphi'(\rho)}}{d\log\rho} - 1\right)$ is a function of $r + s$. Now write $f(\rho) = r + s = \sigma$. Then

$$\sqrt{\varphi'(\rho)} = \frac{d\sigma}{d\log\rho},$$

and consequently

$$m = -\frac{1}{2}\frac{d\log\frac{d\rho}{d\sigma}}{d\sigma}.$$

With the Poisson hypothesis $\varphi(\rho) = a^2\rho^k$, we have

$$f(\rho) = \frac{2a\sqrt{k}}{k-1}\,\rho^{(k-1)/2} + \text{const.}$$

If we choose 0 for the value of the arbitrary constant, then

$$\sqrt{\varphi'(\rho)} + u = \frac{k+1}{2}r + \frac{k-3}{2}s,$$
$$\sqrt{\varphi'(\rho)} - u = \frac{k-3}{2}r + \frac{k+1}{2}s,$$
$$m = \left(\frac{1}{2} - \frac{1}{k-1}\right)\frac{1}{\sigma} = \frac{k-3}{2(k-1)(r+s)}.$$

If we assume Boyle's law $\varphi(\rho) = a^2\rho$, we obtain

$$f(\rho) = a\log\rho,$$
$$\sqrt{\varphi'(\rho)} + u = r - s + a, \quad \sqrt{\varphi'(\rho)} - u = s - r + a,$$
$$m = -\frac{1}{2a},$$

that is, the values obtained from the above if we reduce $f(\rho)$ by the factor $\frac{2a\sqrt{k}}{k-1}$, and thus reduce r and s by $\frac{2a\sqrt{k}}{k-1}$, and then set $k = 1$.

The introduction of r and s as independent variables is, however, only possible where the determinant of these functions of x and t, which is $2\sqrt{\varphi'(\rho)}\,\frac{\partial r}{\partial x}\,\frac{\partial}{\partial}$ does not vanish, that is, when $\frac{\partial r}{\partial x}$, $\frac{\partial s}{\partial x}$ are both nonzero.

If $\frac{\partial r}{\partial x} = 0$, we obtain $dr = 0$ from (1) and

$$x - \left(u - \sqrt{\varphi'(\rho)}\right)t = \text{a function of } s$$

from (2). Consequently the expression (3) is again a complete differential, and w is merely a function of s.

Similarly, if $\frac{\partial s}{\partial x} = 0$, then s also is a constant with respect to t, and $x - \left(u + \sqrt{\varphi'(\rho)}\right)t$ and w are functions of r.

Finally, if $\frac{\partial r}{\partial x}$ and $\frac{\partial s}{\partial x}$ are both 0, then from the differential equation, r, s and w are constants.

3.

In order to solve the problem, we must first of all determine w as a function of r and s in such a way that it satisfies the differential equation

$$\frac{\partial^2 w}{\partial r \partial s} - m\left(\frac{\partial w}{\partial r} + \frac{\partial w}{\partial s}\right) = 0 \tag{1}$$

together with the initial conditions, which will determine w up to an additive constant that can clearly be chosen arbitrarily.

The equation

$$\left(x - \left(u + \sqrt{\varphi'(\rho)}\, t\right)\right) dr - \left(x - \left(u - \sqrt{\varphi'(\rho)}\right) t\right) ds = dw \tag{2}$$

will then yield the time and location where a particular value of r meets a particular value of s. Finally we find u and ρ as functions of x and t by referring to the equations

$$f(\rho) + u = 2r, \;\; f(\rho) - u = 2s. \tag{3}$$

Indeed, provided that neither dr nor ds is zero in a finite interval and consequently neither r nor s is constant, the equations

$$x - \left(u + \sqrt{\varphi'(\rho)}\right) t = \frac{\partial w}{\partial r}, \tag{4}$$

$$x - \left(u - \sqrt{\varphi'(\rho)}\right) t = -\frac{\partial w}{\partial s}, \tag{5}$$

which follow from (2), can be combined with (3) to express u and ρ in terms of x and t.

If, however, r initially has the same value r' in a finite interval, then this interval gradually moves forward to larger values of x. Inside this region where $r = r'$, we cannot derive the value of $x - \left(u + \sqrt{\varphi'(\rho)}\right) t$ from the equation (2), because $dr = 0$. Indeed, in this case the question of where and when this value r' encounters a particular value of s permits no precise answer. Equation (4) is now valid only at the boundaries of this region and yields the values of x, between which the constant value r' of r occurs, when time is specified, or alternatively the time interval during which this value of r occurs, when location is specified. Between these limits, u and ρ are determined as functions of x and t from equations (3) and (5). Similarly we find these functions when s takes the value s' in a finite region, while r varies, or when r, s are both constant. In the latter case, they take constant values, which are found from (3), between the bounds defined via (4) and (5).

4.

Before we set about the solution of equation (1) of the previous section, it seems convenient to note some considerations that are not needed for this purpose. Concerning the function $\varphi(\rho)$, the only hypothesis required is that its derivative does not decrease with increasing ρ, which in physical reality is certainly always the case. We also make a remark here that will be used several times in the following section. The quantity

$$\frac{\varphi(\rho_1) - \varphi(\rho_2)}{\rho_1 - \rho_2} = \int_0^1 \varphi'(\alpha\rho_1 + (1-\alpha)\rho_2) d\alpha,$$

when one of ρ_1 and ρ_2 varies, either remains constant or increases and decreases with this quantity. It follows from this equation that the value of this expression lies between $\varphi'(\rho_1)$ and $\varphi'(\rho_2)$.

We treat first of all the case where the initial disturbance of equilibrium is restricted to a finite region defined by the inequalities $a < x < b$. Thus outside this interval, u and ρ, and consequently r and s, are constant. The values of these quantities for $x < a$ are denoted with suffix 1; for $x > b$, suffix 2. The region in which r is variable gradually moves forward according to Section 1, its lower bound having velocity $\sqrt{\varphi'(\rho_1)} + u_1$, while the upper bound of the region, in which s is variable, moves backward with velocity $\sqrt{\varphi'(\rho_2)} - u_2$. After a time interval

$$\frac{b-a}{\sqrt{\varphi'(\rho_1)} + \sqrt{\varphi'(\rho_2)} + u_1 - u_2},$$

the two regions separate, and between them a gap forms in which $s = s_2$ and $r = r_1$, and consequently the gas particles are again in equilibrium. Thus from the initially disturbed location, two waves issue in opposite directions. In the forward wave, $s = s_2$; accordingly, to a particular value ρ of the density is associated the velocity $u = f(\rho) - 2s_2$, and both values move forward with constant velocity

$$\sqrt{\varphi'(\rho)} + u = \sqrt{\varphi'(\rho)} + f(\rho) - 2s_2.$$

In the wave moving backward, on the other hand, the velocity $-f(\rho) + 2r_1$ is associated to the density ρ, and these two values move backward with velocity $\sqrt{\varphi'(\rho)} + f(\rho) - 2r_1$. The rate of propagation is greater for greater densities, because both $f(\rho)$ and $\sqrt{\varphi'(\rho)}$ increase with ρ.

If we think of ρ as the ordinate of a curve for the abscissa x, then each point of this curve moves forward parallel to the x-axis with constant velocity. Indeed the greater the ordinate, the greater the velocity will be. It is easy to see that, according to this law, points with greater ordinates would finally overtake preceding points with smaller ordinates, so that to a given value of x would correspond more than one value of ρ. Since this cannot occur in physical reality, a condition must enter that renders the law invalid. In fact, the derivation of the differential equation is based on the assumption that u and ρ are continuous functions of x having finite derivatives. However, this assumption ceases to hold as soon as the density curve is perpendicular to the x-axis at some point. From this moment on, a discontinuity appears in this curve, so that a greater value of ρ immediately succeeds a smaller value. This case will be discussed in the next section.

The compression waves, that is, the parts of the wave in which the density increases in the direction of propagation, become ever narrower with their forward progress and finally become compression shocks. However, the width of the expansion waves grows in proportion to elapsed time.

We may easily show, at least under the assumption of Poisson's (or Boyle's) law, that in the case when the initial disturbance of equilibrium is not confined to a finite region, compression shocks must also form in the course of the motion, excluding quite special cases. The velocity with which a value of r moves forward is

$$\frac{k+1}{2}r + \frac{k-3}{2}s$$

with this hypothesis. Thus larger values will, on average, move with greater velocity. A larger value r' must eventually overtake a preceding smaller value r'', unless the value of s corresponding to r'' is, on average, smaller by

$$(r' - r'')\,\frac{1+k}{3-k}$$

than the value of s simultaneously corresponding to r'. In this case, s becomes negatively infinite for positive infinite x, and thus for $x = +\infty$, the velocity u is $+\infty$ (or instead, the density, according to Boyle's law, becomes infinitely small). Thus excluding special cases, it must always transpire that a value of r, larger by a finite amount, follows immediately after a smaller value. Consequently, since $\frac{\partial r}{\partial x}$ becomes infinite, the differential equations lose their validity, and forward moving compression shocks must occur. Likewise, in almost all cases, since $\frac{\partial s}{\partial x}$ tends to infinity, backward moving compression shocks will form.

To determine the time and location for which $\frac{\partial r}{\partial x}$, $\frac{\partial s}{\partial x}$ become infinite and sudden compressions begin, we obtain from equations (1) and (2) of Section 2, on introducing the function w,

$$\frac{\partial r}{\partial x}\left(\frac{\partial^2 w}{\partial r^2}+\left(\frac{d\log\sqrt{\varphi'(\rho)}}{d\log\rho}+1\right)t\right)=1,$$

$$\frac{\partial s}{\partial x}\left(-\frac{\partial^2 w}{\partial s^2}-\left(\frac{d\log\sqrt{\varphi'(\rho)}}{d\log\rho}+1\right)t\right)=1.$$

5.

Since sudden compressions almost always occur, even when density and velocity initially vary continuously everywhere, we must now seek the laws for the propagation of compression shocks.

We suppose that at time t, a jump in u and ρ occurs at $x=\xi$. We denote the value for $x=\xi-0$ of these quantities, and those depending on them, using suffix 1; for $x=\xi+0$, suffix 2. The velocities with which the gas moves relative to the location of the discontinuity, $u_1-\frac{d\xi}{dt}$, $u_2-\frac{d\xi}{dt}$, may be denoted by v_1 and v_2. The mass that passes in the positive direction through an element ω of the plane where $x=\xi$, during the time element dt, is then

$$v_1\rho_1\omega\,dt=v_2\rho_2\omega\,dt;$$

the force acting on the element is $(\varphi(\rho_1)-\varphi(\rho_2))\omega\,dt$ and the resultant increase in velocity is v_2-v_1. Thus

$$(\varphi(\rho_1)-\varphi(\rho_2))\omega\,dt=(v_2-v_1)v_1\rho_1\,\omega\,dt$$

and

$$v_1\rho_1=v_2\rho_2,$$

so that

$$v_1=\mp\sqrt{\frac{\rho_2}{\rho_1}\,\frac{\varphi(\rho_1)-\varphi(\rho_2)}{\rho_1-\rho_2}}.$$

Thus

(1) $$\frac{d\xi}{dt}=u_1\pm\sqrt{\frac{\rho_2}{\rho_1}\,\frac{\varphi(\rho_1)-\varphi(\rho_2)}{\rho_1-\rho_2}}=u_2\pm\sqrt{\frac{\rho_1}{\rho_2}\,\frac{\varphi(\rho_1)-\varphi(\rho_2)}{\rho_1-\rho_2}}.$$

For a compression shock, $\rho_2 - \rho_1$ must have the same sign as v_1 and v_2. For a forward moving shock the sign is negative, for a backward moving shock, positive. In the first case, the upper signs apply and ρ_1 is greater than ρ_2. Thus, by the assumption on $\varphi(\rho)$ at the beginning of the previous section,

$$u_1 + \sqrt{\varphi'(\rho_1)} > \frac{d\xi}{dt} > u_2 + \sqrt{\varphi'(\rho_2)}. \tag{2}$$

Consequently, the location of the discontinuity moves forward more slowly than the succeeding values of r, and more quickly than the preceding values. Thus r_1 and r_2 are determined at a given moment via the differential equations valid on either side of the discontinuity. The same applies for s_2, and consequently for ρ_2 and u_2, since the values of s move backward with velocity $\sqrt{\varphi'(\rho)} - u$; but not for s_1. The values of s_1 and $\frac{d\xi}{dt}$ are uniquely determined from r_1, ρ_2 and u_2 via the equation (1). Indeed, only one value of ρ_1 satisfies the equation

$$2(r_1 - r_2) = f(\rho_1) - f(\rho_2) + \sqrt{\frac{(\rho_1 - \rho_2)(\varphi(\rho_1) - \varphi(\rho_2))}{\rho_1\rho_2}}. \tag{3}$$

For the right side takes each positive value only once when ρ_1 increases to infinity from the value ρ_2, because as well as $f(\rho_1)$, the two factors

$$\sqrt{\frac{\rho_1}{\rho_2}} - \sqrt{\frac{\rho_2}{\rho_1}} \quad \text{and} \quad \sqrt{\frac{\varphi(\rho_1) - \varphi(\rho_2)}{\rho_1 - \rho_2}}$$

into which the last term can be divided, persistently increase, or only the latter factor remains constant. Now when ρ_1 is determined, we obtain fully determined values of u_1 and $\frac{d\xi}{dt}$ via the equation (1).

Entirely similar considerations hold for a backward moving compression shock.

6.

We have found that in a propagating compression shock, the values of u and ρ on either side are always linked by the equation

$$(u_1 - u_2)^2 = \frac{(\rho_1 - \rho_2)(\varphi(\rho_1) - \varphi(\rho_2))}{\rho_1\rho_2}.$$

The question now arises as to what happens when, at a given time and location, arbitrarily given discontinuities occur. In this case, according to the values of u_1, ρ_1, u_2, ρ_2, either two compression shocks run from this location in opposite directions, or one runs one forward, or one runs backward; or, finally, no compression shock actually occurs, in which case the motion follows from the differential equation.

We denote the values that u and ρ take after or between the compression shocks at the first moment of their propagation using an accent. Then in the first case $\rho' > \rho_1$, $\rho' > \rho_2$, and we have

$$(1)\qquad \begin{aligned} u_1 - u' &= \sqrt{\frac{(\rho' - \rho_1)(\varphi(\rho') - \varphi(\rho_1))}{\rho'\rho_1}}, \\ u' - u_2 &= \sqrt{\frac{(\rho' - \rho_2)(\varphi(\rho') - \varphi(\rho_2))}{\rho'\rho_2}}, \end{aligned}$$

$$(2)\qquad \begin{aligned} u_1 - u_2 = &\sqrt{\frac{(\rho' - \rho_1)(\varphi(\rho') - \varphi(\rho_1))}{\rho'\rho_1}} \\ &+ \sqrt{\frac{(\rho' - \rho_2)(\varphi(\rho') - \varphi(\rho_2))}{\rho'\rho_2}}. \end{aligned}$$

Since both terms on the right side of (2) increase together with ρ', $u_1 - u_2$ must be positive and

$$(u_1 - u_2)^2 > \frac{(\rho_1 - \rho_2)(\varphi(\rho_1) - \varphi(\rho_2))}{\rho_1\rho_2}.$$

Conversely, when these conditions are fulfilled, one and only one pair of values of u' and ρ' satisfies the equations (1).

For the latter case to occur, and thus for the motion to be determined via the differential equations, it is necessary and sufficient that $r_1 \leq r_2$ and $s_1 \geq s_2$, that is, $u_1 - u_2$ is negative and

$$(u_1 - u_2)^2 \geq (f(\rho_1) - f(\rho_2))^2.$$

The values r_1 and r_2, s_1 and s_2, then move apart, since the value in front moves forward with greater velocity, so that the discontinuity vanishes.

If neither the former, nor the latter, conditions are fulfilled, then the initial values are satisfied by a single compression wave, indeed one that moves forward or backward according as ρ_1 is larger or smaller than ρ_2.

Indeed, in that case, if $\rho_1 > \rho_2$,

$$2(r_1 - r_2) \text{ or } f(\rho_1) - f(\rho_2) + u_1 - u_2$$

is positive, because $(u_1 - u_2)^2 < (f(\rho_1) - f(\rho_2))^2$, and at the same time it is

$$\leq f(\rho_1) - f(\rho_2) + \sqrt{\frac{(\rho_1 - \rho_2)(\varphi(\rho_1) - \varphi(\rho_2))}{\rho_1 \rho_2}},$$

because

$$(u_1 - u_2)^2 \leq \frac{(\rho_1 - \rho_2)(\varphi(\rho_1) - \varphi(\rho_2))}{\rho_1 \rho_2}.$$

Now we may find a value for the density ρ' behind the compression shock that satisfies condition (3) of the previous section, and this value is $\leq \rho_1$. Consequently, since $s' = f(\rho') - r_1$, $s_1 = f(\rho_1) - r_1$, we also have $s' \leq s_1$, so that the motion proceeds according to the differential equation behind the compression shock.

The other case, $\rho_1 < \rho_2$, is obviously not essentially different from this one.

7.

In order to illustrate the above by a simple example, where the motion can be determined by the methods we have found, we shall assume that the pressure and density depend on one another via Boyle's law and that initially density and velocity have a jump at $x = 0$, but are constant on both sides of this location.

By the above, there are four cases to be distinguished.

I. If $u_1 - u_2 > 0$, so that the two gas masses move together, and $\left(\frac{u_1 - u_2}{a}\right)^2 > \frac{(\rho_1 - \rho_2)^2}{\rho_1 \rho_2}$, then two compression shocks running in opposite directions will form. From (1) of Section 6, if α denotes $\left(\frac{\rho_1}{\rho_2}\right)^{1/4}$ and θ denotes the positive root of the equation

$$\frac{u_1 - u_2}{a(\alpha + \alpha^{-1})} = \theta - \frac{1}{\theta},$$

the density between the compression shocks is $\rho' = \theta^2\sqrt{\rho_1\rho_2}$. From (1) of Section 5, the forward moving compression shock satisfies

$$\frac{d\xi}{dt} = u_2 + a\alpha\theta = u' + \frac{a}{\alpha\theta};$$

the backward moving one satisfies

$$\frac{d\xi}{dt} = u_1 - \frac{a\theta}{\alpha} = u' - \frac{a\alpha}{\theta}.$$

After time t has elapsed, if

$$\left(u_1 - a\frac{\theta}{\alpha}\right)t < x < (u_2 + a\alpha\theta)t,$$

the values of the velocity and density are u' and ρ'. For a smaller x, the values are u_1 and ρ_1; for a larger x, u_2 and ρ_2.

II. If $u_1 - u_2 < 0$, so that the gas masses move apart, and moreover

$$\left(\frac{u_1 - u_2}{a}\right)^2 \geq \left(\log\frac{\rho_1}{\rho_2}\right)^2,$$

then two gradually widening expansion waves issue from the boundary in opposite directions. By Section 4, between them we have $r = r_1$, $s = s_2$, $u = r_1 - s_2$. In the forward moving wave, $s = s_2$ and $x-(u+a)t$ is a function of r, whose value is 0 from the initial values $t = 0$, $x = 0$. For the backward moving wave, on the other hand, we have $r = r_1$ and $x - (u - a)t = 0$. One equation for the determination of u and ρ, for

$$(r_1 - s_2 + a)t < x < (u_2 + a)t,$$

is $u = -a + \frac{x}{t}$; for smaller values of x, it is $r = r_1$ and for larger values, $r = r_2$. The other equation is $u = a + \frac{x}{t}$ for

$$(u_1 - a)t < x < (r_1 - s_2 - a)t.$$

For smaller x, it is $s = s_1$, and for larger x, $s = s_2$.

III. If neither of these two cases holds, and $\rho_1 > \rho_2$, a backward moving expansion wave and a forward moving compression wave occur. For the latter, we find from (3) of Section 5, θ denoting the root of the equation

$$\frac{2(r_1 - r_2)}{a} = 2\log\theta + \theta - \frac{1}{\theta},$$

that $\rho' = \theta^2\rho_2$. From (1) of Section 5,

$$\frac{d\xi}{dt} = u_2 + a\theta = u' + \frac{a}{\theta}.$$

Thus after time t has elapsed, in front of the compression shock, that is for $x > (u_2 + a\theta)t$, we have $u = u_2$, $\rho = \rho_2$. Behind the compression shock, however, we have $r = r_1$ and moreover, if

$$(u_1 - a)t < x < (u' - a)t,$$

we have $u = a + \frac{x}{t}$. For a smaller x, we have $u = u_1$, and for a larger x, $u = u'$.

IV. Finally, if neither of the first two cases holds, and $\rho_1 < \rho_2$, the progress is just as in III, with the direction reversed.

8.

To solve our problem in general, by Section 3 we must determine the function w in such a way that it satisfies the differential equation

$$\frac{\partial^2 w}{\partial r \partial s} - m\left(\frac{\partial w}{\partial r} + \frac{\partial w}{\partial s}\right) = 0 \tag{1}$$

and the initial conditions.

Let us exclude the case where discontinuities occur. Then obviously from Section 1, the location and elapsed time, or the values of x and t, for which a definite value r' of r meets a definite value s' of s, are completely determined once the initial values of r and s are given for the interval between the values r' of r and s' of s, provided that the differential equations (3) of Section 1 are satisfied in the region (S) of values of x which (for any given t) lie between the two values corresponding to $r = r'$ and $s = s'$. Thus the value of w for $r = r'$, $s = s'$ is also fully determined if w satisfies the differential equation (1) in the region (S), and if for the initial values of r and s the values of $\frac{\partial w}{\partial r}$ and $\frac{\partial w}{\partial s}$, and hence, up to an additive constant, that of w, are given. This constant may be chosen arbitrarily.

For these conditions are equivalent to those given above. It also follows from Section 3, that $\frac{\partial w}{\partial r}$ takes different values on either side of a value r'' of r, when this value falls in a finite interval. However, $\frac{\partial w}{\partial r}$ varies continuously with s; likewise, for the variation of $\frac{\partial w}{\partial s}$ with r. The function w itself varies continuously with both r and s.

After these preparations we can proceed to the solution of our problem, the determination of the value of w for two arbitrary values r', s' of r, s.

For visualization, we think of x and t as abscissa and ordinate of a point in a plane, and draw the curves in this plane where r and s have constant values. We may denote the first of these by (r), the second by (s). We treat the direction in which t increases as positive in the curves. The region (S) will then be represented by a part of the plane bounded by the curve (r'), the curve (s'), and the part of the x axis lying between them. The point in question is to determine the value of w at the intersection of the first two curves from the values given in the latter line. We wish to generalize the problem somewhat. We assume that the region (S) is bounded by an arbitrary curve c, rather than the latter line, which cuts neither curve (r), (s) more than once, and that for values of r and s belonging to the curve the values of $\frac{\partial w}{\partial r}$, $\frac{\partial w}{\partial s}$ are given. It will emerge from the solution of the problem that we require for these values of $\frac{\partial w}{\partial r}$, $\frac{\partial w}{\partial s}$ only that they vary continuously with position in the curve: otherwise they may be arbitrary. If the curve c met one of the curves $(r), (s)$ more than once, these values would not be independent of each other.

In order to determine functions that satisfy linear partial differential equations and linear boundary conditions, we can apply a procedure entirely analogous to the solution of a system of simultaneous linear equations, multiplying by undetermined factors, adding and then determining those factors so that the sum of the unknown quantities reduces to a single one.

We consider the part (S) of the plane bounded by the curves (r) and (s) to be cut up into infinitely small parallelograms and denote by $\delta r, \delta s$ the variations that the quantities r and s undergo, when passing along the curve elements that form the sides of the parallelograms in a positive sense. Further denote by v an arbitrary function of r and s that is everywhere continuous and has continuous derivatives. From equation (1),

$$(2) \qquad 0 = \int v \left(\frac{\partial^2 w}{\partial r \partial s} - m \left(\frac{\partial w}{\partial r} + \frac{\partial w}{\partial s} \right) \right) \delta r \, \delta s$$

where the integral is taken over the region (S). The right side of this equation is now arranged, in other words, the integral is transformed via partial integration, so that apart from known quantities it contains only the function sought, not its derivatives. By carrying out this operation, the integral becomes

$$\int w\left(\frac{\partial^2 v}{\partial r\partial s}+\frac{\partial mv}{\partial r}+\frac{\partial mv}{\partial s}\right)\delta r\,\delta s,$$

taken over (S), together with a simple integral that, since $\frac{\partial w}{\partial r}$ varies continuously with s, $\frac{\partial w}{\partial s}$ with r, and w with both r and s, is taken only over the boundary of (S). Denote by dr and ds the variations of r and s in a boundary element. Here the boundary is traversed in the same direction, relative to the interior normal, as the positive direction in the curves (r) with respect to the positive direction in the curves (s). The boundary integral is

$$-\int\left(v\left(\frac{\partial w}{\partial s}-mw\right)ds+w\left(\frac{\partial v}{\partial r}+mv\right)dr\right).$$

The integral over the entire boundary of S is equal to the sum of the integrals over the curves c, (s'), (r') forming this boundary. Denoting their intersection points by $(c,r'),(c,s'),(r',s')$, the integral is

$$\int_{c,r'}^{c,s'}+\int_{c,s'}^{r',s'}+\int_{s',r'}^{c,r'}.$$

Of these three summands, the first contains only known quantities apart from the function v. The second, since $ds=0$ in this integral, contains only the unknown function w itself, not its derivatives. The third summand, however, can be transformed via partial integration into

$$(vw)_{r',s'}-(vw)_{c,r'}+\int_{s',r'}^{c,r'}w\left(\frac{\partial v}{\partial s}+mv\right)ds;$$

again, only the desired function w occurs in this term.

By these transformations, the equation (2) clearly yields the value of the function w at the point (r',s'), expressed in terms of known quantities, if the function v satisfies the following conditions:

$$(3)\quad\begin{aligned}
&1)\ \text{throughout } S, && \frac{\partial^2 v}{\partial r\partial s}+\frac{\partial mv}{\partial r}+\frac{\partial mv}{\partial s}=0;\\
&2)\ \text{for } r=r', && \frac{\partial v}{\partial s}+mv=0;\\
&3)\ \text{for } s=s', && \frac{\partial v}{\partial r}+mv=0;\\
&4)\ \text{for } r=r', s=s', && \text{we have } v=1.
\end{aligned}$$

We then have

$$(4) \qquad w_{r',s'} = (vw)_{c,r'} + \int_{c,r'}^{c,s'} \left(v \left(\frac{\partial w}{\partial s} - mw \right) ds + w \left(\frac{\partial v}{\partial r} + mv \right) dr \right).$$

9.

By the procedure applied above, the problem of determining a function w satisfying a linear differential equation and linear boundary conditions reduces to a similar but much simpler problem for another function v. The determination of this function is usually easiest by treating a special case of the problem via Fourier's method. Here we must be content merely to point out this calculation, and demonstrate the result in another way.

In equation (1) of the previous section, we introduce independent variables $\sigma = r + s$, $u = r - s$ in place of r and s, and choose for c a curve in which σ is constant. Now the problem can be treated by Fourier's principles. By comparing the result with equation (4) of the previous section, we have, writing $r' + s' = \sigma'$, $r' - s' = u'$,

$$v = \frac{2}{\pi} \int_0^\infty \cos \mu(u - u') \frac{d\rho}{d\sigma} (\psi_1(\sigma')\psi_2(\sigma) - \psi_2(\sigma')\psi_1(\sigma)) d\mu.$$

Here $\psi_1(\sigma)$, $\psi_2(\sigma)$ denote particular solutions of the differential equation $\psi'' - 2m\psi' + \mu^2\psi = 0$, for which

$$\psi_1 \psi_2' - \psi_2 \psi_1' = \frac{d\sigma}{d\rho}.$$

If we assume Poisson's law, according to which $m = \left(\frac{1}{2} - \frac{1}{k-1}\right)\frac{1}{\sigma}$, ψ_1 and ψ_2 can be expressed as definite integrals. We recover a triple integral for v, whose reduction yields

$$v = \left(\frac{r' + s'}{r + s} \right)^{\frac{1}{2} - \frac{1}{k-1}} F\left(\frac{3}{2} - \frac{1}{k-1}, \frac{1}{k-1} - \frac{1}{2}, 1, \frac{-(r - r')(s - s')}{(r + s)(r' + s')} \right).$$

We can easily demonstrate the correctness of this expression by showing that it actually satisfies the condition (3) of the previous section.

Let us write $v = \exp\left(-\int_{\sigma'}^{\sigma} m \, d\sigma\right) y$. For y, the conditions become

$$\frac{\partial^2 y}{\partial r \partial s} + \left(\frac{dm}{d\sigma} - m^2 \right) y = 0,$$

and $y = 1$ both for $r = r'$ and $s = s'$. Under Poisson's hypothesis, however, we can satisfy these conditions by taking y to be a function of

$$z = -\frac{(r-r')(s-s')}{(r+s)(r'+s')}.$$

For then, writing $\lambda = \frac{1}{2} - \frac{1}{k-1}$, $m = \frac{\lambda}{\sigma}$, so that $\frac{dm}{d\sigma} - m^2 = -\frac{(\lambda+\lambda^2)}{\sigma^2}$, the equation becomes

$$\frac{\partial^2 y}{\partial s \partial r} = \frac{1}{\sigma^2}\left(\frac{d^2 y}{d(\log z)^2}\left(1 - \frac{1}{z}\right) + \frac{dy}{d\log z}\right).$$

Consequently, $v = \left(\frac{\sigma'}{\sigma}\right)^\lambda y$, and y is a solution of the differential equation

$$(1-z)\,\frac{d^2 y}{d(\log z)^2} - z\,\frac{dy}{d\log z} + (\lambda + \lambda^2)zy = 0.$$

With the notation introduced in my paper on Gauss's series, y is a function

$$P\begin{pmatrix} 0 & -\lambda & 0 & \\ 0 & 1+\lambda & 0 & z \end{pmatrix};$$

in fact, the particular solution equal to 1 for $z = 0$.

According to the transformation principles developed in that paper, y can be expressed not only via the functions $P(0, 2\lambda+1, 0)$, but also via the functions $P\left(\frac{1}{2}, 0, \lambda + \frac{1}{2}\right)$, $P\left(0, \lambda + \frac{1}{2}, \lambda + \frac{1}{2}\right)$. We obtain in this way a large class of representations of y via hypergeometric series and definite integrals. Here we point out only the following:

$$y = F(1+\lambda, -\lambda, 1, z) = (1-z)^\lambda F\left(-\lambda, -\lambda, 1, \frac{z}{z-1}\right)$$
$$= (1-z)^{-1-\lambda} F\left(1+\lambda, 1+\lambda, 1, \frac{z}{z-1}\right),$$

which are adequate in all circumstances.

In order to derive the results that hold for Boyle's law from those found for Poisson's law, we reduce r, s, r', s' by $\frac{a\sqrt{k}}{k-1}$ according to Section 2, and then set $k = 1$. We obtain $m = -\frac{1}{2a}$ and

$$v = e^{\frac{1}{2a}(r-r'+s-s')} \sum_{n=0}^{\infty} \frac{(r-r')^n (s-s')^n}{(n!)^2 (2a)^{2n}}.$$

10.

If we substitute the expression found for v in the previous section into equation (4) of Section 8, we obtain the value of w for $r = r'$, $s = s'$, expressed in terms of the values of w, $\frac{\partial w}{\partial r}$, $\frac{\partial w}{\partial s}$ on the curve c. In our problem, only $\frac{\partial w}{\partial r}$ and $\frac{\partial w}{\partial s}$ are given directly on the curve, while w must be found from these quantities by integration. Hence it is convenient to transform the expression for $w_{r',s'}$ in such a way that only the partial derivatives of w appear under the integral sign.

Denote by P and Σ the integrals of the expression $-mv\,ds + \left(\frac{\partial v}{\partial r} + mv\right) dr$ and $\left(\frac{\partial v}{\partial s} + mv\right) ds - mv\,dr$, which are complete differentials in view of the equation

$$\frac{\partial^2 v}{\partial r \partial s} + \frac{\partial mv}{\partial r} + \frac{\partial mv}{\partial s} = 0.$$

Denote by ω the integral of $P\,dr + \Sigma\,ds$, an expression which is again a complete differential since

$$\frac{\partial P}{\partial s} = -mv = \frac{\partial \Sigma}{\partial r}.$$

We now determine the integration constants in these integrals in such a way that ω, $\frac{\partial \omega}{\partial r}$, $\frac{\partial \omega}{\partial s}$ vanish for $r = r'$, $s = s'$. Now ω satisfies the equations

$$\frac{\partial \omega}{\partial r} + \frac{\partial \omega}{\partial s} + 1 = v, \qquad \frac{\partial^2 \omega}{\partial r \partial s} = -mv$$

and $\omega = 0$ for both $r = r'$ and $s = s'$. Incidentally, ω is completely determined by these boundary conditions and the differential equation

$$\frac{\partial^2 \omega}{\partial r \partial s} + m\left(\frac{\partial \omega}{\partial r} + \frac{\partial \omega}{\partial s} + 1\right) = 0.$$

If we now introduce ω in place of v in the expression for $w_{r',s'}$, we can transform it via partial integration into

$$\text{(1)} \qquad w_{r',s'} = w_{c,r'} + \int_{c,r'}^{c,s'} \left(\left(\frac{\partial \omega}{\partial s} + 1\right)\frac{\partial w}{\partial s}\,ds - \left(\frac{\partial \omega}{\partial r}\right)^2 dr\right).$$

In order to determine the motion of the gas from the initial state, we must take c to be the curve in which $t = 0$. In this curve, $\frac{\partial w}{\partial r} = x$, $\frac{\partial w}{\partial s} = -x$. By repeated partial integration, we obtain

$$w_{r',s'} = w_{c,r'} + \int_{c,r'}^{c,s'} (\omega\,dx - x\,ds).$$

Consequently, by (4), (5) of Section 3,

$$
(2)\qquad
\begin{aligned}
\left(x-\left(\sqrt{\varphi'(\rho)}+u\right)t\right)_{r',s'} &= x_{r'}+\int_{x_{r'}}^{x}\frac{\partial\omega}{\partial r'}\,dx,\\
\left(x+\left(\sqrt{\varphi'(\rho)}-u\right)t\right)_{r',s'} &= x_{s'}-\int_{x_{r'}}^{x_{s'}}\frac{\partial\omega}{\partial s'}\,dx.
\end{aligned}
$$

However, the equations (2) only express the motion as long as

$$
\frac{\partial^2 w}{\partial r^2}+\left(\frac{d\log\sqrt{\varphi'(\rho)}}{d\log\rho}+1\right)t,\ \frac{\partial^2 w}{\partial s^2}+\left(\frac{d\log\sqrt{\varphi'(\rho)}}{d\log\rho}+1\right)t
$$

remain different from 0. As soon as one of these quantities vanishes, a compression shock occurs, and equation (1) is only valid inside a region that lies entirely to one side of this compression shock. Hence the principles developed here are insufficient—at least in general—to determine the motion from the initial state. However, with the help of (1), and the equations that hold for the compression shock by Section 5, we can indeed determine the motion if the location of the compression shock at time t (that is, ξ as a function of t) is given. But we will not pursue this further. Likewise we forgo the treatment of the case when the air is bounded by a fixed wall. The calculation entails no difficulties, and a comparison of the result with experiment is not possible at present.

IX.

Author's announcement for VIII.

(*Göttinger Nachrichten*, 1859, no. 19.)

This investigation makes no claim to produce results useful for experimental research. The author wishes it to be considered only as a contribution to the theory of nonlinear partial differential equations. For the solution of linear partial differential equations, the most fruitful methods have not been found by developing the general idea of the problem, but rather from the treatment of special physical problems. In the same way, the theory of nonlinear partial differential equations seems generally to demand a thorough treatment of particular physical problems, taking into account all the secondary factors. Indeed, the solution of the quite special problem that is the subject of this work requires new methods and concepts, and leads to results that will probably play a role in more general problems.

By the complete solution of this problem, questions discussed in lively fashion some time ago[1] by the English mathematicians Challis, Airy and Stokes, insofar as not already settled by Stokes[2], would be decided more clearly. The same applies to the controversy with regard to another question in the same area in the *K. K. Ges. d. W. zu Wien* between Petzval, Doppler and A. von Ettinghausen[3].

The only empirical law, other than the general law of motion, assumed in this investigation, is the law according to which the pressure of a gas varies with its density when no heat is taken up or released. Poisson already assumed, albeit based on hypotheses resting on very uncertain foundations, that the pressure is proportional to ρ^k, where the density is ρ. Here k is the ratio of the specific heat at constant pressure to the specific heat at constant volume. Poisson's hypothesis can now be based on the experiments of Regnault on the specific heat of gases, and a mechanical principle of the theory of heat. It seems necessary to include this basis for Poisson's law in the introduction, since it still appears to be little known. The value of k found in this way is 1.4101. The velocity of sound in dry air at 0°C, according

[1] *Phil. Mag.*, vols. 33, 34 and 35.

[2] *Phil. Mag.*, vol. 33, p. 349.

[3] *Sitzungsberichte der K. K. Ges. d. W.*, 15 January, 21 May and 1 June, 1852.

to the experiments of Martins and A. Bravais, is found to be 332.37 meters per second, which produces the value 1.4095 for k.

The comparison of the results of our investigation with reality via experiments and observation poses great difficulties, and indeed is scarcely practicable at present. Nevertheless, we discuss the results here as far as possible without being long-winded.

The work considers the motion of air, or a gas, only in the case when, initially and throughout, the whole motion takes place in one direction; and in every plane perpendicular to this direction, velocity and density are constant. It is known that when the initial disturbance of equilibrium is restricted to a finite segment, under the customary assumption that the pressure differences are of smaller order of magnitude than the pressure, two waves issue from the disturbed location in opposite directions and advance with the velocity $\sqrt{\phi'(\rho)}$, which is constant under this hypothesis. Here $\phi(\rho)$ denotes the pressure at density ρ, and $\phi'(\rho)$ the derivative of this function.

A rather similar outcome occurs for the case when the pressure differences are finite. The location where equilibrium is disturbed likewise decomposes after a finite time into two waves progressing in opposite directions. In these waves the velocity, measured in the direction of propagation, is a definite function $\int \sqrt{\phi'(\rho)}\, d\log\rho$ of the density; the constant of integration may differ in the two waves. Thus in each wave, a given value of the density is linked to each value of the velocity; indeed, with a greater density is associated an algebraically greater velocity. Both these values move forward with constant velocity. Their velocity of propagation relative to the gas is $\sqrt{\phi'(\rho)}$. In space, however, the velocity is greater by the amount of the velocity of the gas in the direction of propagation. Under the hypothesis, found in physical reality, that $\phi'(\rho)$ does not decrease with increasing ρ, greater densities move with greater velocity. It follows that the rarefication waves, that is the parts of the wave in which density increases in the direction of propagation, increase in breadth in proportion to time. Likewise, compression waves decrease in breadth and finally become compression shocks. The laws which apply throughout the region of disturbance of equilibrium, either before the separation of the two waves, or within each wave, are not given here, and likewise for the laws of propagation of compression shocks; this would require complicated formulae.

In connection with acoustics, then, this investigation produces the result that in the cases where density differences cannot be treated as infinitely small, a variation in the form of the sound wave, that is to say the tone, appears in the course of the propagation. However, experimental verification of this result seems to be very difficult, in spite of the very recent progress made by Helmholtz and others in the analysis of tones. For within short distances, a variation of tone is not noticeable, while at greater distances it would be difficult to separate the various causes that could modify the tone. An application to meteorology is really not to be considered, since the motions of air investigated here are of the type that propagate with the velocity of sound. To all appearances, currents in the atmosphere move with much smaller velocity.

X.

A contribution to the study of the motion of a homogeneous fluid ellipsoid.

(*Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, vol. 9, 1861.*)

In his last work, edited by Dedekind, Dirichlet investigated the motion of a homogeneous fluid ellipsoid, whose elements are attracted to one another by the law of gravity. His approach is surprising and opens a new path. The continuation of this fine research has a special appeal for mathematicians, quite apart from the question of the form of heavenly bodies which was the occasion for the investigation. Dirichlet himself carried through the solution of the problem completely only in the simplest cases. For the continuation of the investigation, it is convenient to give a form of the differential equation for the motion of a fluid body that is independent of the time-origin chosen. For example, we may study how the variation in size of the principal axes of an ellipsoid affects the motion of the fluid body relative to these axes. In treating the problem here in this way, we presuppose Dirichlet's discussion. To avoid misunderstanding, we note that it was not possible to adhere to his notation without change.

1.

We denote by a, b, c the principal axes of the ellipsoid at time t, and by x, y, z the coordinates of an element of the fluid body at time t; a suffix 0 is used to denote the initial values of these quantities. We assume that the principal axes of the ellipsoid coincide with the coordinate axes at the start time.

It is well known that Dirichlet's investigation starts from the observation that we can satisfy the differential equation for the motion of the fluid by taking the coordinates x, y, z to be linear expressions in their initial values, the coefficients being functions only of time. We write these expressions in the form

$$
(1) \qquad
\begin{aligned}
x &= \ell \frac{x_0}{a_0} + m \frac{y_0}{b_0} + n \frac{z_0}{c_0}, \\
y &= \ell' \frac{x_0}{a_0} + m' \frac{y_0}{b_0} + n' \frac{z_0}{c_0}, \\
z &= \ell'' \frac{x_0}{a_0} + m'' \frac{y_0}{b_0} + n'' \frac{z_0}{c_0}.
\end{aligned}
$$

Denote by ξ, η, ζ the coordinates of the point (x, y, z) with respect to a moving coordinate system, whose axes coincide at each instant with the principal axes of the ellipsoid. Then ξ, η, ζ are known also to be linear expressions in x, y, z,

$$
(2) \qquad
\begin{aligned}
\xi &= \alpha x + \beta y + \gamma z, \\
\eta &= \alpha' x + \beta' y + \gamma' z, \\
\zeta &= \alpha'' x + \beta'' y + \gamma'' z.
\end{aligned}
$$

The coefficients are the cosines of the angles that the axes of one system form with the axes of the other, $\alpha = \cos \xi x$, $\beta = \cos \xi y$, and so on. Between these coefficients, six equations hold which imply that on substituting these expressions, we obtain

$$
\xi^2 + \eta^2 + \zeta^2 = x^2 + y^2 + z^2.
$$

Since the surface is always formed of the same fluid particles, we must have

$$
\frac{\xi^2}{a^2} + \frac{\eta^2}{b^2} + \frac{\zeta^2}{c^2} = \frac{x_0^2}{a_0^2} + \frac{y_0^2}{b_0^2} + \frac{z_0^2}{c_0^2}.
$$

Now let

$$
(3) \qquad
\begin{aligned}
\frac{\xi}{a} &= \alpha_1 \frac{x_0}{a_0} + \beta_1 \frac{y_0}{b_0} + \gamma_1 \frac{z_0}{c_0} \\
\frac{\eta}{b} &= \alpha_1' \frac{x_0}{a_0} + \beta_1' \frac{y_0}{b_0} + \gamma_1' \frac{z_0}{c_0}, \\
\frac{\zeta}{c} &= \alpha_1'' \frac{x_0}{a_0} + \beta_1'' \frac{y_0}{b_0} + \gamma_1'' \frac{z_0}{c_0}.
\end{aligned}
$$

That is, we express $\frac{\xi}{a}, \frac{\eta}{b}, \frac{\zeta}{c}$ in terms of $\frac{x_0}{a_0}, \frac{y_0}{b_0}, \frac{z_0}{c_0}$ by substituting the values (1) into the equations (2), and denote the resulting coefficients by $\alpha_1, \beta_1, \ldots, \gamma_1''$.

Then $\alpha_1, \beta_1, \ldots, \gamma_1''$ are the coefficients of an orthogonal transformation of coordinates: they can be treated as the cosines of the angles that the axes of a moving coordinate system ξ_1, η_1, ζ_1 form with the axes of the fixed coordinate system x, y, z. If we express x, y, z with the help of equations (2) and (3) in terms of $\frac{x_0}{a_0}, \frac{y_0}{z_0}, \frac{z_0}{c_0}$, this yields

$$
\begin{aligned}
\ell &= a\alpha\alpha_1 + b\alpha'\alpha_1' + c\alpha''\alpha_1'',\\
m &= a\alpha\beta_1 + b\alpha'\beta_1' + c\alpha''\beta_1'',\\
n &= a\alpha\gamma_1 + b\alpha'\gamma_1' + c\alpha''\gamma_1'',\\
\ell' &= a\beta\alpha_1 + b\beta'\alpha_1' + c\beta''\alpha_1'',\\
m' &= a\beta\beta_1 + b\beta'\beta_1' + c\beta''\beta_1'',\\
n' &= a\beta\gamma_1 + b\beta'\gamma_1' + c\beta''\gamma_1'',\\
\ell'' &= a\gamma\alpha_1 + b\gamma'\alpha_1' + c\gamma''\alpha_1'',\\
m'' &= a\gamma\beta_1 + b\gamma'\beta_1' + c\gamma''\beta_1'',\\
n'' &= a\gamma\gamma_1 + b\gamma'\gamma_1' + c\gamma''\gamma_1''.
\end{aligned}
\tag{4}
$$

Hence we can treat the position of the fluid particles, or the values of the quantities $\ell, m, \ldots, n''$ at time t, as dependent on the quantities a, b, c and the position of the two moving coordinate systems. We observe that in interchanging the two coordinate systems, in the system of quantities ℓ, m, n the rows are interchanged with the columns. Thus ℓ, m', n'' remain unchanged, while the pairs of quantities m and ℓ', n and ℓ'', n' and m'' interchange. Our next business is to derive the differential equations for the variations of the principal axes, and the motion of these two coordinate systems, from the fundamental equations for the motion of the fluid particles in §1,1 of Dirichlet's work.

2.

In the equations in question, it is obviously permissible to replace derivatives with respect to the initial values of x, y, z, which were denoted there by a, b, c, by derivatives with respect to ξ, η, ζ. The equations formed in this way are written as aggregates of the original set of equations, and conversely. Inserting the values of $\frac{\partial x}{\partial \xi}, \frac{\partial y}{\partial \eta}, \ldots, \frac{\partial z}{\partial \zeta}$, we obtain in this way

$$(1)\qquad \begin{aligned}
\frac{\partial^2 x}{\partial t^2}\,\alpha + \frac{\partial^2 y}{\partial t^2}\,\beta + \frac{\partial^2 z}{\partial t^2}\,\gamma &= \epsilon\,\frac{\partial V}{\partial \xi} - \frac{\partial P}{\partial \xi},\\
\frac{\partial^2 x}{\partial t^2}\,\alpha' + \frac{\partial^2 y}{\partial t^2}\,\beta' + \frac{\partial^2 z}{\partial t^2}\,\gamma' &= \epsilon\,\frac{\partial V}{\partial \eta} - \frac{\partial P}{\partial \eta},\\
\frac{\partial^2 x}{\partial t^2}\,\alpha'' + \frac{\partial^2 y}{\partial t^2}\,\beta'' + \frac{\partial^2 z}{\partial t^2}\,\gamma'' &= \epsilon\,\frac{\partial V}{\partial \zeta} - \frac{\partial P}{\partial \zeta}.
\end{aligned}$$

Here V is the potential, P the pressure at the point x, y, z at time t, and the constant ϵ represents the attraction between two unit masses at unit distance.

Now it is a question of putting the quantities on the left of the equality signs in the form of linear functions of ξ, η, ζ. Some preparations are needed for this.

For brevity, let us write

$$(2)\qquad \begin{aligned}
\frac{\partial x}{\partial t}\,\alpha + \frac{\partial y}{\partial t}\,\beta + \frac{\partial z}{\partial t}\,\gamma &= \xi',\\
\frac{\partial x}{\partial t}\,\alpha' + \frac{\partial y}{\partial t}\,\beta' + \frac{\partial z}{\partial t}\,\gamma' &= \eta',\\
\frac{\partial x}{\partial t}\,\alpha'' + \frac{\partial y}{\partial t}\,\beta'' + \frac{\partial z}{\partial t}\,\gamma'' &= \zeta'.
\end{aligned}$$

By differentiation of equation (2), we obtain

$$\begin{aligned}
\frac{\partial \xi}{\partial t} &= \frac{d\alpha}{dt}\,x + \frac{d\beta}{dt}\,y + \frac{d\gamma}{dt}\,z + \xi',\\
\frac{\partial \eta}{\partial t} &= \frac{d\alpha'}{dt}\,x + \frac{d\beta'}{dt}\,y + \frac{d\gamma'}{dt}\,z + \eta',\\
\frac{\partial \zeta}{\partial t} &= \frac{d\alpha''}{dt}\,x + \frac{d\beta''}{dt}\,y + \frac{d\gamma''}{dt}\,z + \zeta'.
\end{aligned}$$

If we now express x, y, z in terms of ξ, η, ζ, we find that

$$\frac{\partial\xi}{\partial t} = \left(\frac{d\alpha}{dt}\,\alpha + \frac{d\beta}{dt}\,\beta + \frac{d\gamma}{dt}\,\gamma\right)\xi + \left(\frac{d\alpha}{dt}\,\alpha' + \frac{d\beta}{dt}\,\beta' + \frac{d\gamma}{dt}\,\gamma'\right)\eta + \left(\frac{d\alpha}{dt}\,\alpha'' + \frac{d\beta}{dt}\,\beta'' + \frac{d\gamma}{dt}\,\gamma''\right)\zeta + \xi',$$

$$\frac{\partial\eta}{\partial t} = \left(\frac{d\alpha'}{dt}\,\alpha + \frac{d\beta'}{dt}\,\beta + \frac{d\gamma'}{dt}\,\gamma\right)\xi + \left(\frac{d\alpha'}{dt}\,\alpha' + \frac{d\beta'}{dt}\,\beta' + \frac{d\gamma'}{dt}\,\gamma'\right)\eta + \left(\frac{d\alpha'}{dt}\,\alpha'' + \frac{d\beta'}{dt}\,\beta'' + \frac{d\gamma'}{dt}\,\gamma''\right)\zeta + \eta',$$

$$\frac{\partial\zeta}{\partial t} = \left(\frac{d\alpha''}{dt}\,\alpha + \frac{d\beta''}{dt}\,\beta + \frac{d\gamma''}{dt}\,\gamma\right)\xi + \left(\frac{d\alpha''}{dt}\,\alpha' + \frac{d\beta''}{dt}\,\beta' + \frac{d\gamma''}{dt}\,\gamma'\right)\eta + \left(\frac{d\alpha''}{dt}\,\alpha'' + \frac{d\beta''}{dt}\,\beta'' + \frac{d\gamma''}{dt}\,\gamma''\right)\zeta + \zeta'.$$

Differentiating the known equations $\alpha^2 + \beta^2 + \gamma^2 = 1$, $\alpha\alpha' + \beta\beta' + \gamma\gamma' = 0, \ldots$, we obtain

$$\alpha\,\frac{d\alpha}{dt} + \beta\,\frac{d\beta}{dt} + \gamma\,\frac{d\gamma}{dt} = 0,\quad \alpha'\,\frac{d\alpha'}{dt} + \beta'\,\frac{d\beta'}{dt} + \gamma'\,\frac{d\gamma'}{dt} = 0,$$

$$\alpha''\,\frac{d\alpha''}{dt} + \beta''\,\frac{d\beta''}{dt} + \gamma''\,\frac{d\gamma''}{dt} = 0,$$

(3)

$$\frac{d\alpha'}{dt}\,\alpha'' + \frac{d\beta'}{dt}\,\beta'' + \frac{d\gamma'}{dt}\,\gamma'' = -\left(\frac{d\alpha''}{dt}\,\alpha' + \frac{d\beta''}{dt}\,\beta' + \frac{d\gamma''}{dt}\,\gamma'\right),$$

$$\frac{d\alpha''}{dt}\,\alpha + \frac{d\beta''}{dt}\,\beta + \frac{d\gamma''}{dt}\,\gamma = -\left(\frac{d\alpha}{dt}\,\alpha'' + \frac{d\beta}{dt}\,\beta'' + \frac{d\gamma}{dt}\,\gamma''\right),$$

$$\frac{d\alpha}{dt}\,\alpha' + \frac{d\beta}{dt}\,\beta' + \frac{d\gamma}{dt}\,\gamma' = -\left(\frac{d\alpha'}{dt}\,\alpha + \frac{d\beta'}{dt}\,\beta + \frac{d\gamma'}{dt}\,\gamma\right).$$

Consequently, denoting these last three quantities by p, q, r,

(4)

$$\xi' = \frac{\partial\xi}{dt} - r\eta + q\zeta,$$

$$\eta' = r\xi + \frac{\partial\eta}{dt} - p\zeta,$$

$$\zeta' = -q\xi + p\eta + \frac{\partial\zeta}{\partial t}.$$

By a very similar procedure we obtain, from the equation (2),

(5)
$$\begin{aligned}
\frac{\partial^2 x}{\partial t^2}\,\alpha + \frac{\partial^2 y}{\partial t^2}\,\beta + \frac{\partial^2 z}{\partial t^2}\,\gamma &= \frac{\partial \xi'}{\partial t} - r\eta' + q\zeta',\\
\frac{\partial^2 x}{\partial t^2}\,\alpha' + \frac{\partial^2 y}{\partial t^2}\,\beta' + \frac{\partial^2 z}{\partial t^2}\,\gamma' &= r\xi' + \frac{\partial \eta'}{\partial t} - p\zeta',\\
\frac{\partial^2 x}{\partial t^2}\,\alpha'' + \frac{\partial^2 y}{\partial t^2}\,\beta'' + \frac{\partial^2 z}{\partial t^2}\,\gamma'' &= -q\xi' + p\eta' + \frac{\partial \zeta'}{\partial t}.
\end{aligned}$$

From equations (3) of Section 1, denoting by p_1, q_1, r_1 the quantities that depend on the functions $\alpha_1, \beta_1, \ldots, \gamma_1''$ in the same way as p, q, r on the functions $\alpha, \beta, \ldots, \gamma''$,

(6)
$$\begin{aligned}
\frac{\partial}{\partial t}\left(\frac{\xi}{a}\right) &= r_1\,\frac{\eta}{b} - q_1\,\frac{\zeta}{c},\\
\frac{\partial}{\partial t}\left(\frac{\eta}{b}\right) &= p_1\,\frac{\zeta}{c} - r_1\,\frac{\xi}{a},\\
\frac{\partial}{\partial t}\left(\frac{\zeta}{c}\right) &= q_1\,\frac{\xi}{a} - p_1\,\frac{\eta}{b}.
\end{aligned}$$

If we substitute the values of $\frac{\partial \xi}{\partial t}, \frac{\partial \eta}{\partial t}, \frac{\partial \zeta}{\partial t}$ from (6) into (4), we obtain

(7)
$$\begin{aligned}
\xi' &= \frac{da}{dt}\,\frac{\xi}{a} + (ar_1 - br)\,\frac{\eta}{b} + (cq - aq_1)\,\frac{\zeta}{c},\\
\eta' &= (ar - br_1)\,\frac{\xi}{a} + \frac{db}{dt}\,\frac{\eta}{b} + (bp_1 - cp)\,\frac{\zeta}{c},\\
\zeta' &= (cq_1 - aq)\,\frac{\xi}{a} + (bp - cp_1)\,\frac{\eta}{b} + \frac{dc}{dt}\,\frac{\zeta}{c}.
\end{aligned}$$

It is readily apparent that the geometric significance of these quantities is as follows: ξ', η', ζ' are the velocity components of the point x, y, z of the fluid mass parallel to the axes of ξ, η, ζ; $\frac{\partial \xi}{\partial t}, \frac{\partial \eta}{\partial t}, \frac{\partial \zeta}{\partial t}$ the relative velocities, decomposed in the same way, for the coordinate system ξ, η, ζ. Further, in (1) the quantities on the left side are the accelerations, those on the right side the accelerating forces parallel to these axes. Finally, p, q, r are the instantaneous rotations of the coordinate system ξ, η, ζ about its axes and p_1, q_1, r_1 have the same significance for the coordinate system ξ_1, η_1, ζ_1.

3.

We now substitute the values of the quantities ξ', η', ζ' into the equations (5) from (7), and, using the equations (6), express the derivatives of $\frac{\xi}{a}, \frac{\eta}{b}, \frac{\zeta}{c}$ in terms of ξ, η, ζ. The quantities on the left side of the equation (1) take the form of linear expressions in ξ, η, ζ. On the right side, V has the form

$$H - A\xi^2 - B\eta^2 - C\zeta^2,$$

where H, A, B, C depend on the quantities a, b, c in a known way. Accordingly, we can satisfy these equations when the pressure on the surface has the constant value Q, by taking

$$P = Q + \sigma\left(1 - \frac{\xi^2}{a^2} - \frac{\eta^2}{b^2} - \frac{\zeta^2}{c^2}\right)$$

and determining the ten functions a, b, c; p, q, r; p_1, q_1, r_1 and σ of time so that the nine coefficients of ξ, η, ζ are equal on each side, while the incompressibility condition $abc = a_0 b_0 c_0$ is satisfied. Equating coefficients of $\frac{\xi}{a}, \frac{\eta}{b}$ in the first, and of $\frac{\xi}{a}$ in the second equation yields

$$\frac{d^2a}{dt^2} + 2brr_1 + 2cqq_1 - a(r^2 + r_1^2 + q^2 + q_1^2) = 2\,\frac{\sigma}{a} - 2\epsilon\, aA,$$

$$a\,\frac{dr}{dt} - b\,\frac{dr_1}{dt} + 2\,\frac{da}{dt}\,r - 2\,\frac{db}{dt}\,r_1 + apq + bp_1q_1 - 2cpq_1 = 0,$$

$$a\,\frac{dr_1}{dt} - b\,\frac{dr}{dt} + 2\,\frac{da}{dt}\,r_1 - 2\,\frac{db}{dt}\,r + ap_1q_1 + bpq - 2cp_1q = 0.$$

From these equations, we obtain six others by cyclic permutation of the axes, or alternatively by arbitrary interchanges. Here we note that in an interchange of two axes not only the corresponding quantities are interchanged, but the six quantities $p, q, \ldots, r_1$ change signs.

We can give these six equations a more convenient form for the subsequent investigation by introducing, in place of p, p_1; q, q_1; r, r_1, the half-sums and half-differences

$$u = \frac{p + p_1}{2}\,, \quad v = \frac{q + q_1}{2}\,, \quad w = \frac{r + r_1}{2};$$
$$u' = \frac{p - p_1}{2}\,, \quad v' = \frac{q - q_1}{2}\,, \quad w' = \frac{r - r_1}{2}$$

as unknown functions.

In this way the system of equations that the ten unknown functions of time must satisfy becomes

$$(\alpha)\quad \begin{cases} (a-c)v^2+(a+c)v'^2+(a-b)w^2+(a+b)w'^2-\dfrac{1}{2}\dfrac{d^2a}{dt^2}=\epsilon\, aA-\dfrac{\sigma}{a}, \\ (b-a)w^2+(b+a)w'^2+(b-c)u^2+(b+c)u'^2-\dfrac{1}{2}\dfrac{d^2b}{dt^2}=\epsilon\, bB-\dfrac{\sigma}{b}, \\ (c-b)u^2+(c+b)u'^2+(c-a)v^2+(c+a)v'^2-\dfrac{1}{2}\dfrac{d^2c}{dt^2}=\epsilon\, cC-\dfrac{\sigma}{c}, \\ (b-c)\dfrac{du}{dt}+2\dfrac{d(b-c)}{dt}u+(b+c-2a)vw+(b+c+2a)v'w'=0, \\ (b+c)\dfrac{du'}{dt}+2\dfrac{d(b+c)}{dt}u'+(b-c+2a)vw'+(b-c-2a)v'w=0, \\ (c-a)\dfrac{dv}{dt}+2\dfrac{d}{dt}(c-a)v+(c+a-2b)wu+(c+a+2b)w'u'=0, \\ (c+a)\dfrac{dv'}{dt}+2\dfrac{d(c+a)}{dt}v'+(c-a+2b)wu'+(c-a-2b)w'u=0, \\ (a-b)\dfrac{dw}{dt}+2\dfrac{d}{dt}(a-b)w+(a+b-2c)uv+(a+b+2c)u'v'=0, \\ (a+b)\dfrac{dw'}{dt}+2\dfrac{d}{dt}(a+b)w'+(a-b+2c)uv'+(a-b-2c)u'v=0, \\ abc=a_0b_0c_0. \end{cases}$$

The values of A, B, C follow from the known expression for V,

$$V=H-A\xi^2-B\eta^2-C\zeta^2=\pi\int_0^\infty \frac{ds}{\Delta}\left(1-\frac{\xi^2}{a^2+s}-\frac{\eta^2}{b^2+s}-\frac{\zeta^2}{c^2+s}\right),$$

where

$$\Delta=\sqrt{\left(1+\frac{s}{a^2}\right)\left(1+\frac{s}{b^2}\right)\left(1+\frac{s}{c^2}\right)}.$$

After carrying out the solution of these differential equations, we now have to find the general solution $\theta, \theta', \theta''$ of the differential equations

$$(\beta)\qquad \frac{d\theta}{dt}=r\theta'-q\theta'',\quad \frac{d\theta'}{dt}=-r\theta+p\theta'',\quad \frac{d\theta''}{dt}=q\theta-p\theta'$$

in order to determine the functions $\alpha, \beta, \ldots, \gamma''$.

By (3) of Section 2, $\alpha, \alpha', \alpha''$; β, β', β''; $\gamma, \gamma', \gamma''$ are the three particular solutions of (β) having the values 1, 0, 0; 0, 1, 0; 0, 0, 1 at $t = 0$. To determine the functions $\alpha_1, \beta_1, \ldots, \gamma_1''$, we require the general solution of the simultaneous differential equations

$$(\gamma) \qquad \frac{d\theta_1}{dt} = r_1\theta_1' - q_1\theta_1'' \,, \quad \frac{d\theta_1'}{dt} = -r_1\theta_1 + p_1\theta_1'', \,, \quad \frac{d\theta_1''}{dt} = q_1\theta_1 - p_1\theta_1'.$$

4.

We may now ask what methods for the integration of the differential equations $(\alpha), (\beta), (\gamma)$ are presented by the general hydrodynamical principles used by Dirichlet to extract the seven integrals of first order (§1(a)) from the differential equations to be satisfied by the functions $\ell, m, \ldots, n''$. The equations that follow from these may easily be derived with the help of the expressions for ζ', η', ζ' given above.

The principle of conservation of area yields

$$(1) \qquad \begin{aligned} (b-c)^2u + (b+c)^2u' &= g = \alpha g^0 + \beta h^0 + \gamma k^0, \\ (c-a)^2v + (c+a)^2v' &= h = \alpha' g^0 + \beta' h^0 + \gamma' k^0, \\ (a-b)^2w + (a+b)^2w' &= k = \alpha'' g^0 + \beta'' h^0 + \gamma'' k^0, \end{aligned}$$

where the constants g^0, h^0, k^0 are the initial values of g, h, k, and coincide with the constants $\mathfrak{K}, \mathfrak{K}', \mathfrak{K}''$ in Dirichlet's work. It follows that $\theta = g$, $\theta' = h$, $\theta'' = k$ is a solution of the differential equations (β), which may readily be confirmed from the last six differential equations (α).

From Helmholtz's principle of conservation of rotation, we obtain the equations

$$(2) \qquad \begin{aligned} (b-c)^2u - (b+c)^2u' &= g_1 = \alpha_1\gamma_1^0 + \beta_1 h_1^0 + \gamma_1 k_1^0, \\ (c-a)^2v - (c+a)^2v' &= h_1 = \alpha_1' g_1^0 + \beta_1' h_1^0 + \gamma_1' k_1^0,, \\ (a-b)^2w - (a+b)^2w' &= k_1 = \alpha_1'' g_1^0 + \beta_1'' h_1^0 + \gamma_1'' k_1^0. \end{aligned}$$

The constants g_1^0, h_1^0, k_1^0 are equal to the quantities $BC\mathfrak{A}, CA\mathfrak{B}, AB\mathfrak{C}$ in the work mentioned above.

Finally, the principle of conservation of kinetic energy yields an integral of first order of the differential equations (α):

$$\text{(I)} \qquad \left.\begin{cases} \frac{1}{2}\left(\left(\frac{da}{dt}\right)^2 + \left(\frac{db}{dt}\right)^2 + \left(\frac{dc}{dt}\right)^2\right) \\ +(b-c)^2u^2 + (c-a)^2v^2 + (a-b)^2w^2 \\ +(b+c)^2u'^2 + (c+a)^2v'^2 + (a+b)^2w'^2 \end{cases}\right\} = 2\epsilon H + \text{ const.}$$

From the equations (1) and (2), two integrals of the equations (α) follow at once:

$$\text{(II)} \qquad g^2 + h^2 + k^2 = \text{const.} = \omega^2$$

$$\text{(III)} \qquad g_1^2 + h_1^2 + k_1^2 = \text{const.} = \omega_1^2.$$

Further, two integrals of the equations (β) occur,

$$\text{(IV)} \qquad \theta^2 + \theta'^2 + \theta''^2 = \text{const.},$$

$$\text{(V)} \qquad \theta g + \theta' h + \theta'' k = \text{const.},$$

which permit their *general* integration to be reduced to a quadrature. Since the equations (β) are linear and homogeneous, for their general solution, we need only look for two *particular* solutions distinct from g, h, k. To this end, we may choose the arbitrary constants in the two integrated equations so that the calculation simplifies. Assigning both constants the value 0, one has

$$\text{(3)} \qquad \theta' h + \theta'' k = -g\theta.$$

We now square this equation and add to it the equation

$$-\theta'^2 - \theta''^2 = \theta^2$$

multiplied by $h^2 + k^2$, obtaining

$$-(\theta' k - \theta'' h)^2 = \omega^2\theta^2,$$

and consequently

$$\text{(4)} \qquad \theta' k - \theta'' h = \omega i \theta.$$

Solving the linear equations (3), (4), we find that

$$\theta' = \frac{-gh + k\omega i}{h^2 + k^2}\,\theta, \tag{5}$$

$$\theta'' = \frac{-gk - h\omega i}{h^2 + k^2}\,\theta. \tag{6}$$

Substituting these values in the first of the equations (β), we obtain

$$\frac{1}{\theta}\,\frac{d\theta}{dt} = \frac{-g}{h^2 + k^2}\,\frac{dg}{dt} + \frac{rk + qh}{h^2 + k^2}\,\omega i,$$

$$\log\theta = \frac{1}{2}\,\log(h^2 + k^2) + \omega i \int \frac{qh + rk}{h^2 + k^2}\,dt + \text{ const.} \tag{7}$$

From the solution of the differential equations (β) contained in (5), (6), and (7) we obtain a third, on replacing i by $-i$ throughout. Now it is easy to form the expressions for the functions $\alpha, \beta, \ldots, \gamma''$ from the three particular solutions we have found.

The geometric significance of each real solution of the differential equation (β) is that, multiplied by a suitable constant factor, it expresses the cosine of the angles that the axes of ξ, η, ζ make with a fixed line at time t. For the first of the three solutions found above, this fixed line is formed by the normal to the invariant plane of the whole moving body. For the real and imaginary parts of the other two solutions, the fixed lines are two perpendicular lines in this plane. Now the cosines of the angles between the axes and these normals are $\frac{g}{\omega}, \frac{h}{\omega}, \frac{k}{\omega}$. The position of the axes relative to these normals thus emerges by solving the equations (α), without further integration. For the complete determination of their position, a single quadrature suffices, for example the evaluation of $\omega \int_0^t \frac{qh+rk}{h^2+k^2}\,dt$, which gives the rotation of the plane formed by the normal and the ξ axis, around its normal.

Very similar considerations apply to the differential equations (γ). In the same way we can get the general solutions from the two integrals

$$\theta_1^2 + \theta_1'^2 + \theta_1''^2 = \text{const.} \tag{VI}$$

$$\theta_1 g_1 + \theta_1' h_1 + \theta_1'' k_1 = \text{const.} \tag{VII}$$

and so derive the values of the quantities $\alpha_1, \beta_1, \ldots, \gamma_1''$ at time t, only one quadrature being required. Finally, the position of an arbitrary fluid particle at time t emerges from the expressions for the quantities x, y, z and the functions $\ell, m, \ldots, n''$ given in (1), (4) of Section 1.

5.

We would now like to elicit what has been gained for the purposes of integration, by reducing the differential equations between the functions $\ell, m, \ldots, n''$ (equations (a) of §1 in Dirichlet) to our differential equations. The system of differential equations (a) has order sixteen, and we know first order integrals of seven of the equations, which enable us to reduce it to a system of order nine. The system of equations (α) is only of order ten, and we know first order integrals of three of them. By the transformation of these differential equations given here, the order of the system of differential equations still to be integrated is reduced by two, and in place of this we finally only have to carry out two quadratures. Thus the transformation has the same effect as the discovery of two integrals of first order.

We remark explicitly that our form of the differential equations is only to be preferred for integration and actual determination of the motion. For the most general investigations of this motion, on the other hand, this form of the differential equation is less suitable, not just because the derivation is less simple, but also because the case of the equality of two axes requires special treatment. Namely, with the equality of two axes, the special circumstance enters that the information the equations yield about the form of the fluid mass is incompletely determined. It depends in general also on the instantaneous motion, and only remains arbitrary when the motion is such that the axes are equal throughout. While the investigation of this case is always easy and accordingly does not demand elaboration, it can in particular cases take various particular forms. General investigations, for example the demonstration in general of the possibility of the motion (§2 in Dirichlet) become fairly extensive, in view of the number of cases requiring special treatment.

Before we proceed to the treatment of special cases in which the differential equations (α) can be integrated, it is convenient to observe that in a solution of these differential equations, however directly it proceeds from the form of the equations, changes in sign of the functions $u, v, \ldots, w'$ are permissible if they leave uvw, $u'vw'$, $u'vw'$, $u'v'w$ unchanged. Thus, firstly, we can simultaneously change the signs of the functions u', v', w'. In this way the quantities $\alpha, \beta, \ldots, \gamma''$ are interchanged with $\alpha_1, \beta_1, \ldots, \gamma_1''$, and thus in the system of quantities $\ell, m, \ldots, n''$, rows are interchanged with columns. Secondly, two of the pairs of quantities u, u'; v, v'; w, w' can simultaneously be given the opposite signs. This variation may be reduced to a sign change of a coordinate axis that takes a motion into a symmetrically similar one.

This remark contains the reciprocity theorem found by Dedekind.

6.

We now investigate the case in which one of the pairs of quantities u, u'; v, v'; w, w' is zero throughout, for example $u = u' = 0$. The geometric significance of this hypothesis is that the principal axis always lies in the invariant plane of the whole moving body, and the instantaneous axis of rotation is perpendicular to this principal axis.

From the last six differential equations (α), it follows at once that in this case the quantities

$$(\mu) \qquad (c-a)^2 v, (c+a)^2 v', (a-b)^2 w, (a+b)^2 w'$$

are constant, and the equations

$$(\nu) \qquad \begin{aligned} (b+c-2a)vw + (b+c+2a)v'w' &= 0, \\ (b-c+2a)vw' + (b-c-2a)v'w &= 0 \end{aligned}$$

must hold.

For further investigation, we distinguish whether or not a second of the three pairs is zero. In general, we observe that, as a consequence of equations (μ), the quantities h, k, h_1, k_1 are constant. Consequently, the angles between the principal axes and the invariant plane of the whole moving body are also constant. Moreover, the equations between ratios

$$\begin{aligned} g : h : k &= p : q : r, \\ g_1 : h_1 : k_1 &= p_1 : q_1 : r_1 \end{aligned}$$

follow from the differential equations (β) and (γ), which simplifies the solution of these equations.

First case. Only one of the three pairs $u, u'; v, v'; w, w'$ vanishes.

If neither v and v', nor w and w', vanish together, it follows from the equations (μ) and (ν) that

$$(1) \qquad \begin{aligned} \frac{v'^2}{v^2} &= \frac{(2a-b-c)(2a+b-c)}{(2a+b+c)(2a-b+c)} = \left(\frac{a-c}{a+c}\right)^4 \text{ const.}, \\ \frac{w'^2}{w^2} &= \frac{(2a-b-c)(2a-b+c)}{(2a+b+c)(2a+b-c)} = \left(\frac{a-b}{a+b}\right)^4 \text{ const.} \end{aligned}$$

Taking this together with

$$abc = \text{const.},$$

we see that a, b, c are constant, and consequently v, v', w, w' are constant.

Now let

$$(2)\qquad \begin{aligned} \frac{v^2}{(2a+b+c)(2a-b+c)} &= \frac{v'^2}{(2a-b-c)(2a+b-c)} = S, \\ \frac{w^2}{(2a+b+c)(2a+b-c)} &= \frac{w'^2}{(2a-b-c)(2a-b+c)} = T. \end{aligned}$$

From the first three differential equations (α), we obtain the three equations

$$(3)\qquad (4a^2 - b^2 - 3c^2)S + (4a^2 - 3b^2 - c^2)T = \frac{\epsilon A}{2} - \frac{\sigma}{2a^2},$$

$$(4)\qquad \begin{cases} (b^2 - c^2)T = \frac{\epsilon}{2}\, B - \frac{\sigma}{2b^2}, \\ (c^2 - b^2)S = \frac{\epsilon}{2}\, C - \frac{\sigma}{2c^2}. \end{cases}$$

To derive the values of S, T and σ, we form from equations (4) the equations

$$b^2 T + c^2 S = \frac{\epsilon\pi}{2} \int_0^\infty \frac{s\, ds}{\Delta (b^2 + s)(c^2 + s)},$$

$$T + S = \frac{\sigma}{2b^2c^2} - \frac{\epsilon\pi}{2} \int_0^\infty \frac{ds}{\Delta (b^2 + s)(c^2 + s)}.$$

Now substitute these values in the equation (3):

$$(4a^2 - b^2 - c^2)(T + S) - 2(b^2 T + c^2 S) = \frac{\epsilon A}{2} - \frac{\sigma}{2a^2},$$

which yields

$$(5)\qquad \frac{D\sigma}{2a^2b^2c^2} = \frac{\epsilon\pi}{2} \int_0^\infty \frac{ds}{\Delta} \left(\frac{2s + 4a^2 - b^2 - c^2}{(b^2 + s)(c^2 + s)} + \frac{1}{a^2 + s} \right).$$

Here, for brevity, we write

$$(6)\qquad 4a^4 - a^2(b^2 + c^2) + b^2c^2 = D.$$

Substituting the value of σ in the equations (4), we now obtain

$$(7)\qquad \frac{b^2-c^2}{b^2-a^2}DS=\frac{\epsilon\pi}{2}\int_0^\infty\frac{s\,ds}{\Delta(b^2+s)}\left(\frac{4a^2-c^2+b^2}{c^2+s}-\frac{b^2}{a^2+s}\right),$$

$$(8)\qquad \frac{c^2-b^2}{c^2-a^2}DT=\frac{\epsilon\pi}{2}\int_0^\infty\frac{s\,ds}{\Delta(c^2+s)}\left(\frac{4a^2-b^2+c^2}{b^2+s}-\frac{c^2}{a^2+s}\right).$$

It now only remains to investigate what conditions a, b, c must satisfy for real values of v, v', w, w' to emerge from the equations (7), (8) and the equations (2).

For $\left(\frac{v'}{v}\right)^2$ and $\left(\frac{w'}{w}\right)^2$ to be nonnegative, it is necessary and sufficient that

$$(4a^2-(b+c)^2)(4a^2-(b-c)^2)\geq 0.$$

Thus either $a^2\geq\left(\frac{b+c}{2}\right)^2$ or $a^2\leq\left(\frac{b-c}{2}\right)^2$.

If $a\geq\frac{b+c}{2}$, the quantities S, T must be ≥ 0 for the equations (2) to produce real values for v, v', w, w'. However, we can readily show that when $a\geq\frac{b+c}{2}$, D and the two integrals on the right in (7), (8) are always positive. We need only put D in the form

$$a^2(4a^2-(b+c)^2)+bc(2a^2+bc),$$

and the integral in (7) in the form

$$\frac{\epsilon\pi}{2a^2b^2c^2}\int_0^\infty\frac{s\,ds}{\Delta^2}\left((4a^2-c^2)s+a^2(4a^2+b^2-c^2)-b^2c^2\right).$$

We then note the following consequences of $a\geq\frac{b+c}{2}$:

$$4a^2-(b+c)^2\geq 0\ ,\ 4a^2-c^2>0,$$

and further

$$4a^2+b^2-c^2\geq(b+c)^2+b^2-c^2=2b(b+c),$$

so that

$$a^2(4a^2+b^2-c^2)\geq 2b(b+c)a^2\geq\frac{1}{2}b(b+c)^3>b^2c^2.$$

From these inequalities, it follows that both D and the integral in question have only positive terms. The same holds for the integral on the right side of (8) obtained by interchanging b, c. If we let a run over the values from $\frac{b+c}{2}$ to ∞, then if $b > c$, T will always remain positive; S only remains positive as long as $a < b$. Thus the conditions for this case, writing b for the greater of the axes b, c, are

$$\text{(I)} \qquad \frac{b+c}{2} \leq a \leq b.$$

For the investigation of the second case, when $a^2 \leq \left(\frac{b-c}{2}\right)^2$, we shall suppose that b is the greater of the axes b, c, so that $a \leq \frac{b-c}{2}$. For v, v', w, w' to be real, we must have $S \leq 0$, $T \geq 0$. Since it follows from the inequalities

$$b^2 \geq (2a+c)^2 > 4a^2 + c^2$$

that the integral on the right side of (8) is always negative in our case, the latter condition $T \geq 0$ is only fulfilled when $D(c^2 - a^2) \geq 0$. Thus either $c^2 < \frac{a^2(b^2-4a^2)}{b^2-a^2}$ or $c^2 \geq a^2$. This case thus subdivides into two cases. There is a finite gap between them, since $\frac{a^2(b^2-4a^2)}{b^2-a^2} < a^2$, and there is not a continuous transition from one case to the other. Now the integral in equation (7), as long as $c^2 \leq a^2$, can only be positive, because of the inequalities $c^2+s \leq a^2+s$, $4a^2 - c^2 + b^2 > b^2$. So the conditions to be fulfilled reduce in the first of these cases to $a \leq \frac{b-c}{2}$, or

$$\text{(II)} \qquad c \leq b - 2a \quad \text{and} \quad c^2 < \frac{a^2(b^2-4a^2)}{b^2-a^2}.$$

In the second case, the conditions are

$$\text{(III)} \qquad a \leq \frac{b-c}{2} \quad \text{and} \quad \int_0^\infty \frac{s\,ds}{\Delta(b^2+s)} \left(\frac{4a^2 - c^2 + b^2}{c^2+s} - \frac{b^2}{a^2+s} \right) \leq 0.$$

It is easy to see that when a runs through values from 0 to c, the integral on the left side of the last inequality remains negative as long as $a \leq \frac{c}{2}$, while it is positive for $a = c$. We see that the exact determination of the bounds in which this inequality is satisfied depends on the solution of a transcendental equation.

Regarding the sign of σ, which is known to determine whether or not the motion is possible without outside pressure, we may observe that the value for this quantity found above can be put in the form

$$\frac{\epsilon\pi}{D} \int_0^\infty \frac{3s^2 + 6a^2 s + D}{\Delta^3}\,ds.$$

Thus in cases I and III, where $D > 0$, it is certainly positive. However, for a negative value of D, at least if this value is taken to be below a certain bound, it becomes negative.

7.

Second case. Two of the pairs $u, u'; v, v'; w, w'$ vanish.

We now treat the case where two of the pairs u, u'; v, v'; w, w' are 0 throughout, so that there is only a rotation about one principal axis.

If besides u and u', v and v' are 0 throughout, the equations (μ) and (ν) reduce to

$$(a-b)^2 w = \text{const.} = \tau\ ,\ \ (a+b)^2 w' = \text{const.} = \tau'.$$

Now the first three differential equations (α) produce the equations

$$\frac{\tau^2}{(a-b)^3} + \frac{\tau'^2}{(a+b)^3} - \frac{1}{2}\frac{d^2a}{dt^2} = \epsilon a A - \frac{\sigma}{a},$$

(1) $$\frac{\tau^2}{(b-a)^3} + \frac{\tau'^2}{(b+a)^3} - \frac{1}{2}\frac{d^2b}{dt^2} = \epsilon b B - \frac{\sigma}{b},$$

$$-\frac{1}{2}\frac{d^2c}{dt^2} = \epsilon c C - \frac{\sigma}{c},$$

which, together with

$$abc = a_0 b_0 c_0,$$

determine the quantities a, b, c and σ as functions of time. The principle of conservation of kinetic energy yields the integral of first order for these differential equations:

(2) $$\frac{1}{2}\left(\left(\frac{da}{dt}\right)^2 + \left(\frac{db}{dt}\right)^2 + \left(\frac{dc}{dt}\right)^2\right) + \frac{\tau^2}{(a-b)^2} + \frac{\tau'^2}{(a+b)^2} = 2\epsilon H + \text{const.}$$

From this it follows at once that when τ is nonzero, the principal axes a and b cannot be equal.

From the cases already investigated by MacLaurin and Dirichlet, when $a = b$, one can determine an expression for the motion in closed form if a, b, c

are constant. In this case we obtain from (1), by eliminating σ, the two equations

$$
(3)\qquad
\begin{aligned}
\frac{\tau'^2}{(b+a)^3} + \frac{\tau^2}{(b-a)^3} &= \frac{\epsilon\pi}{b}\int_0^\infty \frac{ds}{\Delta}\,\frac{(b^2-c^2)s}{(b^2+s)(c^2+s)} = K,\\
\frac{\tau'^2}{(b+a)^3} - \frac{\tau^2}{(b-a)^3} &= \frac{\epsilon\pi}{a}\int_0^\infty \frac{ds}{\Delta}\,\frac{(a^2-c^2)s}{(a^2+s)(c^2+s)} = L,
\end{aligned}
$$

where K and L denote the integrals on the right sides. We may also write the equations in the form

$$
(4)\qquad w'^2 = \frac{\tau'^2}{(b+a)^4} = \frac{\epsilon\pi}{2}\int_0^\infty \frac{ds}{\Delta}\left(\frac{s+ab}{(a^2+s)(b^2+s)} - \frac{c^2}{ab(c^2+s)}\right),
$$

$$
(5)\qquad w^2 = \frac{\tau^2}{(b-a)^4} = \frac{\epsilon\pi}{2}\int_0^\infty \frac{ds}{\Delta}\left(\frac{s-ab}{(a^2+s)(b^2+s)} + \frac{c^2}{ab(c^2+s)}\right).
$$

We assume, as in the cases treated earlier, that b is the greater of the two axes a, b. These two equations produce positive values for τ^2 and τ'^2 if, and only if, K is positive and greater than the absolute value of L. It is clear that the first condition is fulfilled as long as $c < b$. The second condition is satisfied when $c = a$ and thus $L = 0$; and consequently, since K, L vary continuously with c, is satisfied inside a finite domain on either side of this value. This domain, however, does not reach the values b and 0. For if $c = b$, τ'^2 is negative, while for infinitely small c, τ^2 is negative, since then

$$
\frac{K}{c} = \epsilon\pi\int_0^\infty \frac{ds}{s^{1/2}(1+s)^{3/2}\left(1+\frac{b^2}{a^2}\,s\right)^{1/2}},
$$

$$
\frac{L}{c} = \epsilon\pi\int_0^\infty \frac{ds}{s^{1/2}(1+s)^{3/2}\left(1+\frac{a^2}{b^2}\,s\right)^{1/2}},
$$

and consequently $L > K$. If b tends to infinity while a and c remain finite, then L can only remain smaller than K if $a^2 - c^2$ also decreases infinitely. Both bounds for c then only differ from a by an infinitely small amount. When b approaches infinitely close to its lower bound a, the upper bound for c, where τ'^2 becomes 0, converges to a. The lower bound, however, converges towards a value for which the integral on the right side of (5) vanishes. To determine this value, we write $\frac{c}{a} = \sin\psi$, and recover the equation

$$
(-5 + 2\cos 2\psi + \cos 4\psi)(\pi - 2\psi) + 10\sin 2\psi + 2\sin 4\psi = 0.
$$

This has only one root between $\psi = 0$, $\psi = \pi/2$, giving

$$\frac{c}{a} = 0.303327\ldots.$$

Admittedly for $b = a$, c can take every value between 0 and b, since τ^2 will always be 0 on account of the factor $b - a$. We then obtain the case studied by MacLaurin. The two cases found by Jacobi and Dedekind emerge for $w^2 = w'^2$.

For $b = a$, the case just discussed coincides with case (I) of the previous section. If

$$\frac{w^2}{(b+c+2a)(b-c+2a)} = \frac{w'^2}{(b+c-2a)(b-c-2a)},$$

it coincides with case (III). Among the four cases now found in which the fluid ellipsoid does not change shape during the motion, these three cases vary continuously into one another, while case (II) remains isolated.

8.

The investigation of whether there are cases apart from these four, in which the principal axes remain constant during the motion, leads to a somewhat lengthy calculation. We indicate this only briefly, since it merely produces a negative result.

From the hypothesis that a, b, c are constant, we can readily show that σ is constant. We multiply the first three differential equations (α) by a, b, c respectively and add. Then we use equation I of Section 4, the principle of conservation of kinetic energy.

Differentiating these three equations, and substituting the values of $\frac{du}{dt}$, $\frac{du'}{dt}$, $\ldots$, $\frac{dw'}{dt}$ from the last six equations (α), we obtain the three equations

$$\text{(1)}\qquad \begin{aligned} (b-c)u(vw - v'w') + (b+c)u'(v'w - vw') &= 0,\\ (c-a)v(wu - w'u') + (c+a)v'(w'u - wu') &= 0,\\ (a-b)w(uv - u'v') + (a+b)w'(u'v - uv') &= 0. \end{aligned}$$

Each of these is a consequence of the others.

I. Now suppose none of the six quantities $u, u', \ldots, w'$ is 0. From these equations we infer the equality of the following three pairs, denoting their values by $2a', 2b', 2c'$:

$$(a-c)\frac{v}{v'} + (a+c)\frac{v'}{v} = (a-b)\frac{w}{w'} + (a+b)\frac{w'}{w} = 2a',$$
$$(b-a)\frac{w}{w'} + (b+a)\frac{w'}{w} = (b-c)\frac{u}{u'} + (b+c)\frac{u'}{u} = 2b',$$
$$(c-b)\frac{u}{u'} + (c+b)\frac{u'}{u} = (c-a)\frac{v}{v'} + (c+a)\frac{v'}{v} = 2c'.$$

From these equations, $a'^2 - b'^2 = a^2 - b^2$, $b'^2 - c'^2 = b^2 - c^2$. Now we may let

$$\theta = a^2 - a'^2 = b^2 - b'^2 = c^2 - c'^2,$$

and from the first three differential equations of (α) we have

$$2\pi a' = \text{const.}, 2\chi b' = \text{const.}, 2\rho c' = \text{const.}.$$

Here we write $vv' + ww' = \pi$, $ww' + uu' = \chi$, $uu' + vv' = \rho$ for brevity. We combine these equations with the equation

$$(a^2-b^2)(a^2-c^2)\pi + (b^2-a^2)(b^2-c^2)\chi + (c^2-a^2)(c^2-b^2)\rho$$
$$= \frac{1}{4}(\omega^2 - \omega_1^2),$$

which follows easily from equations II, III. It follows that θ, and thus $u, u', \ldots, w'$, must be constant if we exclude $a = b = c$. However we can now easily show that the last six equations (α) cannot be fulfilled. In this way we establish that $u, u', \ldots, w'$ cannot all be nonzero provided not all three axes are equal.

The case $a = b = c$ would lead to the case of a sphere at rest; u', v', w' turn out to be 0 while u, v, w remain arbitrary. So the position of the axes at each instant can be varied arbitrarily.

II. The only remaining hypothesis is that one of the quantities $u, u', \ldots, w'$ is 0. This always yields, as we shall see, the earlier case where one of the three pairs u, u'; v, v'; w, w' vanishes.

1. When one of u', v', w' is 0, for example $u' = 0$, the equations

$$(b-c)uvw = 0 \ , \ (b-c)uv'w' = 0$$

follow from (1). This leads to one of the following hypotheses: firstly, the case investigated earlier; secondly, $b = c$; thirdly, $v = 0, w' = 0$, or $v' = 0, w = 0$, which is not essentially different

If $b = c$, u remains arbitrary, and can thus be set equal to 0, which leads to the case discussed earlier.

If $v = 0, w' = 0$, we obtain

$$(b - c - 2a)uv'w = 0, \ (c + a - 2b)uv'w = 0, \ (a - b + 2c)uv'w = 0$$

from the differential equations (α). Adding the first two equations,

$$-(a + b)uv'w = 0.$$

So, apart from u', v, w', another of u, v, w' must be 0, leading to the earlier case again.

2. Finally, when one of u, v, w is 0, for example $u = 0$, it follows from the equations (1) that

$$u'v'w = 0 \ , \ u'vw' = 0.$$

These equations lead either to our earlier hypothesis; or to the hypothesis $u = v' = w' = 0$, which does not differ essentially from the case $u' = v = w' = 0$ just considered; or finally, to $u = v = w = 0$. However, under this hypothesis the differential equations (α) give

$$v'w' = w'u' = u'v' = 0.$$

Hence two more of the quantities u', v', w' must be 0, again producing the case treated earlier.

It has emerged that along with the invariance of form is necessarily associated invariance of the state of motion. In other words, whenever the fluid mass persists in forming the same body, the relative motion of all parts of the body remains constant. We can think of the absolute motion in space in these cases as composed of two simpler motions. We impart firstly an inner motion to the fluid mass, with the fluid particles moving in similar parallel ellipses perpendicular to one cross-section. Now impart to the whole system a uniform rotation around an axis lying in this cross-section. If the cross-section, as assumed above, is perpendicular to the principal axis a, the cosines of the angles between the axis of rotation and the principal axes are $0, \frac{h}{\omega}, \frac{k}{\omega}$; the period of rotation is $\frac{2\pi}{\sqrt{q^2+r^2}}$. Moreover $0, b\,\frac{h_1}{\omega_1}, c\,\frac{k_1}{\omega_1}$ are the coordinates of the endpoint of the instantaneous axis of rotation with respect to the principal axes. From the inner motion, the elliptical paths of the fluid particles are parallel to the tangent plane at this point of the ellipsoid, so that their centers lie in this axis of rotation. The particles move on these paths in such a way that radius vectors drawn from the centers sweep out equal areas in equal time periods, and run over their paths in time $\frac{2\pi}{\sqrt{q_1^2+r_1^2}}$.

9.

We now come back to the treatment of the motion of the fluid mass in the case where u, u'; v, v' are 0 throughout, so that there is only a rotation around one principal axis. We observe at once that the equations (1) of Section 7, giving the variation of the principal axes in this case, permit another mechanical interpretation. Namely, one can treat them as the equations of motion of a material particle (a, b, c) of mass 1, constrained to move upon a surface determined by the equation $abc =$ const. and driven by forces whose potential function is

$$\frac{\tau^2}{(a-b)^2} + \frac{\tau'^2}{(a+b)^2} - 2\epsilon H,$$

or the same quantity with the sign changed.

Denote this quantity by G. The equations of both motions may be put in the form

$$\frac{d^2a}{dt^2}\,\delta a + \frac{d^2b}{dt^2}\,\delta b + \frac{d^2c}{dt^2}\,\delta c + \delta G = 0 \tag{1}$$

for any infinitely small values of $\delta a, \delta b, \delta c$ that satisfy the condition $abc =$ const. The principle of conservation of energy yields

$$\frac{1}{2}\left(\left(\frac{da}{dt}\right)^2 + \left(\frac{db}{dt}\right)^2 + \left(\frac{dc}{dt}\right)^2\right) + G = \text{const.}$$

Accordingly, the part of the energy independent of the variation in form of the fluid mass is G.

In order for a, b, c, and consequently the shape and state of motion of the fluid ellipsoid to remain constant, and thus for $\frac{da}{dt}, \frac{db}{dt}, \frac{dc}{dt}$ to vanish, it is clearly necessary and sufficient that the first order variation of the function G of the variables a, b, c, satisfying $abc =$ const., vanishes. For this we may cite equations (3) or (4) and (5) of Section 7. However, this invariance of the state of motion will be unstable if the value of the function is not a minimum. There are then always arbitrarily small variations of the state of the fluid mass that effect a complete change of the motion.

The direct investigation of the second order variation, in the case when the first order variation of the function G vanishes, would be very involved. However, we may decide as follows the question of whether the function has a minimum value in this case.

It can readily be shown that the function, whatever values τ^2, τ'^2 and a, b, c may assume, must have a minimum for a system of values of the independent variables. This is an obvious consequence of the following three facts. Firstly, the function G, in the boundary case of infinitely small or infinitely large axes, must approach a non-negative limiting value. Secondly, values of a, b, c can always be given for which G becomes negative. Thirdly, G can never become negatively infinite. These three properties of G follow from known properties of the function H. The function H takes its greatest value in the case when the fluid mass assumes the shape of a sphere, namely the value $2\pi\rho^2$, where ρ denotes the radius of this sphere, so that $\rho = (abc)^{1/3}$. Further, H becomes infinitely small when one of the axes becomes infinitely large, and consequently at least one of the others becomes infinitely small. However, when b tends to infinity, Hb does not become infinitely small. Consequently in the function G, provided a does not go to infinity at the same time, the negative term finally outweighs the positive term.

If τ^2 is nonzero, then already among the values of a, b, c satisfying the condition $b > a$, there must be a set of values for which the function takes a minimum. For the three conditions above, from which the existence of a minimum follows, are fulfilled for this region, since G is also non-negative in the limiting case $a = b$.

Now we can further investigate how many solutions of equations (3) of Section 7 satisfy the condition of the vanishing of the first order variation. This investigation is easily carried out if we also treat the expressions for τ^2 and τ'^2 that emerge for complex values of a, b, c. However we cannot undertake this investigation in the present work and must be content to state the result that we shall need later.

If τ^2 is not 0, the equations (3) only admit one solution on either side of $b = a$. Thus the first order variation vanishes for only one system of values on each side of this equation, and the function G must take its minimum for this system of values. We denote this minimum by G^*.

If τ^2 is 0, the first order variation always vanishes for $b = a$, and a value of c that is equal to a for $\tau'^2 = 0$, and steadily decreases for increasing τ'^2.

For this system of values, the second order variation can easily be expressed as an aggregate of $(\delta a+\delta b)^2$, $(\delta a-\delta b)^2$, with the coefficient of $(\delta a+\delta b)^2$ always positive. For the function, as we know from the above investigations, takes its smallest value here, among all values it takes when $b=a$.

However, the coefficient of $(\delta a-\delta b)^2$ is

$$\frac{\epsilon\pi}{2}\int_0^\infty \frac{ds}{\Delta}\left(\frac{s-ab}{(a^2+s)(b^2+s)}+\frac{c^2}{ab(c^2+s)}\right)$$

which is only positive if

$$\frac{c}{a}>0.303327\ldots$$

and consequently $\tau'^2<\epsilon\pi\rho^4\,8.64004\ldots$. It becomes negative if $\frac{c}{a}$ goes beyond this value.

Thus the function G only has a minimum G^* for this system of values in the first case. The investigation of the equation (3) shows that in this case the first order variation vanishes only for this system of values. However, in the latter case, G has a saddle point; accordingly there must be two more systems of values for which a minimum G^* occurs. From the investigation of the equations (3), it follows that the first order variation vanishes for only two further systems of values, which can be obtained from each other by interchanging b, a.

From this investigation, then, we find that in the case, known since MacLaurin, of the rotation of a flattened ellipsoid of rotation turning around its shorter axis, the invariance of the state of motion is unstable as soon as the ratio of the smaller axis to the others is less than $0.303327\ldots$. With the smallest variation of the other two axes in this case, the shape and state of motion of the fluid mass would change completely, and a persistent oscillation of state would occur around the state corresponding to the minimum of the function G. This consists of a uniform rotation of an ellipsoid with distinct axes about the shorter axis, combined with a simultaneous inner motion of the particles in ellipses, similar to one another, perpendicular to the axis of rotation. The period of the latter rotation is equal to the period of revolution, so that each particle already returns to its original position after a half rotation of the ellipsoid.

10.

If the total energy of the system,

$$\frac{1}{2}\left(\left(\frac{da}{dt}\right)_0^2 + \left(\frac{db}{dt}\right)_0^2 + \left(\frac{dc}{dt}\right)_0^2\right) + G_0 = \Omega,$$

which obviously cannot be less than G^*, is negative, then the form of the ellipsoid must oscillate within a finite region bounded via the inequality $G \leq \Omega$.

We can readily investigate these oscillations in the case where $\Omega - G^*$ is treated as infinitely small.

We consider the value c in our function G to be taken from the equation $abc = a_0b_0c_0$. Now equation (1) of the previous section yields

$$\frac{d^2a}{dt^2} - \frac{c}{a}\frac{d^2c}{dt^2} + \frac{\partial G}{\partial a} = 0, \quad \frac{d^2b}{dt^2} - \frac{c}{b}\frac{d^2c}{dt^2} + \frac{\partial G}{\partial b} = 0.$$

Now the values of a, b, c can only vary by infinitely small amounts from the values corresponding to the minimum of G. Denoting the variations at time t by $\delta a, \delta b, \delta c$ and neglecting higher order terms, we connect these variations by the equations

$$\frac{\delta a}{a} + \frac{\delta b}{b} + \frac{\delta c}{c} = 0,$$

(1)
$$\frac{d^2\delta a}{dt^2} - \frac{c}{a}\frac{d^2\delta c}{dt^2} + \frac{\partial^2 G}{\partial a^2}\delta a + \frac{\partial^2 G}{\partial a \partial b}\delta b = 0,$$

$$\frac{d^2\delta b}{dt^2} - \frac{c}{b}\frac{d^2\delta c}{dt^2} + \frac{\partial^2 G}{\partial b^2}\delta b + \frac{\partial^2 G}{\partial a \partial b}\delta a = 0.$$

These are known to be satisfied when we take $\frac{d^2\delta a}{dt^2} = -\mu^2\delta a$, $\frac{d^2\delta b}{dt^2} = -\mu^2\delta b$, and thus also $\frac{d^2\delta c}{dt^2} = -\mu^2\delta c$, and then choose the constant μ^2 so that one of the equations is a consequence of the others. The last condition on μ^2 coincides with the condition that the quadratic expression in $\delta a, \delta b$,

$$2\delta^2 G - \mu^2(\delta a^2 + \delta b^2 + \delta c^2)$$

can be expressed as a square of a linear expression in these quantities. This condition is always satisfied by two positive values of μ^2, since $\delta^2 G$ and

$\delta a^2 + \delta b^2 + \delta c^2$ are essentially positive, and the two values become equal when $\delta^2 G$ and $\delta a^2 + \delta b^2 + \delta c^2$ differ only by a constant factor. These two values of μ^2 yield two solutions of the differential equations (1) for which $\delta a, \delta b, \delta c$ vary in proportion to a periodic function of time, of form $\sin(\mu t + \text{const.})$. We obtain the general solution of (1) by combining these solutions.

Each solution individually produces infinitely small periodic oscillations in the shape and state of motion. Admittedly we can only show from this that there are two types of oscillations, which become more nearly periodic, the smaller they are. However, we obtain the existence of finite periodic oscillations as follows.

If Ω is negative, then obviously a must take one and the same value more than once. We treat the motion starting from the instant when a takes such a value for the first time. The motion will be completely determined by the initial values of $\frac{da}{dt}, \frac{db}{dt}$, and b. Thus the values taken by these quantities, when a later on takes this value, are also functions of their initial values. These functions, taken together, we denote by χ. The motion will become periodic when their values are equal to the initial values. Because of the equation $abc = \text{const.}$ and the principle of conservation of kinetic energy, when b and $\frac{da}{dt}$ resume their initial values, then so do c, $\frac{db}{dt}$ and $\frac{dc}{dt}$. Thus there are only two conditions to fulfill. By taking the derivatives of the functions χ for the case of infinitely small oscillations, we can show that these conditions do not contradict one another and have real roots inside a finite domain.

For this case of periodic oscillations, a, b, c can be expressed as functions of time via Fourier series. Apart from the case treated by Dirichlet, the constants in the series can admittedly be determined only approximately. This can be done, for example, by extending the above expansion for the case of infinitely small oscillations to terms of higher order.

It seems worth the trouble to give at least a superficial treatment of this motion, which is the simplest in which shape and state of motion are constant. We now wish to extend the investigation, carried out in the previous section for the case of rotation around only one principal axis, to all the motions satisfying Dirichlet's hypothesis.

11.

For this purpose, we transform the differential equations (α) into a clearer form. In place of $u, v, \ldots, w'$ we introduce the quantities $g, h, \ldots k_1$.

We generalize the meaning of G; we now take G to denote the expression

$$\frac{1}{4}\left\{\left(\frac{g+g_1}{b-c}\right)^2+\left(\frac{h+h_1}{c-a}\right)^2+\left(\frac{k+k_1}{a-b}\right)^2+\left(\frac{g-g_1}{b+c}\right)^2+\left(\frac{h-h_1}{c+a}\right)^2\right.$$
$$\left.+\left(\frac{k-k_1}{a+b}\right)^2\right\}-2\epsilon\pi\int_0^\infty\frac{a_0b_0c_0\,ds}{\sqrt{(a^2+s)(b^2+s)(c^2+s)}},$$

the part of the energy that is independent of the variation in shape.

We now obtain

$$p=\frac{\partial G}{\partial g},\quad q=\frac{\partial G}{\partial h},\quad r=\frac{\partial G}{\partial k},$$
$$p_1=\frac{\partial G}{\partial g_1},\quad q_1=\frac{\partial G}{\partial h_1},\quad r_1=\frac{\partial G}{\partial k_1}.$$

The last six differential equations (α) can be written in the form

$$\frac{dg}{dt}=h\,\frac{\partial G}{\partial k}-k\,\frac{\partial G}{\partial h},\quad \frac{dg_1}{dt}=h_1\,\frac{\partial G}{\partial k_1}-k_1\,\frac{\partial G}{\partial h_1},$$
$$(1)\qquad \frac{dh}{dt}=k\,\frac{\partial G}{\partial g}-g\,\frac{\partial G}{\partial k},\quad \frac{dh_1}{dt}=k_1\,\frac{\partial G}{\partial g_1}-g_1\,\frac{\partial G}{\partial k_1},$$
$$\frac{dk}{dt}=g\,\frac{\partial G}{\partial h}-h\,\frac{\partial G}{\partial g},\quad \frac{dk_1}{dt}=g_1\,\frac{\partial G}{\partial h_1}-h_1\,\frac{\partial G}{\partial g_1}.$$

The first three equations of (α) become

$$(2)\qquad \frac{d^2a}{dt^2}+\frac{\partial G}{\partial a}-2\,\frac{\sigma}{a}=0,\quad \frac{d^2b}{dt^2}+\frac{\partial G}{\partial b}-2\,\frac{\sigma}{b}=0,\quad \frac{d^2c}{dt^2}+\frac{\partial G}{\partial c}-2\,\frac{\sigma}{c}=0.$$

We note also that, from equations II, when $\omega=0$, the three equations

$$g=0,\ h=0,\ k=0$$

follow. That is, these quantities always remain zero if they are zero initially. Naturally the same holds for the quantities g_1, h_1, k_1.

From the differential equations (1), (2) we now see easily that for the vanishing of the first order variation of the function G of the nine variables $a, b, \ldots, k_1$, connected by the three conditions

$$abc=\text{const.},\ g^2+h^2+k^2=\omega^2,\ g_1^2+h_1^2+k_1^2=\omega_1^2,$$

it is necessary and sufficient that

$$\frac{d^2a}{dt^2}, \frac{d^2b}{dt^2}, \frac{d^2c}{dt^2}, \frac{dg}{dt}, \ldots, \frac{dk_1}{dt}$$

are 0. Thus the shape and state of motion of the ellipsoid remain constant if $\frac{da}{dt}, \frac{db}{dt}, \frac{dc}{dt}$ are 0. Earlier we gave a full discussion of the cases in which this occurs. However, we now see readily that the function G must have a minimum for at least one set of values of the independent variables. For in the limiting case when the axes become infinitely large or infinitely small, G tends to a nonnegative limit and, as already seen, there are definite values of the independent variables for which it takes negative, but not negatively infinite, values. For the constant state of motion corresponding to such a minimum, it follows from the principle of conservation of kinetic energy that each infinitely small variation satisfying Dirichlet's hypothesis entails only infinitely small oscillations. In every other case, invariance of the shape and state of motion is unstable. The search for a state of motion corresponding to a minimum of G is important not only for the determination of the possible stable forms of a moving massive fluid body. It would also be basic to the solution of our differential equations via infinite series. Accordingly we now find the cases, among those where the first order variation vanishes, where the function G has a minimum. Now from each of the cases found earlier in which the ellipsoid maintains its form, by interchanging axes and varying the signs of $g, h, \ldots, k_1$ one obtains several sets of values of $a, b, \ldots, k_1$ that cause the variation of first order of G to vanish. We can combine these here, as the function G takes the same values for all these, and all are equivalent for our question.

Before treating the individual cases, we note that when ω or ω_1 is zero the investigation takes a particularly simple form, since g, h, k or g_1, h_1, k_1 drop out of the function G completely. The earlier investigation of constant states of motion gives only two essentially different cases in which one of these two quantities is 0. In the case discussed in Section 6, this can only occur if

$$\frac{w'^2}{w^2} = \frac{(2a-b-c)(2a-b+c)}{(2a+b+c)(2a+b-c)} = \left(\frac{a-b}{a+b}\right)^4.$$

Thus the expression

$$b^2c^2 + a^2b^2 + a^2c^2 - 3a^4, \tag{3}$$

which we denote by E, is 0, and it emerges that ω or ω_1 is actually 0. Solving the equation $E = 0$ for a produces only one positive root, lying between $\frac{b+c}{2}$ and b, and this can happen only in case (I). Apart from this case, there is still the case where ω or ω_1 is 0, examined in Section 7, where $\tau^2 = \tau'^2$.

We can now show at once that in cases (I), (II), (III), the function G cannot take a minimum value, since its value for fixed a, b, c can always be decreased by suitable variations of $g, h, \ldots, k_1$. Since g and g_1 are 0 and h, h_1, k, k_1, excluding the case $E = 0$, are nonzero, the variations of these quantities are connected by the equations

$$\delta g^2 + 2h\delta h + 2k\delta k = 0,\ \delta g_1^2 + 2h_1\delta h_1 + 2k_1\delta k_1 = 0.$$

The variation of G becomes

$$\frac{1}{4}\left(\left(\frac{\delta g + \delta g_1}{b-c}\right)^2 + \left(\frac{\delta g - \delta g_1}{b+c}\right)^2\right) + \frac{\partial G}{\partial h}\delta h + \frac{\partial G}{\partial k}\delta k + \frac{\partial G}{\partial h_1}\delta h_1 + \frac{\partial G}{\partial k_1}\delta k_1.$$

Since

$$\frac{\partial G}{\partial h} : \frac{\partial G}{\partial k} = h : k, \quad \frac{\partial G}{\partial h_1} : \frac{\partial G}{\partial k_1} = h_1 : k_1,$$

we have

$$\delta G = \frac{1}{4}\left(\left(\frac{\delta g + \delta g_1}{b-c}\right)^2 + \left(\frac{\delta g - \delta g_1}{b+c}\right)^2\right) - \frac{1}{2h}\frac{\partial G}{\partial h}\delta g^2 - \frac{1}{2h_1}\frac{\partial G}{\partial h_1}\delta g_1^2. \tag{4}$$

We form the determinant of this quadratic form in δg and δg_1, and substitute the values obtained in (1) of Section 6,

$$\begin{aligned} \frac{2h}{q} &= b^2 + c^2 - 2a^2 \pm \sqrt{(4a^2 - (b+c)^2)(4a^2 - (b-c)^2)}, \\ \frac{2h_1}{q_1} &= b^2 + c^2 - 2a^2 \mp \sqrt{(4a^2 - (b+c)^2)(4a^2 - (b-c)^2)} \end{aligned} \tag{5}$$

so that $\frac{hh_1}{qq_1} = E$. The determinant is found to be

$$\frac{3(a^2 - b^2)(a^2 - c^2)}{4E(b^2 - c^2)^2}.$$

It is positive in case (I), if $E < 0$, and in case (III), but negative in case (I), if $E = 0$, and in case (II). In the first two cases, the expression (4) can take both positive and negative values: in the other two cases, we find either only positive values, or only negative values. For $\delta g_1 = -\delta g$, δG has the value

$$\delta g^2 \left(\frac{1}{(b+c)^2} - \frac{b^2 + c^2 - 2a^2}{2E} \right).$$

Under the assumptions valid in these cases, this is always negative, as we readily see by rewriting it as

$$-\frac{(b^2 + c^2 - 2a^2)(b^2 + 4bc + c^2 + 2a^2) + (4a^2 - (b+c)^2)(4a^2 - (b-c)^2)}{4(b+c)^2 E} \delta g^2,$$

and observing that $b^2 + c^2 - 2a^2$ is always positive when $E \geq 0$.

If one of the two quantities ω or ω_1 vanishes, for example $\omega_1 = 0$, the conditions connecting $\delta g_1, \delta h_1, \delta k_1$ become

$$\delta g_1^2 + \delta h_1^2 + \delta k_1^2 = 0.$$

Consequently the expression for the variation of G reduces to

$$\delta G = \frac{1}{2} \left(\frac{b^2 + c^2}{(b^2 - c^2)^2} - \frac{q}{h} \right) \delta g^2.$$

We obtain from (5), since $\frac{2h_1}{q_1} = 0$,

$$\frac{h}{q} = b^2 + c^2 - 2a^2.$$

Substituting this value, we obtain

$$\delta G = -\frac{(b^2 + c^2)(4a^2 - (b+c)^2) + (b-c)^2(b^2 + 4bc + c^2)}{4(b^2 - c^2)^2(b^2 + c^2 - 2a^2)} \delta g^2.$$

This is negative, since in this case $b^2 + c^2 - 2a^2$ and $4a^2 - (b+c)^2$ are positive.

Thus in all these cases the function G has no minimum value. It remains to treat the case of Section 7, where we can altogether exclude the singular case where $b = a$ and $\tau'^2 > (\epsilon\pi\rho^4)8.64004\ldots$. If either ω^2 or ω_1^2 is 0, this case produces, for given values of the other quantities, only one constant state of motion for which $\tau^2 = \tau'^2$, and for these values G must take its minimum value. For a given pair of nonzero values of ω^2 and ω_1^2, however, this case produces two constant states of motion of the fluid mass, that pass into each other on interchanging τ^2, τ'^2. For we can determine τ^2 and τ'^2 from ω^2 and ω_1^2 using

$$\tau = \frac{\omega + \omega_1}{2}\ , \ \tau' = \frac{\omega - \omega_1}{2},$$

and vary the signs of ω and ω_1 arbitrarily.

We can easily show that, in the single case when ω and ω_1 have the same sign and so τ^2 has the largest value, no minimum of G occurs. The conditions on the variations of $g, h, \ldots, k_1$ are now

$$\delta g^2 + \delta h^2 + 2k\delta k = 0,\ \delta g_1^2 + \delta h_1^2 + 2k_1\delta k_1 = 0,$$

and the variation of G will be

$$\begin{aligned}
&\frac{1}{4}\left\{\left(\frac{\delta g + \delta g_1}{b-c}\right)^2 + \left(\frac{\delta h + \delta h_1}{c-a}\right)^2 + \left(\frac{\delta g - \delta g_1}{b+c}\right)^2 + \left(\frac{\delta h - \delta h_1}{c+a}\right)^2\right\} \\
&\quad - \frac{1}{4}\left\{\left(\frac{1+\frac{\omega_1}{\omega}}{(a-b)^2} + \frac{1-\frac{\omega_1}{\omega}}{(a+b)^2}\right)(\delta g^2 + \delta h^2)\right. \\
&\quad \left. + \left(\frac{1+\frac{\omega}{\omega_1}}{(a-b)^2} + \frac{1-\frac{\omega}{\omega_1}}{(a+b)^2}\right)(\delta g_1^2 + \delta h_1^2)\right\}.
\end{aligned}$$

However, this has a negative value when we take ω and ω_1 with the same sign and $\delta h = \delta h_1 = 0$, $\delta g_1 = -\delta g$. For this produces

$$\delta G = \left\{\frac{1}{(b+c)^2} - \frac{1}{(b+a)^2} + \left(\frac{1}{(b+a)^2} - \frac{1}{(b-a)^2}\right)\frac{(\omega+\omega_1)^2}{4\omega\omega_1}\right\}\delta g^2.$$

Here $\frac{1}{(b+a)^2} < \frac{1}{(b-a)^2}$. Also $\frac{1}{(b+c)^2} < \frac{1}{(b+a)^2}$, since for $c \le a$, (3) of Section 7 gives $\frac{\tau'^2}{(b+a)^2} \ge \frac{\tau^2}{(b-a)^2}$ and consequently $\tau'^2 > \tau^2$, so that τ^2 can only exceed τ'^2 if $c > a$.

Thus the function has no minimum in this case, and consequently must take its minimum in the single remaining case.

Thus this occurs for the motion treated in Section 7, if $\tau^2 \le \tau'^2$ (excluding the above singular case). In all other cases the persistence of the form and state of motion is unstable. However, in this case every infinitely small variation of the form and state of motion of the fluid mass, satisfying Dirichlet's hypothesis, produces only infinitely small oscillations. Admittedly this does not imply that the state of the fluid mass is stable in this case. The investigation of the conditions under which this occurs could indeed be carried out by known methods, since it leads to linear differential equations. We forgo the treatment of this question in the present work. It is merely a development of the beautiful reasoning that is a high point of Dirichlet's scientific activities.

XI.

On the vanishing of θ-functions.

(*Borchardt's Journal für reine und angewandte Mathematik*, vol. 65, 1865.)

The second part of my memoir on the theory of Abelian functions, which was published in the 54th volume of Borchardt's journal of mathematics, contains the proof of a theorem on the vanishing of theta-functions to which I shall refer, assuming the notation used previously to be familiar to the reader. Everything deduced in the memoir in brief indications of the applications of this theorem is based on the concept of defining a function by means of its discontinuities and infinities. As can easily be seen, these applications must form the foundation of the theory of Abelian functions. However in the theorem itself, and in its proof, insufficient account has been taken of the possibility that the theta-function may vanish *identically* (that is, for every value of the variable) when for its variables are substituted integrals of algebraic functions of a single variable. The purpose of the following short memoir is to remedy this defect.

In discussing θ-functions in an indefinite number of variables, the need for an abbreviated notation to represent a sequence such as

$$v_1, v_2, \ldots, v_m$$

makes itself felt as soon as the expression for v_ν via ν is complicated. We could use symbols analogous to summation and product signs, but such a notation would take too much space, and the symbols governed by the function sign would be difficult to print. I prefer therefore to represent

$$v_1, v_2, \ldots, v_m \text{ by } \begin{pmatrix} m & \\ \nu & (v_\nu) \\ 1 & \end{pmatrix}$$

and thus

$$\theta(v_1, v_2, \ldots, v_p) \text{ by } \theta \begin{pmatrix} p & \\ \nu & (v_\nu) \\ 1 & \end{pmatrix}.$$

1.

Suppose that, in the function $\theta(v_1, v_2, \ldots, v_m)$, we substitute for the variables v the p integrals $u_1 - e_1, u_2 - e_2, \ldots, u_p - e_p$ of algebraic functions that branch in the same way as the surface T. We obtain a function of z, which varies continuously throughout the whole surface T outside the lines b, but acquires the factor $\exp(-u_\nu^+ - u_\nu^- + 2e_\nu)$ in crossing from the negative to the positive side of the line b_ν. As proved in §22, this function, if it does not vanish for all values of z, becomes infinitely small of first order at only p points of the surface T. These points were denoted by $\eta_1, \eta_2, \ldots, \eta_p$, and the value of the function u_ν at the point η_μ by $\alpha_\nu^{(\mu)}$. This led to the congruence for the $2p$ systems of moduli of the θ-function:

$$(1)\quad (e_1, e_2, \ldots, e_p) \equiv \left(\sum_1^p \alpha_1^{(\mu)} + K_1, \sum_1^p \alpha_2^{(\mu)} + K_2, \ldots, \sum_1^p \alpha_p^{(\mu)} + K_p\right).$$

The quantities K depend on the (at present arbitrary) additive constants in the functions u, but are independent of the quantities e and the points η.

If we perform the calculation indicated there, we find that

$$(2)\qquad 2K_\nu = \sum \frac{1}{\pi i} \int (u_\nu^+ + u_\nu^-) du_{\nu'} - \epsilon_\nu \pi i - \sum_{\mu=1}^p \epsilon'_\mu a_{\mu,\nu}.$$

In this expression, the integral $\int (u_\nu^+ + u_\nu^-) du_{\nu'}$ is taken in a positive direction over $b_{\nu'}$ and, in the summation, ν' runs through the natural numbers from 1 to p other than ν. Further, $\epsilon_\nu = \pm 1$, depending on whether the end of the line ℓ_ν lies on the positive or negative side of a_ν, while $\epsilon'_\nu = \pm 1$, depending on whether this same point lies on the positive or negative side of b_ν. The determination of the sign is, incidentally, necessary only if the quantities e are to be completely determined from the discontinuities of $\log \theta$ in accordance with the equations of §22. The above congruence (1) remains valid, whatever sign we take.

We shall first retain the simplifying assumption made there, that the additive constants in the functions u are so determined that the quantities K are all 0. In order to free the final results from this restrictive hypothesis, obviously all that will then be needed will be to add back $-K_1, -K_2, \ldots, -K_p$ to the arguments of the theta-function.

If, therefore, the function $\theta(u_1 - e_1, u_2 - e_2, \ldots, u_p - e_p)$ vanishes at each of the p points $\eta_1, \eta_2, \ldots, \eta_p$ *and does not vanish identically for every value* *of* z, then

$$(e_1, e_2, \ldots, e_p) \equiv \left(\sum_1^p \alpha_1^{(\mu)}, \sum_1^p \alpha_2^{(\mu)}, \ldots, \sum_1^p \alpha_p^{(\mu)}\right).$$

This theorem holds for arbitrary values of the quantities e, and we deduced from this, by allowing the point (s, z) to coincide with the point η_p, that

$$\theta\left(-\sum_1^{p-1} \alpha_1^{(\mu)}, -\sum_1^{p-1} \alpha_2^{(\mu)}, \ldots, -\sum_1^{p-1} \alpha_p^{(\mu)}\right) = 0.$$

Since the θ-function is even,

$$\theta\left(\sum_1^{p-1} \alpha_1^{(\mu)}, \sum_1^{p-1} \alpha_2^{(\mu)}, \ldots, \sum_1^{p-1} \alpha_p^{(\mu)}\right) = 0,$$

no matter what the points $\eta_1, \eta_2, \ldots, \eta_{p-1}$ might be.

2.

However, the proof of this theorem requires, for completeness, some amplification due to the circumstance that the function

$$\theta(u_1 - e_1, u_2 - e_2, \ldots, u_p - e_p)$$

may vanish identically (and in fact this can happen for every system of similarly branching algebraic functions, for certain values of the quantities e).

Because of this, we must initially be content to show that the theorem remains valid, when the points η vary independently within finite limits. The general validity of the theorem will then follow by virtue of the principle that a function of a complex variable cannot vanish throughout a finite region without being zero everywhere.

When z is given, the numbers $e_1, e_2, \ldots, e_p$ can always be chosen so that

$$\theta(u_1 - e_1, u_2 - e_2, \ldots, u_p - e_p)$$

does not vanish. For otherwise the function $\theta(v_1, \ldots, v_p)$ would vanish for every value of the variables v, and every coefficient in the development of the function in integral powers of $e^{2v_1}, e^{2v_2}, \ldots, e^{2v_p}$ would be zero, which is not the case. The quantities e can therefore be varied independently from one another within a finite region, without the function

$$\theta(u_1 - e_1, u_2 - e_2, \ldots, u_p - e_p)$$

vanishing for the given value of z. In other words, one can always find a $2p$-dimensional region E, within which the system of quantities e can move without the function

$$\theta(u_1 - e_1, u_2 - e_2, \ldots, u_p - e_p)$$

vanishing for this value of z. It will therefore become infinitely small of first order at only p different positions of (s, z). If we denote these points by $\eta_1, \ldots, \eta_p$, we have

(1) $$(e_1, e_2, \ldots, e_p) \equiv \left(\sum_1^p \alpha_1^{(\mu)}, \sum_1^p \alpha_2^{(\mu)}, \ldots, \sum_1^p \alpha_p^{(\mu)}\right).$$

To every choice of the system of numbers e inside E, or every point of E, corresponds a determination of the points η, whose totality form the points of a region H corresponding to the region E. However, in consequence of the equation (1), to each point of H corresponds only one point of E. If, therefore, H were a region of $2p-1$ or fewer dimensions, then E could not have $2p$ dimensions. Hence H is a $2p$-dimensional region.

The reasoning on which our theorem is based thus remains valid for arbitrary positions of the points η, within a finite region, and the equation

$$\theta\left(-\sum_1^{p-1} \alpha_1^{(\mu)}, -\sum_1^{p-1} \alpha_2^{(\mu)}, \ldots, -\sum_1^{p-1} \alpha_p^{(\mu)}\right) = 0$$

holds for any positions of the points $\eta_1, \eta_2, \ldots, \eta_{p-1}$ in the interior of finite regions, and consequently in general.

3.

It follows that the system of quantities $(e_1, e_2, \ldots, e_p)$ is congruent, in one and only one way, to an expression of the form

$$\begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix} \left(\sum_1^p \alpha_\nu^{(\mu)}\right)\Bigg), \text{ if } \theta \begin{pmatrix} p \\ \nu \quad (u_\nu - e_\nu) \\ 1 \end{pmatrix}$$

does not vanish for every value of z. For if the points $\eta_1, \eta_2, \ldots, \eta_p$ can be determined in more than one way so that the congruence

$$\begin{pmatrix} p \\ \nu \;(e_\nu) \\ 1 \end{pmatrix} \equiv \begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_1^p \alpha_\nu^{(\mu)}\right)$$

holds, then by the result just proved the function $\theta \begin{pmatrix} p \\ \nu \;(u_\nu - e_\nu) \\ 1 \end{pmatrix}$ would vanish for more than p points, without vanishing identically, which is impossible.

If $\theta \begin{pmatrix} p \\ \nu \;(u_\nu - e_\nu) \\ 1 \end{pmatrix}$ vanishes identically, we need, in order to obtain $\begin{pmatrix} p \\ \nu \;(e_\nu) \\ 1 \end{pmatrix}$ in the above form, to consider

$$\theta \begin{pmatrix} p \\ \nu \;(u_\nu + \alpha_\nu^{(1)} - u_\nu^{(1)} - e_\nu) \\ 1 \end{pmatrix}$$

and if this function vanishes identically for every value of z, ζ_1, z_1, to consider

$$\theta \begin{pmatrix} p \\ \mu \;(u_\nu + \sum_1^2 \alpha_\nu^{(\mu)} - \sum_1^2 u_\nu^{(\mu)} - e_\nu) \\ 1 \end{pmatrix}.$$

We suppose that

(1) $$\begin{cases} \theta \begin{pmatrix} p \\ \nu \left(\sum_1^m \alpha_\nu^{(p+2-\mu)} - \sum_1^{m-1} u_\nu^{(p-\mu)} - e_\nu\right) \\ 1 \end{pmatrix} \\ \text{vanishes identically, but that} \\ \theta \begin{pmatrix} p \\ \nu \left(\sum_1^{m+1} \alpha_\nu^{(p+2-\mu)} - \sum_1^{m} u_\nu^{(p-\mu)} - e_\nu\right) \\ 1 \end{pmatrix} \\ \text{does not.} \end{cases}$$

Hence the latter function, regarded as a function of ζ_{p+1}, vanishes not only for $\epsilon_{p-1}, \epsilon_{p-2}, \ldots, \epsilon_{p-m}$, but for $p - m$ other points which we may denote by

$\eta_1, \eta_2, \ldots, \eta_{p-m}$, so that

$$\begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(-\sum_{p-m+1}^{p} \alpha_\nu^{(\mu)} + e_\nu\right)\Bigg) \equiv \begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_{1}^{p-m} \alpha_\nu^{(\mu)}\right)\Bigg) .$$

These points $\eta_1, \eta_2, \ldots, \eta_{p-m}$ can be determined in only one way so that this congruence holds, because otherwise the function would vanish at more than p points. This same function, regarded as a function of z_{p-1}, besides vanishing for

$$\eta_{p+1}, \eta_p, \ldots, \eta_{p-m+1},$$

vanishes for $p-m-1$ other points, which we may denote by

$$\epsilon_1, \epsilon_2, \ldots, \epsilon_{p-m-1}.$$

We then have

$$\begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(-\sum_{p-m}^{p-2} u_\nu^{(\mu)} - e_\nu\right)\Bigg) \equiv \begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_{1}^{p-m-1} u_\nu^{(\mu)}\right)\Bigg) ,$$

and the points $\epsilon_1, \epsilon_2, \ldots, \epsilon_{p-m-1}$ are completely determined by this congruence.

Under the assumptions (1), then, the two congruences

$$\text{(2)} \qquad \begin{pmatrix} p \\ \nu\ (e_\nu) \\ 1 \end{pmatrix} \equiv \begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_{1}^{p} \alpha_\nu^{(\mu)}\right)\Bigg)$$

and

$$\text{(3)} \qquad \begin{pmatrix} p \\ \nu\ (-e_\nu) \\ 1 \end{pmatrix} \equiv \begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_{1}^{p-2} u_\nu^{(\mu)}\right)\Bigg)$$

can both be satisfied by choosing m of the points η and $m-1$ of the points ϵ arbitrarily, after which the remaining points are determined. The converse theorem is also true, that is, the function vanishes if one of these conditions is satisfied. Accordingly, if the congruence (2) can be satisfied in more than one way, then the congruence (3) also has more than one solution, and if m, but not more, of the points η can be chosen arbitrarily, then $m-1$ of the points ϵ can be chosen arbitrarily, after which the others are determined, and conversely.

It follows in the same way that if

$$\theta\begin{pmatrix} p \\ \nu\ (r_\nu) \\ 1 \end{pmatrix} = 0,$$

the congruences

(4)
$$\begin{pmatrix} p \\ \nu\ (r_\nu) \\ 1 \end{pmatrix} \equiv \begin{pmatrix} p \\ \nu \left(\sum_1^{p-1} \alpha_\nu^{(\mu)}\right) \\ 1 \end{pmatrix},$$

(5)
$$\begin{pmatrix} p \\ \nu\ (-r_\nu) \\ 1 \end{pmatrix} \equiv \begin{pmatrix} p \\ \nu \left(\sum_1^{p-1} u_\nu^{(\mu)}\right) \\ 1 \end{pmatrix}$$

always have a solution. In fact, m of the points η and m of the points ϵ can be chosen arbitrarily, after which the remaining $p-1-m$ points are determined, if

$$\theta\begin{pmatrix} p \\ \nu \left(\sum_1^{m} u_\nu^{(\mu)} - \sum_1^{m} \alpha_\nu^{(\mu)} + r_\nu\right) \\ 1 \end{pmatrix}$$

is identically zero, but

$$\theta\begin{pmatrix} p \\ \nu \left(\sum_1^{m+1} u_\nu^{(\mu)} - \sum_1^{m+1} \alpha_\nu^{(\mu)} + r_\nu\right) \\ 1 \end{pmatrix}$$

is not identically zero, the case $m=0$ not being excluded. This theorem has a converse. If m, but not more, of the points η can be chosen arbitrarily, then the hypothesis is satisfied and it follows that m, but not more, of the points ϵ can also be chosen arbitrarily.

4.

(1) Denote by θ'_ν the first derivative of $\theta(v_1, v_2, \ldots, v_p)$ with respect to v_ν, by $\theta''_{\nu,\mu}$ the second derivative with respect to v_ν and v_μ, and so on.

If

$$\theta\left(\begin{matrix}p\\ \nu\ (u_\nu^{(1)} - \alpha_\nu^{(1)} + r_\nu)\\ 1\end{matrix}\right)$$

vanishes identically for every value of z_1 and ζ_1, all of the functions $\theta'\left(\begin{matrix}p\\ \nu\ (r_\nu)\\ 1\end{matrix}\right)$ will be zero. In fact the equation

$$\theta\left(\begin{matrix}p\\ \nu\ (u_\nu^{(1)} - \alpha_\nu^{(1)} + r_\nu)\\ 1\end{matrix}\right) = 0,$$

when s_1 and z_1 are infinitely close to σ_1 and ζ_1, becomes the equation

$$\sum_1^p \theta'_\mu \left(\begin{matrix}p\\ \nu\ (r_\nu)\\ 1\end{matrix}\right) d\alpha_\mu^{(1)} = 0.$$

If we assume that

$$du_\mu = \frac{\phi_\mu(s,z)\,dz}{\dfrac{\partial F}{\partial s}},$$

then, after omission of the factor

$$\frac{d\zeta_1}{\dfrac{\partial F(\sigma_1,\zeta_1)}{\partial \sigma_1}},$$

the equation becomes

$$\sum_1^p \theta'_\mu \left(\begin{matrix}p\\ \nu\ (r_\nu)\\ 1\end{matrix}\right) \phi_\mu(\sigma_1,\zeta_1) = 0.$$

As no linear equation with constant coefficients can hold between the functions ϕ, it follows that all the first order partial derivatives of $\theta(v_1, v_2, \ldots, v_p)$ must vanish for $\begin{matrix}p\\ \nu\\ 1\end{matrix}\ (v_\nu = r_\nu)$.

In order to prove the converse theorem, assume that $\overset{p}{\underset{1}{\nu}}\ (v_\nu = r_\nu)$ and $\overset{p}{\underset{1}{\nu}}\ (v_\nu = t_\nu)$ are two systems of values for which the function θ vanishes, without vanishing identically for $\overset{p}{\underset{1}{\nu}}\ (v_\nu = u_\nu^{(1)} - \alpha_\nu^{(1)} + r_\nu)$ and $\overset{p}{\underset{1}{\nu}}\ (v_\nu = u_\nu^{(1)} - \alpha_\nu^{(1)} + t_\nu)$. We form the expression

$$(2)\qquad \frac{\theta\left(\overset{p}{\underset{1}{\nu}}\ (u_\nu^{(1)} - \alpha_\nu^{(1)} + r_\nu)\right)\theta\left(\overset{p}{\underset{1}{\nu}}\ (\alpha_\nu^{(1)} - u_\nu^{(1)} + r_\nu)\right)}{\theta\left(\overset{p}{\underset{1}{\nu}}\ (u_\nu^{(1)} - \alpha_\nu^{(1)} + t_\nu)\right)\theta\left(\overset{p}{\underset{1}{\nu}}\ (\alpha_\nu^{(1)} - u_\nu^{(1)} + t_\nu)\right)}.$$

If we regard this expression as a function of z_1, then it is an algebraic function of z_1 and indeed a rational function of s_1 and z_1, because the denominator and numerator of the fraction are continuous on T'' and acquire the same factors on crossing the transverse cuts. For $z_1 = \zeta_1$ and $s_1 = \sigma_1$, the denominator and numerator become infinitely small of second order, so that the function remains finite. However, the other values for which the denominator or numerator vanish are, as shown above, completely determined by the values of the quantities r and t, and thus independent of ζ_1. Now since an algebraic function is completely defined, apart from a constant factor, by the points at which it becomes infinite or zero, the expression is a rational function $\chi(s_1, z_1)$ of the variables s_1, z_1 independent of ζ_1, multiplied by a constant, that is, a quantity independent of z_1. As the expression is symmetric in regard to the systems (s_1, z_1) and (σ_1, ζ_1), this constant is $\chi(\sigma_1, \zeta_1)$ multiplied by a number A not depending on ζ_1 either. If we write

$$\sqrt{A}\,\chi(s, z) = \rho(s, z),$$

we obtain, for our expression (2), the value

$$(3)\qquad \rho(s_1, z_1)\,\rho(\sigma_1, \zeta_1)$$

where $\rho(s, z)$ is a rational function of s and z.

To determine this function, we merely need to let $\zeta_1 = z_1, \sigma_1 = s_1$, obtaining

$$\rho(s_1, z_1)^2 = \left\{ \frac{\sum_\mu \theta'_\mu \begin{pmatrix} p \\ \nu\ (r_\nu) \\ 1 \end{pmatrix} du_\mu^{(1)}}{\sum_\mu \theta'_\mu \begin{pmatrix} p \\ \nu\ (t_\nu) \\ 1 \end{pmatrix} du_\mu^{(1)}} \right\}^2$$

or, after extracting the square root, and removing the factor $\frac{dz_1}{\frac{\partial F(s_1, z_1)}{\partial s_1}}$,

$$\rho(s_1, z_1) = \pm \frac{\sum_\mu \theta'_\mu \begin{pmatrix} p \\ \nu\ (r_\nu) \\ 1 \end{pmatrix} \phi_\mu(s_1, z_1)}{\sum_\mu \theta'_\mu \begin{pmatrix} p \\ \nu\ (t_\nu) \\ 1 \end{pmatrix} \phi_\mu(s_1, z_1)}. \tag{4}$$

We now deduce from (3) and (4) the equation

$$\frac{\theta \begin{pmatrix} p \\ \nu\ (u_\nu^{(1)} - \alpha_\nu^{(1)} + r_\nu) \\ 1 \end{pmatrix} \theta \begin{pmatrix} p \\ \nu\ (\alpha_\nu^{(1)} - u_\nu^{(1)} + r_\nu) \\ 1 \end{pmatrix}}{\theta \begin{pmatrix} p \\ \nu\ (u_\nu^{(1)} - \alpha_\nu^{(1)} + t_\nu) \\ 1 \end{pmatrix} \theta \begin{pmatrix} p \\ \nu\ (\alpha_\nu^{(1)} - u_\nu^{(1)} + t_\nu) \\ 1 \end{pmatrix}} \tag{5}$$

$$= \frac{\sum_\mu \theta'_\mu \begin{pmatrix} p \\ \nu\ (r_\nu) \\ 1 \end{pmatrix} \phi_\mu(s_1, z_1) \sum_\mu \theta'_\mu \begin{pmatrix} p \\ \nu\ (r_\nu) \\ 1 \end{pmatrix} \phi_\mu(\sigma_1, \zeta_1)}{\sum_\mu \theta'_\mu \begin{pmatrix} p \\ \nu\ (t_\nu) \\ 1 \end{pmatrix} \phi_\mu(s_1, z_1) \sum_\mu \theta'_\mu \begin{pmatrix} p \\ \nu\ (t_\nu) \\ 1 \end{pmatrix} \phi_\mu(\sigma_1, \zeta_1)}.$$

It follows from this equation that

$$\theta \begin{pmatrix} p \\ \nu\ (u_\nu^{(1)} - \alpha_\nu^{(1)} + r_\nu) \\ 1 \end{pmatrix}$$

must be equal to zero for every value of z_1 and ζ_1, if all first order partial derivatives of the function $\theta(v_1, v_2, \ldots, v_p)$ vanish for $\overset{p}{\underset{1}{\nu}}\ (v_\nu = r_\nu)$.

5.

If

$$\theta\left(\overset{p}{\underset{1}{\nu}}\left(\sum_1^m \alpha_\nu^{(\mu)} - \sum_1^m u_\nu^{(\mu)} + r_\nu\right)\right) \tag{1}$$

vanishes identically, that is for every value of $\overset{m}{\underset{1}{\mu}}\ (\sigma_\mu, \zeta_\mu)$ and $\overset{m}{\underset{1}{\mu}}\ (s_\mu, z_\mu)$, then by taking $\zeta_m = z_m$, $\sigma_m = s_m$, we find—in the manner indicated above—that all first order partial derivatives of the function $\theta(v_1, v_2, \ldots, v_p)$ vanish for

$$\overset{p}{\underset{1}{\nu}}\left(v_\nu = \sum_1^{m-1} \alpha_\nu^{(\mu)} - \sum_1^{m-1} u_\nu^{(\mu)} + r_\nu\right).$$

Further, by allowing $\zeta_{m-1} - z_{m-1}$, $\sigma_{m-1} - s_{m-1}$ to become infinitely small, we find that all second order partial derivatives vanish for

$$\overset{p}{\underset{1}{\nu}}\left(v_\nu = \sum_1^{m-2} \alpha_\nu^{(\mu)} - \sum_1^{m-2} u_\nu^{(\mu)} + r_\nu\right).$$

Obviously we find in general that all nth order partial derivatives vanish for

$$\overset{p}{\underset{1}{\nu}}\left(v_\nu = \sum_1^{m-n} \alpha_\nu^{(\mu)} - \sum_1^{m-n} u_\nu^{(\mu)} + r_\nu\right),$$

whatever the values of the quantities z and ζ.

It follows that, on the present assumption (1), all partial derivatives of the function $\theta(v_1, v_2, \ldots, v_p)$ up to order m are 0 for $\overset{p}{\underset{1}{\nu}}\ (v_\nu = r_\nu)$.

In order to prove that the converse theorem holds, we first show that if

$$\theta\begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_1^{m-1} \alpha_\nu^{(\mu)} - \sum_1^{m-1} u_\nu^{(\mu)} + r_\nu\right)$$

vanishes identically and the quantities

$$\theta^{(m)}\begin{pmatrix} p \\ \nu\ (r_\nu) \\ 1 \end{pmatrix}$$

are all zero, then

$$\theta\begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_1^{m} \alpha_\nu^{(\mu)} - \sum_1^{m} u_\nu^{(\mu)} + r_\nu\right)$$

must also vanish. To this end we generalize equation (5) of §4.

We assume that

$$\theta\begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_1^{m-1} u_\nu^{(\mu)} - \sum_1^{m-1} \alpha_\nu^{(\mu)} + r_\nu\right)$$

vanishes identically, but

$$\theta\begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_1^{m} u_\nu^{(\mu)} - \sum_1^{m} \alpha_\nu^{(\mu)} + r_\nu\right)$$

does not. We retain the earlier assumption in regard to the quantities t, and consider the expression

(2)

$$\frac{\theta\begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_1^{m} u_\nu^{(\mu)} - \sum_1^{m} \alpha_\nu^{(\mu)} + r_\nu\right)\theta\begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_1^{m} \alpha_\nu^{(\mu)} - \sum_1^{m} u_\nu^{(\mu)} + r_\nu\right)\times \\ \times\prod_{\rho,\rho'}\theta\begin{pmatrix} p \\ \nu\ (u_\nu^{(\rho)} - u_\nu^{(\rho')} + t_\nu) \\ 1 \end{pmatrix}\theta\begin{pmatrix} p \\ \nu\ (\alpha_\nu^{(\rho)} - \alpha_\nu^{(\rho')} + t_\nu) \\ 1 \end{pmatrix}}{\prod_{\rho,\rho'=1}^{m}\theta\begin{pmatrix} p \\ \nu\ (u_\nu^{(\rho)} - \alpha_\nu^{(\rho')} + t_\nu) \\ 1 \end{pmatrix}\theta\begin{pmatrix} p \\ \nu\ (\alpha_\nu^{(\rho)} - u_\nu^{(\rho')} + t_\nu) \\ 1 \end{pmatrix}}.$$

In the above expression, the variables ρ and ρ' under the product signs run from 1 to m, omitting terms with $\rho = \rho'$ in the numerator.

If we regard this expression as a function of z_1, we see that the factor that arises on crossing each transverse cut is 1. Consequently it is an algebraic function of z_1. For $z_1 = \zeta_\rho$ and $s_1 = \sigma_\rho$, the denominator and numerator are infinitely small of second order, and the fraction therefore remains finite. For all other values for which the numerator and denominator vanish, the quantities r and t are fully determined by the quantities $\overset{m}{\underset{2}{\mu}}(s_\mu, z_\mu)$, as shown in §3, and hence are independent of the quantities ζ. Now since the expression is a symmetric function of the quantities z, the same is true for every z_μ: it is an algebraic function of z_μ, and the values of z_μ for which this algebraic function becomes infinitely small or infinitely large do not depend on the quantities ζ. The expression is therefore equal to an algebraic function $\chi(z_1, z_2, \ldots, z_m)$ of the variables z (not depending on ζ), multiplied by a factor not depending on the variables z. Since the expression remains unaltered when the quantities z, ζ are interchanged, this factor is $\chi(\zeta_1, \zeta_2, \ldots, \zeta_m)$ multiplied by a constant A independent of the quantities z, ζ. We may therefore, if we write

$$\sqrt{A}\,\chi(z_1, z_2, \ldots, z_m) = \psi(z_1, z_2, \ldots, z_m),$$

give our expression (2) the form

$$\psi(z_1, z_2, \ldots, z_m)\psi(\zeta_1, \zeta_2, \ldots, \zeta_m). \tag{3}$$

Here $\psi(z_1, z_2, \ldots, z_m)$ is an algebraic function of the variables z, independent of the quantities ζ, which, by virtue of its branching type, can be expressed rationally in $\overset{m}{\underset{1}{\mu}}(s_\mu, z_\mu)$.

If we now allow the points η to coincide with the ϵ, so that the quantities $\zeta_\mu - z_\mu$ and $\sigma_\mu - s_\mu$ all become infinitely small, then—using the notation for partial derivatives in §4, (1)—we have

$$\psi(z_1, z_2, \ldots, z_m) = \frac{\pm \sum\limits_{\nu_1, \ldots, \nu_m = 1}^{p} \theta^{(m)}_{\nu_1, \nu_2, \ldots, \nu_m} \begin{pmatrix} p \\ \rho\ (r_\rho) \\ 1 \end{pmatrix} du^{(1)}_{\nu_1} du^{(2)}_{\nu_2} \cdots du^{(m)}_{\nu_m}}{\prod\limits_{\mu=1}^{m} \sum\limits_{\nu=1}^{p} \theta'_\nu \begin{pmatrix} p \\ \rho\ (r_\rho) \\ 1 \end{pmatrix} du^{(\mu)}_\nu}. \tag{4}$$

It is hardly necessary to point out that the choice of sign is irrelevant, since it has no bearing on the value of

$$\psi(z_1, z_2, \ldots, z_m)\psi(\zeta_1, \zeta_2, \ldots, \zeta_m),$$

and that $du_1^{(\mu)}, du_2^{(\mu)}, \ldots, du_p^{(\mu)}$ can be replaced simultaneously in the numerator and denominator by quantities $\phi_1(s_\mu, z_\mu), \phi_2(s_\mu, z_\mu), \ldots, \phi_p(s_\mu, z_\mu)$ proportional to them.

From the equation contained in (2), (3), and (4), which has been proved for the case in which

$$\theta\begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_1^{m-1} u_\nu^{(\mu)} - \sum_1^{m-1} \alpha_\nu^{(\mu)} + r_\nu\right)$$

is zero, and

$$\theta\begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_1^{m} u_\nu^{(\mu)} - \sum_1^{m} \alpha_\nu^{(\mu)} + r_\mu\right)$$

is not, it follows that

$$\theta\begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_1^{m} u_\nu^{(\mu)} - \sum_1^{m} \alpha_\nu^{(\mu)} + r_\nu\right)$$

cannot be different from zero, if the functions $\theta^{(m)}\begin{pmatrix} p \\ \nu\ (r_\nu) \\ 1 \end{pmatrix}$ all vanish.

Thus, if the functions $\theta^{(m+1)}\begin{pmatrix} p \\ \nu\ (r_\nu) \\ 1 \end{pmatrix}$ all vanish, the validity of the equation

$$\theta\begin{pmatrix} p \\ \nu \\ 1 \end{pmatrix}\left(\sum_1^{n} u_\nu^{(\mu)} - \sum_1^{n} \alpha_\nu^{(\mu)} + r_\nu\right) = 0$$

for $n = m$ implies its validity for $n = m + 1$. Accordingly, if the equation holds for $n = 0$, or if $\theta\begin{pmatrix} p \\ \nu\ (r_\nu) \\ 1 \end{pmatrix} = 0$, and all partial derivatives of the function $\theta\begin{pmatrix} p \\ \nu\ (v_\nu) \\ 1 \end{pmatrix}$ up to order m vanish for $\overset{p}{\underset{1}{\nu}}\ (v_\nu = r_\nu)$, but the $(m+1)$-th order partial derivatives do not all vanish, then the equation also holds for all values of n up to $n = m$, but not for $n = m+1$. For as we found previously, it would follow from

$$\theta\begin{pmatrix} p \\ \nu \left(\sum\limits_{1}^{m+1} u_\nu^{(\mu)} - \sum\limits_{1}^{m+1} \alpha_\nu^{(\mu)} + r_\nu \right) \\ 1 \end{pmatrix} = 0$$

that all the quantities $\theta^{(m+1)}\begin{pmatrix} p \\ \nu\ (r_\nu) \\ 1 \end{pmatrix}$ must vanish.

6.

Combining what we have just proved with the earlier results, we have the following result:

If $\theta(r_1, r_2, \ldots, r_p) = 0$, then $p-1$ points $\eta_1, \eta_2, \ldots, \eta_{p-1}$ can be found, such that

$$(r_1, r_2, \ldots, r_p) \equiv \left(\sum_{1}^{p-1} \alpha_1^{(\mu)}, \sum_{1}^{p-1} \alpha_2^{(\mu)}, \ldots, \sum_{1}^{p-1} \alpha_p^{(\mu)} \right);$$

and the converse is true.

If the function $\theta(v_1, v_2, \ldots, v_p)$ and its partial derivatives up to order m vanish for $v_1 = r_1, v_2 = r_2, \ldots, v_p = r_p$, but the $(m+1)$-th partial derivatives are not all zero, then m of these points η can be chosen arbitrarily without altering the quantities r, and then the remaining $p-1-m$ points are completely determined.

Conversely if m, but not more, of the points η can be chosen arbitrarily without altering the quantities r, then the function $\theta(v_1, v_2, \ldots, v_p)$ and its partial derivatives up to order m vanish for $v_1 = r_1, v_2 = r_2, \ldots, v_p = r_p$, but the $(m+1)$-th partial derivatives are not all zero.

Complete investigation of all the particular cases which can arise in regard to the vanishing of a theta-function was necessary, not so much because of the particular systems of similarly branching algebraic functions to which the different cases correspond; but rather, because without this examination, gaps arise in the proof of the theorem. The proof can now be based on our theorem on the vanishing of a θ-function.

XII.

On the representation of a function by a trigonometric series.

(*Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen*, vol. 13.)

The following essay on trigonometric series consists of two essentially different parts. The first part contains a history of the research and opinions on arbitrary (graphically given) functions and their representation by trigonometric series. In its composition I was guided by some hints of the famous mathematician, to whom the first fundamental work on this topic was due. In the second part, I examine the representation of a function by a trigonometric series including cases that were previously unresolved. For this, it was necessary to start with a short essay on the concept of a definite integral and the scope of its validity.

History of the question of the representation of an arbitrary function by a trigonometric series.

1.

The trigonometric series named after Fourier, that is, the series of the form

$$\begin{aligned} &a_1 \sin x + a_2 \sin 2x + a_3 \sin 3x + \cdots \\ +\frac{1}{2}\, &b_0 + b_1 \cos x + b_2 \cos 2x + b_3 \cos 3x + \cdots \end{aligned}$$

play a significant role in those parts of mathematics where arbitrary functions occur. Indeed, there is reason to assert that the most substantial progress in this part of mathematics, that is so important for physics, has depended on a clear insight into the nature of these series. As soon as mathematical research first led to consideration of arbitrary functions, the question arose whether an arbitrary function could be expressed by a series of the above form.

This occurred in the middle of the eighteenth century during the study of vibrating strings, a topic in which the most prominent mathematicians of the time were interested. Their insights about our topic would probably not be represented were it not for the investigation of this problem.

As is well known, under certain hypotheses that conform approximately to reality, the shape of a string under tension that is vibrating in a plane is determined by the partial differential equation

$$\frac{\partial^2 y}{\partial t^2} = \alpha^2 \frac{\partial^2 y}{\partial x^2}$$

where x is the distance of an arbitrary one of its points from the origin and y is the distance from the rest position at time t. Furthermore α is independent of t, and also of x for a string of uniform thickness.

D'Alembert was the first to give a general solution to this differential equation.

He showed[1] that each function of x and t, which when set in the equation for y yields an identity, must have the form

$$f(x + \alpha t) + \phi(x - \alpha t).$$

This follows by introducing the independent variables $x + \alpha t$, $x - \alpha t$ instead of x and t, whereby

$$\frac{\partial^2 y}{\partial x^2} - \frac{1}{\alpha^2}\frac{\partial^2 y}{\partial t^2} \quad \text{changes into} \quad 4\,\frac{\partial \frac{\partial y}{\partial(x+\alpha t)}}{\partial(x - \alpha t)}.$$

Besides the partial differential equation, which results from the general laws of motion, y must also satisfy the condition that it is always 0 at the endpoints of the string. Thus, if one of these points is at $x = 0$ and the other at $x = \ell$, we have

$$f(\alpha t) = -\phi(-\alpha t), \qquad f(\ell + \alpha t) = -\phi(\ell - \alpha t)$$

and consequently

$$f(z) = -\phi(-z) = -\phi(\ell - (\ell + z)) = f(2\ell + z),$$
$$y = f(\alpha t + x) - f(\alpha t - x).$$

After d'Alembert had succeeded in finding the above for the general solution of the problem, he treated, in a sequel[2] to his paper, the equation

[1] *Mémoires de l'académie de Berlin*, 1747, p. 214.
[2] Ibid. p. 220.

$f(z) = f(2\ell + z)$. That is, he looked for analytic expressions that remained unchanged if z is increased by 2ℓ.

In the next issue of *Mémoires de l'académie de Berlin*[3], Euler made a basic advance, giving a new presentation of d'Alembert's work and recognizing more exactly the nature of the conditions which the function $f(x)$ must satisfy. He noted that, by the nature of the problem, the movement of the string is completely determined, if at some point in time the shape of the string and the velocity are given at each point (that is, y and $\frac{\partial y}{\partial t}$). He showed that if one thinks of the two functions as being determined by arbitrarily drawn curves, then the d'Alembert function $f(z)$ can always be found by a simple geometric construction. In fact, if one assumes that $y = g(x)$ and $\frac{\partial y}{\partial t} = h(x)$ when $t = 0$, then one obtains

$$f(x) - f(-x) = g(x) \quad \text{and} \quad f(x) + f(-x) = \frac{1}{\alpha}\int h(x)\,dx$$

for values of x between 0 and ℓ, and hence obtains the function $f(z)$ between $-\ell$ and ℓ. From this, however, the values of $f(z)$ can be derived for all other values of z using the equation

$$f(z) = f(2\ell + z).$$

This is, represented in abstract but now generally accepted concepts, Euler's determination of the function $f(z)$.

D'Alembert at once protested against this extension of his methods by Euler[4], since it assumed that y could be expressed analytically in t and x.

Before Euler replied to this, Daniel Bernoulli[5] presented a third treatment of this topic, which was quite different from the previous two. Even prior to d'Alembert, Taylor[6] had seen that $y = \sin\frac{n\pi x}{\ell}\cos\frac{n\pi\alpha t}{\ell}$, where n is an integer, satisfies $\frac{\partial^2 y}{\partial t^2} = \alpha^2\frac{\partial^2 y}{\partial x^2}$ and always equals 0 for $x = 0$ and $x = \ell$. From this he explained the physical fact that a string, besides its fundamental tone, can also give the fundamental tone of a string that is $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ as

[3] *Mémoires de l'académie de Berlin*, 1748, p. 69.

[4] *Mémoires de l'académie de Berlin*, 1750, p. 358. 'Indeed, it seems to me, one can only express y analytically in a more general fashion by supposing it is a function of t and x. But with this assumption one only finds a solution of the problem for the case where the different graphs of the vibrating string can be contained in a single equation.'

[5] *Mémoires de l'academie de Berlin*, 1753, p. 147.

[6] Taylor, *De methode incrementorum.*

long (but otherwise similarly constituted). He took his particular solutions as general: he thought that if the pitch of the tone was determined by the integer n, then the vibration of the string would always be as expressed by the equation, or at least very nearly. The observation that a string could simultaneously sound different notes now led Bernoulli to the remark that the string (by the theory) could also vibrate in accordance with the equation

$$y = \sum a_n \sin \frac{n\pi x}{\ell} \cos \frac{n\pi\alpha}{\ell} (t - \beta_n).$$

Further, since all observed modifications of the phenomenon could be explained by this equation, he considered it the most general solution.[7] In order to support this opinion, he examined the vibration of a massless thread under tension, which was weighted at isolated points with finite masses. He showed that the vibrations can be decomposed into a number of vibrations that is always equal to the number of points, each vibration being of the same duration for all masses.

This work of Bernoulli prompted a new paper from Euler, which was printed immediately following it in the *Mémoires de l'académie de Berlin.*[8] He maintained, in opposition to d'Alembert[9], that the function $f(z)$ could be completely arbitrary between $-\ell$ and ℓ. Euler[10] noted that Bernoulli's solution (which he had previously represented as particular) is general if and only if the series

$$\begin{aligned} a_1 \sin \frac{x\pi}{\ell} + a_2 \sin \frac{2x\pi}{\ell} + \cdots \\ +\frac{1}{2} b_0 + b_1 \cos \frac{x\pi}{\ell} + b_2 \cos \frac{2x\pi}{\ell} + \cdots \end{aligned}$$

can represent the ordinate of an arbitrary curve for the abcissa x between 0 and ℓ. Now no one doubted at that time that all transformations which could be made with an analytic expression (finite or infinite) would be valid for each value of the variable, or only inapplicable in very special cases. Thus it seemed impossible to represent an algebraic curve, or in general a nonperiodic analytically given curve, by the above expression. Hence Euler thought that the question must be decided against Bernoulli.

[7]Loc. cit., p. 157 section XIII.
[8]*Mémoires de l'académie de Berlin*, 1753, p. 196.
[9]Loc. cit., p. 214
[10]Loc. cit., sections III–X.

The disagreement between Euler and d'Alembert was still unresolved by this. This induced the young, and then little known, mathematician Lagrange to seek the solution of the problem in a completely new way, by which he reached Euler's results. He undertook[11] to determine the vibration of a massless thread which is weighted with an indeterminate finite number of equal masses that are equally spaced. He then examined how the vibrations change when the number of masses grows towards infinity. Although he carried out the first part of this investigation with much dexterity and a great display of analytic ingenuity, the transition from the finite to the infinite left much to be desired. Hence d'Alembert could continue to vindicate the reputation of his solution as the most general by making this point in a note in his *Opuscules Mathématiques.* The opinions of the prominent mathematicians of this time were, and remained, divided on the matter; for in later work everyone essentially retained his own point of view.

In order to finally arrange his views on the problem of arbitrary functions and their representation by trigonometric series, Euler first introduced these functions into analysis, and supported by geometrical considerations, applied infinitesimal analysis to them. Lagrange[12] considered Euler's results (his geometric construction for the course of the vibration) to be correct, but he was not satisfied with Euler's geometric treatment of the functions. D'Alembert,[13] on the other hand, acceded to Euler's way of obtaining the differential equation and restricted himself to disputing the validity of his result, since one could not know for an arbitrary function whether its derivatives were continuous. Concerning Bernoulli's solution, all three agreed not to consider it as general. While d'Alembert,[14] in order to explain Bernoulli's solution as less general than his own, had to assert that an analytically given periodic function cannot always be represented by a trigonometric series, Lagrange[15] believed it possible to prove this.

2.

Almost fifty years had passed without a basic advance having been made in the question of the analytic representation of an arbitrary function. Then

[11] *Miscellanea Taurinensia*, vol. I. Recherches sur la nature et la propagation du son.
[12] *Miscellanea Taurinensia*, vol. II, *Pars math.*, p. 18.
[13] *Opuscules Mathématiques*, d'Alembert. Vol. 1, 1761, p. 16, Sections VII—XX.
[14] *Opuscules Mathématiques*, vol. I, p. 42, Section XXIV.
[15] *Misc. Taur. vol. III, Pars math.*, p. 221, Section XXV.

a remark by Fourier threw a new light on the topic. A new epoch in the development of this part of mathematics began, which soon made itself known in a wonderful expansion of mathematical physics. Fourier noted that in the trigonometric series

$$f(x) = \begin{cases} a_1 \sin x + a_2 \sin 2x + \cdots \\ +\frac{1}{2}\, b_0 + b_1 \cos x + b_2 \cos 2x + \cdots , \end{cases}$$

the coefficients can be determined by the formulae

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \sin nx \, dx, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \cos nx \, dx.$$

He saw that the method can also be applied if the function $f(x)$ is arbitrary. He used a so-called discontinuous function for $f(x)$ (with ordinate a broken line for the abscissa x) and obtained a series which in fact always gives the value of the function.

Fourier, in one of his first papers on heat, which was submitted to the French academy[16] (December 21, 1807) first announced the theorem, that an arbitrary (graphically given) function can be expressed as a trigonometric series. This claim was so unexpected to the aged Lagrange that he opposed it vigorously. There should[17] be another note about this in the archives of the Paris academy. Nevertheless, Poisson refers,[18] whenever he makes use of trigonometric series to represent arbitrary functions, to a place in Lagrange's work on the vibrating string where this method of representation can be found. In order to refute this claim, which can only be explained by the well known rivalry[19] between Fourier and Poisson, we must once again return to Lagrange's treatise, since nothing can be found that is published about these facts by the academy.

In fact, one finds in the place cited[20] by Poisson the formula:

$$\text{‘}y = 2\int Y \sin X\pi \, dX \times \sin x\pi + 2\int Y \sin 2X\pi \, dX \times \sin 2x\pi$$
$$+ 2\int Y \sin 3X\pi \, dX \times \sin 3x\pi + \text{etc.} + 2\int Y \sin nX\pi \, dX \times \sin nx\pi,$$

[16] *Bulletin des sciences p. la soc. philomatique*, vol I, p. 112.

[17] From a verbal report of Professor Dirichlet.

[18] Among others, in the expanded *Traité de mécanique* No. 323, p. 638

[19] The review in the *Bulletin des Sciences* on the paper submitted by Fourier to the academy was written by Poisson.

[20] *Misc. Taur.*, vol. III, *Pars math.*, p. 261.

so that when $x = X$, one has $y = Y$, Y being the ordinate corresponding to the abscissa X'.

This formula looks so much like a Fourier series that is easy to confuse them with just a quick glance. However, this appearance arises only because Lagrange uses $\int dX$ where today we would use $\sum \Delta X$. It gives the solution to the problem of determining the finite sine series

$$a_1 \sin x\pi + a_2 \sin 2x\pi + \cdots + a_n \sin nx\pi$$

so that it has given values when x equals

$$\frac{1}{n+1}, \frac{2}{n+1}, \ldots, \frac{n}{n+1}.$$

Lagrange denotes the variable by X. If Lagrange had let n become infinitely large in this formula, then certainly he would have obtained Fourier's result. However, if we read through his paper, we see that he was far from believing that an arbitrary function could actually be represented by an infinite sine series. Rather, he had undertaken the whole work because be believed that an arbitrary function could not be expressed by a formula. Concerning trigonometric series, he thought they could be used to represent any analytically given periodic function. Admittedly, it now seems scarcely possible that Lagrange did not obtain Fourier's series from his summation formula. However, this can be explained in that the dispute between Euler and d'Alembert had predisposed him towards a particular opinion about the proper method of proceeding. He thought that the vibration problem, for an indeterminate finite number of masses, must be fully solved before applying limit considerations. This necessitated a rather extensive investigation[21], which was unnecessary if he had been acquainted with the Fourier series.

The nature of the trigonometric series was recognized perfectly correctly by Fourier.[22] Since then these series have been applied many times in mathematical physics to represent arbitrary functions. In each individual case it was easy to convince oneself that the Fourier series really converged to the value of the function. However, it was a long time before this important theorem would be proved in general.

[21] *Misc. Taur.*, vol III, *Pars math.*, p. 251.

[22] *Bulletin d. sc.* vol. I, p. 115. 'The coefficients $a, a', a'', \ldots$, being then determined', etc.

The proof which Cauchy[23] read to the Paris academy on February 27, 1826, is inadequate, as Dirichlet[24] has shown. Cauchy assumed that if x is replaced by the complex argument $x + yi$ in an arbitrary periodic function $f(x)$, then the function is finite for each value of y. However, this only occurs if the function is a constant. It is easy to see that this hypothesis was unnecessary for the later conclusions. It suffices that a function $\phi(x + yi)$ exists which is finite for all positive values of y, whose real part is equal to the given periodic function $f(x)$ when $y = 0$. If one assumes this theorem, which is in fact true,[25] then Cauchy's method certainly leads to the goal; conversely, this theorem can be derived from the Fourier series.

3.

The question of the representation by trigonometric series of everywhere integrable functions with finitely many maxima and minima was first settled rigorously by Dirichlet[26] in a paper of January 1829.

The recognition of the proper way to attack the problem came to him from the insight that infinite series fall into two distinct classes, depending on whether or not they remain convergent when all the terms are made positive. In the first class the terms can be arbitrarily rearranged; in the second, on the other hand, the value is dependent on the ordering of the terms. Indeed, if we denote the positive terms of a series in the second class by

$$a_1, a_2, a_3, \ldots,$$

and the negative terms by

$$-b_1, \ -b_2, \ -b_3, \ldots,$$

then it is clear that $\sum a$ as well as $\sum b$ must be infinite. For if they were both finite, the series would still be convergent after making all the signs the same. If only one were infinite, then the series would diverge. Clearly now an arbitrarily given value C can be obtained by a suitable reordering of the terms. We take alternately the positive terms of the series until the sum is greater than C, and then the negative terms until the sum is less than C. The deviation from C never amounts to more than the size of the term at

[23] *Mémoires de l'ac. d. sc. de Paris*, vol. VI, p. 603.
[24] *Crelle's Journal für die Mathematik*, vol IV, pp. 157 & 158.
[25] The proof can be found in the inaugural dissertation of the author.
[26] *Crelle's Journal*, vol. IV, p. 157.

the last place the signs were switched. Now, since the numbers a as well as the numbers b become infinitely small with increasing index, so also are the deviations from C. If we proceed sufficiently far in the series, the deviation becomes arbitrarily small, that is, the series converges to C.

The rules for finite sums only apply to the series of the first class. Only these can be considered as the aggregates of their terms; the series of the second class cannot. This circumstance was overlooked by mathematicians of the previous century, most likely, mainly on the grounds that the series which progress by increasing powers of a variable generally (that is, excluding individual values of this variable) belong to the first class.

Clearly the Fourier series do not necessarily belong to the first class. The convergence cannot be derived, as Cauchy futilely attempted,[27] from the rules by which the terms decrease. Rather, it must be shown that the finite series

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(\alpha)\sin\alpha\,d\alpha\,\sin x + \frac{1}{\pi}\int_{-\pi}^{\pi} f(\alpha)\sin 2\alpha\,d\alpha\sin 2x + \cdots$$
$$+\frac{1}{\pi}\int_{-\pi}^{\pi} f(\alpha)\sin n\alpha\,d\alpha\,\sin nx$$
$$+\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\alpha)\,d\alpha + \frac{1}{\pi}\int_{-\pi}^{\pi} f(\alpha)\cos\alpha\,d\alpha\cos x$$
$$+\frac{1}{\pi}\int_{-\pi}^{\pi} f(\alpha)\cos 2\alpha\,d\alpha\,\cos 2x + \cdots + \frac{1}{\pi}\int_{-\pi}^{\pi} f(\alpha)\cos n\alpha\,d\alpha\cos nx,$$

or, what is the same, the integral

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\alpha)\,\frac{\sin\frac{2n+1}{2}(x-\alpha)}{\sin\frac{x-\alpha}{2}}\,d\alpha,$$

approaches the value $f(x)$ infinitely closely when n increases infinitely.

Dirichlet based this proof on two theorems:

1) If $0 < c \leq \pi/2$, then $\int_0^c \phi(\beta)\frac{\sin(2n+1)\beta}{\sin\beta}\,d\beta$ tends to $\frac{\pi}{2}\,\phi(0)$ as n increases to infinity.

2) If $0 < b < c \leq \pi/2$, then $\int_b^c \phi(\beta)\frac{\sin(2n+1)\beta}{\sin\beta}d\beta$ tends to 0, as n increases to infinity.

[27] Dirichlet in *Crelle's Journal*, vol IV, p. 158. 'Quoi qu'il en soit de cette première observation, . . . à mesure que n croit.'

It is assumed in both cases that the function $\phi(\beta)$ is either always increasing or always decreasing between the limits of integration.

If the function f does not change from increasing to decreasing, or from decreasing to increasing, infinitely often, then using the above theorems the integral

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\alpha)\,\frac{\sin\frac{2n+1}{2}(x-\alpha)}{\sin\frac{x-\alpha}{2}}\,d\alpha$$

can clearly be split into a finite number of parts, one of which tends[28] to $\frac{1}{2}f(x+0)$, another to $\frac{1}{2}f(x-0)$, and the others to 0, as n increases to infinity.

It follows from this that a periodic function of period 2π, which

1. is everywhere integrable,

2. does not have infinitely many maxima and minima, and

3. assumes the average of the two one-sided limits when the value changes by a jump,

can be represented by a trigonometric series.

It is clear that a function satisfying the first two properties but not the third cannot be represented by a trigonometric series. A trigonometric series representing such a function, except at the discontinuities, would deviate from it at the discontinuities. Dirichlet's research leaves undecided, whether and when functions can be represented by a trigonometric series that do not satisfy the first two conditions.

Dirichlet's work gave a firm foundation for a large amount of important research in analysis. He succeeded in bringing light to a point where Euler was in error. He settled a question that had occupied many distinguished mathematicians for over 70 years (since 1753). In fact, for all cases of nature, the only cases concerned in that work, it was completely settled. For however great our ignorance about how forces and states of matter vary for infinitely small changes of position and time, surely we may assume that the functions which are not included in Dirichlet's investigations do not occur in nature.

[28]It is easy to prove that the value of a function f, which does not have infinitely many maxima or minima, for increasing or decreasing values of the argument with limit x, either approaches fixed limits $f(x+0)$ and $f(x-0)$ (using Dirichlet's notation in Dove's *Repertorium der Physik*, vol. 1, p. 170); or must become infinitely large.

Nevertheless, there are two reasons why those cases unresolved by Dirichlet seem to be worthy of consideration.

First, as Dirichlet noted at the end of his paper, the topic has a very close connection with the principles of infinitesimal calculus, and can serve to bring greater clarity and rigor to these principles. In this regard the treatment of the topic has an immediate interest.

Secondly, however, the applications of Fourier series are not restricted to research in the physical sciences. They are now also applied with success in an area of pure mathematics, number theory. Here it is precisely the functions whose representation by a trigonometric series was not examined by Dirichlet that seem to be important.

Admittedly Dirichlet promised at the conclusion of his paper to return to these cases later, but that promise still remains unfulfilled. The works by Dirksen and Bessel on the cosine and sine series did not supply this completion. Rather, they take second place to Dirichlet in rigor and generality. Dirksen's paper,[29] (almost simultaneous with Dirichlet's, and clearly written without knowledge of it) was, indeed, in a general way correct. However, in the particulars it contained some imprecisions. Apart from the fact that he found an incorrect result in a special case[30] for the sum of a series, he relied in a secondary consideration on a series expansion[31] that is only possible in particular cases. Hence the proof is only complete for functions whose first derivatives are everywhere finite. Bessel[32] tried to simplify Dirichlet's proof. However, the changes in the proof did not give any essential simplification, but at most clothed it in more familiar concepts, at the expense of rigor and generality.

Hence, until now, the question of the representation of a function by a trigonometric series is only settled under the two hypotheses, that the function is everywhere integrable and does not have infinitely many maxima and minima. If the last hypothesis is not made, then the two integral theorems of Dirichlet are not sufficient for deciding the question. If the first is discarded, however, the Fourier method of determining the coefficients is not applicable. In the following, when we examine the question without any particular assumptions on the nature of the function, the method employed, as we will see, is constrained by these facts. An approach as direct as Dirichlet's is not

[29] *Crelle's Journal*, vol IX, p. 170.
[30] *Loc. cit.*, formula 22.
[31] *Loc. cit.*, section 3.
[32] Schumacher, *Astronomische Nachrichten*, 374 (vol. 16, p. 229.)

possible by the nature of the case.

On the concept of a definite integral and the range of its validity.

4.

Vagueness still prevails in some fundamental points concerning the definite integral. Hence I provide some preliminaries about the concept of a definite integral and the scope of its validity.

Hence first: What is one to understand by $\int_a^b f(x)\,dx$?

In order to establish this, we take a sequence of values $x_1, x_2, \ldots, x_{n-1}$ between a and b arranged in succession, and denote, for brevity, $x_1 - a$ by δ_1, $x_2 - x_1$ by δ_2, ..., $b - x_{n-1}$ by δ_n, and a positive fraction less than 1 by ϵ. Then the value of the sum

$$S = \delta_1 f(a + \epsilon_1\delta_1) + \delta_2 f(x_1 + \epsilon_2\delta_2) + \delta_3 f(x_2 + \epsilon_3\delta_3) + \cdots + \delta_n f(x_{n-1} + \epsilon_n\delta_n)$$

depends on the selection of the intervals δ and the numbers ϵ. If this now has the property, that however the δ's and ϵ's are selected, S approaches a fixed limit A when the δ's become infinitely small together, this limiting value is called $\int_a^b f(x)\,dx$.

If we do not have this property, then $\int_a^b f(x)\,dx$ is undefined. In some of these cases, attempts have been made to assign a meaning to the symbol, and among these extensions of the concept of a definite integral there is one recognized by all mathematicians. Namely, if the function $f(x)$ becomes infinitely large when the argument approaches an isolated value c in the interval (a, b), then clearly the sum S, no matter what degree of smallness one may prescribe for δ, can reach an arbitrarily given value. Thus it has no limiting value, and by the above $\int_a^b f(x)\,dx$ would have no meaning. However if

$$\int_a^{c-\alpha_1} f(x)\,dx + \int_{c+\alpha_2}^b f(x)\,dx$$

approaches a fixed limit when α_1 and α_2 become infinitely small, then one understands this limit to be $\int_a^b f(x)\,dx$.

Other hypotheses by Cauchy on the concept of the definite integral in the cases where the fundamental concepts do not give a value may be appropriate in individual classes of investigation. These are not generally established, and are hardly suited for general adoption in view of their great arbitrariness.

5.

Let us examine now, secondly, the range of validity of the concept, or the question: In which cases can a function be integrated, and in which cases can it not?

We consider first the concept of integral in the narrow sense, that is, we suppose that the sum S converges if the δ's together become infinitely small. We denote by D_1 the greatest fluctuation of the function between a and x_1, that is, the difference of its greatest and smallest values in this interval, by D_2 the greatest fluctuation between x_1 and $x_2, \ldots$, by D_n that between x_{n-1} and b. Then

$$\delta_1 D_1 + \delta_2 D_2 + \cdots + \delta_n D_n$$

must become infinitely small when the δ's do. We suppose further, that Δ is the greatest value this sum can reach, as long as all of the δ's are smaller than d. Then Δ will be a function of d, which is decreasing with d and becomes infinitely small with d. Now, if the total length of the intervals, in which the fluctuation is greater than σ, is s, then the contribution of these intervals to the sum $\delta_1 D_1 + \delta_2 D_2 + \cdots + \delta_n D_n$ is clearly $\geq \sigma s$. Therefore one has

$$\sigma s \leq \delta_1 D_1 + \delta_2 D_2 + \cdots + \delta_n D_n \leq \Delta, \text{ hence } s \leq \frac{\Delta}{\sigma}.$$

Now, if σ is given, Δ/σ can always be made arbitrarily small by a suitable choice of d. The same is true for s, which yields:

In order for the sum S to converge whenever all the δ's become infinitely small, in addition to $f(x)$ being finite, it is necessary that the total length of the intervals, in which the fluctuations exceed σ, can be made arbitrarily small for any given σ by a suitable choice of d.

This theorem also has a converse:

If the function $f(x)$ is always finite, and by infinitely decreasing the δ's together, the total length s of the intervals in which the fluctuation of the function is greater than a given number σ always becomes infinitely small, then the sum S converges as the δ's become infinitely small together.

For those intervals in which the fluctuations are $> \sigma$ make a contribution to the sum $\delta_1 D_1 + \cdots + \delta_n D_n$, less than s times the largest fluctuation of the function between a and b, which is finite (by agreement). The contribution of the remaining intervals is $< \sigma(b-a)$. Clearly one can now choose σ arbitrarily small and then always determine the size of the intervals (by agreement) so that s is also arbitrarily small. In this way the sum $\delta_1 D_1 + \delta_2 D_2 + \cdots + \delta_n D_n$

can be made as small as desired. Consequently the value of the sum S can be enclosed between arbitrarily narrow bounds.

Thus we have found necessary and sufficient conditions for the sum S to be convergent when the quantities δ tend together to zero, or equivalently, for the existence of the integral of $f(x)$ between a and b in the narrow sense.

If we now extend the integral concept as above, then it is clear that for the integration to be possible everywhere, the second of the two conditions established is still necessary. In place of the condition, that the function is always finite, will enter the condition, that the function becomes infinite only on the approach of the argument to isolated values, and that a definite limiting value arises, if the limits of integration tend to these values.

6.

Having examined the conditions for integrability in general, that is, without special assumptions on the nature of the function to be integrated, this investigation will be applied and also carried further, in special cases. First we consider functions which are discontinuous infinitely often between any two numbers, no matter how close.

Since these functions have never been considered before, it is well to start from a particular example. Designate, for brevity, (x) to be the excess of x over the closest integer, or if x lies in the middle between two (and thus the determination is ambiguous) the average of the two numbers $1/2$ and $-1/2$, hence zero. Furthermore, let n be an integer and p an odd integer, and form the series

$$f(x) = \frac{(x)}{1} + \frac{(2x)}{4} + \frac{(3x)}{9} + \cdots = \sum_{n=1}^{\infty} \frac{(nx)}{n^2}.$$

It is easy to see that the series converges for each value of x. When the argument continuously decreases to x, as well as when it continuously increases to x, the value always approaches a fixed limit. Indeed, if $x = \frac{p}{2n}$ (where p and n are relatively prime)

$$f(x+0) = f(x) - \frac{1}{2n^2}\left(1 + \frac{1}{9} + \frac{1}{25} + \cdots\right) = f(x) - \frac{\pi^2}{16n^2},$$
$$f(x-0) = f(x) + \frac{1}{2n^2}\left(1 + \frac{1}{9} + \frac{1}{25} + \cdots\right) = f(x) + \frac{\pi^2}{16n^2};$$

in all other cases $f(x+0) = f(x)$ and $f(x-0) = f(x)$.

Hence this function is discontinuous for each rational value of x, which in lowest terms is a fraction with even denominator. Thus, while f is discontinuous infinitely often between any two bounds, the number of jumps greater than a fixed number is always finite. The function is everywhere integrable. Besides finiteness, it has the two properties, that for each value of x it has limiting values $f(x+0)$ and $f(x+0)$ on both sides, and that the number of jumps greater than or equal to a given number σ is always finite. Applying our above investigation, as an obvious consequence of these two conditions, d can be taken so small that in the intervals which do not contain these jumps, the fluctuations are smaller then σ, and the total length of the intervals which do contain these jumps will be arbitrarily small.

It is worthwhile to note that functions which do not have infinitely many maxima and minima (to which the example just considered does not belong), where they do not become infinite, always have those two properties, and hence permit an integration everywhere where they are not infinite. This is also easy to show directly.

Now consider the case where the function $f(x)$ to be integrated has a single infinite value. We assume this occurs at $x = 0$, so that for decreasing positive values of x its value eventually grows over any given bound.

It can easily be shown that $xf(x)$ cannot always remain larger then finite number c as x decreases from a finite bound a. For then we would have

$$\int_x^a f(x)\,dx > c\int_x^a \frac{dx}{x},$$

thus larger than $c\left(\log\frac{1}{x} - \log\frac{1}{a}\right)$, which increases to infinity with decreasing x. Thus if $xf(x)$ does not have infinitely many maxima and minima in a neighborhood of $x = 0$, then $xf(x)$ must become infinitely small with x if $f(x)$ can be integrated. On the other hand, if

$$f(x)x^\alpha = \frac{f(x)\,dx\,(1-\alpha)}{d(x^{1-\alpha})},$$

for a value $\alpha < 1$, becomes infinitely small with x, then it is clear that the integral converges as the lower limit tends to 0.

In the same way one finds that in the cases where the integral converges, the functions

$$f(x)x\log\frac{1}{x}=\frac{f(x)\,dx}{-d\log\log\frac{1}{x}},\ f(x)x\log\frac{1}{x}\log\log\frac{1}{x}=\frac{f(x)\,dx}{-d\log\log\log\frac{1}{x}},\dots,$$
$$f(x)x\log\frac{1}{x}\log\log\frac{1}{x}\cdots\log^{n-1}\frac{1}{x}\log^{n}\frac{1}{x}=\frac{f(x)\,dx}{-d\log^{1+n}\frac{1}{x}}$$

cannot remain always larger than a finite number as x decreases from a finite bound. Thus if they do not have infinitely many maxima and minima, these functions must become infinitely small with x. On the other hand, the integral $\int f(x)\,dx$ converges as the lower limit of integration tends to 0, if

$$f(x)x\log\frac{1}{x}\cdots\log^{n-1}\frac{1}{x}\left(\log^{n}\frac{1}{x}\right)^{\alpha}=\frac{f(x)\,dx\,(1-\alpha)}{-d\left(\log^{n}\frac{1}{x}\right)^{1-\alpha}}$$

becomes infinitely small with x, for $\alpha>1$.

However, if $f(x)$ has infinitely many maxima and minima, then nothing can be determined about the order at which it becomes infinite. In fact, given the absolute value of f, and thereby given the order of infinity of f at 0, by a suitable determination of the sign one can always make the integral $\int f(x)\,dx$ converge when the lower limit of integration tends to 0. The function

$$\frac{d\left(x\cos e^{1/x}\right)}{dx}=\cos e^{1/x}+\frac{1}{x}e^{1/x}\sin e^{1/x}$$

serves as an example of a function which becomes infinite in such a way that its order (taking the order of $\frac{1}{x}$ as one) is infinitely large.

The above discussion, on the principles of a topic belonging to another area, suffices. We now proceed to our actual problem, a general investigation of the representation of a function by a trigonometric series.

Investigation of the representation of a function by a trigonometric series without particular assumptions on the nature of the function.

7.

The previous work on this topic served the purpose of proving the Fourier series for the cases occurring in nature. Hence the proofs could start for an arbitrary function, and later for the purposes of the proof one could impose arbitrary restrictions on the function, when they did not impair the goal.

For our purposes we only impose conditions necessary for the representation of the function. Hence we must first look for necessary conditions for the representation, and from these select sufficient conditions for the representation. While the previous work showed: 'If a function has this or that property then it is represented by a Fourier series', we must start from the converse question: If a function is represented by a Fourier series, what are the consequences for the function, regarding the changes of its values with a continuous change of the argument?

Hence we consider the series

$$a_1 \sin x + a_2 \sin 2x + \cdots$$
$$+\frac{1}{2}\, b_0 + b_1 \cos x + b_2 \cos 2x + \cdots$$

as given. For brevity, set

$$\frac{1}{2}\, b_0 = A_0,\ \ a_1 \sin x + b_1 \cos x = A_1,\ \ a_2 \sin 2x + b_2 \cos 2x = A_2, \ldots;$$

the series becomes

$$A_0 + A_1 + A_2 + \cdots .$$

We denote this expression by Ω and its value by $f(x)$, so that this function is defined only for values of x where the series converges.

For the series to converge, it is necessary that the terms eventually become infinitely small. If the coefficients a_n and b_n diminish infinitely with increasing n, then the terms of the series Ω eventually become infinitely small for each value of x. Otherwise convergence can only occur for particular values of x. It is necessary to treat both cases separately.

8.

Hence we suppose, first of all, that the terms of the series Ω eventually become arbitrarily small for each x.

Under this assumption, the series

$$C + C'x + A_0 \frac{x^2}{2} - A_1 - \frac{A_2}{4} - \frac{A_3}{9} \ldots = F(x)$$

converges for each value of x. The series is obtained by integrating each term of Ω twice with respect to x. The value $F(x)$ changes continuously with x, and consequently this function F of x is everywhere integrable.

In order to establish both the convergence of the series and the continuity of $F(x)$, one denotes the sum of the terms to $-\frac{A_n}{n^2}$ inclusive by N, the remainder of the series, that is, the series

$$-\frac{A_{n+1}}{(n+1)^2} - \frac{A_{n+2}}{(n+2)^2} - \cdots$$

by R; and the greatest value of A_m for $m > n$ by ϵ. Then, no matter how far one continues the series, the absolute value of R clearly remains

$$< \epsilon \left(\frac{1}{(n+1)^2} + \frac{1}{(n+2)^2} + \cdots \right) < \frac{\epsilon}{n},$$

and R can be enclosed within arbitrarily small bounds if n is sufficiently large. Hence the series converges. Furthermore, the function $F(x)$ is continuous, that is, its variation can be made as small as we wish, if one imposes a sufficiently small corresponding change of x. For the combined changes of $F(x)$ consists of the change in R and in N. Clearly one can first assume that n is so large that R is arbitrarily small whatever x may be, and consequently also the change of R will be arbitrarily small for any change in x. Then assume the change of x is so small that the change in N also becomes arbitrarily small.

It is well to place here some results about the function $F(x)$, whose proofs would otherwise break the thread of the investigation.

Theorem 1 *If the series Ω converges, then*

$$\frac{F(x+\alpha+\beta) - F(x+\alpha-\beta) - F(x-\alpha+\beta) + F(x-\alpha-\beta)}{4\alpha\beta},$$

converges to the same value as Ω when α and β become infinitely small while their ratio remains finite.

Indeed, we have

$$\frac{F(x+\alpha+\beta) - F(x+\alpha-\beta) - F(x-\alpha+\beta) + F(x-\alpha-\beta)}{4\alpha\beta}$$
$$= A_0 + A_1 \frac{\sin\alpha}{\alpha}\frac{\sin\beta}{\beta} + A_2 \frac{\sin 2\alpha}{2\alpha}\frac{\sin 2\beta}{2\beta} + A_3 \frac{\sin 3\alpha}{3\alpha}\frac{\sin 3\beta}{3\beta} + \cdots .$$

In order to settle the simplest case $\beta = \alpha$ first,

$$\frac{F(x+2\alpha) - 2F(x) + F(x-2\alpha)}{4\alpha^2} = A_0 + A_1 \left(\frac{\sin\alpha}{\alpha}\right)^2 + A_2 \left(\frac{\sin 2\alpha}{2\alpha}\right)^2 + \cdots$$

If the infinite series converges,

$$A_0 + A_1 + A_2 + \ldots = f(x),$$

and we write

$$A_0 + A_1 + \cdots + A_{n-1} = f(x) + \epsilon_n,$$

then for an arbitrarily given number δ, there must exist an integer m so that if $n > m$ we have $\epsilon_n < \delta$. Now, assume α is so small that $m\alpha < \pi$. We use the substitution

$$A_n = \epsilon_{n+1} - \epsilon_n,$$

to put $\sum_{n=0}^{\infty} \left(\frac{\sin n\alpha}{n\alpha}\right)^2 A_n$ in the form

$$f(x) + \sum_{n=1}^{\infty} \epsilon_n \left\{ \left(\frac{\sin(n-1)\alpha}{(n-1)\alpha}\right)^2 - \left(\frac{\sin n\alpha}{n\alpha}\right)^2 \right\},$$

and separate this last infinite series into three parts, in which we put together

1. the terms of index 1 to m inclusive,
2. from index $m+1$ up to the largest integer s less than $\frac{\pi}{\alpha}$,
3. from $s+1$ to infinity.

Then the first part consists of a finite number of continuously varying terms, and therefore approaches its limiting value 0 arbitrarily closely when one lets α become sufficiently small. The second part, since the factor of ϵ_n is always positive, has absolute value

$$< \delta \left\{ \left(\frac{\sin m\alpha}{m\alpha}\right)^2 - \left(\frac{\sin s\alpha}{s\alpha}\right)^2 \right\}.$$

In order to enclose the third part within bounds, one breaks up the general term into

$$\epsilon_n \left\{ \left(\frac{\sin(n-1)\alpha}{(n-1)\alpha}\right)^2 - \left(\frac{\sin(n-1)\alpha}{n\alpha}\right)^2 \right\}$$

and

$$\epsilon_n \left\{ \left(\frac{\sin(n-1)\alpha}{n\alpha} \right)^2 - \left(\frac{\sin n\alpha}{n\alpha} \right)^2 \right\} = -\epsilon_n \frac{(\sin(2n-1)\alpha)\sin\alpha}{(n\alpha)^2}.$$

Hence clearly it is

$$< \delta \left\{ \frac{1}{(n-1)^2\alpha^2} - \frac{1}{n^2\alpha^2} \right\} + \delta \frac{1}{n^2\alpha}$$

and consequently the sum from $n = s + 1$ to ∞ is

$$< \delta \left\{ \frac{1}{(s\alpha)^2} + \frac{1}{s\alpha} \right\}.$$

For an infinitely small α, that number becomes

$$\delta \left\{ \frac{1}{\pi^2} + \frac{1}{\pi} \right\}.$$

The series

$$\sum \epsilon_n \left\{ \left(\frac{\sin(n-1)\alpha}{(n-1)\alpha} \right)^2 - \left(\frac{\sin n\alpha}{n\alpha} \right)^2 \right\}$$

therefore approaches a limiting value, as α decreases, that cannot be larger than

$$\delta \left\{ 1 + \frac{1}{\pi} + \frac{1}{\pi^2} \right\},$$

hence must be zero. Consequently

$$\frac{F(x+2\alpha) - 2F(x) + F(x-2\alpha)}{4\alpha^2},$$

which is equal to

$$f(x) + \sum \epsilon_n \left\{ \left(\frac{\sin(n-1)\alpha}{(n-1)\alpha} \right)^2 - \left(\frac{\sin n\alpha}{n\alpha} \right)^2 \right\},$$

converges to $f(x)$ as α tends to 0. This proves our theorem for the case $\beta = \alpha$.

In order to prove the general case, let

$$F(x+\alpha+\beta)-2F(x)+F(x-\alpha-\beta)=(\alpha+\beta)^2(f(x)+\delta_1)$$
$$F(x+\alpha-\beta)-2F(x)+F(x-\alpha+\beta)=(\alpha-\beta)^2(f(x)+\delta_2),$$

from which

$$F(x+\alpha+\beta)-F(x+\alpha-\beta)-F(x-\alpha+\beta)+F(x-\alpha-\beta)$$
$$=4\alpha\beta f(x)+(\alpha+\beta)^2\delta_1-(\alpha-\beta)^2\delta_2.$$

As a consequence of the above result, δ_1 and δ_2 become infinitely small when α and β do. Then

$$\frac{(\alpha+\beta)^2}{4\alpha\beta}\delta_1-\frac{(\alpha-\beta)^2}{4\alpha\beta}\delta_2$$

will also be infinitely small if the coefficients of δ_1 and δ_2 do not become infinitely large, which does not occur since β/α remains finite. Consequently,

$$\frac{F(x+\alpha+\beta)-F(x+\alpha-\beta)-F(x-\alpha+\beta)+F(x-\alpha-\beta)}{4\alpha\beta}$$

converges to $f(x)$, as we wished to prove.

Theorem 2

$$\frac{F(x+2\alpha)+F(x-2\alpha)-2F(x)}{2\alpha}$$

tends to 0 with α for all x.

In order to prove this, we split the series

$$\sum A_n\left(\frac{\sin n\alpha}{n\alpha}\right)^2$$

into three parts. The first contains all terms up to a fixed index m, from which term on the A_n are always smaller than ϵ. The second contains all of the following terms for which $n\alpha \le$ a fixed number c. Then the third includes the rest of the series. It is then easy to see that if α decreases infinitely, the sum of the first finite part remains finite, that is, $<$ a fixed number Q; the second $< \epsilon\,\frac{c}{\alpha}$; and the third

$$<\epsilon\sum_{c<n\alpha}\frac{1}{n^2\alpha^2}<\frac{\epsilon}{\alpha c}.$$

Consequently

$$\frac{F(x+2\alpha)+F(x-2\alpha)-2F(x)}{2\alpha} = 2\alpha \sum A_n \left(\frac{\sin n\alpha}{n\alpha}\right)^2$$
$$< 2\left(Q\alpha + \epsilon\left(c + \frac{1}{c}\right)\right)$$

from which the theorem follows.

Theorem 3 *Let b and c denote two arbitrary constants with $c > b$. Let $\lambda(x)$ denote a function which is always continuous together with its first derivative between b and c, is 0 at the boundaries, and for which the second derivative does not have infinitely many maxima and minima. Then the integral*

$$\mu^2 \int_b^c F(x) \cos \mu(x-a)\, \lambda(x)\, dx,$$

is eventually less than any given number, if μ grows to infinity.

If one replaces $F(x)$ by its series expression, then one obtains for

$$\mu^2 \int_b^c F(x) \cos \mu(x-a)\, \lambda(x)\, dx$$

the series (Φ)

$$\mu^2 \int_b^c \left(C + C'x + A_0 \frac{x^2}{2}\right) \cos \mu(x-a)\, \lambda(x)\, dx$$
$$- \sum_{n=1}^{\infty} \frac{\mu^2}{n^2} \int_b^c A_n \cos \mu(x-a)\, \lambda(x)\, dx.$$

Now $A_n \cos \mu(x-a)$ can clearly be expressed as an aggregate of

$$\cos(\mu+n)(x-a),\ \sin(\mu+n)(x-a),\ \cos(\mu-n)(x-a),\ \sin(\mu-n)(x-a).$$

Denote the sum of the first two terms by $B_{\mu+n}$, and the sum of the last two terms by $B_{\mu-n}$. Then $A_n \cos \mu(x-a) = B_{\mu+n} + B_{\mu-n}$,

$$\frac{d^2 B_{\mu+n}}{dx^2} = -(\mu+n)^2 B_{\mu+n}, \quad \frac{d^2 B_{\mu-n}}{dx^2} = -(\mu-n)^2 B_{\mu-n},$$

and, with increasing n, $B_{\mu+n}$ and $B_{\mu-n}$ become infinitely small, whatever x is.

Thus the general term of the series (Φ),

$$-\frac{\mu^2}{n^2}\int_b^c A_n \cos\mu(x-a)\,\lambda(x)\,dx,$$

is equal to

$$\frac{\mu^2}{n^2(\mu+n)^2}\int_b^c \frac{d^2B_{\mu+n}}{dx^2}\,\lambda(x)\,dx+\frac{\mu^2}{n^2(\mu-n)^2}\int_b^c \frac{d^2B_{\mu-n}}{dx^2}\,\lambda(x)\,dx.$$

After two integrations by parts, in which one first considers $\lambda(x)$ and then $\lambda'(x)$ as constant, we obtain

$$\frac{\mu^2}{n^2(\mu+n)^2}\int_b^c B_{\mu+n}\,\lambda''(x)\,dx+\frac{\mu^2}{n^2(\mu-n)^2}\int_b^c B_{\mu-n}\,\lambda''(x)\,dx,$$

since $\lambda(x)$ and $\lambda'(x)$, and hence also the terms standing outside the integral sign, will be 0 at the limits.

It is now easy to convince ourselves that $\int_b^c B_{\mu\pm n}\,\lambda''(x)\,dx$ becomes infinitely small when μ grows to infinity, whatever n may be. For this expression is equal to an aggregate of the integrals

$$\int_b^c \cos(\mu\pm n)(x-a)\,\lambda''(x)\,dx,\qquad \int_b^c \sin(\mu\pm n)(x-a)\,\lambda''(x)\,dx,$$

and if $\mu\pm n$ becomes infinitely large, these integrals tend to 0. However, if $\mu\pm n$ does not become infinitely large because n is infinitely large, their coefficients in these expressions are infinitely small.

Clearly, to prove our theorem it therefore suffices that the sum

$$\sum\frac{\mu^2}{(\mu-n)^2n^2}$$

extended over all values of n which satisfy $n<-c'$, $c''<n<\mu-c'''$, $\mu+c^{IV}<n$, remains finite when μ becomes infinitely large for any choice of quantities c. For, except for the terms for which $-c'<n<c''$, $\mu-c'''<n<\mu+c^{IV}$, which clearly become infinitely small and are of finite number, the series (Φ) clearly remains smaller than the sum multiplied by the largest value of $\int_b^c B_{\mu\pm n}\,\lambda''(x)\,dx$, which becomes infinitely small.

However, if $c > 1$, the sum

$$\sum \frac{\mu^2}{(\mu-n)^2 n^2} = \frac{1}{\mu} \sum \frac{\frac{1}{\mu}}{\left(1-\frac{n}{\mu}\right)^2 \left(\frac{n}{\mu}\right)^2},$$

within the limits above, is smaller than

$$\frac{1}{\mu} \int \frac{dx}{(1-x)^2 x^2},$$

taken from

$$-\infty \text{ to } -\frac{c'-1}{\mu}, \quad \frac{c''-1}{\mu} \text{ to } 1-\frac{c'''-1}{\mu}, \quad 1+\frac{c^{IV}-1}{\mu} \text{ to } \infty.$$

For if we decompose the whole interval from $-\infty$ to ∞, starting from 0, into intervals of length $1/\mu$, and replace the function under the integral sign by the smallest value in each interval, we obtain all the terms of the series, since this function does not have maxima anywhere between the integration limits.

If the integration is carried out, we obtain,

$$\frac{1}{\mu} \int \frac{dx}{x^2(1-x)^2} = \frac{1}{\mu} \left(-\frac{1}{x} + \frac{1}{1-x} + 2\log x - 2\log(1-x) \right) + \text{const.},$$

and consequently between the above limits a number that does not become infinitely large with μ.

9.

We use these theorems to determine the following about the representation of a function by a trigonometric series whose terms tends to 0 for each value of the argument.

I. For a periodic function of period 2π to be represented by a trigonometric series whose terms eventually become infinitely small for each value of x, there must exist a continuous function $F(x)$ for which

$$\frac{F(x+\alpha+\beta) - F(x+\alpha-\beta) - F(x-\alpha+\beta) + F(x-\alpha-\beta)}{4\alpha\beta},$$

converges to $f(x)$ as α and β become infinitely small with their ratio remaining finite.

Furthermore, with increasing μ,

$$\mu^2 \int_b^c F(x) \cos \mu(x-a)\, \lambda(x)\, dx$$

must eventually become infinitely small as μ increases, if $\lambda(x)$ and $\lambda'(x)$ are 0 at the integration limits and always continuous between them, and $\lambda''(x)$ does not have infinitely many maxima and minima.

II. Conversely, if both these conditions are satisfied, then there is a trigonometric series in which the coefficients eventually become infinitely small and which represents the function, wherever it converges.

For the proof, determine the numbers C' and A_0 so that

$$F(x) - C'x - A_0 \frac{x^2}{2}$$

is a periodic function of period 2π, and expand this by Fourier's method into a trigonometric series

$$C - \frac{A_1}{1} - \frac{A_2}{4} - \frac{A_3}{9} - \cdots .$$

Here we let

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left(F(t) - C't - A_0 \frac{t^2}{2} \right) dt = C,$$

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \left(F(t) - C't - A_0 \frac{t^2}{2} \right) \cos n(x-t)\, dt = -\frac{A_n}{n^2}.$$

Then, by agreement,

$$A_n = -\frac{n^2}{\pi} \int_{-\pi}^{\pi} \left(F(t) - C't - A_0 \frac{t^2}{2} \right) \cos n(x-t)\, dt$$

must eventually become infinitely small with increasing n. It follows by Theorem 1 of the preceding section that the series

$$A_0 + A_1 + A_2 + \cdots$$

converges to the function $f(x)$, wherever it converges.

III. Let $b < x < c$, and $\varrho(t)$ be a function such that $\varrho(t)$ and $\varrho'(t)$ are 0 for $t = b$ and $t = c$ and are continuous between those values, and such that

$\varrho''(t)$ does not have infinitely many maxima and minima, and, furthermore, such that for $t = x$, $\varrho(t) = 1$, $\varrho'(t) = 0$, $\varrho''(t) = 0$, $\varrho'''(t)$ and $\varrho^{IV}(t)$ are finite and continuous. Then the difference between the series

$$A_0 + A_1 + \cdots + A_n$$

and the integral

$$\frac{1}{2\pi}\int_b^c F(t)\,\frac{d^2\frac{\sin\frac{2n+1}{2}(x-t)}{\sin\frac{(x-t)}{2}}}{dt^2}\,\varrho(t)\,dt$$

tends to zero with increasing n. Hence the series

$$A_0 + A_1 + A_2 + \cdots$$

will converge or not, depending on whether

$$\frac{1}{2\pi}\int_b^c F(t)\frac{d^2\frac{\sin\frac{2n+1}{2}(x-t)}{\sin\frac{(x-t)}{2}}}{dt^2}$$

approaches a limit with increasing n, or not.

In fact,

$$A_1 + A_2 + \ldots + A_n = \frac{1}{\pi}\int_{-\pi}^{\pi}\left(F(t) - C't - A_0\,\frac{t^2}{2}\right)\sum_{k=1}^{n} -k^2\cos k(x-t)\,dt.$$

Since

$$2\sum_{k=1}^{n} -k^2\cos k(x-t) = 2\sum_{k=1}^{n}\ \frac{d^2\cos k(x-t)}{dt^2} = \frac{d^2\frac{\sin\frac{2n+1}{2}(x-t)}{\sin\frac{(x-t)}{2}}\frac{(x-t)}{2}}{dt^2},$$

$$A_1 + A_2 + \cdots + A_n = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(F(t) - C't - A_0\,\frac{t^2}{2}\right)\frac{d^2\frac{\sin\frac{2n+1}{2}(x-t)}{\sin\frac{(x-t)}{2}}}{dt^2}\,dt.$$

Now, by Theorem 3 of the preceding section,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(F(t) - C't - A_0\,\frac{t^2}{2}\right)\frac{d^2\frac{\sin\frac{2n+1}{2}(x-t)}{\sin\frac{(x-t)}{2}}}{dt^2}\lambda(t)\,dt$$

tends to 0 with infinitely increasing n if $\lambda(t)$ along with its first derivative is continuous, $\lambda''(t)$ does not have infinitely many maxima and minima, and for $t = x$, $\lambda(t) = 0$, $\lambda'(t) = 0$, $\lambda''(t) = 0$, $\lambda'''(t)$ and $\lambda^{IV}(t)$ are finite and continuous.

Set $\lambda(t)$ equal to 1 outside the boundaries b, c and $1 - \varrho(t)$ within those boundaries, which is clearly allowable. It follows that the difference between the series $A_1 + \cdots + A_n$ and the integral

$$\frac{1}{2\pi}\int_b^c \left(F(t) - C't - A_0\frac{t^2}{2}\right) \frac{d^2 \frac{\sin\frac{2n+1}{2}(x-t)}{\sin\frac{(x-t)}{2}}}{dt^2}\,\varrho(t)\,dt$$

tends to 0 with increasing n. We easily see, by integration by parts, that

$$\frac{1}{2\pi}\int_b^c \left(C't + A_0\frac{t^2}{2}\right) \frac{d^2 \frac{\sin\frac{2n+1}{2}(x-t)}{\sin\frac{(x-t)}{2}}}{dt^2}\,\varrho(t)\,dt,$$

converges to A_0 when n becomes infinitely large, and we obtain the above theorem.

10.

It has emerged from the investigation that if the coefficients of the series Ω tend to 0, then the convergence of the series for a particular value of x depends only on the behavior of the function $f(x)$ in the immediate neighborhood of this value.

Whether the coefficients of the series eventually become infinitely small, will in many cases not be decided by their expression as a definite integral, but in other ways. One case should be emphasized where the determination can be made immediately from the nature of the function. Namely, suppose the function $f(x)$ is everywhere finite and integrable.

In this case, we split the whole interval $-\pi$ to π into a sequence of pieces of length

$$\delta_1,\ \delta_2,\ \delta_3, \ldots$$

and denote by D_1 the greatest fluctuation of the function in the first, by D_2 the greatest fluctuation in the second, and so on. Then

$$\delta_1 D_1 + \delta_2 D_2 + \delta_3 D_3 + \cdots$$

must become infinitely small when the δ's become infinitely small together.

Consider the integral $\int_{-\pi}^{\pi} f(x)\sin n(x-a)\,dx$, which, apart from the factor $1/\pi$, gives the coefficients of the series, or what is the same thing, $\int_a^{a+2\pi} f(x)\sin n(x-a)\,dx$. We split this integral beginning at $x=a$, into integrals of range $2\pi/n$. Then each integral contributes to the sum a quantity less than $2/n$ multiplied by the greatest fluctuation in its interval, and their sum is hence smaller than a number, which by assumption must become infinitely small with $2\pi/n$.

In fact, these integrals have the form

$$\int_{a+\frac{s}{n}2\pi}^{a+\frac{s+1}{n}2\pi} f(x)\sin n(x-a)\,dx.$$

The sine is positive in the first half, and negative in the second. Denoting the largest value of $f(x)$ in the interval of integration by M and the smallest by m, it is obvious that the integral is bigger if we replace $f(x)$ by M in the first half and by m in the second. The integral is smaller if $f(x)$ is replaced by m in the first half and M in the second. In the first case we obtain the value $\frac{2}{n}(M-m)$; in the other $\frac{2}{n}(m-M)$. Hence the absolute value of the integral is smaller than $\frac{2}{n}(M-m)$, and the integral

$$\int_a^{a+2\pi} f(x)\sin n(x-a)\,dx$$

is smaller than

$$\frac{2}{n}(M_1-m_1)+\frac{2}{n}(M_2-m_2)+\frac{2}{n}(M_3-m_3)+\cdots,$$

where M_s denotes the largest value of $f(x)$ in the s-th interval and m_s the smallest. However, if $f(x)$ is integrable, this sum must become infinitely small as n goes to infinity, and the lengths of the intervals $2\pi/n$ become infinitely small.

In the case under discussion, then, the coefficients of the series become infinitely small.

11.

We must still examine the case where the terms of the series Ω eventually become infinitely small for an argument value x, without this occurring for each value of the argument. This case can be reduced to the previous one.

Namely, adding the terms of equal rank in the series for values $x+t$ and $x-t$, we obtain the series

$$2A_0 + 2A_1 \cos t + 2A_2 \cos 2t + \cdots .$$

In this series the terms for each value of t eventually become infinitely small, and the previous analysis can be applied.

For this purpose, denote the value of the infinite series

$$C + C'x + A_0 \frac{x^2}{2} + A_0 \frac{t^2}{2} - A_1 \frac{\cos t}{1} - A_2 \frac{\cos 2t}{4} - A_3 \frac{\cos 3t}{9} - \cdots$$

by $G(t)$, so that wherever the series $F(x+t)$ and $F(x-t)$ converge,

$$\frac{F(x+t) + F(x-t)}{2} = G(t).$$

We have the following:

I. If the terms of the series Ω tend to 0 for an argument value x, then

$$\mu^2 \int_b^c G(t) \cos \mu(t-a)\, \lambda(t)\, dt,$$

must eventually become infinitely small with increasing μ, where λ is a function as designated in §9. The value of the integral consists of the components

$$\mu^2 \int_b^c \frac{F(x+t)}{2} \cos \mu(t-a)\, \lambda(t)\, dt \quad \text{and} \quad \mu^2 \int_b^c \frac{F(x-t)}{2} \cos \mu(t-a)\, \lambda(t)\, dt,$$

provided that these expressions have a value. Hence the integral tends to 0 because of the behavior of the function F at two places lying symmetrically on both sides of x. It should be noted, however, that the positions must be situated where each component is not itself infinitely small. For then the terms of the series would eventually become infinitely small for each value of the argument. Thus the contribution of the positions situated symmetrically on both sides of x must cancel in such a way that their sum becomes infinitely small for an infinite μ. It follows from this that the series Ω can converge only for those values of x at the midpoint of places where

$$\mu^2 \int_b^c F(x) \cos \mu(x-a)\, \lambda(x)\, dx$$

does not become infinitely small for an infinite μ. Clearly the number of those places must be infinitely large if the trigonometric series whose coefficients are not infinitely decreasing is to converge for an infinite number of argument values.

Conversely,

$$A_n = -n^2 \frac{2}{\pi} \int_0^{\pi} \left(G(t) - A_0 \frac{t^2}{2} \right) \cos nt \, dt$$

and thus A_n tends to 0 with increasing n, if

$$\mu^2 \int_b^c G(t) \cos \mu(t-a)\, \lambda(t) \, dt$$

always becomes infinitely small for infinite μ.

II. If the terms of the series Ω eventually become infinitely small for an argument value x, then whether or not the series converges depends only on the behavior of the function $G(t)$ for infinitely small t. Indeed, the difference between

$$A_0 + A_1 + \cdots + A_n$$

and the integral

$$\frac{1}{\pi} \int_0^b G(t) \frac{d^2 \frac{\sin \frac{2n+1}{2} t}{\sin \frac{t}{2}}}{dt^2} \varrho(t) \, dt$$

tends to 0 with increasing n, where b is a constant, however small, between 0 and π, and $\varrho(t)$ denotes a function such that $\varrho(t)$ and $\varrho'(t)$ are everywhere continuous and zero for $t = b$, $\varrho''(t)$ does not have infinitely many maxima and minima and for $t = 0$, $\varrho(t) = 1$, $\varrho'(t) = 0$, $\varrho''(t) = 0$, and $\varrho'''(t)$, $\varrho^{IV}(t)$ are finite and continuous.

12.

The conditions for the representation of a function by a trigonometric series can certainly be restricted a little further. Hence our examination can be extended somewhat further without special hypothesis on the nature of the functions. For example, in the last theorem the condition $\varrho''(t) = 0$ can be omitted if in the integral

$$\frac{1}{\pi} \int_0^b G(t) \frac{d^2 \frac{\sin \frac{2n+1}{2} t}{\sin \frac{t}{2}}}{dt^2} \varrho(t) \, dt,$$

$G(t)$ is replaced by $G(t) - G(0)$. However, nothing essential is gained.

Therefore we turn to the consideration of particular cases. We will first examine a function which does not have infinitely many maxima and minima. We seek to give a complete solution for this case, which is possible by the work of Dirichlet.

It is noted above that such a function is everywhere integrable where it is not infinite, and clearly that can only occur for a finite number of argument values. Also by Dirichlet's proof, in the integral expressions for the nth term of the series and for the sum of the first n terms, the contribution from all intervals eventually become infinitely small with increasing n, with the exception of those where the function becomes infinite and the infinitesimal interval enclosing the argument of the series. Further, by Dirichlet's proof,

$$\int_x^{x+b} f(t) \frac{\sin \frac{2n+1}{2}(x-t)}{\sin \frac{x-t}{2}} dt$$

will converge to $\pi f(x+0)$ as n tends to infinity, if $0 < b < \pi$ and $f(t)$ is not infinite between the integration limits. Indeed nothing more is needed when one omits the unnecessary hypothesis that the function is continuous. Hence it only remains to examine, for this integral, in which cases the contribution of the places where the function becomes infinite tends to 0 with increasing n. This investigation is still incomplete. But Dirichlet showed in passing that this takes place if the function to be represented is integrable. This hypothesis is unnecessary.

We have seen above that if the terms of the series Ω tend to zero for each value of x, the function $F(x)$ whose second derivative is $f(x)$ must be finite and continuous and that

$$\frac{F(x+\alpha) - 2F(x) + F(x-\alpha)}{\alpha}$$

always becomes infinitely small with α. Now, if $F'(x+t) - F'(x-t)$ does not have infinitely many maxima and minima, then as t tends to zero it must converge to a limit L, or become infinitely large. It is clear that likewise,

$$\frac{1}{\alpha}\int_0^{\alpha} (F'(x+t) - F'(x-t))\, dt = \frac{F(x+\alpha) - 2F(x) + F(x-\alpha)}{\alpha}$$

must converge to L or to infinity and hence can only become infinitely small if $F'(x+t) - F'(x-t)$ converges to zero. Therefore $f(a+t) + f(a-t)$ must always be integrable up to $t = 0$ if $f(x)$ is infinitely large for $x = a$. This suffices for

$$\left(\int_b^{a-\epsilon} + \int_{a+\epsilon}^c\right) dx(f(x)\cos n(x-a))$$

to converge with decreasing ϵ, and to tend to 0 with increasing n. Furthermore, since $F(x)$ is finite and continuous, then $F'(x)$ can be integrated up to $x = a$ and $(x-a)F'(x)$ becomes infinitely small with $x - a$, if this function does not have infinitely many maxima and minima. It follows that

$$\frac{d(x-a)F'(x)}{dx} = (x-a)f(x) + F'(x),$$

and hence $(x-a)f(x)$, can be integrated up to $x = a$. Therefore $\int f(x)\sin n(x-a)\,dx$ can be integrated up to $x = a$. For the coefficients of the series eventually to become infinitely small, clearly it is only necessary that

$$\int_b^c f(x)\sin n(x-a)\,dx, \quad \text{where} \quad b < a < c,$$

tends to 0 with increasing n. If one sets

$$f(x)(x-a) = \phi(x),$$

then for an infinite n, if this function does not have infinitely many maxima and minima,

$$\int_b^c f(x)\sin n(x-a)\,dx = \int_b^c \frac{\phi(x)}{x-a}\sin n(x-a)\,dx = \pi\frac{\phi(a+0)+\phi(a-0)}{2},$$

as Dirichlet has shown. Therefore

$$\phi(a+t) + \phi(a-t) = f(a+t)t - f(a-t)t$$

must tend to 0 with t. Since

$$f(a+t) + f(a-t)$$

is integrable up to $t = 0$ and consequently

$$f(a+t)t + f(a-t)t$$

also becomes infinitely small with decreasing t, then $f(a+t)t$ as well as $f(a-t)t$ tend to 0 with decreasing t. Apart from the functions which have infinitely many maxima and minima, it is necessary and sufficient for the representation of a function $f(x)$ by a trigonometric series whose coefficients tend to 0, that if f become infinite for $x = a$, then $f(a+t)t$ and $f(a-t)t$ tend to 0 with t and $f(a+t) + f(a-t)$ is integrable up to $t = 0$.

A function $f(t)$ which does not have infinitely many maxima and minima can be represented only for finitely many values of the argument by a trigonometric series whose coefficients do not eventually tend to 0. For

$$\mu^2 \int_b^c F(x) \cos \mu(x-a)\, \lambda(x)\, dx$$

fails to tend to 0 as μ becomes infinite, at only a finite number of values. Hence it is unnecessary to consider this further.

13.

Concerning functions with infinitely many maxima and minima, it is probably not superfluous to note that there exists such a function $f(x)$, everywhere integrable, that cannot be represented by a Fourier series. This occurs, for example, if

$$f(x) = \frac{d\left(x^\nu \cos \frac{1}{x}\right)}{dx}, \quad \text{for } 0 \le x \le 2\pi, \text{ and } 0 < \nu < 1/2.$$

For the contribution in the integral $\int_0^{2\pi} f(x) \cos n(x-a)\, dx$ with increasing n of those places where x is close to $\sqrt{\frac{1}{n}}$ is, generally speaking, eventually infinitely large, so that the ratio of this integral to

$$\frac{1}{2} \sin\left(2\sqrt{n} - na + \frac{\pi}{4}\right) \sqrt{\pi} n^{-\frac{1-2\nu}{4}}$$

converges to 1, as we find by the method just described. In order to generalize the the example, and bring out the essence of the matter, let

$$\int f(x)\, dx = \phi(x) \cos \psi(x)$$

and assume that $\phi(x)$ is infinitely small for an infinitely small x, and $\psi(x)$ becomes infinitely large, and elsewhere these functions together with their derivatives are continuous and do not have infinitely many maxima and minima. Then

$$f(x) = \phi'(x)\cos\psi(x) - \phi(x)\psi'(x)\sin\psi(x),$$

and

$$\int f(x)\cos n(x-a)\,dx$$

is the sum of the four integrals

$$\frac{1}{2}\int \phi'(x)\cos(\psi(x) \pm n(x-a))\,dx,$$
$$-\frac{1}{2}\int \phi(x)\psi'(x)\sin(\psi(x) \pm n(x-a))\,dx.$$

Taking $\psi(x)$ positive, we consider the term

$$-\frac{1}{2}\int \phi(x)\psi'(x)\sin(\psi(x) + n(x-a))\,dx$$

and examine in this integral the place where the changes of sign of the sine follow one another most slowly. Let

$$\psi(x) + n(x-a) = y,$$

then this occurs where $\frac{dy}{dx} = 0$. Thus $x = \alpha$ with

$$\psi'(\alpha) + n = 0.$$

We therefore examine the behavior of the integral

$$-\frac{1}{2}\int_{\alpha-\epsilon}^{\alpha+\epsilon} \phi(x)\psi'(x)\sin y\,dx$$

in the case that ϵ becomes infinitely small for an infinite n, and introduce y as a variable. Let

$$\psi(\alpha) + n(\alpha - a) = \beta,$$

then for sufficiently small ϵ

$$y = \beta + \psi''(\alpha)\frac{(x-\alpha)^2}{2} + \cdots$$

and, indeed, $\psi''(\alpha)$ is positive, since $\psi(x)$ tends to $+\infty$ as x tends to 0. Furthermore,

$$\frac{dy}{dx} = \psi''(\alpha)(x-\alpha) = \pm\sqrt{2\psi''(\alpha)(y-\beta)},$$

depending on whether $x - \alpha > 0$ or < 0; and

$$\begin{aligned} -\frac{1}{2}\int_{\alpha-\epsilon}^{\alpha+\epsilon} & \phi(x)\psi'(x)\sin y\,dx \\ &= \frac{1}{2}\left(\int_{\beta+\psi''(\alpha)\frac{\epsilon^2}{2}}^{\beta} - \int_{\beta}^{\beta+\psi''(\alpha)\frac{\epsilon^2}{2}}\right)\frac{\phi(\alpha)\psi'(\alpha)}{\sqrt{2\psi''(\alpha)}}\left(\sin y\frac{dy}{\sqrt{y-\beta}}\right) \\ &= -\int_0^{\psi''(\alpha)\frac{\epsilon^2}{2}} \sin(y+\beta)\,\frac{\phi(\alpha)\psi'(\alpha)}{\sqrt{2\psi''(\alpha)}}\,\frac{dy}{\sqrt{y}}. \end{aligned}$$

Let ϵ decrease with increasing n so that $\psi''(\alpha)\epsilon^2$ becomes infinitely large. If

$$\int_0^{\infty} \sin(y+\beta)\,\frac{dy}{\sqrt{y}},$$

which is known to be $\sin(\beta+\pi/4)\sqrt{\pi}$, is not zero, then disregarding quantities of lower order,

$$-\frac{1}{2}\int_{\alpha-\epsilon}^{\alpha+\epsilon} \phi(x)\psi'(x)\sin(\psi(x)+n(x-a))\,dx = -\sin\left(\beta+\frac{\pi}{4}\right)\frac{\sqrt{\pi}\phi(\alpha)\psi'(\alpha)}{\sqrt{2\psi''(\alpha)}}.$$

Hence, if the last quantity does not become infinitely small, its ratio to

$$\int_0^{2\pi} f(x)\cos n(x-a)\,dx$$

converges to 1 with an infinite increase of n, since the remaining contributions become infinitely small.

Assume that $\phi(x)$ and $\psi'(x)$ are of the same order as powers of x for infinitely small x, with $\phi(x)$ of the order of x^{ν} and $\psi'(x)$ of the order of $x^{-\mu-1}$, where we must have $\nu > 0$ and $\mu \geq 0$. Then for infinite n,

$$\frac{\phi(\alpha)\psi'(\alpha)}{\sqrt{2\psi''(\alpha)}}$$

has the same order as $\alpha^{\nu-\frac{\mu}{2}}$ and hence is not infinitely small when $\mu \geq 2\nu$. In general however, if $x\psi'(x)$ or, what is the same thing, if $\frac{\psi(x)}{\log x}$ is infinitely large for an infinitely small x, $\phi(x)$ can be taken so that $\phi(x)$ tends to 0 with x, while

$$\phi(x)\,\frac{\psi'(x)}{\sqrt{2\psi''(x)}} = \frac{\phi(x)}{\sqrt{-2\frac{d}{dx}\,\frac{1}{\psi'(x)}}} = \frac{\phi(x)}{\sqrt{-2\lim\frac{1}{x\psi'(x)}}}$$

will be infinitely large. Consequently $\int_x f(x)\,dx$ can be extended to $x = 0$, while

$$\int_0^{2\pi} f(x)\cos n(x-a)\,dx$$

does not become infinitely small for an infinite n. We see that the increases in the integral $\int_x f(x)\,dx$ as x tends to 0 cancel out because of the rapid changes of sign of the function $f(x)$, although their variation increases very rapidly in ratio to the change of x. However, the introduction here of the factor $\cos n(x-a)$ results in this increase being summable.

Just as, in the above, the Fourier series does not converge for a function in spite of the overall integrability, and the terms themselves eventually become infinitely large, it can happen that, despite the overall non-integrability of $f(x)$, between each two values of x, no matter how close, there are infinitely many values for which the series Ω converges.

An example is given by the function defined by the series

$$\sum_{n=1}^{\infty}\frac{(nx)}{n}$$

which exists for each rational value of x, where the meaning of (nx) is taken as in §6. This can be represented by the trigonometric series

$$\sum_{n=1}^{\infty}\frac{\Sigma^{\theta}-(-1)^{\theta}}{n\pi}\sin 2nx\pi,$$

where θ runs over the divisors of n. The function is not bounded in any interval, no matter how small, and hence is nowhere integrable.

Another example is obtained if in the series

$$\sum_{n=0}^{\infty}c_n\cos n^2x, \qquad \sum_{n=1}^{\infty}c_n\sin n^2x$$

$c_0, c_1, c_2, \ldots$ are positive numbers which always decrease and tend to 0, while $\sum_{s=1}^{n} c_s$ becomes infinitely large with n. For if the ratio of x to 2π is rational and in lowest terms has denominator m, then clearly the series converges or tends to infinity depending on whether

$$\sum_{n=0}^{m-1} \cos n^2 x, \qquad \sum_{n=0}^{m-1} \sin n^2 x$$

are zero or not. Both cases arise, by a well known theorem[33] on partitioning the circle, for infinitely many values of x between any two bounds, no matter how close.

The series Ω can converge in a range just as large, without the value of the series

$$C' + A_0 x - \sum \frac{1}{n^2} \frac{dA_n}{dx},$$

which one obtains by termwise integration of Ω, being integrable on any interval, however small.

For example, we expand the expression

$$\sum_{n=1}^{\infty} \frac{1}{n^3}(1 - q^n) \log \left(\frac{-\log(1 - q^n)}{q^n} \right),$$

where the logarithms are taken so that they vanish for $q = 0$, by increasing powers of q, and replace q by e^{xi}. The imaginary part is a trigonometric series whose second derivative with respect to x converges infinitely often on any interval, while its first derivative becomes infinite infinitely often.

In the same range, that is, between any two argument values no matter how close, a trigonometric series can also converge infinitely often when its coefficients do not tend to 0. A simple example of such a series is given by $\sum_{n=1}^{\infty} \sin(n! x\pi)$, where as usual,

$$n! = 1 \cdot 2 \cdot 3 \cdots n.$$

This not only converges for each rational value of x, for which it changes into a finite sum, but also for an infinite number of irrationals, of which the simplest are $\sin 1$, $\cos 1$, $2/e$ and their multiples, odd multiples of e, $\frac{e - \frac{1}{e}}{4}$, and so on.

[33]Disquis. ar. p. 636, §356. (Gauss, *Werke*, vol. I, p. 442.)

Contents.

XIII.

The hypotheses on which geometry is based.

(*Königlichen Gesellschaft der Wissenschaften zu Göttingen*, vol. 13.)

Plan of the investigation.

It is known that geometry takes for granted the notion of space as well as the fundamental first principles used in constructions carried out in space. Only nominal definitions are given of these basic concepts, while the essential role in determining their properties is played by the axioms. The relationships between the assumptions embodied in these axioms, however, remain obscure. It is not clear whether, and if so to what extent, they are necessarily linked; or whether, *a priori*, they are even possible.

This obscurity has existed from Euclid to Legendre, to name the most famous of recent geometers, but neither the mathematicians nor the philosophers who have concerned themselves with this problem have dispelled it. The reason for this may well be that the general concept of quantities having several dimensions, which incorporates the notions that we use in geometry, remains a field in which no work has ever been done. I have accordingly set myself as a first task the construction of the concept of a *multi-dimensional quantity* from the general concept of *magnitude*. It emerges from this study that a multi-dimensional object is capable of being measured in different ways and that space is only a particular example of the case of a three-dimensional quantity. It necessarily follows from this that the theorems of geometry cannot be deduced from the general notion of magnitude alone, but only from those properties which distinguish space from other conceivable three-dimensional entities, and these properties can only be found *experimentally*. This raises the problem of seeking out the simplest facts which enable us to determine the metric relationships of space—a problem which by its very nature can never be completely decided, because there are several different systems of "simple facts" which suffice to determine the metric relationships of space. The most important of these for our present purposes is the system of axioms laid down by Euclid. These, like all facts based on observation, are not *necessary truths*, they have only empirical certainty and are indeed hypotheses. We can therefore investigate their probability, which is undoubtedly very great within the limits of observation, and from there form a judgment as to the admissibility of extending them outside the limits of observation, in the realms of both the immeasurably great and the immeasurably small.

I.
The notion of an n-dimensional quantity.

As I try to solve the first of these problems, that of developing the concept of multi-dimensional entities, I feel obliged to ask for indulgence on the part of my listeners, because I am unused to dealing with matters of a philosophical nature where the difficulties reside in the concepts themselves rather than in their constructs. Apart from a few brief indications mentioned by Privy Councillor Gauss in his second memoir on biquadratic residues published in the *Göttingenschen gelehrte Anzeigen*, in his *Jubiläumschrift*, and some philosophical researches of Herbart, there has been no previous work which I could use.

1.

The idea that an entity has a magnitude is possible only where the entity falls under some more general concept which allows its size to be defined in a variety of different ways. Depending on whether or not it is possible to make the transition from one mode of determination to another in a continuous fashion, the modes of determination constitute a continuous or discrete manifold. In the former case, these individual modes of determination are called *points* and in the latter case, the *elements* of the manifold. Concepts whose modes of determination form a discrete manifold are so frequent that in all the more highly developed languages it is always possible to find one which includes them. Mathematicians were therefore, in the theory of discrete magnitudes, able to regard given things as belonging to the same class, without hesitation. On the other hand, occasions which give rise to notions whose measurement involves the consideration of continuous manifolds are so rarely encountered in everyday life that the location of material objects perceived through the senses, and colors, are perhaps the only simple examples of concepts whose modes of determination constitute a multi-dimensional manifold. Not until we enter the realm of higher mathematics does the need to create and develop such concepts make itself felt.

Well-defined parts of a manifold, distinguished by some characteristic feature or boundary, are called quanta. Their quantitative comparison is made, in the discrete case, by counting, and in the continuous case, by measuring. Measurement is done by superimposing one of the quantities to be measured on the other. To measure something, therefore, requires a means of transporting one of the quantities which can act as a scale against which the other can be compared. If this means is not available, we can compare two quantities only when one of them is a part of the other, and even then we can only compare greater and lesser; we cannot say by how much. The investigation of this case forms a branch of the theory of quantities independent of measurement. Here quantities do not exist independently of their position in space, and cannot be expressed via an unit but must be regarded as regions in a manifold. Such studies have become a necessity in various parts of mathematics, notably in the treatment of multi-valued analytic functions. The lack of these studies may well be one of the main reasons why Abel's famous theorem and the contributions of Lagrange, Pfaff, and Jacobi to the general theory of differential equations have remained unfruitful for so long. From this general theory of multi-dimensional entities, in which nothing more is assumed than is inherent in its axiomatic definition, it will suffice for our present purpose to emphasize two points. The first concerns the creation of the concept of a multi-dimensional manifold. The second relates to the reduction of the determination of position within a given manifold to quantitative determinations, and will make clear the essential character of an n-dimensional entity.

2.

If, in a concept whose modes of determination constitute a continuous manifold, we pass from one mode of determination to another in a well-defined manner, then the modes of determination utilized form a simply-extended manifold whose essential feature is that a continuous progression from one point to the next is possible in only two different directions, forwards or backwards. Now imagine that this manifold is moved into a completely different one, again in a well-defined manner (that is, in such a way that each point of the one is transformed into a well-defined point of the other). Then the various modes of determination so obtained constitute a two-dimensional manifold. In a similar fashion we obtain a three-dimensional manifold if we imagine a two-dimensional manifold moving into a completely different one by some specified transformation; and it is easy to see how this process can be extended. If, instead of taking the notion of a quantity that can be determined, we treat the conceived object as the variable, then we could describe this construction as the composition of a variety of $n+1$ dimensions from a variety of n dimensions and a one-dimensional variety.

3.

I shall now show how, conversely, a variety in a given domain can be broken down into a variety of fewer dimensions and a one-dimensional variety. For this purpose, let us visualize a variable segment of a one-dimensional manifold (measured from a fixed point, so that the lengths of these variable segments can all be compared one to another) and suppose that the points of this given manifold each have a well-defined associated value which varies in a continuous fashion. In other words, we assume that a continuous function of position within the manifold has been defined; in fact, a function which is never constant along any part of this manifold. Every set of points for which the function has a constant value then defines a continuous manifold which has a smaller number of dimensions than the given manifold. Whenever the function changes, so does this corresponding manifold and in a continuous manner. Thus we can assume that from one of these manifolds all the others can be derived continuously, and, generally speaking, the correspondence will be such that every point of one manifold is transformed into a definite point of the other. There are exceptions to this general rule, which need careful investigation, but we may ignore these for the present. By this means the determination of position in the given manifold is reduced to the determination of a parameter and the determination of position in a manifold of lesser dimensions. It is now easy to show that this latter manifold must be one of $n-1$ dimensions, when the given manifold has n dimensions. By repeating the process n times, the determination of position in an n-dimensional manifold can be reduced to n quantitative determinations. Accordingly, the determination of the position of an element in a given manifold, when this is possible, can be reduced to the determination of a finite number of quantities. There are also manifolds in which the determination of position requires not merely a finite number of quantitative determinations but rather an infinite series or even a continuous manifold of them. Such manifolds are, for example, the possible determinations of a function in a given domain, the possible shapes of a figure in space, and so on.

II.
Metric relationships which can exist in a manifold of n dimensions, on the assumption that the lines have a length which is independent of their situation, so that every line can be measured by every other line.

After having constructed the concept of an n-dimensional variety, and having found as its essential characteristic that defining the position of its elements can be reduced to the determination of n quantities, we turn to the second of the problems posed earlier. This is the study of the metric relationships which can hold in such a manifold and the conditions which suffice to determine these relations. These metric relations can only be studied by using symbols representing abstract quantities and representing their dependence by means of formulae. Under certain hypotheses, however, the relationships can be broken down into simpler relations which can be individually interpreted geometrically, so that the result of the calculation can be expressed in geometric language. In order to remain on *terra firma* we cannot avoid working with abstract formulae, but the results of the investigation can be expressed in a geometric form. The foundations for these two aspects of the question were laid down by Privy Councillor Gauss in his celebrated memoir on curved surfaces.

1.

Measurement requires that the measure of the entities being measured must be independent of their location, and this can be the case in more than one way. The assumption which first suggests itself, and which I intend to pursue here, is that the length of lines is independent of their position, so that every line can be measured by comparing it with any other line. If the determination of the position of a point in a given n-dimensional manifold is reduced to the determination of n variables $x_1, x_2, x_3, \ldots, x_n$, then a line may be defined by the statement that the quantities x are given functions of a single variable. The problem then is to find a mathematical expression for the length of a line, and for this purpose we need to consider the quantities x as expressible in terms of units. I shall handle this problem only under certain restrictions, and confine myself in the first place to lines in which the relations between the quantities dx—the associated variations of the variables x—vary in continuous fashion. We can then visualize the line as being divided up into elements, within which the ratios of the increments dx can be regarded as constant, and the problem reduces to finding a general expression for the line element ds starting from a given point, which will involve the variables x as well as the variables dx. Secondly, I shall assume that the length of the line element, disregarding quantities of the second order of magnitude, remains unchanged if all its points undergo the same infinitesimal displacement. This assumption implies that if all the variables dx are proportionally increased in size, the line element will increase in the same proportion. Under these assumptions, the line element could be an arbitrary homogeneous function of the first degree in the variables dx, that remains unaltered by changing the sign of all dx, and such that the arbitrary constants are continuous functions of the variables x. In order to find the simplest cases, I first seek an expression for the $(n-1)$-dimensional manifolds which are everywhere equidistant from the point of origin of the line element; that is, I look for a continuous function of position which can be used to distinguish one such manifold from another. This function will have the property that it either continually decreases or else continually increases as we move away from the origin in any direction. I shall suppose that it increases in every direction, so that it has a minimum at the origin. It must then, if its first and second derivatives are finite, be a function whose first derivative vanishes and whose second derivative is never negative. I shall suppose that it is always positive. This differential expression of the second order therefore remains constant when ds remains constant and increases in the same proportion as the square of dx if the variables dx and hence ds are all increased proportionately. Thus the function can be written as const. ds^2, and consequently ds is the square root of an everywhere positive quadratic form in the variables dx whose coefficients are continuous functions of the variables x. For space, when the position of a point is expressed in rectangular coordinates, we have

$$ds = \sqrt{\sum (dx)^2}.$$

Thus space falls under this simplest case. The next simplest case would comprise manifolds in which the line element can be expressed by the fourth root of a quartic differential expression. The study of this more general class would not, it is true, require any essentially new principles, but would be time-consuming and probably throw relatively little new light on the theory of space, particularly since the results would not lend themselves to geometric form. I therefore restrict myself to manifolds in which the line element is expressed as the square root of a quadratic differential expression. Such an expression can be transformed into a similar one by substituting, for the n independent variables, functions of n new independent variables. Not every such expression, however, can be transformed into any other in this way; for clearly the expression contains $n(n+1)/2$ coefficients which are arbitrary functions of the independent variables. By introducing n new variables, we can only satisfy n different relations and thus can only make n of the coefficients assume specified values. The other $n(n-1)/2$ coefficients are already completely determined by the nature of the manifold to be represented, and to define its metric properties $n(n-1)/2$ functions of position are required. The manifolds in which, as in the plane and in space, the line element can be expressed in the form $\sqrt{\sum(dx)^2}$ thus constitute a special case of the manifolds under investigation here. They merit a special name, and I designate those manifolds in which the square of the infinitesimal line element can be brought into the form of a sum of the squares of the individual differentials as *flat*. In order to review the essential differences between the diverse flat manifolds, we need to put aside differences which arise from the particular mode of representation. This will be achieved by choosing the variables in accordance with a certain principle.

2.

To this end, imagine that given an arbitrary point of the manifold, a system of shortest lines be constructed emanating from this origin. The position of an indeterminate point will then be defined by the initial direction of the particular shortest line on which it lies, and its distance from the origin measured along this shortest line. This position can therefore be expressed by the ratios of the quantities dx^0, that is, the initial values of dx at the origin of this shortest line, and the length s of the line. We now introduce, instead of dx^0, new linear expressions formed from them, $d\alpha$, such that the initial value of the square of the line element is equal to the sum of the squares of the $d\alpha$. The independent variables are now: the length s, and the ratios of the $d\alpha$. Finally we replace the $d\alpha$ by quantities $x_1, x_2, \ldots x_n$, proportional to them, which have s^2 as the sum of their squares. If we introduce these quantities as the new variables, then for infinitely small values of x, the square of the line element will be $\sum(dx)^2$, but the term of the next higher order of magnitude in the expression will be a quadratic form in the $n(n-1)/2$ quantities $x_1dx_2 - x_2dx_1$, $x_1dx_3 - x_3dx_1, \ldots$, that is, an infinitesimal of the fourth dimension. Thus we obtain a finite number if this expression is divided by the square of the area of the infinitesimal triangle at whose vertices the values of the variables are $(0, 0, 0, \ldots), (x_1, x_2, x_3, \ldots)$ and $(dx_1, dx_2, dx_3, \ldots)$. This number remains unchanged as long as the quantities x and the quantities dx are contained in the same binary linear forms, or in other words, as long as the shortest line from $(0, 0, 0, \ldots)$ to $(x_1, x_2, x_3, \ldots)$ and the shortest line from $(0, 0, 0, \ldots)$ to $(dx_1, dx_2, dx_3, \ldots)$ remain in the same surface element, and thus the number depends only on the position and orientation of this element. Its value will obviously be zero when the manifold is flat, that is, when the the square of the infinitesimal line element is reducible to the form $\sum(dx)^2$, and can therefore be regarded as a measure of the departure from flatness of the manifold at the point in question in the surface-direction concerned. When multiplied by $-3/4$, it is equal to the quantity called by Privy Councillor Gauss the curvature [*Krümmungsmass*] of the surface. To determine the metric relations in an n-dimensional manifold susceptible of representation in the assumed form, we found earlier that $n(n-1)/2$ functions of position were needed. If, therefore, the curvature at each point is specified for each of $n(n-1)/2$ surface-directions, then all the metric relations in the manifold can be deduced, unless there are some identical relations between the specified values which, generally speaking, will not be the case. The metric relationships of those manifolds in which the line element is represented by the square root of a quadratic differential expression can thus be expressed in a form which is completely independent of the variables. An entirely analogous method can be employed to achieve the same goal in the case of manifolds in which the line element has a less simple form, for example, the fourth root of a quartic differential expression. Generally speaking, the line element would then no longer be capable of being expressed as the square root of a sum of squares of differential expressions. Consequently in the expression for the square of the line element, the departure from flatness would be an infinitesimal quantity of degree 2, rather than degree 4, as in the case previously considered. This characteristic property of the latter manifolds might well be called *planarity in the smallest parts.* For our present purposes, the most important property of the manifolds which we have considered so far, and indeed the main motivation for this investigation, is that the metric relationships in 2-dimensional manifolds can be interpreted geometrically by surfaces, and those in higher dimensional manifolds can be reduced to relations in the surfaces which are contained in them. This calls for a brief further explanation.

3.

In apprehending the concept of a surface, besides the intrinsic properties which yield the lengths of paths in the surface, there are extrinsic properties which concern the positions of the points in the surface in relation to points outside it. We can, however, abstract these latter properties by considering at the same time as the given surface all those surfaces into which it can be moved by a transformation which leaves unaltered the lengths of lines lying wholly within the surface, and regarding such surfaces as equivalent. We think of such surfaces as being derived by bending but not stretching. The surface of a cone or a cylinder are examples of surfaces which are equivalent to a plane. For they can be formed from a plane merely by bending, which does not alter metric relationships within the surface, and the results of planimetry remain valid. The situation is essentially different in the case of the surface of a sphere, as it cannot be transformed into a plane surface without stretching. From the foregoing investigations, it is clear that in a 2-dimensional manifold, where, as is the case with surfaces, the line element is the square root of a quadratic differential expression, the intrinsic metric relationships at each point are characterized by the curvature. This number has a simple intuitive interpretation in the case of surfaces; it is the product of the two principal curvatures of the surface at the point. Alternatively, the product of this number by the area of an infinitesimal triangle whose sides are lines of shortest length in the surface is half the excess of the sum of its angles over two right angles. The first definition assumes the theorem that the product of the two principal radii of curvature at any point on a surface remains unaltered by a mere bending operation. The second definition assumes that at this point the total angle, less π, of an infinitesimal triangle is proportional to its area. In order to obtain an easily grasped interpretation of the curvature in the case of an n-dimensional manifold for a given point and a given surface direction, we begin from the idea that a shortest line emanating from the given point is fully determined by its initial direction. With the help of this notion, a well-defined surface is obtained when all the lines emanating from the given point, in every possible initial direction (and lying within the given infinitesimal surface element), are prolonged into lines of shortest length. The surface formed in this way has at each point a definite curvature which is the curvature of the n-dimensional manifold at the point in question and in the given surface direction.

4.

Before applying the foregoing to space, it will be necessary to consider a few matters concerning flat manifolds in general, that is, manifolds in which the square of the line element is expressible as the sum of squares of complete differentials.

In a flat n-dimensional manifold, the curvature at every point and in every direction is zero. In accordance with our earlier investigations, in order to specify the metric relations, it suffices to know that at every point of the surface, the curvature is zero in $n(n-1)/2$ different surface directions (whose curvatures are independent). The manifolds whose curvature is everywhere zero can be regarded as a particular case of those whose curvature is everywhere constant. Manifolds with constant curvature can also be characterized by the property that geometric figures within the variety concerned can be freely moved around without stretching. For it is obvious that figures within the variety could not be translated and rotated unless the curvature at each point of the surface and in each direction were the same. On the other hand, the curvature completely determines the metric relations of the manifold, so that all the metric relations within a manifold with constant curvature at every point and in every direction are completely defined by this constant and are the same at each point in all directions. As the same constructions can be carried out in each direction starting from a given point, the figure can be moved so that it occupies an arbitrary position in the manifold. The metric relations of these manifolds depend only on the value of the curvature. As far as their analytical expression is concerned, it may be noted that the expression for the line element can be given the form

$$\frac{1}{1+\frac{\alpha}{4}\sum x^2}\sqrt{\sum (dx)^2},$$

where α is the curvature.

5.

To elucidate the foregoing in a geometric context, the case of surfaces with constant curvature may serve as an example. It is easily seen that the surfaces with a constant positive curvature can always be applied without stretching on the surface of a sphere whose radius is the reciprocal of the square root of the curvature. However, in order to embrace the whole class of surfaces of constant curvature, let us give to one of them the shape of a sphere, and to the others the shape of the surface of a solid of revolution, touching the sphere along its equator. The surfaces whose curvature exceeds that of the sphere will then be in contact with the sphere on the inside, and in the neighborhood of the line of contact will have a shape similar to the external surface of a torus at the points furthest from its axis. They could be applied along zones of a sphere of smaller radius, but would cover these zones more than once. The surfaces with a smaller positive curvature can be obtained by cutting out of the surface of a sphere of larger radius, a piece of surface bounded by two halves of great circles, and fitting together the lines of the cuts. The surface of zero curvature is the curved surface of a circular cylinder whose base is the equator of the sphere. The surfaces with a negative constant curvature will touch the cylinder externally and have a shape similar to the surface of a torus nearest the axis. If we regard all these surfaces as regions in which a piece of surface can be moved around from one position to another, in the same way as space is a region in which rigid bodies can be moved from one position to another, then in all these surfaces, the pieces of surface are freely mobile without stretching. The surfaces of positive curvature can moreover always be given a form such that the movable surface elements can be moved arbitrary without undergoing any bending, and this form will be that of a sphere. This will no longer be possible for surfaces of negative constant curvature. In addition to this independence of surface elements from position, there is, in the case of surfaces of zero curvature, an independence of orientation from position which does not hold for other surfaces.

III.
Application to space.

1.

After this investigation into the determination of the metric relations between n-dimensional entities, it is now possible to indicate a set of necessary and sufficient conditions for the determination of these relations in space. We make the hypotheses that the lengths of lines are independent of position, and that the length of the infinitesimal line element is expressible as the square root of a quadratic differential expression, so that flatness in the smallest parts is assumed.

Now in the first place, this property can be expressed by saying that the curvature at any point in three different surface directions is zero. The metric relations in the space are, in particular, fully determined if the sum of the angles of a triangle is always equal to two right-angles.

If, secondly, we assume, as did Euclid, an existence not only of line segments whose lengths are independent of their position in space but also of solids whose dimensions are likewise independent of their position in space, it follows that the curvature is everywhere constant, and the sum of the internal angles of a triangle is determined by its value for any given triangle.

Lastly, we could, instead of assuming that the length of line segments is independent of their position and direction, assume that their length and direction is independent of their position. If we adopt this point of view, changes of position or differences in position are expressions in three independent units.

2.

In the course of our presentation we have taken care to separate the topological relations [*Ausdehnungs- oder Gebietsverhältnisse*] from the metric relations. We found that different measurement systems are conceivable for one and the same topological structure, and we have sought to find a simple system of measurements which allows all the metric relations in this space to be fully determined and all metric theorems applying to this space to be deduced as a necessary consequence. It now only remains to discuss the question of whether and to what extent the assumptions which we have made are confirmed by experience. In this connection there is an essential difference between topological and metric relations. The various possible cases of the former constitute a discrete manifold where the facts revealed by experience can be expressed by statements of whose truth one can never be certain but which are at least not inexact. However, the latter relations form a continuous manifold where the measurements revealed by experience must necessarily be inexact, no matter how great the probability of the measurement being correct. This circumstance becomes important when empirical determinations based on experimental observations are extrapolated outside the limits of observation, to the regions of the immeasurably large and the immeasurably small. It is obvious that within the limits of observation the measurements in the latter case become less and less precise, but not in the former case.

When we extend constructions in space to the immeasurably large, a distinction has to be made between the unlimited and the infinite; the first applies to relations of a topological nature, the second to metric relations. That space is an unlimited three-dimensional manifold is a hypothesis that is applied in all our conceptions of the external world, at all times allowing us to complete our perception of the universe by extending the domain of which we are truly aware, and to construct the possible positions of some particular object. It is a hypothesis which is constantly being verified in its applications. The property of the unboundedness of space possesses therefore a greater empirical certainty than any other fact established by observation. However, the infiniteness of space does not in any way follow from this. On the contrary, if one assumes that the size of solid bodies does not depend on their position, and in consequence ascribes to space a constant curvature, then space is necessarily finite, whenever this constant is positive, no matter how small. If we were to extend the infinitesimal line elements in each initial direction in an element of surface into line segments along paths of shortest length, the lines would lie in an unbounded surface of constant curvature. This surface in a flat three-dimensional manifold would take the form of a sphere, so that it would be finite.

3.

Questions relating to immeasurably large quantities have no relevance to the elucidation of natural phenomena. Questions relating to the immeasurably small are, however, another matter. The degree of accuracy with which we can pursue phenomena toward the infinitely small has a profound effect on our knowledge of the causal relationships between them. The advances during the last few centuries in our knowledge of the mechanism of Nature have been almost entirely due to the accuracy of the models which have been constructed, following the discovery of infinitesimal analysis, and the simple basic concepts found by Archimedes, Galileo, and Newton, which are used in present-day physics. In those natural sciences where the simple concepts needed to construct such models are not yet available, we try to extend the examination of phenomena into ever smaller regions of space as far as the use of the microscope permits, in order to understand the causal relationships involved. Hence questions about metric relations of space concerning the immeasurably small are not superfluous.

If we assume that bodies have a size independent of their position, the curvature is everywhere constant, and it follows from astronomical measurements that it cannot be other than zero. At any rate, the reciprocal of the curvature would correspond to a surface compared with whose radius the range of our present telescopes would be negligibly small. If, however, the independence of the size of bodies from their position is not assumed, no deduction can be made as to the metric relations in the small from those in the large. At every point in space the curvature in three directions could have arbitrary values, provided that the curvature of every measurable portion of space did not differ markedly from zero. Even more complicated relationships could obtain if we suppose that the line element cannot be represented as the square root of a quadratic differential expression. Now it seems that the empirical concepts on which the metric determination of extent are based, the concept of a rigid body, and the concept of a light ray, cease to be meaningful in the realm of the infinitely small. It is therefore quite conceivable that metric relations in the infinitesimal domain do not obey the usual axioms of geometry, and we ought to modify our assumptions accordingly, if thereby observed phenomena can be explained more simply.

The question of the validity of the axioms of geometry in the infinitely small is bound up with the fundamentals of metric relations in space. The latter question can be regarded as belonging to the theory of abstract spaces, and the remark made earlier applies. In any discrete manifold, the principle used for the measurement of the size is already present in the definition of the manifold; but where the variety is continuous, this principle must come from somewhere else. Accordingly, either the physical reality on which space is founded must be a discrete variety, or else the foundation of its metric relations must be sought from some outside source in the forces which bind together its elements.

The answer to these questions can be found only by starting from the existing theories, whose foundations were laid by Newton and which were derived from a study of phenomena. These theories and concepts must then be gradually modified whenever facts are encountered which cannot be explained on the existing basis. Investigations such as the present one, which proceed from concepts of a general nature, can only be helpful in preventing this work from being hindered by too narrow a view of the possibilities, so that advances in our understanding of the universe are not hampered by traditional prejudices.

This takes us into the realm of another science—physics—which the nature of today's occasion does not allow us to explore.

Summary.

[1] Article I serves equally as a preface for contributions to analysis situs.

[2] The investigation into the possible metric relations of an n-dimensional manifold is very far from complete, but probably sufficient for our present purposes.

[3]§3 of part III requires reworking and further development.

XIV.

A contribution to electrodynamics.

(*Poggendorff's Annalen der Physik und Chemie*, vol. 131.)

I take the liberty of communicating a remark to the Royal Society that brings the theory of electricity and magnetism into a closer connection with that of light and radiant heat. I have found that the electrodynamic influence of galvanic currents may be explained if we suppose that the influence of one electrical mass on another is not instantaneous. Rather, it propagates towards it with a constant velocity (equal to the velocity of light within the limits of observational error). With this hypothesis, the differential equation for the propagation of electrical force will be the same as that for the propagation of light and radiant heat.

Let S and S' be two conductors, through which constant galvanic currents flow, that are not in relative motion. Let ϵ be a particle of electrical mass in the conductor S whose position at time t is (x, y, z), ϵ' an electrical particle of S' with position (x', y', z') at time t'. Now consider the motion of the electrical particles with opposing effect for positive and negative electricity in each small portion of the conductor. I assume that these motions are distributed at any given instant so that the sums

$$\sum \epsilon f(x, y, z), \quad \sum \epsilon' f(x', y', z'),$$

taken over all particles of the conductor, are negligible compared to the corresponding sums taken over either positive or negative particles, provided that f and its partial derivatives are continuous.

This hypothesis can be fulfilled in a variety of ways. For example, suppose the conductor is crystalline in the smallest parts. Relative to these parts, the distribution of electricity repeats periodically at definite distances infinitely small compared to the dimensions of the conductor. Denoting by β the length of such a period, each sum becomes infinitely small like $c\beta^n$ if f and its partial derivatives up to order $n-1$ are continuous, and like $e^{-c/\beta}$ if all these derivatives are continuous.

Empirical law of electrodynamic influences.

Let the specific current intensities per unit mass at time t at the point (x, y, z), parallel to the three axes, be u, v, w, and at the point (x', y', z') be (u', v', w'). Denote by r the distance between these points, and by c the constant determined by Kohlrausch and Weber. Empirically, the potential of the forces of S acting on S' is

$$-\frac{2}{c}\iint \frac{uu' + vv' + ww'}{r}\, dSdS';$$

the integral is taken over all elements dS and dS' of the conductors S and S'. We introduce the product of the velocities with specific densities in place of the specific current intensities, and then take the product of these with the masses contained in the volume elements. The expression becomes

$$\sum \frac{\epsilon\epsilon'}{c^2}\,\frac{1}{r}\,\frac{dd'(r^2)}{dt^2},$$

if the variation of r^2 during time dt' arising from the motion of ϵ is denoted by d, and that from the motion of ϵ' by d'.

On neglecting

$$\frac{d\sum\sum \frac{\epsilon\epsilon'}{c^2}\,\frac{1}{r}\,\frac{d'(r^2)}{dt}}{dt},$$

which vanishes via the summation over ϵ, this expression becomes

$$-\sum\sum \frac{\epsilon\epsilon'}{c^2}\,\frac{d\left(\frac{1}{r}\right)}{dt}\,\frac{d'(r^2)}{dt}.$$

On adding

$$\frac{d'\sum\sum \frac{\epsilon\epsilon' r^2}{c^2}\,\frac{d\left(\frac{1}{r}\right)}{dt}}{dt}$$

which vanishes via the summation over ϵ', this is transformed into

$$\sum \epsilon\epsilon' \frac{r^2}{c^2}\frac{dd'\left(\frac{1}{r}\right)}{dt^2}.$$

Derivation of this law from the new theory.

With the previous assumptions on the electrostatic influence, the potential function of arbitrarily distributed electrical masses having density ρ at the point (x, y, z) is determined by the condition

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} - 4\pi\rho = 0.$$

We have the further condition that U is continuous and is constant at infinite distance from the masses exerting influence. A particular integral of the equation

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = 0,$$

continuous except at the points (x', y', z'), is

$$\frac{f(t)}{r}.$$

This function forms the potential function arising from the point (x', y', z') if the quantity located there at time t is $-f(t)$.

Instead of this, I assume that the potential function is determined by the condition

$$\frac{\partial^2 U}{\partial t^2} - \alpha^2 \left(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial x^2}\right) + \alpha^2 4\pi\rho = 0.$$

Thus the potential function arising from the point (x', y', z') is

$$\frac{f\left(t - \frac{r}{\alpha}\right)}{r}$$

if the quantity $-f(t)$ is found there at time t.

We denote the coordinates of the quantity ϵ at time t by x_t, y_t, z_t and those of ϵ' at time t' by $x'_{t'}, y'_{t'}, z'_{t'}$. For brevity, let

$$\left((x_t - x'_{t'})^2 + (y_t - y'_{t'})^2 + (z_t - z'_{t'})^2\right)^{-\frac{1}{2}} = \frac{1}{r(t, t')} = f(t, t').$$

With this hypothesis the potential of ϵ on ϵ' at time t is

$$-\epsilon\epsilon' F\left(t - \frac{r}{\alpha}, t\right).$$

The potential of the totality of forces from the masses ϵ of the conductor S, acting on the masses ϵ' of the conductor S' from time 0 up to time t, will then be

$$P = -\int_0^t \sum\sum \epsilon\epsilon' F\left(t - \frac{r}{\alpha}, \tau\right) d\tau,$$

the sums being taken over all masses of both conductors.

The motion is opposite for opposing electrical charges in each small portion of the conductor. Hence the function $f(t, t')$, on differentiation with respect to t, undergoes a sign change with ϵ, and, on differentiation with respect to t', a sign change with ϵ'. By our hypothesis on the distribution of electricity, the quantity

$$\sum \epsilon\epsilon' F_{n'}^{(n)}(\tau, \tau),$$

taken over all electrical masses, will be infinitely small compared to the corresponding sum over electrical masses of one type if n, n' are both odd. Here differentiation with respect to t is denoted by an upper accent, with a lower accent for t'.

We now suppose that the electrical masses travel only a very small distance during the propagation of force from one conductor to the other. Consider the influence during a time interval relative to which the time of propagation vanishes. In the expression for P, we can begin by replacing

$$F\left(\tau - \frac{r}{\alpha}, \tau\right)$$

by

$$f(\tau - \frac{r}{\alpha}, \tau) - f(\tau, \tau) = -\int_0^{\frac{r}{\alpha}} F'(\tau - \sigma, \tau) d\sigma.$$

For $\sum \epsilon\epsilon' f(\tau, \tau)$ may be neglected. We obtain

$$P = \int_0^t d\tau \sum\sum \epsilon\epsilon' \int_0^{r/\alpha} F'(\tau - \sigma, \tau) d\sigma.$$

If we now interchange the order of integration and substitute $\tau + \sigma$ for τ,

$$P = \sum\sum \epsilon\epsilon' \int_0^{r/\alpha} d\sigma \int_{-\sigma}^{t-\sigma} d\tau F'(\tau, \tau + \sigma).$$

If we transform the limits of the inner integral into 0 and t, then at the upper limit the expression

$$H(t) = \sum\sum \epsilon\epsilon' \int_0^{\frac{r}{\alpha}} d\sigma \int_{-\sigma}^{0} d\tau F'(t + \tau, t + \tau + \sigma)$$

is added; at the lower limit, the value of this expression for $t = 0$ is subtracted. Thus we have

$$P = \int_0^t d\tau \sum\sum \epsilon\epsilon' \int_0^{\frac{r}{\alpha}} d\sigma F'(\tau, \tau + \sigma) - H(t) + H(0).$$

In this expression, we may replace $F'(\tau, \tau + \sigma)$ by $F'(\tau, \tau + \sigma) - F'(\tau, \tau)$, since

$$\sum\sum \epsilon\epsilon' \frac{r}{\alpha} F'(\tau, \tau)$$

may be neglected. In this way we obtain as a factor of $\epsilon\epsilon'$ an expression which changes sign with both ϵ and ϵ', in such a way that the terms do not cancel each other out on summation, while infinitely small fractions of the individual terms may be neglected. If we now replace $F'(\tau, \tau + \sigma)$ by $F'(\tau, \tau + \sigma) - F'(\tau, \tau)$, since

$$\sum\sum \epsilon\epsilon' \frac{r}{\alpha} F'(\tau, \tau)$$

may be neglected. In this way we obtain as a factor of $\epsilon\epsilon'$ an expression which changes sign with both ϵ and ϵ', in such a way that the terms do not cancel each other out on summation, while infinitely small fractions of the individual terms may be neglected. If we now replace

$$F'(\tau, \tau + \sigma) - F'(\tau, \tau) \text{ by } \frac{\sigma dd'\left(\frac{1}{r}\right)}{d\tau^2},$$

and integrate with respect to σ, this yields

$$P = \int_0^t \sum\sum \epsilon\epsilon' \frac{r^2}{2\alpha^2} \frac{dd'\left(\frac{1}{r}\right)}{d\tau^2} d\tau - H(t) + H(0)$$

up to a negligible quantity.

It is easy to see that $H(t)$ and $H(0)$ may be neglected. For

$$F'(t + \tau, t + \tau + \sigma) = \frac{d}{dt}\left(\frac{1}{r}\right) + \frac{d^2\left(\frac{1}{r}\right)}{dt^2} \tau + \frac{dd'\left(\frac{1}{r}\right)}{dt^2} (\tau + \sigma) + \ldots,$$

so that

$$H(t) = \sum\sum \epsilon\epsilon' \left(\frac{r^2}{2\alpha^2} \frac{d\left(\frac{1}{r}\right)}{dt} - \frac{r^3}{6\alpha^3} \frac{d^2\left(\frac{1}{r}\right)}{dt^2} + \frac{r^3}{6\alpha^3} \frac{dd'\left(\frac{1}{r}\right)}{dt^2} + \ldots \right).$$

However, only the first term of the factor of $\epsilon\epsilon'$ in this expression is of the same order of magnitude as the first term of P, and will be negligible in comparison on account of the summation over ϵ'.

The value of P that emerges from our theory coincides with the empirical quantity

$$P = \int_0^t \sum\sum \epsilon\epsilon' \frac{r^2}{c^2} \frac{dd'\left(\frac{1}{r}\right)}{d\tau^2}\, d\tau$$

if we take $\alpha^2 = \frac{1}{2}c^2$.

The value determined by Weber and Kohlrausch is

$$c = 439450.10^6 \text{ millimeters per second,}$$

giving 41949 nautical miles per second for α. The velocity of light found by Busch from Bradley's aberration observations is 41994 miles per second, and Fizeau's value found via direct measurement is 41882 miles per second.

XV.

A proof of the theorem that a single-valued periodic function of n variables cannot be more than $2n$-fold periodic.

(Extract from a letter from Riemann to Weierstrass.)

(*Borchardt's Journal für reine und angewandte Mathematik*, vol. 71.)

... I feel that I probably did not make myself quite clear in our conversation about the proof of the proposition, to which you have recently lent your support, that the existence of periodic functions of n variables that are more than $2n$-fold periodic is an impossibility. As I mentioned only briefly the basic underlying ideas, I should like to return to the subject here.

Let f be a $2n$-fold periodic function of n variables $x_1, x_2, \ldots, x_n$, and—if I may use my terminology, with which you are familiar—let the modulus of periodicity of x_ν for the μ-th period be a_μ^ν. It is then known that the variables x can be expressed in the form[1]

$$x_\nu = \sum_{\mu=1}^{2n} a_\mu^\nu \xi_\mu \qquad (\nu = 1, 2, \ldots, n)$$

in such a way that the quantities ξ are real.

Suppose now that the variables ξ run through all real values between 0 and 1, with the exception of one of these two limits. Then the resulting region of $2n$-dimensional quantities has the property that every set of values of the n variables is congruent, with respect to the $2n$ systems of moduli, to one and only one set of values lying within this region. In order to express myself more concisely later, I shall call this region the *periodic recurring region* for these $2n$ systems of moduli.

Suppose now that the function has a $(2n+1)$-th system of moduli, which is not a combination of the first $2n$ systems of moduli. Any of the points [*Grössensysteme*] congruent modulo this system of moduli can be reduced to a congruent point, modulo the first $2n$ systems, in this region. Clearly we obtain arbitrarily many points in this region, congruent modulo the $2n+1$ systems, provided two points congruent modulo the $(2n+1)$-th system are

[1]This is not always the case, but only when the $2n$ equations which determine the quantities ξ are independent of each other. The exceptions are easily handled.

not also congruent modulo the first $2n$ systems. In that case there would have to be equations of the form

$$\sum_{\mu=1}^{2n+1} a_\mu^\nu m_\mu = 0$$

between the $2n+1$ systems in which the numbers m are integers, and consequently, as I shall show later, the $(2n+1)$-th system would be a combination of the first $2n$ systems of moduli.

For each of the quantities ξ, we divide the interval from 0 to 1 into q equal parts, so that the periodic recurring region for the first $2n$ systems of moduli breaks up into q^{2n} regions, each of which the quantities ξ vary only up to $\frac{1}{q}$. Clearly any set of more than q^{2n} points in the larger region which are congruent with respect to the $(2n+1)$-th system must contain two points in one of the smaller subregions, so that the values of the quantities ξ corresponding to these two points never differ by more than $\frac{1}{q}$. The value of the function therefore remains unchanged, while the ξ are varied by amounts which do not exceed $\frac{1}{q}$, and it follows from this—since q can be chosen to be arbitrarily large—that the function (if it is continuous) is a function of fewer than n linear expressions in the quantities x.

It remains to be shown that $2n+1$ systems of moduli, connected by the n equations

$$\sum_{\mu=1}^{2n+1} a_\mu^\nu m_\mu = 0,$$

are combinations of $2n$ systems of moduli.

Firstly, it can easily be shown that if for one system of moduli we have

$$\sum_{\mu=1}^{2n} a_\mu^\nu m_\mu = b_1^\nu,$$

where the m are integers without common factor, then we can always find $2n-1$ other systems of moduli $b_2, b_3, \ldots, b_{2n}$, such that congruence with respect to the systems is equivalent to congruence with respect to the systems b. Let θ_1 be the greatest common divisor of m_1 and m_2 and let α, β be two integers satisfying the equation

$$\beta m_1 - \alpha m_2 = \theta_1.$$

If we now put

$$a_1^\nu m_1 + a_2^\nu m_2 = c_1^\nu \theta_1$$

and

$$\alpha a_1^\nu + \beta a_2^\nu = b_{2n}^\nu,$$

then we have

$$a_1^\nu = \beta c_1^\nu - \frac{m_2}{\theta_1}\, b_{2n}^\nu,\ a_2^\nu = -\alpha c_1^\nu + \frac{m_1}{\theta_1}\, b_{2n}^\nu.$$

The systems of moduli a_1 and a_2 are therefore conversely combinations of the systems b_{2n} and c_1, and congruence with respect to either system has the same meaning. Thus one can substitute the systems of moduli c_1 and b_{2n} for the systems of moduli a_1 and a_2. Similarly, if θ_2 is the greatest common divisor of θ_1 and m_2, the systems c_1 and a_3 can be replaced by the system

$$\frac{1}{\theta_2}\left(\theta_1 c_1^\nu + m_3 a_3^\nu\right) = c_2^\nu$$

and the system b_{2n-1}. By repeating this process we clearly obtain the theorem to be proved. The volume of the periodic recurring region is the same for the new systems of moduli b as for the old.

With the help of this theorem, the first $2n$ systems of moduli in the equations

$$\sum_1^{2n+1} a_\mu^\nu m_\mu = 0$$

can be replaced by $2n$ new systems $b_1, b_2, \ldots, b_{2n}$ in such a way that these equations take the form

$$p b_1^\nu - q a_{2n+1}^\nu = 0$$

where p and q are relatively prime integers. If now γ, δ are two integers satisfying the equation

$$p\delta + q\gamma = 1,$$

then obviously the two systems b_1 and a_{2n+1} can be replaced by the single system

$$\gamma b_1^\nu + \delta a_{2n+1}^\nu = \frac{a_{2n+1}^\nu}{p} = \frac{b_1^\nu}{q}.$$

All systems of moduli which are combinations of the systems $a_1, a_2, \ldots, a_{2n+1}$, are also therefore combinations of the systems $\frac{b_1}{q}, b_2, b_3, \ldots, b_{2n}$ and conversely.

The volume of the periodic recurring region for these $2n$ systems of moduli amounts to only $\frac{1}{q}$ of that for the first $2n$ systems of moduli a. If therefore the function possessed, in addition to these systems of moduli, another one related to these systems by similar linear equations with integral coefficients, then it would be possible to find $2n$ new systems of moduli of which all of these systems are combinations. The volume of the new periodic recurring region would thus be reduced to a fraction of the old one. If this region becomes infinitely small, the function must be a function of fewer than n linear expressions in the variables. In fact, the number of these expressions will be $n-1$ or $n-2$ or $n-m$, depending on whether one, or two, or m dimensions of this region become infinitely small. If this is not the case, then the process must terminate and thus one obtains $2n$ systems of moduli, of which all systems of moduli of the function are combinations.

Göttingen, 26th October 1859.

XVI.

Extract from a letter written in Italian on January 21, 1864 to Professor Enrico Betti.

(*Annali di Matematica*, Ser. 1, vol. 7.)

Dearest friend,

... To find the attraction due to any homogeneous right ellipsoidal cylinder, I consider the infinite cylinder whose points, in rectangular Cartesian co-ordinates x, y, z, are those satisfying the inequality

$$1 - \frac{x^2}{a^2} - \frac{y^2}{b^2} > 0,$$

filled with matter whose constant density is $+1$, when $z < 0$, and -1, when $z > 0$. If we then, as usual, denote by V the potential at the point x, y, z and write

$$\frac{\partial V}{\partial x} = X, \ \frac{\partial V}{\partial y} = Y, \ \frac{\partial V}{\partial z} = Z$$

then, for $z = 0, V = 0, X = 0, Y = 0$.

Now Z is the potential of the ellipse

$$1 - \frac{x^2}{a^2} - \frac{y^2}{b^2} > 0$$

of density 2, and can be found by Dirichlet's method. Denote by σ the larger root of the equation

$$1 - \frac{x^2}{a^2 + s} - \frac{y^2}{b^2 + s} - \frac{z^2}{s} = F = 0$$

and write D for

$$\sqrt{\left(1 + \frac{s}{a^2}\right)\left(1 + \frac{s}{b^2}\right) s}.$$

Then Z is given by

$$4 \int_\sigma^\infty \frac{\sqrt{F}}{D} \, ds.$$

Now X and Y can be determined from the equation

$$\frac{\partial X}{\partial z} = \frac{\partial Z}{\partial x}, \ \frac{\partial Y}{\partial z} = \frac{\partial Z}{\partial y}$$

and the conditions

$$X = 0, \; Y = 0 \quad \text{for } z = 0.$$

To carry out the calculation, it is convenient to substitute for the line integral $4\int_\sigma^\infty$ the contour integral $2\int_\infty^\infty$, taken around a contour in the s-plane which includes the point σ, but no other point of discontinuity or branch point of the integrand. Denote the roots of the equation $F = 0$ in order of magnitude by $\sigma, \sigma', \sigma''$. Then the singularities of the integrand are all real and, in order of magnitude, are $\sigma, 0, \sigma', -b^2, \sigma'', -a^2$ with

$$\sigma > 0 > \sigma' > -b^2 > \sigma'' > -a^2.$$

Let

$$F = t - z^2/s,$$

so that

$$Z = 2\int_\infty^\infty \frac{\sqrt{ts - z^2}}{D\sqrt{s}}\, ds,$$

$$\frac{\partial X}{\partial z} = \frac{\partial Z}{\partial x} = \int_\infty^\infty \frac{s\,\frac{\partial t}{\partial x}\,(ts - z^2)^{-1/2}}{D\sqrt{s}}\, ds.$$

Now

$$\int_0^z (ts - z^2)^{-1/2} dz = \int_0^{\frac{z}{\sqrt{ts}}} (1 - \xi^2)^{-1/2} d\xi = \int_0^{\frac{z}{\sqrt{ts}}} \left(\frac{1}{\xi^2} - 1\right)^{-1/2} d\log\xi,$$

and

$$\frac{s\,\frac{\partial t}{\partial x}\, ds}{D\sqrt{s}} = -2abx(a^2 + s^2)^{-3/2}(b^2 + s^2)^{-1/2} ds = \frac{4abx}{b^2 - a^2}\, d\sqrt{\frac{b^2 + s}{a^2 + s}}.$$

Hence, after integration by parts:

$$X = \frac{2abxz}{b^2 - a^2}\int_\infty^\infty \sqrt{\frac{b^2 + s}{a^2 + s}}\,(ts - z^2)^{-1/2} d\log ts.$$

If the same path of integration is taken as in the expression for Z, the value of the integral will always satisfy the condition

$$\frac{\partial X}{\partial z} = \frac{\partial Z}{\partial x}.$$

However, the integral may differ by a function of x and y, because the integrand is also discontinuous for $t = 0$. It is therefore necessary to choose another contour of integration.

In the expression of $\frac{\partial X}{\partial z} = \frac{\partial Z}{\partial x}$, the integrand is continuous for $s = 0$; and so the part of the s-plane inside the contour has to contain the point $s = \sigma$, and may or may not contain the point $s = \sigma$, but must not contain any of the other singularities listed above. In the expression for X, the interior of the contour must be determined in such a way that $X = 0$ when $z = 0$. In order for this to occur, it will also—since it has to include $s = \sigma$—have to contain the largest root of the equation $ts = 0$. This is the largest root of $t = 0$ if

$$1 - \frac{x^2}{a^2} - \frac{y^2}{b^2} < 0,$$

but 0 if

$$1 - \frac{x^2}{a^2} - \frac{y^2}{b^2} > 0.$$

The interior must contain no other root of $ts = 0$. This is because, when $z = 0$, the roots of the equation $F = 0$ coincide with those of the equation $ts = 0$. If the path of integration were to pass through two points of discontinuity which coincide when $z = 0$, it would have to do so in such a manner that although the integral in the expression for X would become infinite, the value of X would still remain finite, thanks to the factor z.– –

Your affectionate friend Riemann.

XVII.

On the surface of least area with a given boundary.

(*Königlichen Gesellschaft der Wissenschaften zu Göttingen*, vol. 13.)

1.

A surface can be specified, in the sense of analytic geometry, if we represent the rectilinear coordinates x, y, z of an arbitrary point as well-defined functions of two independent variables p and q. If p and q assume a particular constant value, that combination always corresponds to a single point of the surface. The independent variables p and q can be chosen in a great many different ways. For a simply connected surface this can be done expediently as follows. We shrink the surface along the whole boundary by discarding a strip of surface whose width everywhere is infinitely small of the same order. By continuation of this process the surface will inexorably shrink until it becomes a point. The series of boundary curves arising from this are closed curves separated from each other. We can distinguish among them by associating to each curve a particular constant value of the variable p which takes an infinitely small increase or decrease depending on whether we pass to a neighboring enclosing or enclosed curve. The function p then has a constant maximal value on the boundary of the surface and a minimal value at the point in the interior to which the gradually shrinking surface reduces. We can think of producing the passage from one boundary of the shrinking surface to the next by replacing each point of the curve (p) with a particular infinitely close point of the curve $(p + dp)$. The path of the individual points then form a second system of curves, which form a ray running from the point of minimal value of p to the boundary of the surface. In each of these curves we assign a particular constant value to the variable q, which is smallest at an arbitrarily chosen initial curve and grows continuously if we pass from one curve of the second system to another when going along a curve (p) in a particular direction, chosen for this purpose. In passing from the last curve (q) to the first curve, q jumps by a finite constant.

In order to treat a multiply connected surface, we can decompose it into a simply connected one by transverse cuts.

In this manner, any point on the surface can be understood as the intersection of a particular curve of the system (p) with a particular curve of the system (q). The normals erected at the point (p, q) run out from the surface in two opposite directions, positive and negative. To distinguish them, we have to determine the mutual positions of the increasing positive normal, increasing p and increasing q. If nothing else is stipulated, we assume that, looking down the positive x-axis, the shortest rotation that takes the positive y-axis into the positive z-axis is from right to left. The direction of the increasing positive normal lies with respect to the direction of increasing p and increasing q as the positive x-axis lies with respect to the positive y and z axes. The side on which the positive normal lies will be called the positive side of the surface.

2.

Let an integral be taken over the region of the surface whose element is the element $dp\,dq$ multiplied by a functional determinant, thus

$$\iint \left(\frac{\partial f}{\partial p}\frac{\partial g}{\partial q} - \frac{\partial f}{\partial q}\frac{\partial g}{\partial p}\right) dp\,dq.$$

For brevity we write this as

$$\iint (df\,dg).$$

If we consider f and g as new independent variables, the integral becomes $\iint df\,dg$, and the integration can be carried out with respect to f or g. However, to actually establish f and g as independent variables presents difficulties, or at least extensive splitting into cases, if the same combination of values for f and g exist at several points of the surface or in a line. It is quite impossible when f and g are complex.

Hence it is expedient to carry out the integration with respect to f or g by the procedure of Jacobi (*Crelle's Journal* vol. 27, p. 208), in which p and q will be retained as independent variables. In order to integrate with respect to f, we bring the functional determinant to the form

$$\frac{\partial\left(f\frac{\partial g}{\partial q}\right)}{\partial p} - \frac{\partial\left(f\frac{\partial g}{\partial p}\right)}{\partial q}$$

and obtain first of all

$$\int \frac{\partial\left(f\frac{\partial g}{\partial p}\right)}{\partial q}\,dq = 0,$$

since the integration is taken over a closed curve.

On the other hand,

$$\int \frac{\partial\left(f\,\frac{\partial g}{\partial q}\right)}{\partial p}\,dp$$

is to be taken in the direction of increasing p, that is from the minimal point in the interior along a curve (q) to the boundary. We obtain $f\,\frac{\partial g}{\partial q}$ and indeed the value that this expression assumes on the boundary, since at the lower limit of the integral $\frac{\partial g}{\partial q} = 0$. Consequently

$$\iint (df\,dg) = \int f\,\frac{\partial g}{\partial q}\,dq = \int f\,dg$$

and the integral on the right is taken in the direction of increasing q along the boundary. On the other hand we have, from the notation, $(df\,dg) = -(dg\,df)$. Hence

$$\iint (df\,dg) = -\iint (dg\,df) = -\int g\,df,$$

where the integral on the right is also to be taken in the direction of increasing q along the boundary of the surface.

3.

The surface whose points are given by the system of curves (p), (q) will be mapped in the following way to a sphere of radius 1. At the point (p, q) of the surface whose rectilinear coordinates are x, y, z, we draw the positive normal and place its parallel at the center of the ball. The endpoint of this parallel on the ball's surface is the image of the point (x, y, z). If the point (x, y, z) runs along an arc in the continuously bending surface, the image will be an arc on the sphere. In the same manner we see that the image of a piece of the surface will be a piece of surface on the sphere. The image of the whole surface is a surface that covers the sphere, or part of it, one or more times.

The point on the sphere which is in the direction of the positive x-axis will be chosen as the pole, and the 0-meridian is put through the point which corresponds to the positive y-axis. The image of a point (x, y, z) on the sphere will then be determined by its distance r from the pole and the angle ϕ between its meridian and the 0-meridian. To determine the sign of ϕ, let the point corresponding to the positive z-axis have coordinates $r = \frac{\pi}{2}$, $\phi = +\frac{\pi}{2}$.

4.

From the above, we obtain

$$\cos r\, dx + \sin r \cos\phi\, dy + \sin r \sin\phi\, dz = 0, \tag{1}$$

the differential equation of the surface.

If y and z are the independent variables, we have the equations

$$\cos r = \frac{1}{\pm\sqrt{1+\left(\frac{\partial x}{\partial y}\right)^2+\left(\frac{\partial x}{\partial z}\right)^2}},$$

$$\sin r \cos \phi = \frac{\frac{\partial x}{\partial y}}{\mp\sqrt{1+\left(\frac{\partial x}{\partial y}\right)^2+\left(\frac{\partial x}{\partial z}\right)^2}},$$

$$\sin r \sin \phi = \frac{\frac{\partial x}{\partial z}}{\mp\sqrt{1+\left(\frac{\partial x}{\partial y}\right)^2+\left(\frac{\partial x}{\partial z}\right)^2}}$$

for r and ϕ, in which either the upper or the lower signs are simultaneously valid.

A parallelogram on the positive side of the surface, bounded by the curves (p) and $(p+dp)$, (q) and $(q+dq)$, projects on the yz-plane into an element of surface whose area equals the absolute value of $(dy\, dz)$. The sign of this functional determinant depends on whether the positive normal at the point (p,q) subtends the x-axis in an acute or obtuse angle. In the first case, the projections of dp and dq into the yz-plane lie with respect to each other just as do the positive y-axis and the positive z-axis. In the second case this is reversed. Thus the functional determinant is positive in the first case and negative in the second. The expression

$$\frac{1}{\cos r}\,(dy\, dz)$$

is always positive. It gives the area of an infinitely small parallelogram on the surface. Thus to obtain the area of the surface itself, we take the double integral

$$S = \iint \frac{1}{\cos r}\,(dy\,dz)$$

over the whole surface.

For this area to be a minimum, the first variation of this double integral is set equal to 0. We obtain

$$\iint \frac{\dfrac{\partial x}{\partial y}\dfrac{\partial \delta x}{\partial y} + \dfrac{\partial x}{\partial z}\dfrac{\partial \delta x}{\partial z}}{\pm\sqrt{1+\left(\dfrac{\partial x}{\partial y}\right)^2+\left(\dfrac{\partial x}{\partial z}\right)^2}}(dy\,dz) = 0,$$

and the upper or lower sign holds for the root depending on whether $(dy\,dz)$ is positive or negative. The left side may be written

$$\begin{aligned}&\iint \frac{\partial}{\partial y}(-\sin r\cos\phi\,\delta x)(dy\,dz)\\ &+\iint \frac{\partial}{\partial z}(-\sin r\sin\phi\,\delta x)(dy\,dz)\\ &-\iint \delta x\,\frac{\partial}{\partial y}(-\sin r\cos\phi)(dy\,dz)\\ &-\iint \delta x\,\frac{\partial}{\partial z}(-\sin r\sin\phi)(dy\,dz).\end{aligned}$$

The first two integrals reduce to a line integral taken in the direction of increasing q around the boundary of the surface, namely

$$\int \delta\,x(-\sin r\cos\phi\,dz + \sin r\sin\phi\,dy).$$

The value is 0 since $\delta\,x = 0$ in the boundary. Thus the condition for minimality appears as

$$\iint \delta\,x\left(\frac{\partial(\sin\,r\cos\,\phi)}{\partial y} + \frac{\partial(\sin\,r\sin\,\phi)}{\partial z}\right)(dy\,dz) = 0.$$

It is satisfied if

$$(2)\qquad -\sin r\sin\phi\,dy + \sin r\cos\phi\,dz = d\mathfrak{x}$$

is a complete differential.

5.

The coordinates r and ϕ on the sphere may be replaced by a complex quantity $\eta = \tan\frac{r}{2}\, e^{\phi i}$, whose geometric significance is easy to see. If we place a tangent plane to the sphere at the pole whose positive side is opposite from the sphere and draw from the antipodal pole a line through the point (r, ϕ), then this meets the tangent plane at a point which is represented by the complex quantity 2η. Then $\eta = 0$ corresponds to the pole and $\eta = \infty$ to the antipodal pole. We have $\eta = +1$ and $\eta = +i$ for the points in the direction of the positive y and z axes respectively.

If we also introduce the complex quantities

$$\eta' = \tan\frac{r}{2}\, e^{-\phi i}, \quad s = y + zi, \quad s' = y - zi,$$

then equations (1) and (2) become

$$(1^*) \qquad (1 - \eta\eta')dx + \eta'\, ds + \eta\, ds' = 0,$$

$$(2^*) \qquad (1 + \eta\eta')d\mathfrak{x}i \; - \eta'\, ds + \eta\, ds' = 0.$$

These may be combined by addition and subtraction. We obtain

$$x + \mathfrak{x}i = 2X, \quad x - \mathfrak{x}i = 2X',$$

so that conversely $x = X + X'$. Then the problem can be expressed analytically by the two equations

$$(3) \qquad ds - \eta dX + \frac{1}{\eta'}\, dX' = 0,$$

$$(4) \qquad ds' + \frac{1}{\eta}\, dX - \eta' dX' = 0.$$

If we consider X, X' as independent variables and impose the conditions that ds and ds' are complete differentials, then

$$\frac{\partial \eta}{\partial X'} = 0, \quad \frac{\partial \eta'}{\partial X} = 0,$$

that is, η depends only on X, η' only on X', and hence, conversely, X is a function only of η and X' only of η'.

Consequently the problem reduces to determining η as a function of the complex variable X, or conversely X as a function of the complex variable η, while also satisfying the boundary conditions. If we have η as a function of X, then η' is given by changing each complex number in the expression of η to its conjugate. At this point we only have to integrate (3) and (4) in order to attain expressions for s and s'. Finally we obtain through elimination of η an equation between x, y, z, the equation of the minimal surface.

6.

If equations (3) and (4) are integrated, the area of the minimal surface can easily be found, namely

$$S = \iint \frac{1}{\cos r}\,(dy\,dz) = \iint \frac{1+\eta\eta'}{1-\eta\eta'}\,(dy\,dz).$$

The functional determinant $(dy\,dz)$ is rewritten in the following way:

$$\begin{aligned}(dy\,dz) &= \left(\frac{\partial y}{\partial s}\frac{\partial z}{\partial s'} - \frac{\partial y}{\partial s'}\frac{\partial z}{\partial s}\right)(ds\,ds') \\ &= \frac{i}{2}\,(ds\,ds') \\ &= \frac{i}{2}\left(\eta\eta' - \frac{1}{\eta\eta'}\right)\frac{\partial x}{\partial \eta}\frac{\partial x}{\partial \eta'}\,(d\eta\,d\eta').\end{aligned}$$

We obtain

$$\begin{aligned}2iS &= \iint \left(2 + \eta\eta' + \frac{1}{\eta\eta'}\right)\frac{\partial x}{\partial \eta}\frac{\partial x}{\partial \eta'}\,(d\eta\,d\eta') \\ &= \iint \left(2\frac{\partial x}{\partial \eta}\frac{\partial x}{\partial \eta'} + \frac{\partial s}{\partial \eta}\frac{\partial s'}{\partial \eta'} + \frac{\partial s}{\partial \eta'}\frac{\partial s'}{\partial \eta}\right)(d\eta\,d\eta') \\ &= 2\iint \left(\frac{\partial x}{\partial \eta}\frac{\partial x}{\partial \eta'} + \frac{\partial y}{\partial \eta}\frac{\partial y}{\partial \eta'} + \frac{\partial z}{\partial \eta}\frac{\partial z}{\partial \eta'}\right)(d\eta\,d\eta').\end{aligned}$$

For another transformation of this expression, we construct y from Y and Y' and z from Z and Z' just as x was constructed from X and X'. Then the equations

$$\begin{aligned}
X &= \int \frac{\partial x}{\partial \eta} d\eta, \quad X' = \int \frac{\partial x}{\partial \eta'} d\eta', \\
Y &= \int \frac{\partial y}{\partial \eta} d\eta, \quad Y' = \int \frac{\partial y}{\partial \eta'} d\eta', \\
Z &= \int \frac{\partial z}{\partial \eta} d\eta, \quad Z' = \int \frac{\partial z}{\partial \eta'} d\eta', \\
x &= X + X', \quad \mathfrak{x} i = X - X', \\
y &= Y + Y', \quad \mathfrak{y} i = Y - Y', \\
z &= Z + Z', \quad \mathfrak{z} i = Z - Z'
\end{aligned}$$

hold. We obtain finally:

$$\begin{aligned}
(5) \qquad S &= -i \iint [(dX\, dX') + (dY\, dY') + (dZ\, dZ')] \\
&= \frac{1}{2} \iint [(dx\, d\mathfrak{x}) + (dy\, d\mathfrak{y}) + (dz\, d\mathfrak{z})].
\end{aligned}$$

7.

The minimal surface and its images on the sphere and on the planes whose points are represented respectively by the complex quantities η, X, Y, Z are similar to one another in the smallest parts. We see this easily by expressing the squares of the linear elements in these surfaces:

on the sphere	$\sin r^2\, d\log\eta\, d\log\eta'$,
in the plane of η	$d\eta\, d\eta'$,
in the plane of X	$\frac{\partial x}{\partial \eta}\frac{\partial x}{\partial \eta'}\, d\eta\, d\eta'$,
in the plane of Y	$\frac{\partial y}{\partial \eta}\frac{\partial y}{\partial \eta'}\, d\eta\, d\eta'$,
in the plane of Z	$\frac{\partial z}{\partial \eta}\frac{\partial z}{\partial \eta'}\, d\eta\, d\eta'$,

on the minimal surface itself

$$\begin{aligned} dx^2 + dy^2 + dz^2 &= (dX + dX')^2 + (dY + dY')^2 + (dZ + dZ')^2 \\ &= 2(dX\,dX' + dY\,dY' + dZ\,dZ') \\ &= 2\left(\frac{\partial x}{\partial \eta}\frac{\partial x}{\partial \eta'} + \frac{\partial y}{\partial \eta}\frac{\partial y}{\partial \eta'} + \frac{\partial z}{\partial \eta}\frac{\partial z}{\partial \eta'}\right) d\eta\, d\eta'. \end{aligned}$$

By equations (3) and (4), if η and η' are considered as independent variables, we see that

$$\begin{aligned} \eta\,\frac{dX}{d\eta} &= \frac{\partial s}{\partial \eta} = -\eta^2\,\frac{\partial s'}{\partial \eta}, \\ \eta'\,\frac{dX'}{d\eta'} &= \frac{\partial s'}{\partial \eta'} = -\eta'^2\,\frac{\partial s}{\partial \eta'}, \end{aligned}$$

and therefore

$$\begin{aligned} dX^2 + dY^2 + dZ^2 &= 0, \\ dX'^2 + dY'^2 + dZ'^2 &= 0. \end{aligned}$$

The ratio of any two of the above squares of linear elements is independent of $d\eta$ and $d\eta'$, that is, the direction of the elements. The similarity of the mapping in the smallest parts rests on this. Since the linear expansion by the mapping at any point is the same in all directions, we obtain the surface magnification as the square of the linear magnification. The square of the linear element in the minimal surface, however, is twice the sum of the square of the corresponding linear elements in the planes of X, Y and Z. Hence the surface element in the minimal surface is also twice the sum of the corresponding surface elements in those planes. The same holds on the entire surface and its images in the planes of X, Y, and Z.

8.

An important consequence can be drawn from the theorem on similarity in the smallest parts. We introduce a new complex variable η_1 which puts the pole at an arbitrary point $(\eta = \alpha)$ and chose the 0-meridian arbitrarily. If η_1 has the same meaning for the new coordinates as η had for the old, we can now map an infinitely small triangle on the sphere to the plane of η, just as to the plane of η_1. The two images are then also mappings of each other and are similar in the smallest parts. In the case of direct similarity, it follows at once that $\frac{d\eta_1}{d\eta}$ is independent of the direction of the shifting of η. That is, η_1 is a function of the complex variable η. In the case of inverse (symmetrical) similarity we can go back to the preceding, where instead of η_1 we take the conjugate complex quantity. Now, in order to express η_1 as a function of η, we must observe that $\eta_1 = 0$ at the point on the sphere for which $\eta = \alpha$, and $\eta_1 = \infty$ at the antipodal point, that is, for $\eta = -1/\alpha'$. It follows from this that $\eta_1 = c\,\frac{\eta-\alpha}{1+\alpha'\eta}$. To determine the constant c, we remark that if $\eta_1 = \beta$ when $\eta = 0$, then we find that $\eta_1 = -1/\beta'$ when $\eta = \infty$. Thus $\beta = -c\alpha$ and $-1/\beta' = c/\alpha'$, that is, $\beta = -\alpha/c'$. Hence $cc' = 1$, and $c = e^{\theta i}$ for some real θ. Arbitrary values can be chosen for the quantities α and θ: α depends on the position of the new pole and θ on the position of the new 0-meridian. This new coordinate system on the sphere corresponds to the directions of the axes of a new rectangular coordinate system. In the new coordinate system x_1, s_1, s_1' can be given the same interpretation as x, s, s' in the old. Then the transformation formulae are

$$(6)\qquad \begin{aligned} \eta_1 &= \frac{\eta-\alpha}{1+\alpha'\eta}e^{\theta i},\\ (1+\alpha\alpha')x_1 &= (1-\alpha\alpha')x + \alpha's + \alpha s',\\ (1+\alpha\alpha')s_1e^{-\theta i} &= -2\alpha x + s - \alpha^2 s',\\ (1+\alpha\alpha')s_1'e^{\theta i} &= -2\alpha' x - \alpha'^2 s + s'. \end{aligned}$$

9.

From the transformation formulae (6), we find that

$$\left(\frac{d\eta_1}{d\eta}\right)^2 \frac{\partial x_1}{\partial \eta_1} = \frac{\eta_1}{\eta}\frac{\partial x}{\partial \eta}$$

or

$$(d\log\eta_1)^2\,\frac{\partial x_1}{\partial \log\eta_1} = (d\log\eta)^2\,\frac{\partial x}{\partial\log\eta}.$$

Accordingly, it is expedient to define a new complex quantity u by the equation

$$(7)\qquad u = \int\sqrt{i\,\frac{\partial x}{\partial\log\eta}}\,d\log\eta,$$

which is independent of the position of the coordinate system (x, y, z). If we succeed in determining u as a function of η, we obtain

$$(8)\qquad x = -i\int\left(\frac{du}{d\log\eta}\right)^2 d\log\eta + i\int\left(\frac{du'}{d\log\eta'}\right)^2 d\log\eta'.$$

Here x is the distance from the point on the minimal surface, corresponding to η, to a plane which goes through the origin of coordinates perpendicular to the direction $\eta = 0$. If we replace η by $\frac{\eta-\alpha}{1+\alpha'\eta}e^{\theta i}$ in (8), we obtain the distance of the same point of the minimal surface from the plane through the origin of the coordinates which is at right angles to the direction $\eta = \alpha$. In particular, then, for $\alpha = 1$ and $\alpha = i$,

$$(9)\qquad \begin{aligned} y = &-\frac{i}{2}\int\left(\frac{du}{d\log\eta}\right)^2\left(\eta - \frac{1}{\eta}\right) d\log\eta \\ &+\frac{i}{2}\int\left(\frac{du'}{d\log\eta'}\right)^2\left(\eta' - \frac{1}{\eta'}\right) d\log\eta', \end{aligned}$$

$$(10)\qquad \begin{aligned} z = &-\frac{1}{2}\int\left(\frac{du}{d\log\eta}\right)^2\left(\eta + \frac{1}{\eta}\right) d\log\eta \\ &-\frac{1}{2}\int\left(\frac{du'}{d\log\eta'}\right)^2\left(\eta' + \frac{1}{\eta'}\right) d\log\eta'. \end{aligned}$$

10.

The quantity u is to be determined as a function of η, that is, as a single-valued function of position in that surface, spread over the η-plane, to which the minimal surface is mapped similarly in the smallest parts. Thus it depends primarily on the discontinuities and branchings in this mapping. In investigating these we must distinguish points in the interior of the surface from boundary points.

To deal with a point in the interior of the minimal surface, put the origin of the coordinate system (x, y, z) at the point and place the x-axis in the positive normal. Consequently the yz-plane is tangential. Then in the expansion of x there are no constant term and no terms multiplied by y and z. By a suitable choice of directions for the y and z axes, we can also make the yz term vanish. Under these conditions, the partial differential equation of the minimal surface reduces for infinitely small values of y and z to $\frac{\partial^2 x}{\partial y^2} + \frac{\partial^2 x}{\partial z^2} = 0$.

Hence the Gaussian curvature is negative, and the principal radii of curvature are equal and opposite. The tangent plane divides the surface into four quadrants, when the radii of curvature are not ∞. These quadrants lie alternately over and under the tangent plane. If the expansion of x starts with terms of degree n ($n > 2$), then the radii of curvature are ∞ and the tangent planes divide the surface into $2n$ sectors, which alternately lie above and below the plane, and will be bisected by the curvature lines.

Now, if we regard X as a function of the complex variable Y, then in the case of four sectors

$$\log X = 2\log Y + \text{funct. cont.},$$

in the case of $2n$ sectors

$$\log X = n\log Y + \text{f.c.}$$

Since $\frac{dX}{dY} = \frac{-2\eta}{1-\eta^2}$ follows from (8) and (9), the expansion of η in the first case starts with the first power of Y, and in the second with the $(n-1)$th power of Y. Conversely, if Y were considered as a function of η, the expansion in the first case progresses in integer powers of η, in the second in integer powers of $\eta^{\frac{1}{n-1}}$. That is, the mapping on the η-plane has at the relevant point either no branch point or an $(n-2)$-fold branch point, depending on whether the first or second case occurs.

With respect to u, we get $\frac{du}{d\log Y} = \frac{du}{d\log\eta}\,\frac{d\log\eta}{d\log Y}$. Hence with the aid of equation (9),

$$\left(\frac{du}{d\log Y}\right)^2 = -2i\,\frac{dY}{d\eta}\,\frac{\eta^2}{1-\eta^2}\left(\frac{d\eta}{dY}\right)^2\frac{Y^2}{\eta^2}.$$

Consequently, at a $(n-2)$-fold branch point of the mapping into the η-plane, either

$$\log\frac{du}{d\log Y} = \frac{n}{2}\log Y + \text{f.c.}$$

or

$$\log\frac{du}{dY} = \left(\frac{n}{2} - 1\right)\log Y + \text{f.c.}$$

11.

Further investigation will initially be restricted to the case where the given boundary consists of straight lines. Then, the mapping of the boundary onto the η-plane can actually be constructed. All of the normals erected at points of one of the straight boundary lines lie in parallel planes. Hence the image on the sphere is an arc of a great circle.

In order to examine a point in the interior of a straight boundary line, we place, as previously, the origin of the coordinates at this point, and the positive x-axis in the positive normal. Then the entire boundary line lies in the yz-plane. Therefore the real part of X is 0 on the whole boundary line. If we move around the origin of the coordinates, within the interior of the minimal surface, from a preceding to a succeeding boundary point, then the argument of X changes by $n\pi$, an integer multiple of π. The argument of Y changes at the same time by π. Hence we have, as previously,

$$\log X = n \log Y + \text{f.c.},$$
$$\log \eta = (n-1) \log Y + \text{f.c.},$$
$$\log \frac{du}{dY} = \left(\frac{n}{2} - 1\right) \log Y + \text{f.c.}$$

The boundary points considered correspond to an $(n-2)$-fold branch point in the mapping to the η-plane. In this mapping, the points of the succeeding boundary segment make an angle of $(n-1)\pi$ with those of the preceding one.

12.

In passing from one boundary line to the next, we must distinguish two cases. Either they meet at a finite point, or they stretch to infinity.

In the first case, let $\alpha\pi$ be the angle in the interior of the minimal surface between the two boundary lines. Place the origin of the coordinates at the vertex in question, with the positive x-axis in the positive normal, then in both boundary lines the real part of X is 0. In passing from the first boundary line to the next, the argument of X changes by $m\pi$, an integer multiple of π, the argument of Y by $\alpha\pi$. Hence we have

$$\frac{\alpha}{m} \log X = \log Y + \text{f.c.},$$
$$\left(1 - \frac{\alpha}{m}\right) \log X = \log \eta + \text{f.c.},$$
$$\log \frac{du}{dY} = \left(\frac{m}{2\alpha} - 1\right) \log Y + \text{f.c.}$$

If the surface between two consecutive boundary lines extends to infinity, we set the positive x-axis in the shortest line connecting them, parallel to the positive normal at infinity. Let the length of the shortest connecting line be A, and let $\alpha\pi$ be the angle which the projection of the minimal surface fills in the yz-plane. Then the real parts of X and $i \log \eta$ remain finite and continuous at infinity and take constant values on the boundary lines. It follows (for $y = \infty$, $z = \infty$) that

$$X = -\frac{Ai}{2\alpha\pi} \log \eta + \text{f.c.},$$

$$u = \sqrt{\frac{A}{2\alpha\pi}} \log \eta + \text{f.c.},$$

$$Y = -\frac{Ai}{4\alpha\pi} \frac{1}{\eta} + \text{f.c.}$$

We set the x_1-axis of a coordinate system in one of the straight boundary lines, the x_2-axis of another coordinate system in the second straight boundary line, ... Then in the first line, $\log \eta_1$ is pure imaginary, in the second, $\log \eta_2$ is pure imaginary, and so on, since the normals are perpendicular to the respective axes of x_1, $x_2, \ldots$ Thus $i \frac{\partial x_1}{\partial \log \eta_1}$ is real in the first boundary line, $i \frac{\partial x_2}{\partial \log \eta_2}$ in the second, and so on. However, for any coordinate system (x, y, z),

$$\sqrt{i \frac{\partial x}{\partial \log \eta}} \; d \log \eta = \sqrt{i \frac{\partial x_1}{\partial \log \eta_1}} \; d \log \eta_1 = \sqrt{i \frac{\partial x_2}{\partial \log \eta_2}} \; d \log \eta_2 = \cdots,$$

so that on each straight boundary line

$$du = \sqrt{i \frac{\partial x}{\partial \log \eta}} \; d \log \eta$$

has either real or pure imaginary values.

13.

The minimal surface is determined as soon as one of the quantities u, η, X, Y, or Z has been expressed in terms of one of the others. This succeeds in many cases. Particular notice is deserved for those cases in which $\frac{du}{d \log \eta}$ is an algebraic function of η. For this, it is necessary and sufficient that the mapping on the sphere and its symmetric and congruent extensions form a closed surface, which covers the whole sphere one or more times.

In general, however, it is difficult to express one of the quantities u, η, X, Y, or Z directly in terms of one of the others. Instead of this we can determine each of them as functions of a new, expediently selected, independent variable. We introduce such an independent variable t, so that the image of the surface in the t-plane covers the half-infinite plane once, and indeed, the half for which the imaginary part of t is positive. In fact, it is always possible to determine t as a function of u (or any one of the other quantities η, X, Y, Z) in the surface, so that the imaginary part is 0 in the boundary, and is infinite of first order at an arbitrary boundary point ($u = b$). That is,

$$t = \frac{\text{const.}}{u-b} + \text{f.c.} \qquad (u = b).$$

The argument of the factor of $\frac{1}{u-b}$ is determined by the condition that the imaginary part of t is 0 on the boundary and positive in the interior of the surface. Hence in the expression for t only the modulus of this factor and an additive constant remain arbitrary.

Let $t = a_1, a_2, \ldots$, for the branch points in the interior of the image on the η-plane, $t = b_1, b_2, \ldots$, for the branch points on the boundary which are not vertices, $t = c_1, c_2, \ldots$, for the vertices and $t = e_1, e_2, \ldots$, for those in the unbounded sectors. We assume for simplicity that all the quantities a, b, c, e lie in the finite region of the t-plane. Then we have

$$\begin{aligned}
\log \frac{du}{dt} &= \left(\frac{n}{2} - 1\right) \log(t-a) + \text{f.c.} &\qquad& \text{for } t = a,\\
\log \frac{du}{dt} &= \left(\frac{n}{2} - 1\right) \log(t-b) + \text{f.c.} && \text{for } t = b,\\
\log \frac{du}{dt} &= \left(\frac{m}{2} - 1\right) \log(t-c) + \text{f.c.} && \text{for } t = c,\\
u &= \sqrt{\frac{A\alpha}{2\pi}} \log(t-e) + \text{f.c.} && \text{for } t = e.
\end{aligned}$$

The investigation can be restricted to the case $n = 3$, $m = 1$, that is, to a simple branch point. The general case can be deduced from this by allowing several simple branch points to coincide.

In order to form the expression for $\frac{du}{dt}$ we must observe that dt is always real along the boundary, and du is either real or pure imaginary. Hence $\left(\frac{du}{dt}\right)^2$

is real when t is real. This function can be continuously extended over the line of real values of t by the condition that, for conjugate values t and t' of the variable, the function will also have conjugate values. Then $\left(\frac{du}{dt}\right)^2$ is determined for the whole t-plane and turns out to be single-valued.

Let $a_1', a_2' \ldots$ be the conjugate values to $a_1, a_2, \ldots$ and designate the product $(t-a_1)(t-a_2)\cdots$ by $\Pi(t-a)$. Then

$$(11) \qquad u = \text{const.} + \int \sqrt{\frac{\Pi(t-a)\Pi(t-a')\Pi(t-b)}{\Pi(t-c)}}\,\frac{\text{const.}\,dt}{\Pi(t-e)}.$$

The constants $a, b, c, \ldots$ must be determined so that, when $t = e$,

$$u = \sqrt{\frac{A\alpha}{2\pi}}\log(t-e) + \text{f.c.}$$

For u to be continuous and finite for all values of t except a, b, c or e, a relation must exist among the numbers of the latter values. Namely, the difference between the numbers of vertices, and branch points lying in the boundary is 4 larger then twice the difference between the numbers of interior branch points and infinite sectors. We use the abbreviations

$$\Pi(t-a)\Pi(t-a')\Pi(t-b) = \phi(t),$$
$$\Pi(t-c)\Pi(t-e)^2 = \chi(t),$$

that is,

$$\frac{du}{dt} = \text{const.}\sqrt{\frac{\phi(t)}{\chi(t)}}.$$

Then the polynomial $\phi(t)$ is of degree $\nu - 4$ if $\chi(t)$ is of degree ν. Here ν denotes the number of vertices plus twice the number of infinite sectors.

14.

We still need to express η as a function of t. We only succeed in doing this directly in the simplest cases. In general the following method is adopted. Let v be another function of t, yet to be determined, but assumed to be known. In equations (8), (9), (10), v appears essentially in $\frac{du}{d\log\eta}$, which can also be written as $\frac{du}{dv}\frac{dv}{d\log\eta}$. The last factor may be regarded as the product of the factors

$$(12) \qquad k_1 = \sqrt{\frac{dv}{d\eta}}, \quad k_2 = \eta\sqrt{\frac{dv}{d\eta}},$$

that satisfy the first order differential equation

(13) $$k_1 \frac{dk_2}{dv} - k_2 \frac{dk_1}{dv} = 1,$$

as well as the second order differential equation

(14) $$\frac{1}{k_1}\frac{d^2k_1}{dv^2} = \frac{1}{k_2}\frac{d^2k_2}{dv^2}.$$

If we can express one or the other side of this last equation as a function of t, then a second order linear homogeneous differential equation can be set up for which k_1 and k_2 are particular solutions. Let k be the complete integral. We replace $\frac{d^2k}{dv^2}$ by its equivalent

$$\frac{\frac{dv}{dt}\frac{d^2k}{dt^2} - \frac{dk}{dt}\frac{d^2v}{dt^2}}{\left(\frac{dv}{dt}\right)^3}$$

and obtain the differential equation for k

(15) $$\frac{dv}{dt}\frac{d^2k}{dt^2} - \frac{d^2v}{dt^2}\frac{dk}{dt} - \left(\frac{dv}{dt}\right)^3 \left\{\frac{1}{k_1}\frac{d^2k_1}{dv^2}\right\} k = 0.$$

By equation (15) there are two independent particular solutions K_1 and K_2, whose quotient $K_2 : K_1 = H$ gives an image bounded by arcs of great circles of the positive t-half-plane on the sphere. The same occurs for each expression of the form

(16) $$\eta = e^{\theta i}\frac{H-\alpha}{1+\alpha' H},$$

where θ is real and α, α' are conjugate complex quantities.

The function v is to be chosen so that, for finite values of t, the discontinuities of $\frac{1}{k}\frac{d^2k}{dv^2}$ only occur among the points a, a', b, c, e.

If we set

(17) $$\frac{dv}{dt} = \frac{1}{\sqrt{\phi(t)\chi(t)}} = \frac{1}{\sqrt{f(t)}},$$

then the function $\frac{1}{k}\frac{d^2k}{dv^2}$ will have discontinuities at finite values only for the points a, a', b, c, and, indeed, is infinite of the first order for each of these.

Namely, for $t = c$, we have

$$v - v_c = \frac{2\sqrt{t-c}}{\sqrt{f'(c)}},$$
$$\eta - \eta_c = \text{const.}(t-c)^{\gamma}.$$

Consequently

$$k_1 = \sqrt{\frac{dv}{d\eta}} = \text{const.}(v - v_c)^{(1/2)-\gamma},$$

so that

$$\frac{1}{k}\frac{d^2k}{dv^2} = \frac{1}{4}\frac{\left(\gamma^2 - \frac{1}{4}\right) f'(c)}{t-c}.$$

We obtain corresponding expressions for $t = a, a', b$, in which c is replaced by a, a', b respectively and γ by 2.

A similar observation shows that for $t = e$ the function $\frac{1}{k}\frac{d^2k}{dv^2}$ remains continuous.

For $t = \infty$, we obtain

$$\frac{1}{k}\frac{d^2k}{dv^2} = \left(-\frac{\nu}{2} + 2\right)\left(\frac{\nu}{2} - 1\right) t^{2\nu-6}.$$

Hence the expression for $\frac{1}{k}\frac{d^2k}{dv^2}$ appears as follows:

$$\frac{1}{k}\frac{d^2k}{dv^2} = \frac{1}{4}\sum \frac{\left(\gamma^2 - \frac{1}{4}\right) f'(g)}{t-g} + F(t).$$

The sum is over all the points $g = a, a', b, c$ and for a,a' and b, γ is replaced by 2. Here $F(t)$ is a polynomial of degree $2\nu - 6$ in which the first two coefficients may be determined in the following way. We bring dv into the form

$$dv = \frac{t^{-\nu+4}\frac{dt}{t^2}}{\sqrt{f(t)t^{-2\nu+4}}} = t^{-\nu+4}dv_1,$$

or briefly, αdv_1.

Differentiation yields

$$\frac{d^2}{dv^2}\left[\left(\frac{d\eta}{dv}\right)^{-1/2}\right] = \alpha^{-3/2}\frac{d^2}{dv_1^2}\left[\left(\frac{d\eta}{dv_1}\right)^{-1/2}\right] + \left(\frac{d\eta}{dv_1}\right)^{-1/2}\frac{d^2(\alpha^{1/2})}{dv^2}.$$

Hence

$$\left(\frac{d\eta}{dv}\right)^{1/2}\frac{d^2}{dv^2}\left[\left(\frac{d\eta}{dv}\right)^{-1/2}\right]$$
$$=\alpha^{-2}\left(\frac{d\eta}{dv_1}\right)^{1/2}\frac{d^2}{dv_1^2}\left[\left(\frac{d\eta}{dv_1}\right)^{-1/2}\right]+\alpha^{-1/2}\frac{d^2(\alpha^{1/2})}{dv^2},$$

or

$$\left(\frac{d\eta}{dv_1}\right)^{1/2}\frac{d^2}{d{v_1}^2}\left[\left(\frac{d\eta}{dv_1}\right)^{-1/2}\right]=t^{-2\nu+8}\,\frac{1}{k}\,\frac{d^2k}{dv^2}-\alpha^{3/2}\frac{d^2(\alpha^{1/2})}{dv^2},$$

or

$$\left(\frac{d\eta}{dv_1}\right)^{1/2}\frac{d^2}{d{v_1}^2}\left[\left(\frac{d\eta}{dv_1}\right)^{-1/2}\right]$$
$$=t^{-2\nu+8}\sum\frac{1}{4}\,\frac{\left(\gamma^2-\frac{1}{4}\right)f'(g)}{t-g}+t^{-2\nu+8}F(t)-\alpha^{3/2}\frac{d^2(\alpha^{1/2})}{dv^2}.$$

The function on the left side is finite for $t=\infty$. Hence we have to equate the coefficients of t^2, respectively t, on the right in the expansions of $t^{-2\nu+8}F(t)$ and of $\alpha^{3/2}\frac{d^2(\alpha^{1/2})}{dv^2}$. The expansion of $\alpha^{3/2}\frac{d^2(\alpha^{1/2})}{dv^2}$, by a simple calculation, is

$$\alpha^{3/2}\,\frac{d^2(\alpha^{1/2})}{dv^2}=\frac{1}{2}\left(-\frac{\nu}{2}+2\right)t^{-\nu+5}\,\frac{d(t^{-\nu+2}f(t))}{dt}.$$

Now $F(t)$ still has $2\nu-7$ undetermined coefficients. However, it is important to note that these must be real. For we found in §12 that du is real or pure imaginary on all straight boundary lines of the minimal surface and hence also everywhere on the boundary of the images. By (17), the same holds for dv. We can prove from this that $\frac{1}{k}\,\frac{d^2k}{dv^2}$ must have real values for real values of t.

In order to carry out the proof, we consider the image on the sphere of radius 1 and take any piece of the boundary, hence the arc of a particular great circle. We place the tangent plane at the pole of this great circle and call it the plane of η_1. Then the constants α_1, α_1', θ_1 can be determined so that

$$\eta_1=e^{\theta_1 i}\,\frac{H-\alpha_1}{1+\alpha_1'H},$$

and we obtain two functions $k_1' = \sqrt{\frac{dv}{d\eta_1}}$ and $k_2' = \eta_1\sqrt{\frac{dv}{d\eta_1}}$ which are particular solutions of the differential equation (15). Consequently we have

$$\frac{1}{k}\,\frac{d^2k}{dv^2} = \frac{1}{k_1}\,\frac{d^2k_1'}{dv^2}.$$

The part of the boundary just considered is mapped on the η_1-plane by the equation

$$\eta_1 = e^{\phi_1 i},$$

and if we insert this into k_1', it is easy to see that $\frac{1}{k_1'}\frac{d^2k_1'}{dv^2}$ turns out to be real on the piece of boundary under consideration. The same follows for $\frac{1}{k}\frac{d^2k}{dv^2}$. Since this observation can be made for each piece of the boundary, $\frac{1}{k}\frac{d^2k}{dv^2}$ is real on the entire boundary.

Now, however, for dv real or pure imaginary, the function $\frac{1}{k_1'}\frac{d^2k_1'}{dv^2}$ is also real, if we set in general

$$\eta_1 = \varrho_1 e^{\phi_1 i}$$

and take the modulus ϱ_1 to be constant. For the real t axis to be mapped, in the sphere of radius 1, onto an arc of a great circle, we must have $\varrho_1 = 1$ for each part of the boundary. This gives just as many conditions as there are individual boundary lines.

For this investigation, as in the preceding paragraphs, the values a, b, c, e are all assumed to be finite. If this were not the case, the treatment would need a small adjustment.

Remark: The problem is fully formulated above. In particular cases it comes down to merely setting up and solving the differential equation (15). Incidentally, it is important to note that the number of arbitrary real constants appearing in the solution is exactly the number of conditions which must be satisfied by the nature of the question and the data of the problem. We denote the number of points a, b, c, e respectively by A, B, C, E and note that $2A + B + 4 = C + 2E = \nu$. In the differential equation (15), $2A + B + 4C + 5E - 10$ arbitrary real constants appear. Namely, the angles γ, whose number is C; the $2\nu - 7$ constants for the function $F(t)$; the real quantities b, c, e, of which one may assign three values arbitrarily by making a linear substitution with real coefficients for t; the real and imaginary parts of the quantities a. To these arbitrary constants we add 10 more by integration, namely, if $\eta = \frac{\alpha k_1 + \beta k_2}{\gamma k_1 + \delta k_2}$, the three complex ratios $\alpha : \beta : \gamma : \delta$ account for six real constants, a (real or pure imaginary) factor of du, and a real additive constant in the expressions for each of x, y, z. The constants, however, must still be subject to conditions which must be satisfied if our formulas are to really represent a minimal surface. Of these conditions, $2A + B$ correspond to the points a, a', b. They assert that no logarithm appears in the expansion of the solution of the differential equation (15) valid in a neighborhood of these points. Further $C + E$ equations signify that the pieces of the real t axis which lie between the individual points c, e are mapped into $C + E$ arcs of great circles on the sphere of radius 1. Thus the number of remaining undetermined constants in the solution is $3C + 4E$.

The data of the problem consists of the coordinates of the vertices and in the angles which give the directions of the boundary lines that go to infinity. These data are expressed in $3C + 4E$ equations, for the fulfillment of which we have the correct number of constants available.

Examples.

15.

The boundary consists of two infinite straight lines that do not lie in a plane. Let A be the length of the shortest line segment between the lines, and let $\alpha\pi$ be the angle which the projection of the surface fills in on the plane perpendicular to that shortest line segment.

We take the shortest connecting line as the x-axis, hence x has a constant value in each of the boundary lines. Similarly, ϕ is constant on each of the two boundary lines. At infinity the positive normal for one sector is parallel to the positive x-axis, for the other sector parallel to the negative x-axis. The boundary is mapped on the sphere into two great circles through the poles $\eta = 0$ and $\eta = \infty$ that enclose an angle $\alpha\pi$.

Hence

$$X = -\frac{iA}{2\alpha\pi}\log\eta,$$
$$s = -\frac{iA}{2\alpha\pi}\left(\eta - \frac{1}{\eta'}\right),$$
$$s' = -\frac{iA}{2\alpha\pi}\left(\frac{1}{\eta} - \eta'\right),$$

and consequently

$$x = -i\,\frac{A}{2\alpha\pi}\,\log\left(\frac{\eta}{\eta'}\right)$$

(a)
$$= -i\,\frac{A}{2\alpha\pi}\,\log\left(-\frac{s}{s'}\right),$$

which we recognize as the equation of the helicoid.

16.

The boundary consists of three straight lines, of which two intersect and the third runs parallel to the plane of the first two.

Place the origin of the coordinates at the intersection of the first two lines and the positive x-axis in the negative normal, then the intersection point maps to the point $\eta = \infty$ on the sphere. The images of the first two lines are great half circles running from $\eta = \infty$ to $\eta = 0$. Let their angle be $\alpha\pi$. The image of the third line is an arc of the great circle which runs from $\eta = 0$ to a certain point, reverses itself and returns to $\eta = 0$. This arc forms angles of $-\beta\pi$ and $\gamma\pi$ with the first two great half circles, where β and γ are positive numbers, $\beta + \gamma = \alpha$. In order to obtain the mapping onto the t-half-plane, we specify that $t = \infty$ when $\eta = \infty$, that $t = b$ will correspond to the infinite sector between the first and third lines, that $t = c$ will correspond to the infinite sector between the second and the third lines, and that $t = a$ will correspond to the reversing point of the normals on the third line. Then a, b and c are real and $c > a > b$. These specifications correspond to $\eta = (t-b)^\beta (t-c)^\gamma$. The value of a depends on b and c. We have, namely,

$$\frac{d\log\eta}{dt} = \frac{\beta(t-c) + \gamma(t-b)}{(t-b)(t-c)}$$

and this must be 0 at the reversing point, thus $a = \frac{c\beta + b\gamma}{\beta+\gamma}$. Furthermore, from sections 12 and 13,

$$du = \sqrt{\frac{A(c-b)(\beta+\gamma)}{2\pi}}\,\frac{(t-a)^{1/2}dt}{(t-b)(t-c)},$$

or, if we take $c - b = \frac{2\pi}{A}$,

$$du = \sqrt{\beta + \gamma}\,\frac{(t-a)^{1/2}dt}{(t-b)(t-c)},$$
$$\frac{du}{d\log\eta} = \frac{1}{\sqrt{(\beta+\gamma)(t-a)}},$$
$$\left(\frac{du}{d\log\eta}\right)^2 d\log\eta = \frac{dt}{(t-b)(t-c)}.$$

Consequently,

$$\text{(b)}\quad \begin{aligned} x &= -i\int\frac{dt}{(t-b)(t-c)} + i\int\frac{dt'}{(t'-b)(t'-c)},\\ y &= -\frac{i}{2}\int\frac{(t-b)^{\beta}(t-c)^{\gamma} - (t-b)^{-\beta}(t-c)^{-\gamma}}{(t-b)(t-c)}dt\\ &\quad + \frac{i}{2}\int\frac{(t'-b)^{\beta}(t'-c)^{\gamma} - (t'-c)^{-\beta}(t'-c)^{-\gamma}}{(t'-b)(t'-c)}dt',\\ z &= -\frac{1}{2}\int\frac{(t-b)^{\beta}(t-c)^{\gamma} + (t-b)^{-\beta}(t-c)^{-\gamma}}{(t-b)(t-c)}dt\\ &\quad - \frac{1}{2}\int\frac{(t'-b)^{\beta}(t'-c)^{\gamma} + (t'-b)^{-\beta}(t'-c)^{-\gamma}}{(t'-b)(t'-c)}dt'. \end{aligned}$$

17.

The boundary consists of three skew lines, whose shortest distances apart are A, B, C. Between each two boundary lines the surface stretches to infinity. Let $\alpha\pi$, $\beta\pi$, $\gamma\pi$ be the angles of the directions in which the boundary lines of the first, the second, and the third sectors run towards infinity. We specify that the quantity t will be respectively $0, \infty$, and 1 at infinity for the three sectors of the minimal surface. Thus we obtain

$$\frac{du}{dt} = \frac{\sqrt{\phi(t)}}{t(1-t)}.$$

Here $\phi(t)$ is a polynomial of degree 2. Its coefficients are determined by

$$\frac{du}{d\log t} = \sqrt{\frac{A\alpha}{2\pi}} \qquad \text{for } t = 0,$$

$$\frac{du}{d\log t} = \sqrt{\frac{B\beta}{2\pi}} \qquad \text{for } t = \infty,$$

$$\frac{du}{d\log(1-t)} = \sqrt{\frac{C\gamma}{2\pi}} \qquad \text{for } t = 1.$$

From this it follows that

$$\phi(t) = \frac{A\alpha}{2\pi}(1-t) + \frac{C\gamma}{2\pi}t - \frac{B\beta}{2\pi}t(1-t).$$

Depending on whether the roots of the equation $\phi(t) = 0$ are imaginary or real, the image on the sphere has a branch point in the interior or two reversing points of the normals on the boundary.

The functions $k_1 = \sqrt{\frac{dv}{d\eta}}$ and $k_2 = \eta\sqrt{\frac{dv}{d\eta}}$ will only be discontinuous for the three sectors, if we take $\frac{dv}{d\eta} = \phi(t)$. Indeed, the discontinuity of k_1 is such that

$$\begin{aligned} &t^{-\frac{1}{2}+\frac{\alpha}{2}}k_1, && \text{for } t = 0,\\ &t^{-\frac{3}{2}-\frac{\beta}{2}}k_1, && \text{for } t = \infty,\\ &(1-t)^{-\frac{1}{2}+\frac{\gamma}{2}}\,k_1, && \text{for } t = 1 \end{aligned}$$

are single valued, with finite nonzero limits. The functions k_1 and k_2 are particular solutions of a homogeneous linear second order differential equation, which is obtained if we represent $\frac{1}{k}\frac{d^2k}{dv^2}$, except at its discontinuities, as a function of t, and substitute t instead of v as the independent variable in $\frac{d^2k}{dv^2}$. If we have found the particular solution k_1, then k_2 is found from the first order differential equation

$$\text{(c)} \qquad k_1\frac{dk_2}{dt} - k_2\frac{dk_1}{dt} = \phi(t).$$

The complete integral of the second order homogeneous linear differential equation will be denoted by

$$\text{(d)} \qquad k = Q\begin{Bmatrix} \frac{1}{2}-\frac{\alpha}{2} & -\frac{3}{2}-\frac{\beta}{2} & \frac{1}{2}-\frac{\gamma}{2} & \\ & & & t \\ \frac{1}{2}+\frac{\alpha}{2} & -\frac{3}{2}+\frac{\beta}{2} & \frac{1}{2}+\frac{\gamma}{2} & \end{Bmatrix}.$$

This function satisfies essentially the same conditions as in the definition of the P-function in the paper on the Gauss series $F(\alpha, \beta, \gamma, x)$.[1] It differs from the P-function in that the sum of the exponents is -1 instead of $+1$ as for P.

We can express the function Q with the help of a P-function and its first derivative. First,

$$k = t^{\frac{1}{2}-\frac{\alpha}{2}}(1-t)^{\frac{1}{2}-\frac{\gamma}{2}} Q \begin{Bmatrix} 0 & \frac{-\alpha-\beta-\gamma-1}{2} & 0 & \\ & & & t \\ \alpha & \frac{-\alpha+\beta-\gamma-1}{2} & \gamma & \end{Bmatrix}.$$

If we set

$$\sigma = P \begin{Bmatrix} 0 & \frac{-\alpha-\beta-\gamma+1}{2} & 0 & \\ & & & t \\ \alpha & \frac{-\alpha+\beta-\gamma+1}{2} & \gamma & \end{Bmatrix},$$

the constants a, b, c can be determined so that

$$\text{(e)} \qquad k = t^{\frac{1}{2}-\frac{\alpha}{2}}(1-t)^{\frac{1}{2}-\frac{\gamma}{2}} \left((a+bt)\sigma + ct(1-t)\frac{d\sigma}{dt} \right).$$

In fact, we have only to substitute this expression into the differential equation (c) and consider the second order differential equation for σ in order to obtain the equations

$$\phi(t) = t^{1-\alpha}(1-t)^{1-\gamma}\left(\sigma_1 \frac{d\sigma_2}{dt} - \sigma_2 \frac{d\sigma_1}{dt}\right) F(t),$$

$$F(t) = a(a+c\alpha)(1-t) + (a+b)(a+b-c\gamma)t$$
$$- t(1-t)\left(b - \frac{\alpha+\beta+\gamma-1}{2}c\right)\left(b - \frac{\alpha-\beta+\gamma-1}{2}c\right).$$

By the properties of the function σ, we can set

$$t^{1-\alpha}(1-t)^{1-\gamma}\left(\sigma_1 \frac{d\sigma_2}{dt} - \sigma_2 \frac{d\sigma_1}{dt}\right) = 1,$$

and hence $F(t) = \phi(t)$. From this we obtain three equations for a, b, c, which take a very simple form if we set

$$a + \frac{\alpha}{2}c = p, \quad b - \frac{\alpha+\gamma-1}{2}c = q, \quad a+b-\frac{\gamma}{2}c = -r.$$

[1] Contributions to the theory of the functions represented by the Gauss series $F(\alpha, \beta, \gamma, x)$ (**IV** in this collection).

The equations are then

$$p^2 - \alpha^2(p+q+r)^2 = \frac{A\alpha}{2\pi},$$
$$q^2 - \beta^2(p+q+r)^2 = \frac{B\beta}{2\pi},$$
$$r^2 - \gamma^2(p+q+r)^2 = \frac{C\gamma}{2\pi}.$$

With the help of the function

$$\lambda = P\begin{Bmatrix} -\frac{\alpha}{2} & -\frac{\beta}{2} & \frac{1}{2}-\frac{\gamma}{2} & \\ & & & t \\ \frac{\alpha}{2} & \frac{\beta}{2} & \frac{1}{2}+\frac{\gamma}{2} & \end{Bmatrix},$$

whose branches λ_1 and λ_2 satisfy the differential equation

$$\lambda_1 \frac{d\lambda_2}{d\log t} - \lambda_2 \frac{d\lambda_1}{d\log t} = 1,$$

we can express k even more simply, namely,

$$\text{(f)} \qquad k = t^{1/2}\left((p+qt)\lambda + ct(1-t)\frac{d\lambda}{dt}\right).$$

It would not be hard to produce the individual branches of the function k in the form of definite integrals. The method is shown in §7 of the paper on the P-function.

In the particular case where the three boundary straight lines run parallel to the coordinate axes, $\alpha = \beta = \gamma = \frac{1}{2}$. Then we obtain

$$\lambda = P\begin{pmatrix} -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & \\ & & & t \\ \frac{1}{4} & \frac{1}{4} & \frac{3}{4} & \end{pmatrix} = \left(\frac{t-1}{t}\right)^{1/4} P\begin{pmatrix} 0 & -\frac{1}{4} & 0 & \\ & & & t \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{2} & \end{pmatrix}.$$

The branch λ_1 of this function is

$$\left(\frac{t-1}{t}\right)^{1/4}\sqrt{t^{1/2}+(t-1)^{1/2}}\ \text{const.},$$

so that

$$k_1 = \sqrt{2}\,t^{1/4}(t-1)^{1/4}\sqrt{t^{1/2}+(t-1)^{1/2}}\left\{p+qt-\frac{c}{4}-\frac{c}{4}\sqrt{t(t-1)}\right\},$$
$$k_2 = -\sqrt{2}\,t^{1/4}(t-1)^{1/4}\sqrt{t^{1/2}-(t-1)^{1/2}}\left\{p+qt-\frac{c}{4}+\frac{c}{4}\sqrt{t(t-1)}\right\}.$$

With the aid of these two functions, dX, dY, dZ can be expressed in the following way:

$$dX = -ik_1k_2 \frac{dt}{t^2(1-t)^2},$$
$$dY = -\frac{i}{2}\left(k_2^2 - k_1^2\right)\frac{dt}{t^2(1-t)^2},$$
$$dZ = -\frac{1}{2}\left(k_2^2 + k_1^2\right)\frac{dt}{t^2(1-t)^2},$$

$$\text{(g)}\qquad \begin{aligned} iX &= (p+q-r)^2\sqrt{\frac{t}{t-1}} + (-p+q+r)^2\sqrt{\frac{t-1}{t}} \\ &\quad + \frac{1}{2}(p+3q+r)(p-q+r)\log\frac{t^{1/2}+(t-1)^{1/2}}{t^{1/2}-(t-1)^{1/2}}, \\ iY &= -(p-q+r)^2 t^{1/2} - (-p+q+r)^2 t^{-1/2} \\ &\quad - \frac{1}{2}(p+q+3r)(p+q-r)\log\frac{1+t^{1/2}}{1-t^{1/2}}, \\ iZ &= (p-q+r)^2(1-t)^{1/2} + (p+q-r)^2(1-t)^{-1/2} \\ &\quad + \frac{1}{2}(3p+q+r)(-p+q+r)\log\frac{1+\sqrt{1-t}}{1-\sqrt{1-t}}. \end{aligned}$$

If p, q and r are real, then twice the coefficient of i in the three quantities on the right gives the rectilinear coordinates of a point in the surface.

18.

The boundary consists of four intersecting line segments which we obtain by removing two edges that do not touch from the edges of an arbitrary tetrahedron. The image on the sphere is a spherical quadrilateral, whose angles may be taken to be $\alpha\pi$, $\beta\pi$, $\gamma\pi$, and $\delta\pi$. It follows that

$$du = \frac{C\,dt}{\sqrt{(t-a)(t-b)(t-c)(t-d)}} = \frac{C\,dt}{\sqrt{\Delta(t)}},$$

where the real values $t = a, b, c, d$ designate the points on the t-plane into which the vertices of the quadrilateral are mapped.

If the methods developed in §14 are applied to determine η, then we have here, in particular, $\phi(t) = 1$, $\chi(t) = \Delta(t)$. Consequently $v = \frac{u}{C}$ and

$$k_1 = \sqrt{\frac{dv}{d\eta}}, \quad k_2 = \eta\sqrt{\frac{dv}{d\eta}}.$$

The functions k_1 and k_2 satisfy the differential equation

$$k_1 \frac{dk_2}{dv} - k_2 \frac{dk_1}{dv} = 1$$

and are particular solutions of the second order differential equation

$$\frac{4}{k}\frac{d^2k}{dv^2} = \frac{\left(\alpha^2 - \frac{1}{4}\right)\Delta'(a)}{t-a} + \frac{\left(\beta^2 - \frac{1}{4}\right)\Delta'(b)}{t-b} + \frac{\left(\gamma^2 - \frac{1}{4}\right)\Delta'(c)}{t-c} + \frac{\left(\delta^2 - \frac{1}{4}\right)\Delta'(d)}{t-d} + h.$$

The polynomial $F(t)$ of §14 is of degree 2 here, but the coefficients of t^2 and t are zero. Thus h is a constant. On the left side of the last equation, we introduce t as an independent variable and obtain

$$\text{(h)} \qquad \frac{4}{k}\left(\Delta(t)\frac{d^2k}{dt^2} + \frac{1}{2}\Delta'(t)\frac{dk}{dt}\right)$$
$$= \frac{\left(\alpha^2 - \frac{1}{4}\right)\Delta'(a)}{t-a} + \frac{\left(\beta^2 - \frac{1}{4}\right)\Delta'(b)}{t-b} + \frac{\left(\gamma^2 - \frac{1}{4}\right)\Delta'(c)}{t-c} + \frac{\left(\delta^2 - \frac{1}{4}\right)\Delta'(d)}{t-d} + h$$

as the second order linear differential equation which k must satisfy.

If x, y, z are actually expressed as functions of t, then 16 undetermined real coefficients enter into the solution. These are the four quantities a, b, c, d of which, as above, three can be taken arbitrarily, the four quantities $\alpha, \beta, \gamma, \delta$, the quantity h, six more real constants in the expressions for η, a constant factor in du and an additive constant in each of x, y, z. To determine these 16 numbers there are 16 equations at hand. Namely, 4 equations which express that the 4 boundary lines in the η-plane map into great circles on the sphere, and 12 equations which state that x, y, z have given values at the four vertices.

In the special case of a regular tetrahedron, the image on the sphere is a regular quadrilateral in which each angle is $\frac{2}{3}\pi$. The diagonals bisect it and are perpendicular to each other. The points that are diametrically opposite to the vertices on the sphere are the vertices of a congruent quadrilateral. Between the two lie four quadrilaterals congruent to the original, each of which has two vertices in the original and two in its opposite. These six quadrilaterals fill the surface of the sphere once. Thus $\frac{du}{d\log\eta}$ is an algebraic function of η.

The minimal surface we are seeking can be continuously extended across the original boundaries by rotation of 180° about each of the original boundaries, using those boundaries as rotation axes. Along such a boundary line the original surface and the extension have a common normal. By repetition of this construction on the new pieces of surface, we can extend the original surface arbitrarily far. Whichever extension we consider, however, always maps on the sphere in one of the six congruent quadrilaterals. In fact, the image of two pieces of surface either have a side in common, or lie opposite to one another, depending on whether the surface pieces themselves adjoin at a boundary line, or abut opposite boundary lines of a piece of surface lying between them. In the last case the relevant surface pieces can be made to coincide by parallel translation. Hence $\left(\frac{du}{d\log\eta}\right)^2$ must remain unchanged if η and $-1/\eta$ are interchanged.

If we set the pole ($\eta = 0$) at the midpoint of a quadrilateral, and the 0-meridian through the middle of a side, then for the vertices of this quadrilateral,

$$\eta = \left(\tan\frac{c}{2}\right)e^{\pm\pi i/4}, \quad \left(\tan\frac{c}{2}\right)e^{\pm 3\pi i/4}$$

and

$$\tan\frac{c}{2} = \frac{\sqrt{3}-1}{\sqrt{2}}.$$

Points which correspond to opposite values of η have the same x-coordinates. Thus $\left(\frac{du}{d\log\eta}\right)^2$ must remain invariant on interchanging η and $-\eta$. We obtain

$$\left(\frac{du}{d\log\eta}\right)^2 = \frac{C_1}{\sqrt{\eta^4+\eta^{-4}+14}}.$$

The constant C_1 must be real, since du^2 has real values on the boundary. We get the same result in the following way. The substitution

$$\left\{\frac{\eta^2+\eta^{-2}-2\sqrt{3}i}{\eta^2+\eta^{-2}+2\sqrt{3}i}\right\}^3 = \left(\frac{t^2-1}{t^2+1}\right)^2$$

yields an image on the t-plane which is bounded by a closed continuously bending line. The computation shows that $d\log t$ is pure imaginary on the boundary. Consequently the image of the boundary in the t-plane is a circle with center at $t = 0$. The radius of this circle is 1. The vertices

$$\eta = \pm\left(\tan\frac{c}{2}\right) e^{\pi i/4}$$

correspond to $t = \pm 1$; the vertices

$$\eta = \pm\left(\tan\frac{c}{2}\right) e^{-\pi i/4}$$

correspond to $t = \pm i$. If we pass, within the interior of the minimal surface, at any of these four places from one boundary line to the next, then the argument of dt changes by π. Thus, as in §13, we can set here

$$\frac{du}{dt} = \frac{C_2}{\sqrt{(t^2-1)(t^2+1)}},$$

and C_2^2 must be pure imaginary since du^2 is real on the boundary. It turns out that $C_1 = 3\sqrt{3}\,C_2^2 i$.

This expression agrees with the previous one for $\left(\frac{du}{d\log\eta}\right)^2$. For further simplification we take

$$\left(\frac{t^2-1}{t^2+1}\right)^2 = \omega^3, \quad \eta^2 + \eta^{-2} = 2\lambda$$

and observe that

$$\left(\frac{du}{d\log\eta}\right)^2 d\log\eta = \left(\frac{du}{d\lambda}\right)^2 \frac{d\lambda}{d\log\eta}\, d\lambda.$$

Then a very simple calculation gives

$$X = -i\int\left(\frac{du}{d\log\eta}\right)^2 d\log\eta = C\int\frac{d\omega}{\sqrt{\omega(1-\varrho\omega)(1-\varrho^2\omega)}},$$

$$\text{(i)}\quad Y = -\frac{i}{2}\int\left(\frac{du}{d\log\eta}\right)^2\left(\eta-\frac{1}{\eta}\right) d\log\eta = C\varrho^2\int\frac{d\omega}{\sqrt{\omega(1-\omega)(1-\varrho^2\omega)}},$$

$$Z = -\frac{1}{2}\int\left(\frac{du}{d\log\eta}\right)^2\left(\eta+\frac{1}{\eta}\right) d\log\eta = C\varrho\int\frac{d\omega}{\sqrt{\omega(1-\omega)(1-\varrho\omega)}},$$

where $\varrho = -\frac{1}{2}(1-i\sqrt{3})$ denotes a cube root of 1. The real constant $C = \frac{1}{8}C_1$ is determined from the the given length of the edges of the tetrahedron.

19.

Finally we will consider the problem of the minimal surface for the case where the boundary consists of two arbitrary circles that lie in parallel planes. Then one does not know the direction of the normals on the boundary. Hence these can not be mapped to the sphere. However, we succeed in solving the problem by the assumption that all cross-sections parallel to the plane of the boundary circles are circles. It will be shown that under this assumption the minimality conditions can be satisfied.

Place the x-axis at right angles to the planes of the boundary circles, then the equation of a cross section by a parallel plane is

$$\text{(k)} \qquad F = y^2 + z^2 + 2\alpha y + 2\beta z + \gamma = 0.$$

and α, β, γ are to be determined as functions of x. For brevity we set

$$\sqrt{\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2 + \left(\frac{\partial F}{\partial z}\right)^2} = \frac{1}{n},$$

so that

$$\cos r = n\,\frac{\partial F}{\partial x}, \quad \sin r \cos\phi = n\,\frac{\partial F}{\partial y}, \quad \sin r \sin\phi = n\,\frac{\partial F}{\partial z}.$$

Then the minimality condition may be brought into the form

$$\frac{\partial\left(n\,\frac{\partial F}{\partial x}\right)}{\partial x} + \frac{\partial\left(n\,\frac{\partial F}{\partial y}\right)}{\partial y} + \frac{\partial\left(n\,\frac{\partial F}{\partial z}\right)}{\partial z} = 0,$$

or after carrying out the differentiation,

$$4\,\frac{\partial^2 F}{\partial x^2}(F + \alpha^2 + \beta^2 - \gamma) + 4\left(\frac{\partial F}{\partial x}\right)^2 - 4\,\frac{\partial F}{\partial x}\,\frac{\partial}{\partial x}(F + \alpha^2 + \beta^2 - \gamma)$$
$$+4 \cdot 2(F + \alpha^2 + \beta^2 - \gamma) = 0.$$

If we write $\alpha^2 + \beta^2 - \gamma = -q$ and observe that $F = 0$, the last equation becomes

$$\text{(l)} \qquad q\,\frac{\partial^2 F}{\partial x^2} - \frac{\partial F}{\partial x}\,\frac{\partial q}{\partial x} + 2q = 0$$

and gives after one integration

$$\frac{1}{q}\frac{\partial F}{\partial x}+2\int\frac{dx}{q}+\text{const.}=0.$$

The integration constant is independent of x. If on the other hand we take $\int\frac{dx}{q}$ independent of y and z, then the integration constant must be a linear function of y and z, since $\frac{1}{q}\frac{\partial F}{\partial x}$ is. Thus we have

$$\frac{1}{q}\frac{\partial F}{\partial x}+2\int\frac{dx}{q}+2ay+2bz+\text{const.}=0.$$

If we compare this with the result of the direct differentiation of F, namely

$$\frac{\partial F}{\partial x}=2y\frac{d\alpha}{dx}+2z\frac{d\beta}{dx}+\frac{d\gamma}{dx},$$

we obtain

$$\frac{d\alpha}{dx}=-aq,\quad \frac{d\beta}{dx}=-bq,$$

and if we set $\int q\,dx=m$:

$$\alpha=-am+d,\quad \beta=-bm+e.$$

Accordingly we have

$$\begin{aligned}\frac{\partial F}{\partial x}&=-\,2aqy-2bqz+\frac{d\gamma}{dx},\\ \frac{\partial^2 F}{\partial x^2}&=-\,2ay\frac{dq}{dx}-2bz\frac{dq}{dx}+\frac{d^2\gamma}{dx^2},\end{aligned}$$

and these expressions are inserted into equation (1). After simplification we obtain

$$q\frac{d^2\gamma}{dx^2}-\frac{dq}{dx}\frac{d\gamma}{dx}+2q=0,$$

an equation which can be further simplified if we observe that

$$\begin{aligned}\gamma=q+\alpha^2+\beta^2=q+f(m)&=\frac{dm}{dx}+f(m),\\ f(m)=(a^2+b^2)m^2-2(ad+be)m+d^2+e^2.\end{aligned}$$

If we take $\frac{d\gamma}{dx}$ and $\frac{d^2\gamma}{dx^2}$ from these expressions, then the differential equation which expresses the condition of minimality becomes

$$\text{(m)} \qquad q\,\frac{d^2q}{dx^2} - \left(\frac{dq}{dx}\right)^2 + 2q + 2(a^2+b^2)q^3 = 0.$$

To carry out the integration, we set $\frac{dq}{dx} = p$ and consider q as an independent variable. In this way, we obtain for p^2 as a function of q a linear first order differential equation, namely,

$$\frac{1}{2}q\,\frac{d(p^2)}{dq} - p^2 + 2q + 2(a^2+b^2)q^3 = 0$$

or

$$\frac{q^2 d(p^2) - p^2 d(q^2)}{q^4} = -\left(\frac{4}{q^2} + 4(a^2+b^2)\right)dq.$$

The solution is

$$\text{(n)} \qquad \frac{p^2}{q^2} = \frac{4}{q} - 4(a^2+b^2)q + 8c.$$

Replacing p by $\frac{dq}{dx}$, we obtain

$$dx = \frac{dq}{2\sqrt{q + 2cq^2 - (a^2+b^2)q^3}},$$
$$dm = \frac{q\,dq}{2\sqrt{q + 2cq^2 - (a^2+b^2)q^3}}.$$

Hence

$$\text{(o)} \qquad \begin{aligned} x &= \int \frac{dq}{2\sqrt{q + 2cq^2 - (a^2+b^2)q^3}},\\ m &= \int \frac{q\,dq}{2\sqrt{q + 2cq^2 - (a^2+b^2)q^3}},\\ y &= am - d + \sqrt{-q}\,\cos\psi,\\ z &= bm - e + \sqrt{-q}\,\sin\psi. \end{aligned}$$

Hence we have expressed x, y, z as functions of two real variables q and ψ. The expressions are, except for algebraic terms, elliptic integrals with upper limit q. By the general method developed above, we have obtained x, y, z as sums of two conjugate functions of two conjugate complex variables. This suggests the conjecture that these complex expressions can be combined into a single integral expression with the variable q, using the addition theorem for elliptic functions.

This is easy to confirm. Namely, from the formulas for the direction coordinates r and ϕ of the normals, we have

$$\frac{\eta}{\eta'} = e^{2\phi i} = \frac{\frac{\partial F}{\partial y} + \frac{\partial F}{\partial z} i}{\frac{\partial F}{\partial y} - \frac{\partial F}{\partial z} i} = \frac{y + zi + \alpha + \beta i}{y - zi + \alpha - \beta i} = e^{2\psi i}.$$

If we combine this with the defining equation of q, namely:

$$(y + zi + \alpha + \beta i)(y - zi + \alpha - \beta i) = -q,$$

we obtain

$$(y + zi) + (\alpha + \beta i) = (-q)^{1/2} \eta^{1/2} \eta'^{-1/2},$$
$$(y - zi) + (\alpha - \beta i) = (-q)^{1/2} \eta^{-1/2} \eta'^{1/2}.$$

Furthermore, we have

$$\cot r = \frac{\frac{\partial F}{\partial x}}{\sqrt{\left(\frac{\partial F}{\partial y}\right)^2 + \left(\frac{\partial F}{\partial x}\right)^2}} = \frac{1}{2\sqrt{-q}} \{p - 2aq(y + \alpha) - 2bq(z + \beta)\}$$

or

$$\frac{1}{\sqrt{\eta\eta'}} - \sqrt{\eta\eta'} = \frac{\cos^2 \frac{r}{2} - \sin^2 \frac{r}{2}}{\sin \frac{r}{2} \cos \frac{r}{2}} = \frac{1}{\sqrt{-q}} \{p - 2aq(y + \alpha) - 2bq(z + \beta)\}.$$

On the right side, the expressions in η and η' that we just found are inserted for $y + \alpha$ and $z + \beta$. The equation becomes

$$\frac{p}{q} = (-q)^{1/2} \left[(a + bi) \left(\frac{\eta'}{\eta}\right)^{1/2} + (a - bi) \left(\frac{\eta}{\eta'}\right)^{1/2} \right]$$
$$+ (-q)^{-1/2} \left(\sqrt{\eta\eta'} + \frac{1}{\sqrt{\eta\eta'}} \right).$$

We square both sides of this equation and replace $\frac{p^2}{q^2}$ by its value from (n). After simplifying, we obtain

$$(\text{p})\quad \begin{aligned}(-q)\left[(a+bi)\left(\frac{\eta'}{\eta}\right)^{1/2}-(a-bi)\left(\frac{\eta}{\eta'}\right)^{1/2}\right]^2+\frac{1}{(-q)}\left[\sqrt{\eta\eta'}+\frac{1}{\sqrt{\eta\eta'}}\right]^2\\ =8c-2(a+bi)\left(\eta'-\frac{1}{\eta}\right)-2(a-bi)\left(\eta-\frac{1}{\eta'}\right).\end{aligned}$$

The equation so obtained, which gives the relation between q, η, and η', can be considered as the solution of a differential equation in η and η', and q can be understood as the integration constant. The differential equation is transformed by a direct differentiation into the form

$$\begin{aligned}0=\frac{d\eta}{\eta}\Bigg[&\frac{1}{\sqrt{-q}}\left(\sqrt{\eta\eta'}+\frac{1}{\sqrt{\eta\eta'}}\right)\\ &-\sqrt{-q}\left((a+bi)\left(\frac{\eta'}{\eta}\right)^{1/2}-(a-bi)\left(\frac{\eta}{\eta'}\right)^{1/2}\right)\Bigg]\\ +\frac{d\eta'}{\eta'}\Bigg[&\frac{1}{\sqrt{-q}}\left(\sqrt{\eta\eta'}+\frac{1}{\sqrt{\eta\eta'}}\right)\\ &+\sqrt{-q}\left((a+bi)\left(\frac{\eta'}{\eta}\right)^{1/2}-(a-bi)\left(\frac{\eta}{\eta'}\right)^{1/2}\right)\Bigg].\end{aligned}$$

With the aid of the primitive equation (p), the factors of $\frac{d\eta}{\eta}$ and $\frac{d\eta'}{\eta'}$ can be expressed differently. We merely use the left side from (p) in two ways to form a complete square, in which we insert the missing doubled product one time as positive and the other as negative. We obtain

$$\begin{aligned}&\frac{1}{\sqrt{-q}}\left(\sqrt{\eta\eta'}+\frac{1}{\sqrt{\eta\eta'}}\right)+\sqrt{-q}\left((a+bi)\sqrt{\frac{\eta'}{\eta}}-(a-bi)\sqrt{\frac{\eta}{\eta'}}\right)\\ &\quad=\pm2\sqrt{\left[2c+(a+bi)\frac{1}{\eta}-(a-bi)\eta\right]},\end{aligned}$$

$$\frac{1}{\sqrt{-q}}\left(\sqrt{\eta\eta'}+\frac{1}{\sqrt{\eta\eta'}}\right)-\sqrt{-q}\left((a+bi)\sqrt{\frac{\eta'}{\eta}}-(a-bi)\sqrt{\frac{\eta}{\eta'}}\right)$$
$$=\pm 2\sqrt{\left[2c+(a-bi)\,\frac{1}{\eta'}-(a+bi)\eta'\right]}.$$

If we take the square roots with the same signs, then the differential equation becomes

$$\text{(q)}\qquad \begin{aligned} 0 = & \frac{d\,\eta}{2\eta\sqrt{2c+(a+bi)\,\frac{1}{\eta}-(a-bi)\eta}} \\ & + \frac{d\,\eta'}{2\eta'\sqrt{2c+(a-bi)\frac{1}{\eta'}-(a+bi)\eta'}}. \end{aligned}$$

Its solution in algebraic form is expressed in equation (p), or what amounts to the same thing, in the two equations

$$\text{(r)}\qquad \begin{aligned} \frac{1}{\sqrt{-q}}\,(1+\eta\eta') = & \sqrt{\eta'[(a+bi)+2c\eta-(a-bi)\eta^2]} \\ & + \sqrt{\eta[(a-bi)+2c\eta'-(a+bi)\eta'^2]}, \\ \sqrt{-q}\Big((a+bi)\eta' - (a-bi)\eta\Big) = & \sqrt{\eta'[(a+bi)+2c\eta-(a-bi)\eta^2]} \\ & - \sqrt{\eta[(a-bi)+2c\eta'-(a+bi)\eta'^2]}. \end{aligned}$$

In transcendental form the solution is

$$\text{(s)}\qquad \begin{aligned} \text{const.} = & \int \frac{d\,\eta}{2\sqrt{\eta[(a+bi)+2c\eta-(a-bi)\eta^2]}} \\ & + \int \frac{d\,\eta'}{2\sqrt{\eta'[(a-bi)+2c\eta'-(a+bi)\eta'^2]}}, \end{aligned}$$

and the integration constant can be expressed via

$$\text{const.} = \int \frac{d\,q}{2\sqrt{q[1+2cq-(a^2+b^2)q^2]}}.$$

This follows easily from equation (r), if we take η or η' to be the constant 0. We recognize here the addition theorem for elliptic integrals of the first kind.

XIX.

An attempt to generalize integration and differentiation.

In the following essay an attempt is made to define an operation which yields from any given function of a single variable, another function of this variable, whose dependence on the original function can be expressed by a numerical parameter, and in the case where this number is a positive integer, a negative integer, or zero, the new function coincides respectively with a repeated derivative, a repeated integral, or the original function. The known results of differential and integral calculus will be taken as a foundation. It will not be assumed that results for derivatives and integrals valid in the integer case remain valid for fractional orders. Such results serve on the one hand as the foundation for the operation indicated above, and on the other hand as a guide to finding it.

With the latter aim in view, let us consider rather more closely the sequence of successive derivatives. Clearly we cannot simply utilise the usual definition based on a recursive law of formation, because this cannot yield anything other than the terms corresponding to integers. We are thus obliged to look for an independent determination. One possible means is the development derived from the original function by increasing its argument by an arbitrary increment, and expanding in positive integer powers of the increment. For, since the well-known expansion,

$$(1) \qquad z_{(x+h)} = \sum_{p=0}^{\infty} \frac{1}{1.2\ldots p} \frac{d^p z}{dx^p} h^p$$

(where $z_{(x+h)}$ denotes the result of replacing x in the function $z_{(x)}$ by $x+h$) is valid for arbitrary values of h, the coefficients must have well-defined values. Thus these coefficients can be used to define the derivatives. Accordingly we propose the following definition:

> the nth derivative of the function $z_{(x)}$ is equal to the coefficient of h^n in the expansion of $z_{(x)}$ in positive powers of h, multiplied by a constant depending only on n, namely $1.2.\ldots n$.

This interpretation of the derivatives leads very easily to the definition of a general operation which includes, as particular cases, differentiation and integration. We shall denote the operation, in whose definition the limit of ratio of vanishingly small quantities plays no part, by ∂_x^ν, and call this a *derivation*. Here we follow Lagrange, who used the term 'fonctions derivées'.

Thus we shall take $\partial_x^\nu z$, or the expression "the νth derivation of $z_{(x)}$ with respect to x" to mean the coefficient of h^ν in the expansion of $z_{(x+h)}$ as an infinite series in powers of h ascending and descending by integer amounts, multiplied by constants depending only on ν. That is, we define $\partial_x^\nu z$ by the equation

$$z_{(x+h)} = \sum_{\nu=-\infty}^{\infty} k_\nu \partial_x^\nu z\, h^\nu. \tag{2}$$

In this definition the factors k_ν depending only on ν must naturally be defined in such a way that when the exponents of h are integers, the series (2) becomes the series (1). Only then will the derivatives truly be included as particular cases of derivations. If this were not possible, our aim of defining an operation which includes differentiation as a special case would not be achieved, and we should have to look for another approach.

However, before attempting to determine this factor, we make some preliminary observations concerning series of this form. We see that they form the foundation of this whole attempt at a theory of derivations.

It is generally accepted that no firm conclusions can be drawn from infinite series, unless the values assigned to the variables are such that the series converges. That is, the value of the sum of the series can be found (or at least approximated) by an actual numerical addition. Assuming—as we always shall here—that the coefficients obey some definite law, we can calculate the precise value of each term. The series is consequently a quantity determined in all its parts, hence a definite quantity. I see no reason why the rules of numerical addition, although insufficient to yield an exact value, should not be applied to the series, and the results regarded as correct.

To show by an example that one can find a value for the sum of a series of form (2), we use a process applicable in many cases for this purpose. We expand the function x^μ in fractional powers of $x - b$; we need this in any case in the course of our investigation.

Let the series which represents x^μ, which for brevity we denote by z, be

$$\sum_{\alpha=-\infty}^{\infty} c_\alpha (x-b)^\alpha.$$

If x^{μ}, then

$$\frac{dz}{dx} = \mu x^{\mu-1}$$

and consequently

$$\mu z - x\,\frac{dz}{dx} = 0.$$

Hence

$$\sum[(\mu-\alpha)c_\alpha - b(\alpha+1)c_{\alpha+1}](x-b)^\alpha = 0$$

must hold.

Obviously this condition is satisfied whenever

$$(\mu-\alpha)c_\alpha - b(\alpha+1)c_{\alpha+1} = 0.$$

Now all the expressions which satisfy the differential equation are contained in the different values of kx^μ. The series z obeying the law

$$(\mu-\alpha)c_\alpha - b(\alpha+1)c_{\alpha+1} = 0$$

must necessarily be one of these. In order to find it, we write

$$\begin{aligned} \cdots + c_{\alpha-1}(x-b)^{\alpha-1} + c_\alpha(x-b)^\alpha &= p, \\ p' = c_{\alpha+1}(x-b)^{\alpha+1} + c_{\alpha+2}(x-b)^{\alpha+2} &+ \cdots, \end{aligned}$$

so that

$$p + p' = z = kx^\mu,$$

and consequently

$$\mu p - x\,\frac{dp}{dx} = (\mu-\alpha)c_\alpha(x-b)^\alpha = X, \quad \mu p' - x\,\frac{dp'}{dx} = -X.$$

These differential equations have as their general solution

$$\int Xx^{-\mu-1}dx + k_1 = px^{-\mu} = c_\alpha(x-b)^{\alpha-\mu}x^{-\mu} + c_{\alpha-1}(x-b)^{\alpha-1}x^{-\mu} + \cdots,$$

$$\int Xx^{-\mu-1}dx + k_2 = p'x^{-\mu} = c_{\alpha+1}(x-b)^{\alpha+1}x^{-\mu} + c_{\alpha+2}(x-b)^{\alpha+2}x^{-\mu} + \cdots$$

If we substitute the value of X, and set $x = b/y$, we obtain

$$\begin{aligned} px^{-\mu} &= c_\alpha(\mu-\alpha)b^{\alpha-\mu}\int y^{\mu-\alpha-1}(1-y)^\alpha dy + k_1 \\ &= c_\alpha b^{\alpha-\mu}(1-y)^\alpha y^{\mu-\alpha} + c_{\alpha-1}b^{\alpha-1-\mu}(1-y)^{\alpha-1}y^{\mu-\alpha+1} + \cdots \end{aligned}$$

and

$$p'x^{-\mu} = -c_\alpha(\mu-\alpha)b^{-\alpha-\mu}\int y^{\mu-\alpha-1}(1-y)^\alpha dy + k_2$$
$$= c_{\alpha+1}b^{\alpha+1-\mu}(1-y)^{\alpha+1}y^{\mu-\alpha-1} + c_{\alpha+2}b^{\alpha+2-\mu}(1-y)^{\alpha+2}y^{\mu-\alpha-2} + \cdots$$

Now in the case where $\mu > \alpha > -1$, the right sides of these equations vanish for $y = 0, 1$ respectively. Thus the two integrals, the first between 0 and y, and the second between 1 and y, have the same value provided they are continuous between these limits. It might seem that this condition would be violated as soon as some or all of the terms of the series were to increase or decrease beyond all bounds; but this would not prevent us from finding a definite value for the series by actual addition, since the terms could cancel each other. As we shall not draw the conclusion that, in such cases, the series has no definite value, we can only decide the continuity or discontinuity of the series $px^{-\mu}$ and $p'x^{-\mu}$ by considering the corresponding integrals.[1] It is known, however, that an expression can only become discontinuous when its derivative becomes infinite; the expression $(1-y)^{\mu-\alpha-1}y^\alpha$ has a finite value for all finite values of y when the exponents $\mu-\alpha-1$ and α are positive. Hence the integrals vary continuously, and by considering the singular integrals for $y = 1$ and $y = 0$ we see that this is true as long as both exponents remain greater than -1.

Hence, in the case where $\mu > \alpha > -1$ and y is finite[2],

$$k = zx^{-\mu} = px^{-\mu} + p'x^{-\mu} = (\mu-\alpha)c_\alpha b^{\alpha-\mu}\int_0^1 (1-y)^{\mu-\alpha-1}y^\alpha dy$$
$$= c_\alpha b^{\alpha-\mu}\frac{\Pi(\alpha)\Pi(\mu-\alpha)}{\Pi(\mu)}$$

(where Π denotes the well-known definite integral). This result holds, as noted above, only when $\mu > \alpha > -1$, but it can be extended to all values

[1] If one deals with the integrals before the substitution of x by b/y, then they are discontinuous for $x = 0$. However, it is easily seen that in this form the associated constants of integration must have the same value for positive and negative values of x, because the value of the integrals changes in a continuous fashion as x passes from $+\infty$ to ∞.

[2] In the case where $y = \pm\infty$, and therefore $x = 0$, both integrals have the value ∞. Consequently $k = \infty - \infty$, that is, k is arbitrary, which is obvious from a direct consideration of this case.

of μ and α if (as we assume throughout) we extend the definition of Π to negative numbers by the rule $\Pi(n) = \frac{1}{n+1}\Pi(n+1)$. For firstly, by the law assumed to hold between the terms of the series, the result must be valid for every value of α for which one of the inequalities $\alpha < \mu$, $\alpha > -1$ holds. Thus, if $\mu > 0$,

$$kx^\mu = \sum_{\alpha=-\infty}^{\infty} k \frac{\Pi(\mu)}{\Pi(\alpha)\Pi(\mu-\alpha)} b^{\mu-\alpha}(x-b)^\alpha$$

or

$$\frac{x^\mu}{\Pi(\mu)} = \sum_{\alpha=-\infty}^{\infty} \frac{b^{\mu-\alpha}}{\Pi(\mu-\alpha)} \frac{(x-b)^\alpha}{\Pi(\alpha)}.$$

Differentiating n times, we obtain

$$\frac{x^{\mu-n}}{\Pi(\mu-n)} = \sum \frac{b^{\mu-\alpha}}{\Pi(\mu-\alpha)} \frac{(x-b)^{\alpha-n}}{\Pi(\alpha-n)}$$

which proves that the rule is also valid when μ is negative.

We have thus shown in general that

$$\frac{x^\mu}{\Pi(\mu)} = \sum_{\alpha=-\infty}^{\infty} \frac{b^{\mu-\alpha}}{\Pi(\mu-\alpha)} \frac{(x-b)^\alpha}{\Pi(\alpha)}. \tag{3}$$

It is worth noting that this formula does not yield a series for x^μ when μ is a negative integer, because the expression on the left side becomes zero. This is a point to which we return later. We also observe that there are series of this form which have the value zero, or are constant, for every value of x.

After this protest against the condemnation which has been pronounced on divergent series, we pursue the above method of defining the notion of derivation. The goal which we have set, namely that differentiation should be a particular case of derivation, will be achieved provided that the function k_ν has the value $\frac{1}{1.2\ldots\nu}$ for every positive integer value of ν, and 0 for every negative integer; for then the series (2) becomes the series (1). This condition can obviously be satisfied by infinitely many different functions of ν. Moreover, we cannot assume that there exists only one single expansion of such a function in powers of h; that is, that only one system of coefficients of a series of the specified form gives a definite value. Rather, we must assume that infinitely many distinct systems are possible. Without detriment to our objective, we must choose k_ν from among the different possible functions of ν, at the same time making a choice from the different possible systems of coefficients. Clearly it is best to make the choice so that the derivations obey several laws which otherwise would be valid only for integer indices.

The following treatment accords with this.

Since we wish the expression $\sum k_\nu \partial_x^\nu z h^\nu$ to include all possible developments of $z_{(x+h)}$ in this form, it follows that

$$\frac{d \sum k_\nu \partial_x^\nu z h^\nu}{dh} = \sum k_\nu \nu \partial_x^\nu \, z h^{\nu-1}$$

must include all possible developments of $\frac{dz_{(x+h)}}{dh}$ in this form. Similarly,

$$\frac{d \sum k_\nu \partial_x^\nu z h^\nu}{dx} = \sum k_\nu \, \frac{d\partial_x^\nu z}{dx} \, h^\nu$$

must include all possible developments of $\frac{dz_{(x+h)}}{dx}$. Now $\frac{dz_{(x+h)}}{dh}$ and $\frac{dz_{(x+h)}}{dx}$ are known to be identical; therefore both expressions include exactly the same set of series. Consequently $k_{\nu+1}(\nu+1)\partial_x^{\nu+1} z$ and $k_\nu \frac{d\partial_x^\nu z}{dx}$ have exactly the same values, or in other words are equal. If we set $k_{\nu+1}(\nu+1) = k_\nu$, which clearly does not contradict the basic assumption, because this relation must hold for integer values of ν, we obtain, even for derivation with fractional indices,

$$\partial_x^{\nu+1} z = \frac{d\partial_x^\nu z}{dx}.$$

Consequently, for all natural numbers n,

$$\partial_x^{\nu+n} z = \frac{d^n \partial_x^\nu z}{dx^n}. \tag{4}$$

From the law which we have assumed for k_ν, it follows that

$$\Pi(\nu) k_\nu = \Pi(\nu+1) k_{\nu+1}.$$

Hence the function $\Pi(\nu)k_\nu$, which we denote by ℓ_ν, has the same value for all arguments ν differing from each other by integers. We would thus make the most suitable choice of the function ℓ_ν, not from the consideration of a single form of development, but from a combination of distinct conclusions. Accordingly we investigate whether we can make our choice so that

$$\partial_x^{\nu}\partial_x^{\mu} z = \partial_x^{\nu+\mu} z.$$

To this end, let the variable x in the formula (2) increase again, and denote the increment by k, so that

$$(\alpha) \qquad z_{(x+h+k)} = \sum_{\mu=-\infty}^{\infty}\sum_{\nu=-\infty}^{\infty} \ell_\mu \ell_\nu \partial_x^{\mu}\partial_x^{\nu} z \frac{k^\mu}{\Pi(\mu)}\frac{h^\nu}{\Pi(\nu)},$$

where the expression represents all possible developments of $z_{(x+h+k)}$ in powers of h and k. But then

$$(\beta) \qquad z_{(x+h+k)} = \sum_{\mu+\nu=-\infty}^{\infty} \ell_{\mu+\nu}\partial_x^{\mu+\nu} z \frac{(h+k)^{\mu+\nu}}{\Pi(\mu+\nu)} = \sum_{\mu=-\infty}^{\infty}\sum_{\nu=-\infty}^{\infty} \ell_{\mu+\nu}\partial_x^{\mu+\nu} z \frac{h^\nu k^\mu}{\Pi(\nu)\Pi(\mu)},$$

by virtue of (3).

Now this last expression (β) does not represent all possible representations of $z_{(x+h+k)}$ of this form, because the equation (3) gives only one expansion of

$$\frac{(h+k)^{\mu+\nu}}{\Pi(\mu+\nu)}$$

and this is not necessarily the only possible one. Nevertheless all the possible developments it includes are included in (α) . If we stipulate that the function ℓ satisfies

$$\ell_{\mu+\nu} = \ell_\mu \ell_\nu$$

then all values of $\partial_x^{\mu+\nu}$ will also be values of $\partial_x^{\mu}\partial_x^{\nu} z$, even though the latter expression may also have other values.

Hence

$$(5) \qquad \partial_x^{\mu}\partial_x^{\nu} z = \partial_x^{\mu+\nu} z$$

under the restriction stated.

It follows, however, from $\ell_{\mu+\nu} = \ell_\mu \ell_\nu$ that

$$\ell_{\mu+\nu+\pi} = \ell_{\mu+\nu}\ell_\pi = \ell_\mu \ell_\nu \ell_\pi$$

and more generally that the product of any ℓ-numbers is equal to the ℓ-number of their sum. Taking the individual factors equal,

$$\ell_{m\nu} = \ell_\nu^m$$

for a natural number m. If we write $\frac{m\nu}{n} = \pi$, then

$$\ell_{m\nu} = \ell_{n\pi} = \ell_\nu^m = \ell_\pi^n$$

or

$$\ell_{\frac{m\nu}{n}} = \ell_\nu^{m/n}.$$

Thus the rule $\ell_{\mu\nu} = \ell_\nu^\mu$ is valid for all rational values of μ and consequently—by the known law for interpolation—valid in general. Now since for integer ν, $\ell_\nu = 1$, it follows that $\ell_\nu = 1^\nu$.

If therefore the laws (4) and (5) are to hold for derivations in general, and differentiation is to be included as a particular case of derivation, we must choose the derivations among those functions of x which satisfy the equation

$$z_{(x+h)} = \sum \frac{1^\nu h^\nu}{\Pi(\nu)} \partial_x^\nu z = \sum \frac{h^\nu}{\Pi(\nu)} \partial_x^\nu z.$$

The most convenient choice is one that is most flexible for calculations. If we examine the expansions of a few functions of $x+h$ in fractional powers of h, we see that the simplest and easiest expansions in series of this type are those in which the coefficient of

$$\frac{h^{\nu+1}}{\Pi(\nu+1)}$$

is the derivative of the coefficient of

$$\frac{h^\nu}{\Pi(\nu)}.$$

Thus we restrict the operation of derivation, assuming that the symbol $\partial_x^\nu z$ does not represent the coefficient of $\frac{h^\nu}{\Pi(\nu)}$ in every possible expansion of $z_{(x+h)}$, but only those for which the coefficient of $\frac{h^{\nu+1}}{\Pi(\nu+1)}$ is the derivative of the coefficient[3] of $\frac{h^\nu}{\Pi(\nu)}$.

[3]It follows, in fact, from (4), that if $\sum \partial_x^\nu z \frac{h^\nu}{\Pi(\nu)}$ is an expansion of $z_{(x+h)}$, then so is $\sum \frac{d\partial_x^\nu z}{dx} \frac{h^{\nu+1}}{\Pi(\nu+1)}$, but not that these two expansions are identical. By making the assumption described above, we also achieve the result that derivations with a negative integer index, to which no meaning has so far been ascribed, coincide with integrals, as we prove later.

It follows immediately that to one value of $\partial_x^\nu z$ belongs only one expansion; for if any particular value, say p_ν, of $\partial_x^\nu z$ had two expansions, a and b, then these two expansions would coincide in all their subsequent terms, because these would be derived from p_ν through differentiation. If we denote by $p_{\nu-1}, p_{\nu-2}, \ldots$ the preceding terms in the expansion a, and by $q_{\nu-1}, q_{\nu-2}, \ldots$ those in b, then $p_{\nu-1}$ and $q_{\nu-1}$ must each have derivative p_ν. Hence they differ by a constant, that is,

$$q_{\nu-1} = p_{\nu-1} + K_1,$$

and similarly

$$q_{\nu-2} = p_{\nu-2} + K_1 x + K_2, \quad q_{\nu-3} = p_{\nu-3} + K_1 \frac{x^2}{\Pi(2)} + K_2 x + K_3.$$

The expansion b is thus

$$a + \sum_{m=1}^{\infty} K_m \sum_{n=0}^{\infty} \frac{x^n}{\Pi(n)} \frac{h^{\nu-n-m}}{\Pi(\nu-n-m)} = a + \sum_{m=1}^{\infty} K_m \frac{(x+h)^{\nu-m}}{\Pi(\nu-m)}.$$

However, for all values of $x+h$ we require $a = b$, and as is well-known this implies that all the constants are 0. Hence the two expansions are identical.

If p_ν is a value of $\partial_x^\nu z$, then so is $p_\nu + \frac{Kx^{-\nu-n}}{\Pi(-\nu-n)}$ (where n is a positive integer and K a finite constant). For

$$\sum \left(p_\nu + \frac{Kx^{-\nu-n}}{\Pi(-\nu-n)}\right) \frac{h^\nu}{\Pi(\nu)} = \sum p_\nu \frac{h^\nu}{\Pi(\nu)} + \frac{K(x+h)^{-n}}{\Pi(-n)}$$
$$= \sum p_\nu \frac{h^\nu}{\Pi(\nu)} = z_{(x+h)},$$

and the following law holds:

$$\frac{d}{dx}\left(p_\nu + K \frac{x^{-\nu-n}}{\Pi(\nu-n)}\right) = p_{\nu+1} + K \frac{x^{-\nu-n-1}}{\Pi(-\nu-n-1)}.$$

We shall call a set comprising the different values of $\partial_x^\nu z$, obtained from one another by addition of expressions of the form $\frac{Kx^{-\nu-n}}{\Pi(-\nu-n)}$, a system of values. Thus all the values of $\partial_x^\nu z$ belonging to the same system are contained in the expression

(6) $$p_\nu + \sum_{n=1}^{\infty} K_n \frac{x^{-\nu-n}}{\Pi(-\nu-n)}$$

(where the K_n denote finite constants).

We now seek to determine one value of $\partial_x^\nu z$.

It is well-known that

$$z_{(x)} = z_{(k)} + \left(\frac{dz}{dx}\right)_{(k)} (x-k) + \left(\frac{d^2z}{dx^2}\right)_{(k)} \frac{(x-k)^2}{1.2} + \cdots$$

provided that $z_{(k)}$ is continuous between the limits x and k. If $x+h$ is substituted for x and the terms are expanded in powers of h using (3), then

$$z_{(x+h)} = \sum_{\mu=-\infty}^{\infty} \frac{h^\mu}{\Pi(\mu)} \left(z_{(k)} \frac{(x-k)^{-\mu}}{\Pi(-\mu)} + \left(\frac{dz}{dx}\right)_{(k)} \frac{(x-k)^{-\mu+1}}{\Pi(-\mu+1)} + \left(\frac{d^2z}{dx^2}\right)_{(k)} \frac{(x-k)^{-\mu+2}}{\Pi(-\mu+2)} + \cdots \right)$$

and in this series the coefficient of $h^\mu/\Pi(\mu)$ is the derivative of the coefficient of $h^{\mu-1}/\Pi(\mu-1)$. It is consequently a value of $\partial_x^\mu z$, which we denote by p_μ. Differentiation with respect to k yields

$$\frac{dp_\mu}{dk} = -z_{(k)} \frac{(x-k)^{-\mu-1}}{\Pi(-\mu-1)}; \text{ consequently } p_\mu = \int -z_{(k)} \frac{(x-k)^{-\mu-1}}{\Pi(-\mu-1)} dk.$$

Now all the terms of the above series vanish for $k=x$; and so the integral from k to x is equal to p_μ if it is continuous between these limits. This is obviously the case, because z is assumed to be continuous between x and k and $-\mu-1 > -1$. Thus

$$(7) \qquad \int_x^k -z_{(k)} \frac{(x-k)^{-\mu-1}}{\Pi(-\mu-1)} dk = \frac{1}{\Pi(-\mu-1)} \int_k^x (x-t)^{-\mu-1} z_{(t)} dt$$

is a value of $\partial_x^\mu z$ if z is continuous between x and k and μ is negative. The value of $\partial_x^{\mu-n} z$ associated to this same expansion is equal to

$$\frac{1}{\Pi(-\mu+n-1)} \int_k^x (x-t)^{-\mu+n-1} z_{(t)} dt.$$

It is easily seen that different expansions of $z_{(x+h)}$ can be derived from this one by giving different values to k, but they all belong to the same system. For, from the value

$$\frac{1}{\Pi(-\mu-1)} \int_k^x (x-t)^{-\mu-1} z_{(t)} dt$$

we obtain

$$\frac{1}{\Pi(-\mu-1)}\int_{k_1}^{x}(x-t)^{-\mu-1}z_{(t)}dt$$

by the addition of

$$\frac{1}{\Pi(-\mu-1)}\int_{k_1}^{k}(x-t)^{-\mu-1}z_{(t)}dt=\sum_{n=0}^{\infty}\frac{x^{-\mu-1-n}}{\Pi(-\mu-1-n)}\int_{k_1}^{k}\frac{(-t)^n}{\Pi(n)}z_{(t)}dt.$$

Because z is continuous between x and k_1 and thus between k and k_1, all the integrals are finite and moreover constant with respect to x. The process accordingly always leads to the same system of values. Thus, if we restrict the concept of derivations to this system of values, we reduce their determination to known values. With the help of this definition, we can find their properties and determine the derivations of given functions.

Accordingly, we have

$$1. \qquad \partial_x^\nu z=\frac{1}{\Pi(-\nu-1)}\int_{k}^{x}(x-t)^{-\nu-1}z_{(t)}dt+\sum_{n=1}^{\infty}K_n\frac{x^{-\nu-n}}{\Pi(-n-\nu)}$$

where the K_n are arbitrary constants[4], ν is negative, and z is continuous between x and k. For any $\nu \geq 0$, $\partial_x^\nu z$ denotes the expression[5] which is obtained by differentiating m times with respect to x (where $m > \nu$) the expression for $\partial_x^{\nu-m}$. This value always satisfies the equation[6]

$$2. \quad z_{(x+h)}=\sum_{n=1}^{\infty}\frac{h^{\nu-n}}{\Pi(\nu-n)}\int^{(n)}\partial_x^\nu z\,dx^n+\frac{h^\nu}{\Pi(\nu)}\partial_x^\nu z+\sum_{n=1}^{\infty}\frac{h^{\nu+n}}{\Pi(\nu+n)}\frac{d^n\partial_x^\nu z}{dx^n}.$$

[4]We denote these arbitrary functions by ϕ_ν; we note that for positive integer n, each function ϕ_ν is also a function $\phi_{\nu-n}$.

[5]The definition

$$\partial_x^\nu z=\sum_{n=0}^{\infty}\left(\frac{d^n z_{(x)}}{dx^n}\right)_k\frac{(x-k)^{n-\nu}}{\Pi(n-\nu)}+\phi_\nu$$

which is identical with that given, would certainly be valid for all values of ν. We have chosen the above because of its greater flexibility.

[6]The question of whether formula 1 includes all the values which can satisfy the equation clearly depends on whether the functions ϕ_ν are the only ones which, when substituted for $\partial_x^\nu z$, make the series 2 vanish. Now it can be shown without difficulty that there are no algebraic functions not contained in the ϕ_ν that do this; but I have not so far been able to decide whether there is no such function at all.

Hence

3. $$\partial_x^{-m} z = \int_k^{x(m)} z_{(t)} dt^m + \sum_{n=1}^{m} K_n \frac{x^{-n+m}}{\Pi(-n+m)},$$

4. $$\partial_x^0 z = z,$$

5. $$\partial_x^m z = \frac{d^m z}{dx^m},$$

and furthermore

6. $$\partial_x^\mu \partial_x^\nu z = \partial_x^{\mu+\nu} z + \phi_\mu.$$

Thus every value of $\partial_x^{\nu+\mu} z$ is also a value of $\partial_x^\mu \partial_x^\nu z$.

The converse, however, is true only when μ is a positive integer or ν is a negative integer. In this case the two expressions are identical. It also follows from the definition (c denoting a constant) that

7. $$\partial_x^\nu (p+q) = \partial_x^\nu p + \partial_x^\nu q,$$

8. $$\partial_x^\nu (cp) = c \partial_x^\nu p,$$

9. $$\partial_{x+c}^\nu z = \partial_x^\nu z,$$

10. $$\partial_{cx}^\nu z = \partial_x^\nu z \, c^{-\nu}.$$

Two values of $\partial_x^\nu z$ and $\partial_x^\mu z$, with the same constants $K, K_1, \ldots$ are said to be corresponding values. All the values associated to the same expansion of $z_{(x+h)}$ are corresponding values.

We now turn to the determination of the derivations of some specific functions of x. This naturally amounts to finding one value of one derivation. From this value, the general value can immediately be deduced by adding a function ϕ. Generally, after using the transformation 1 of the expression, we obtain a simpler expression, that is an explicit function of x in finite form. Generally, the transformation amounts to taking x outside the integral sign.

Consider first of all the function x^μ.

If μ is positive, x^μ is continuous for all values of x. Hence

$$\frac{1}{\Pi(-\nu-1)} \int_0^x (x-t)^{-\nu-1} t^\mu dt$$

is always a value of $\partial_x^\nu(x^\mu)$. This integral, however, is equal to

$$\frac{1}{\Pi(-\nu-1)}\int_0^1 x^{\mu-\nu}(1-y)^{-\nu-1}y^\mu dy = \frac{\Pi(\mu)}{\Pi(\mu-\nu)}\,x^{\mu-\nu}.$$

Since the mth derivative of this last expression is

$$\frac{\Pi(\mu)}{\Pi(\mu-\nu-m)}\,x^{\mu-\nu-m} = \partial_x^{\nu+m}(x^\mu)$$

by (4), it follows that, for every value of ν,

$$\partial_x^\nu(x^\mu) = \frac{\Pi(\mu)}{\Pi(\mu-\nu)}\,x^{\mu-\nu} + \phi_\nu.$$

If μ is negative, x^μ is discontinuous for $x = 0$, but continuous for all other values of x; and so x and k must always have the same sign in the expression 1. Now integration by parts m times yields

$$\frac{1}{\Pi(-\nu-1)}\int_k^x (x-t)^{-\nu-1}t^\mu dt$$
$$= \frac{\Pi(\mu)}{\Pi(-\nu-1-m)\Pi(\mu+m)}\int_k^x (x-t)^{-\nu-1-m}t^{\mu+m}dt + \phi_\nu$$

as long as $-\nu-m > 0$. In this way, when $-\nu > -\mu$, those integrals where $\mu < -1$ can be reduced to integrals in which the exponent of t is ≥ -1. If the exponent is > -1, then

$$\int_0^k (x-t)^{-\nu-1-m}t^{\mu+m}dt$$

belongs to the set of functions ϕ_ν. Hence

$$\frac{\Pi(\mu)}{\Pi(-\nu-1-m)\Pi(\mu+m)}\int_0^x (x-t)^{-\nu-1-m}t^{\mu+m}dt = \frac{\Pi(\mu)}{\Pi(\mu-\nu)}\,x^{\mu-\nu}$$

is a value of $\partial_x^\nu(x^\mu)$, when $-\nu > -\mu$, and this result must be valid for all ν by the law

$$\partial_x^{\nu+1}z = \frac{d\partial_x^\nu z}{dx}.$$

If, however, $\mu + m = -1$, then

$$\begin{aligned}\int_k^x (x-t)^{-\nu-1-m} t^{\mu+m} dt \\ &= (\log x)x^{\mu-\nu} - (\log k)x^{\mu-\nu} + \int_k^x \frac{(x-t)^{\mu-\nu} - x^{\mu-\nu}}{t}\, dt \\ &= (\log x)x^{\mu-\nu} + \int_0^x \frac{(x-t)^{\mu-\nu} - x^{\mu-\nu}}{t}\, dt + \phi_\nu \\ &= (\log x)x^{\mu-\nu} + x^{\mu-\nu}\int_0^1 \frac{y^{\mu-\nu} - y}{1-y}\, dt \\ &= (\log x)x^{\mu-\nu} - (\Psi(\mu-\nu) - \Psi(0))x^{\mu-\nu}.\end{aligned}$$

If we generalize this result by differentiation, we obtain the following value for $\partial_x^\nu(x^\mu)$:

11. $$\partial_x^\nu(x^\mu) = \frac{\Pi(\mu)}{\Pi(\mu-\nu)}\, x^{\mu-\nu}$$

when μ is not a negative integer;

12. $$\partial_x^\nu(x^\mu) = \frac{\Pi(\mu)}{\Pi(-1)\Pi(\mu-\nu)} \left[(\log x)x^{\mu-\nu} - (\Psi(\mu-\nu) - \Psi(0))x^{\mu-\nu}\right]$$

when μ is a negative integer.

It should be noted that the formula 11 follows from formula 12, provided that those constants which become expressions of the form ∞/∞ are treated appropriately, which will also be required when $\mu-\nu$ and μ are both negative integers. It is easy to see that the values given by these formulae for different values of ν are corresponding values. This is also the reason why, in 12, we are unable to absorb the term containing $x^{\mu-\nu}$ in the function ϕ_ν, as in the case where μ is a negative integer.

If we apply a similar process to the function e^x, we obtain

13.
$$\partial_x^\nu(e^x) = \frac{1}{\Pi(-\nu-1)}\int_{-\infty}^x e^t(x-t)^{-\nu-1}dt = \frac{1}{\Pi(-\nu-1)}\, e^x \int_0^\infty e^{-y}y^{-\nu-1}dy = e^x.$$

The derivations of $\log x$ can be found by the same method, or more easily and even for all values of ν, from 6 and 12:

14. $$\partial_x^\nu(\log x) = \partial_x^\nu \partial_x^{-1} x^{-1} = \frac{1}{\Pi(-\nu)}\left((\log x)x^{-\nu} - [\Psi(-\nu) - \Psi(0)]x^{-\nu}\right).$$

By applying rules 7 to 10, the derivatives of $\sin x$, $\cos x$, $\tan x$ and arc tan x can be deduced with the greatest of ease from 13 and 14.

Finally we remark that the theory set out above can, with equal trustworthiness, be extended to the case where the quantities concerned are assigned imaginary values.

XX.

New theory of residual charge in apparatus for static charge.

1.

Preliminary remarks.

Professor Kohlrausch has succeeded in subjecting the formation of residual charge in apparatus for static charge to precise measurements. This was the basis of a theory that accords with these observations published in *Poggendorff's Annalen*[1]. The exactitude of these measurements stimulated me to test against them a law for the motion of electricity, that is plausible on other grounds. In the form of the law given for this purpose, it applies to the motion of electricity in all ponderable bodies, but only under the hypothesis that the bodies considered are relatively at rest, and that no noticeable thermal and magnetic (or electrically induced) effects and influences occur. For the purpose of unrestricted applicability, some reworking and supplementation are required, which I will take up elsewhere.

In the following article, drawn from a written communication to Professor Kohlrausch, this new theory of residual change is not self-contained, but developed in conjunction with his theory. I did not attempt to reduce phenomena directly to that theory. Accordingly I have used the concepts used by Professor Kohlrausch in his paper (electrical moment of an isolated wall, tension, total charge, disposable charge, residual charge) to express the underlying ideas here. I have further taken into consideration his method of treatment in several respects.

2.

The law on which the calculation is based.

Let t denote time, x, y, z rectangular coordinates, ρ the density of electric charge at time t at the point (x, y, z). Let u be $(4\pi)^{-1}$ times the (Gaussian) potential of all electrical quantities acting at (x, y, z) at time t,

$$u = \frac{1}{4\pi}\int \frac{\rho'\, dx'\, dy'\, dz'}{\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}},$$

[1] 1. vol. 91, p. 56.

where $\rho' dx' dy' dz'$ denotes the electric charge of the element $dx' dy' dz'$ at time t. Then

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = -\rho.$$

The laws to be applied here for the motion of electricity interior to a homogeneous ponderable body, with the conditions under discussion, are as follows.

I. The electromotive force at point (x, y, z) at time t is made up of two terms: one conforming with Coulomb's law, with components proportional to

$$-\frac{\partial u}{\partial x}, \ -\frac{\partial u}{\partial y}, \ -\frac{\partial u}{\partial z},$$

and another with components proportional to

$$-\frac{\partial \rho}{\partial x}, \ -\frac{\partial \rho}{\partial y}, \ -\frac{\partial \rho}{\partial z}.$$

Together the components become

$$-\frac{\partial u}{\partial x} - \beta^2 \frac{\partial \rho}{\partial x}, \ -\frac{\partial u}{\partial y} - \beta^2 \frac{\partial \rho}{\partial y}, \ -\frac{\partial u}{\partial z} - \beta^2 \frac{\partial \rho}{\partial z}.$$

Here β depends only on the nature of the ponderable body.

II. The current intensity is proportional to the electromotive force. Thus

$$-\frac{\partial u}{\partial x} - \beta^2 \frac{\partial \rho}{\partial x} = \alpha\xi, \ -\frac{\partial u}{\partial y} - \beta^2 \frac{\partial \rho}{\partial y} = \alpha\eta, \ -\frac{\partial u}{\partial z} - \beta^2 \frac{\partial \rho}{\partial z} = \alpha\zeta.$$

Here α is a constant depending on the nature of the ponderable body, and ξ, η, ζ are the components of the current intensity.

With the inclusion of the kinematic law

$$\frac{\partial \rho}{\partial t} + \frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial y} + \frac{\partial \zeta}{\partial z} = 0,$$

we obtain the equations

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = -\rho,$$

$$\alpha \frac{\partial \rho}{\partial t} + \rho - \beta^2 \left(\frac{\partial^2 \rho}{\partial x^2} + \frac{\partial^2 \rho}{\partial y^2} + \frac{\partial^2 \rho}{\partial z^2} \right) = 0$$

for[2] u.

If we take the length β and the time α as unity, we have

$$\frac{\partial \rho}{\partial t} + \rho - \left(\frac{\partial^2 \rho}{\partial x^2} + \frac{\partial^2 \rho}{\partial y^2} + \frac{\partial^2 \rho}{\partial z^2}\right) = 0.$$

This gives a partial differential equation for u of first degree with respect to t, and of fourth degree with respect to the space coordinates. The determination of u interior to the ponderable body at a given time requires, in addition to this equation, one condition at each point of the body at the start time, and two conditions at each point of the surface at all subsequent times.

3.

Plausible interpretation of the above law.

In the previous section the law of motion of electricity is expressed in concepts that are customary in our present theory of electricity. However, this interpretation admits a reworking, apparently leading to a somewhat more accurate and complete picture of the actual context.

---

[2]Accordingly the equations for equilibrium (in an electronically isolated conductor) are

$$-\frac{\partial u}{\partial x} - \beta^2 \frac{\partial \rho}{\partial x} = 0, \; -\frac{\partial u}{\partial y} - \beta^2 \frac{\partial \rho}{\partial y} = 0, \; -\frac{\partial u}{\partial z} - \beta^2 \frac{\partial \rho}{\partial z} = 0,$$

or

$$u - \beta^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) = \text{const.}$$

For current equilibrium or moving equilibrium in a constant closed circuit, the equations are

$$\frac{\partial \rho}{\partial t} = 0$$

or

$$\rho - \beta^2 \left(\frac{\partial^2 \rho}{\partial x^2} + \frac{\partial^2 \rho}{\partial y^2} + \frac{\partial^2 \rho}{\partial z^2}\right) = 0.$$

If the length β is very small compared to the dimensions of the body, then $u -$ const. (in the first case) and ρ (in the second) decreases very rapidly away from the surface and is very small throughout the interior. Indeed, these quantities change with distance p from the surface approximately of order $e^{-p/\beta}$, as long as the radius of curvature remains very large compared to β. This must be assumed to be the case for metallic conductors.

Instead of assuming a natural cause that strives to move positive electricity at the point (x, y, z) in the direction of the three axes with force

$$-\beta^2 \frac{\partial \rho}{\partial x}, \ -\beta^2 \frac{\partial \rho}{\partial y}, \ -\beta^2 \frac{\partial \rho}{\partial z}$$

and negative electricity in the opposite sense, we may assume a natural cause that strives to decrease positive electrical charge, and increase negative electrical charge, with intensity $\beta^2 \rho$ at (x, y, z). One can regard this cause as a resistance of ponderable bodies to the holding of static electrical charge, or resistance to an electrified state.

Likewise, the elecromotive force with components

$$-\frac{\partial u}{\partial x}, \ -\frac{\partial u}{\partial y}, \ -\frac{\partial u}{\partial z}$$

can be replaced by a cause, having intensity u at the point (x, y, z), which strives to reduce the density of electricity having the same sign, and to increase density in the case of opposite signs.

However, we need not assume the existence of two different kinds of electrical charge, and treat $\rho \, dx \, dy \, dz$ as the excess of positive over negative charge for the element $dx \, dy \, dz$, in order to give the quantity ρ physical significance. Rather one can essentially go back to Franklin's interpretation of electrical phenomena, perhaps in the simplest form via the following hypothesis.

The ponderable body that is the location of the electrical charge fills space continuously and with uniform electrical capacity inversely proportional to its resistance, which the density of the actual electrical charge contained in it steadily approaches indefinitely closely. With an excess or a defect of electricity (positive or negative electrical charge) the ponderable body takes on a positive or negative electrical state according to which it strives to increase or decrease the density of the electricity it contains, and does so with a force equal to the density of electrical charge ρ multiplied by a factor depending on the nature of the body (its antielectric force). For its own part the electricity takes on a state, tension, with the appearance of electrical charge, according to which it strives to decrease its density (or for negative tension, to increase it). This tension is measured by a quantity u depending on all elements of electrical charge according to the formula

$$u = \frac{1}{4\pi} \int \frac{\rho' dx' dy' dz'}{\sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}},$$

or also via the law

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = -\rho,$$

together with the condition that u remains infinitely small at infinite distance from electrical charge. Electricity moves relative to the ponderable body with a velocity which is equal, at a given instant, to the electromotive force arising from this cause.

However, this law of motion of electricity, if we take account of its relation to heat and magnetism in the calculation, would have to be modified and recast, and a modified interpretation of these phenomena would be needed.

4.

Treatment of the problem of formation of residual charge. Expression of the quantities to be determined in terms of the potential.

I now turn to the study of the formation of residual charge. I concern myself first with expressing the quantities to be determined in terms of the potential or rather, in order to simplify the calculation, in terms of the function u proportional to the potential. For greater convenience of physicists less accustomed to the abstract approach, I have treated the potential as the measure u of a force, the tension, which at the point (x, y, z) strives to decrease the density of electricity. Thus at (x, y, z), the components of the electromotive force arising from the potential are

$$-\frac{\partial u}{\partial x}, \ -\frac{\partial u}{\partial y}, \ -\frac{\partial u}{\partial z}.$$

We must take as unit of tension that which occurs interior to a sphere of radius 1 due to the distribution of electricity on its surface with density 1; or, as the unit of electrical force, the force produced by the quantity 4π of electricity at unit distance. To simplify the calculation, we further introduce α as unit time, β as unit length. If we let the unit of electromotive force depend on the unit of electricity in the fashion assumed here, then α and β^2 are measures of electrical resistance $\left(\frac{\text{electromotive force}}{\text{current intensity}}\right)$ and contraelectric force $\left(\frac{\text{force of the body}}{\text{density of electric charge}}\right)$ of the ponderable location.

As a discussion of the above observations, the solution of the following problem will suffice. We determine the variations of electrical charge in the interior of a uniformly thick homogeneous wall when the surfaces, covered by perfect conductors, take up equal quantities of electricity of opposite signs and possess no electromotive force (no contact effects are found in them). We further suppose that the dimensions of the surfaces may be treated as infinitely large compared to the thickness of the wall (that is, the influence of the boundary and the curvature are negligible).

We take the origin of coordinates in the center of the wall, the x-axis perpendicular to its surfaces, and denote its half thickness by a. The expression for the wall is $a > x > -a$; u is a function of x only, and

$$\rho = -\frac{\partial^2 u}{\partial x^2}.$$

Hence

$$\int_{x'}^{x''} \rho\, dx = \left(\frac{\partial u}{\partial x}\right)_{x'} - \left(\frac{\partial u}{\partial x}\right)_{x''}.$$

The quantity of electricity per unit area contained between two values of x, expressed geometrically, is equal to the difference between the slopes of the tangents to the tension curve, that is, the curve whose ordinate is u for abscissa x. If no static charge exists, this curve is piecewise linear. It is continuous, convex from above (that is, for locations with greater ordinates) where the static charge is positive, and convex from below where negative. It changes direction for values of x comprising a finite set.

Thus the tension produced by charging, or destroyed by a discharge, will always be represented by a curve of form A (see Figure 1). That is, if its value at the plates is u_a, u_{-a}, and therefore in the center is

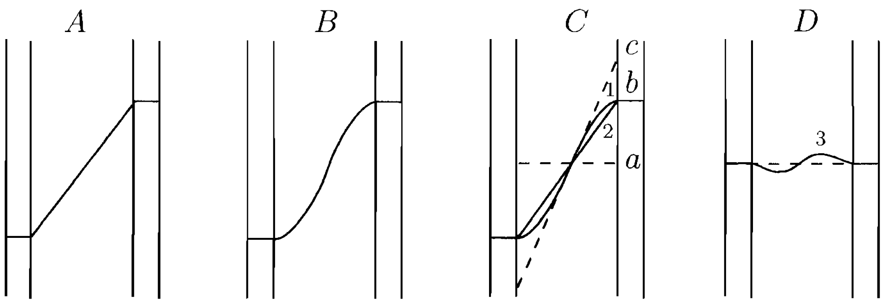
$$\frac{u_a + u_{-a}}{2} = u_0,$$

then in the interior the value is

$$u_0 + \frac{x}{a}\,(u_a - u_0).$$

With the penetration of electricity into the interior, the tension curve takes the form B. The combination of opposite electricities per unit area is equal to the slopes of the tangent to B in the center,

$$\left(\frac{\partial u}{\partial x}\right)_0.$$



1) Tension curve of the total charge.

2) Tension curve of the disposable charge.

3) Tension curve of the residual charge.

Total charge = ac, disposable charge = ab, residual charge = bc.

Figure 1: tension curves.

The electrical moment is

$$\int_{-a}^{a} \rho\, x\, dx = u_a - u_{-a} - a\left(\left(\frac{\partial u}{\partial x}\right)_a + \left(\frac{\partial u}{\partial x}\right)_{-a}\right) = u_a - u_{-a},$$

that is, the tension difference between the surfaces.

The tension at the plates is neutralized by a discharge. The tension destroyed is thus u_a, u_{-a} at the plates, and

$$u_0 + \frac{x}{a}\,(u_a - u_0)$$

in the interior. The disposable charge is

$$\frac{1}{a}\,(u_a - u_0)$$

per unit area. The tension remaining in the interior is

$$u - u_0 - \frac{x}{a}\,(u_a - u_0)$$

and the latent residual charge per unit area is

$$\left(\frac{\partial u}{\partial x}\right)_0 - \frac{1}{a}\,(u_a - u_0).$$

The quantity of electricity passed on at the surface $x = a$ with the discharge is

$$-\frac{1}{a}\,(u_a - u_0).$$

5.

Solution of the problem in the simplest case, where no outflow and inflow occurs through the surfaces.

After this outline, and geometrical representation of the quantities to be found, I pass to their calculation according to the stated laws. First I treat the case where initially no free electricity exists in the interior, and a unit quantity per area passes through the surfaces, while afterwards there is no outflow or inflow through the surfaces.

The conditions for the determination of u are:

$$\begin{aligned}
&\frac{\partial^2 u}{\partial x^2} = -\rho,\ \frac{\partial \rho}{\partial t} + \rho - \frac{\partial^2 \rho}{\partial x^2} = 0 \quad \text{for } t > 0, a > x > -a;\\
&\frac{\partial u}{\partial x} = 1 \quad \text{for } t = 0, a > x > -a;\\
&\frac{\partial u}{\partial x} = 0,\ \frac{\partial u}{\partial x} + \frac{\partial \rho}{\partial x} = 0 \quad \text{for } t > 0, x = \pm a.
\end{aligned}$$

The last condition expresses the property that in the surfaces the quantity of electricity, as well as the flow, and hence the electromotive force, vanishes.

These conditions are satisfied by two expressions, one useful for small values, and the other for large values, of t.

For brevity, let

$$\int_\lambda^\infty e^{-\lambda^2} d\lambda = \phi(\lambda)$$

and

$$\int_\lambda^\infty \phi(\lambda) d\lambda = \frac{1}{2}\, e^{-\lambda^2} - \lambda\phi(\lambda) = \psi(\lambda).$$

Then the expressions in question are firstly

$$u - u_0 = e^{-t}\left[x + \frac{4\sqrt{t}}{\sqrt{\pi}} \sum_{n=1}^{\infty} \left(\psi\left(\frac{a(2n-1) - x}{2\sqrt{t}}\right) - \psi\left(\frac{a(2n-1) + x}{2\sqrt{t}}\right)\right)\right],$$

secondly

$$u - u_0 = e^{-t} \sum \frac{(-1)^{n-1} 2a}{\pi^2 \left(n - \frac{1}{2}\right)^2} e^{-\left(n-\frac{1}{2}\right)^2 \pi^2 t/a^2} \sin\left(n - \frac{1}{2}\right) \frac{x\pi}{a}.$$

The determinations that ensue are:

for the distribution of electricity,

$$\begin{aligned} \rho = -\frac{\partial^2 u}{\partial x^2} &= \frac{e^{-t}}{\sqrt{\pi t}} \sum (-1)^{n-1} \left(e^{-(a(2n-1)-x)^2/4t} - e^{-(a(2n-1)+x)^2/4t}\right) \\ &= \frac{2e^{-t}}{a} \sum (-1)^{n-1} e^{-(n-1/2)^2\pi^2 t/a^2} \sin(n-1/2)x\pi/a; \end{aligned}$$

for the total charge

$$\begin{aligned} Q_t^* = \left(\frac{\partial u}{\partial x}\right)_0 &= e^{-t}\left(1 + \frac{4}{\sqrt{\pi}} \sum (-1)^n \phi\left(\frac{\left(n-\frac{1}{2}\right)a}{\sqrt{t}}\right)\right) \\ &= e^{-t} \frac{(-1)^{n-1} 2}{\left(n - \frac{1}{2}\right)\pi} e^{-(n-1/2)^2\pi^2 t/a^2}; \end{aligned}$$

for the disposable charge,

$$\begin{aligned} L_t^* = \frac{u_a - u_{-a}}{2a} &= e^{-t}\left\{1 - \frac{2\sqrt{t}}{a\sqrt{\pi}}\left(1 + 4\sum (-1)^n \psi\left(\frac{an}{\sqrt{t}}\right)\right)\right\} \\ &= e^{-t} \sum \frac{2}{\pi^2 \left(n-\frac{1}{2}\right)^2} e^{-\left(n-\frac{1}{2}\right)^2 \pi^2 t/a^2}; \end{aligned}$$

for the residual charge,

$$\begin{aligned} r_t^* &= \left(\frac{\partial u}{\partial x}\right)_0 - \frac{u_a - u_{-a}}{2a} \\ &= \frac{2\sqrt{t}\, e^{-t}}{a\sqrt{\pi}}\left\{1 + 4\sum (-1)^n \left(\psi\left(\frac{an}{\sqrt{t}}\right) + \frac{a}{2\sqrt{t}}\phi\left(\frac{\left(n-\frac{1}{2}\right)a}{\sqrt{t}}\right)\right)\right\} \\ &= e^{-t} \sum \frac{2}{\pi\left(n-\frac{1}{2}\right)}\left((-1)^{n-1} - \frac{1}{\pi\left(n-\frac{1}{2}\right)}\right) e^{-(n-1/2)^2\pi^2 t/a^2}. \end{aligned}$$

6.

Reduction of the general problem to this simplest case.

We now reduce the general case, where outflow and inflow occur through the surfaces, to this simplest case. Let $\chi(t)$ be the expression for the tension difference $u - u_0$ at time t in the simplest case; for negative t, let $\chi(t) = 0$.

If we now determine the tension produced when the charge $\pm\mu$ passes through the surfaces $x = \pm a$ at time 0, then $\pm\mu'$ at time t', $\pm\mu''$ at time t'', ..., we have

$$u - u_0 = \mu\chi(t) + \mu'\chi(t - t') + \mu''\chi(t - t'') + \cdots$$

since this value satisfies all the conditions specified for its determination.

In the case of continuous outflow and inflow of electricity, we have

$$u - u_0 = \int_0^t \chi(t - \tau)\,\frac{d\mu}{d\tau}\,d\tau$$

where $\pm\frac{d\mu}{d\tau}$ denotes the flow of electricity into the interior through the surfaces $x = \pm a$ during the time element $d\tau$.

One can combine the two expressions into the expression

$$u - u_0 = \int_0^t \chi(t - \tau)d\mu$$

where $\pm d\mu$ denotes the electricity arriving at the surfaces $x = \pm a$ in the time element $d\tau$, where this is either a finite quantity, or proportional to $d\tau$, depending on whether a sudden charge or discharge, or a continuous outflow or inflow, occurs.

From these expressions for the tension, it follows that

$$Q_t = \int_0^t Q^*_{t-\tau}d\mu,\ L_t = \int_0^t L^*_{t-\tau}d\mu,\ r_t = \int_0^t r^*_{t-\tau}d\mu.$$

In these formula, the time units have duration α, the length is in units β. In order to introduce familiar units, we need only replace a and x by $\frac{a}{\beta}$, $\frac{x}{\beta}$; t and τ by $\frac{t}{\alpha}$, $\frac{\tau}{\alpha}$.

7.

Comparison of the calculation with observations.

We now compare the formula obtained with the actual course of residual charge formation, as established very precisely in Professor Kohlrausch's measurements published in Poggendorff's Annalen. Perhaps it is most convenient to proceed from the fact that the charge curve approximates a parabola with gradually decreasing parameter; that is, the quantity $\frac{L_0-L_t}{\sqrt{t}}$ gradually decreases.

By virtue of the formula derived for L_t, $L_0 - L_t$ is proportional to $\sqrt{t}$ for very small values of t, indeed,

$$\frac{L_0 - L_t}{\sqrt{t}} = L_0 \frac{2}{\sqrt{\pi}} \left(\frac{\beta^2}{a^2\alpha}\right)^{1/2}.$$

As a consequence of the measurements, one must assume that this proportionality still occurs approximately throughout the observations.

Accordingly we can give a rough approximation to the time $\frac{a^2}{\beta^2}\alpha$ from observations. In fact,

$$\frac{L_0^* - e^{t/\alpha}L_t^*}{\sqrt{t}} = L_0^* \frac{2}{\sqrt{\pi}} \left(\frac{\beta^2}{a^2\alpha}\right)^{1/12} \left(1 - 4\psi\left(\left(\frac{a^2\alpha}{\beta^2 t}\right)^{1/2}\right) + 4\psi\left(2\left(\frac{a^2\alpha}{\beta^2 t}\right)^{1/2}\right)\right.$$
$$\left. -4\psi\left(3\left(\frac{a^2\alpha}{\beta^2 t}\right)^{1/2}\right) + \cdots\right)$$

is then a function that decreases slowly with increasing t. Nevertheless, $\frac{L_0-L_t}{\sqrt{t}}$ would increase with t if one assigned a suitable value to $\frac{1}{\alpha}$. The same thing also appears to occur if one assumes a considerable loss through the air, at least if one takes Coulomb's law as a basis for this.

Hence for the first treatment of the observations, we take the time α (that is, the resistance of glass for the electromotive force conforming to the Coulomb law) to be infinitely large, neglect the loss through the air and at first restrict ourselves to the investigation of how far the corresponding determination of $a^2\alpha/\beta^2$ satisfies the observations.

Once we have convinced ourselves that the hypotheses of the calculation are approximately correct, a more precise comparison with observations is wasted labor if we do not have the opportunity to find the source of the differences between calculation and experimental observation in order to make the necessary corrections to the calculation due to departures from our assumptions. Since I lack the means for experimental study of the subject, I must forego its further pursuit for the present.

8.

Relation of this problem to electrometry and the theory of related phenomena.

The quantity $\frac{\beta^2}{a^2\alpha}$, approximately $\frac{1}{2000}$ for the bottle b, gives the quotient $\frac{\text{contraelectric force}}{\text{resistance}}$ of the glass of the bottle in absolute measure, if we take the thickness of the glass as unit of length and the second as unit of time. For this determination it does not matter how the unit of electromotive force is made to depend on the unit of electrical charge. The constants α and β^2 would, however, give the resistance and the contraelectric force in different units from those of Weber, where the unit of electromotive force is specified via the influence of a unit charge according to Ampère's law.

For comparison of the case investigated here with phenomena on good conductors, the treatment of the steady state with tension difference held constant at the surfaces (or, constant inflow) serves. For this, we have

the density in the interior: $\rho = -\dfrac{\partial^2 u}{\partial x^2} = e^x - e^{-x}$,

the tension: $u = u_0 - e^x + e^{-x} + x(e^a + e^{-a})$,

the difference in tension of the surfaces:

$$u_a - u_{-a} = 2(a(e^a + e^{-a}) - (e^a - e^{-a})),$$

the total charge: $\left(\dfrac{\partial u}{\partial x}\right)_0 = e^a + e^{-a} - 2$,

the residual charge: $\left(\dfrac{\partial u}{\partial x}\right)_0 - \dfrac{u_a - u_{-a}}{2a} = \dfrac{e^a - e^{-a}}{a} - 2$,

and the quantity of flow in unit time:

$$\frac{\partial u}{\partial x} + \frac{\partial \rho}{\partial x} = -(e^a + e^{-a})$$

or the proportionate quantities, where for simplicity we have as above taken α as unit of time, β as unit of length, and the unit of tension as the tension interior to a ball of radius 1 when electricity is distributed on the surface with density 1.

It seems to me particularly important to test the proposed law, or alternatively to determine the constants α and β, for gases. The observations of Riess[3] and Kohlrausch[4], according to which the electrical loss in air in a closed space does not obey Coulomb's law, can perhaps serve as a starting point for this investigation. For this purpose it would certainly be desirable to have a set of measurements on the electrical loss interior to a somewhat regular closed space.

[3] *Pogg. Ann.*, vol. 71, p. 359.
[4] *Pogg. Ann.*, vol. 72, p. 374.

XXI.

Two general theorems on linear differential equations with algebraic coefficients.

(20th February 1857.)

It is well-known that every solution of a linear homogeneous differential equation of order n can be expressed as a linear combination, with constant coefficients, of n independent particular solutions. If the coefficients of the differential equations are rational functions of the independent variable x, then each branch of the (generally) multi-valued functions which satisfy the equation can be expressed for each x as a linear combination of n single-valued functions, though admittedly these must be discontinuous along certain systems of lines. If, however, the coefficients are algebraic functions of x, which can be expressed rationally in terms of x and a μ-valued algebraic function of x, then there will be, for each branch of this μ-valued algebraic function, a group of n independent particular solutions. In this case, every branch of a solution of the differential equation can be expressed as a linear combination of at most μn single-valued functions, but of these only n can belong to the same group. After these preliminary remarks it will be clear that the following theorems apply to all linear differential equations with algebraic coefficients, because every non-homogeneous linear differential equation can easily be transformed into a homogeneous equation of the next higher order.

Let $y_1, y_2, \ldots, y_n$ be functions of x which for all complex values of this variable are single-valued and finite except at $a, b, c, \ldots, g$, which, when x encircles one of these branch points, become linear combinations with constant coefficients of their former values.

In order to determine these combinations more precisely, let us separate the complex domain into two regions by means of a closed line passing successively through all the branch-points $(g, \ldots, c, b, a)$ so that within each of these regions the individual functions vary continuously. Suppose that the values of the functions in the region on the positive side of the line have been given. A circuit of x around a now takes y_1 to $\sum_{i=1}^{n} A_i^{(1)} y_i$; y_2 to $\sum_{i=1}^{n} A_i^{(2)} y_i; \ldots$; y_n to $\sum_{i=1}^{n} A_i^{(n)} y_i$. A circuit around b takes y_ν to $\sum B_i^{(\nu)} y_i, \ldots$; a circuit around g takes y_ν to $\sum G_i^{(\nu)} y_i$.

For brevity, we write (y) for the system $(y_1, y_2, \ldots, y_n)$, (A) for the system of n^2 coefficients

$$\begin{matrix} A_1^{(1)} & A_2^{(1)} & \cdots & A_n^{(1)} \\ A_1^{(2)} & A_2^{(2)} & \cdots & A_n^{(2)} \\ \ldots & \ldots & \ldots & \\ A_1^{(n)} & A_2^{(n)} & \cdots & A_n^{(n)} \end{matrix}$$

and $(B), \ldots, (G)$ for the corresponding systems of B's, $\ldots$, G's. Write $(A)(y) = (A)(y_1, \ldots, y_n)$ for the system of values $\sum A_i^{(1)} y_i, \sum A_i^{(2)} y_i, \ldots, \sum A_i^{(n)} y_i$ formed from (y) via the system of coefficients (A). Thus the following equation holds between these coefficient systems:

$$(1) \qquad (G)(F) \ldots (B)(A) = (0).$$

Here (0) denotes a coefficient system which changes nothing, that is the coefficients in the diagonal from top left to bottom right are 1, and the others are 0. In fact, if x travels along the positive edge of the boundary between one branch-point and the next, and then encircles in a positive direction the branch-point itself, the functions (y) successively become $(G)(y)$, $(G)(F)(y), \ldots$, and finally $(G)(F) \ldots (B)(A)(y)$. The same result is achieved if x moves along the negative edge of the boundary, or runs over the whole boundary of the region on the negative side, which must take $(y_1, \ldots, y_n)$ into its former value since it is single-valued in this region.

A system of n functions having the above properties will be denoted by

$$Q \begin{pmatrix} a & b & c & \cdots & g & \\ A & B & C & \cdots & G & x \end{pmatrix}.$$

We shall now consider as belonging to the same class all systems whose branch-points, and corresponding linear substitutions satisfying the equation (1), take given values. As will soon become clear, there are infinitely many systems in a class. By an easily proved theorem often used by Jacobi, every linear substitution can, generally speaking, be decomposed into the product of three substitutions, of which the last is the inverse of the first, and the middle one has zero coefficients except on the diagonal. The effect of the diagonal substitution, applied to given variables, is simply to multiply each by a factor. For example one can write

$$(A) = (\alpha) \begin{Bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{Bmatrix} (\alpha)^{-1}$$

where $(\alpha)^{-1}$ is the inverse of (α).

The quantities λ are the n roots of an equation of degree n, determined by (A). In the case where this equation has equal roots, the middle substitution has to be given a modified form. For simplicity we exclude this special case for now, and assume that it does not arise in the decomposition of the substitutions $(A), (B), \dots, (G)$. The substitution (α) can be transformed into

$$(\alpha) \begin{Bmatrix} \ell_1 & 0 & \dots & 0 \\ 0 & \ell_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \ell_n \end{Bmatrix}$$

by post-multiplication by a diagonal substitution. In this form, all possible values are included, as is clear from the equations that determine the substitution.

When x makes a positive circuit around the branch-point a, the values of the functions y change from $(p_1, p_2, \dots, p_n)$ into $(A)(p)$. The values of the functions derived from (y) by the substitution $(\alpha)^{-1}$,

$$(z_1, z_2, \dots, z_n) = (\alpha)^{-1}(y),$$

change from $(\alpha)^{-1}(p)$ to

$$(\alpha)^{-1}(A)(p) = \begin{Bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{Bmatrix} (\alpha)^{-1}(p)$$

or, in other words, from $(z_1, z_2, \dots, z_n)$ to $(\lambda_1 z_1, \lambda_2 z_2, \dots, \lambda_n z_n)$.

If a function z is multiplied by a constant factor λ when x makes a positive circuit around a branch point a, then z can be transformed through multiplication by a power of $x - a$ into a function of $x - a$ that is single-valued in the neighborhood of a. In fact, $(x - a)^\mu$ acquires the factor $e^{\mu 2\pi i}$

when x makes a positive circuit around a. Thus if one determines μ so that $e^{\mu 2\pi i} = \lambda$, taking $\mu = \frac{\log \lambda}{2\pi i}$, then $z(x-a)^{-\mu}$ will be a single-valued function for $x = a$. This function can be developed as a series in integral powers of $x - a$, and the function z itself in powers that differ from μ by integers.

Thus $z_1, z_2, \ldots, z_n$ can be developed as series in powers of $x - a$ with exponents of the form

$$\frac{\log \lambda_1}{2\pi i} + m, \ \frac{\log \lambda_2}{2\pi i} + m, \ldots, \frac{\log \lambda_n}{2\pi i} + m,$$

where m denotes an integer.

We shall now assume that the functions y are never infinite of infinite order, so that these series have only a finite number of terms with a negative exponent. Denote by $\mu_1, \mu_2, \ldots, \mu_n$ the lowest exponents in these series, so that

$$z_1(x-a)^{-\mu_1}, \ldots, z_n(x-a)^{-\mu_n}$$

have finite non-zero values. Obviously the difference between any two of the quantities $\mu_1, \mu_2, \ldots, \mu_n$ can never be an integer because the values of $\lambda_1, \ldots, \lambda_n$ are all distinct. On the other hand, the values of the corresponding exponents in two systems belonging to the same class must differ by integers, because the quantities $\lambda_1, \ldots, \lambda_n$ are determined by (A). These exponents therefore serve to distinguish the different systems of functions belonging to the same class, or rather to arrange them in groups, and it is sufficient, when the system is known, to give the substitution (α) rather than (A), because the quantities $\lambda_1, \ldots, \lambda_n$ are already determined by (A). We therefore use the following notation to characterise the system $(y_1, y_2, \ldots, y_n)$ more exactly:

$$Q \left\{ \begin{array}{lllll} a & b & \cdots & g & \\ (\alpha) & (\beta) & \cdots & (\theta) & \\ \mu_1 & \nu_1 & \cdots & \rho_1 & x \\ \cdots & \cdots & \cdots & \cdots & \\ \mu_n & \nu_n & \cdots & \rho_n & \end{array} \right\}.$$

Here the quantities in the columns have the meanings for $b, \ldots, y$ corresponding to those in the first column for a. It is readily seen that every system can be regarded as a particular case of another system, in which some or all of the corresponding exponents are smaller.

It is not difficult to show now that, between any $n+1$ systems belonging to the same class, there must exist a linear homogeneous equation whose coefficients are polynomials in x. We distinguish the corresponding quantities in these $n+1$ systems by upper indices. Suppose that the following n equations connect them:

$$(2)\qquad \begin{aligned} a_0 y_1 + a_1 y_1^{(1)} + \cdots + a_n y_1^{(n)} &= 0, \\ a_0 y_2 + a_1 y_2^{(1)} + \cdots + a_n y_2^{(n)} &= 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_0 y_n + a_1 y_n^{(1)} + \cdots + a_n y_n^{(n)} &= 0. \end{aligned}$$

Then the quantities $a_0, a_1, \ldots, a_n$ must be proportional to the determinants of the systems obtained by omitting from the $n(n+1)$ quantities y the first, second, ..., $(n+1)$-th columns respectively. A determinant

$$\sum \pm y_1^{(1)} y_2^{(2)} \ldots y_n^{(n)}$$

of this kind acquires a factor Det A when x makes a positive circuit around a, and cannot become infinite of infinite order for $x = a$. It can therefore be expressed as a series in powers of $x - a$ that increase in steps of 1. To determine the smallest exponent in this expansion, the determinant can be put into the form

$$\mathrm{Det}(\alpha) \sum \pm z_1^{(1)} z_2^{(2)} \ldots z_n^{(n)}.$$

In the last determinant, the first term $z_1^{(1)} z_2^{(2)} \ldots z_n^{(n)}$ is

$$(x-a)^{\mu_1^{(1)} + \mu_2^{(2)} + \cdots + \mu_n^{(n)}}$$

multiplied by a function which has a non-zero finite value for $x = a$. Thus the smallest exponent in the expansion of this term in powers of $x - a$ is

$$\mu_1^{(1)} + \mu_2^{(2)} + \cdots + \mu_n^{(n)}.$$

By permuting the upper indices, the terms with the lowest exponents in the expansions of the other terms can be found. It is clear that, generally speaking, the exponent sought is the smallest of these values, and in any case cannot be smaller. If we denote by $\bar{\mu}$ the smallest of these values, and by $\bar{\nu}, \ldots, \bar{\rho}$ the corresponding values for the other branch points, then

$$\sum \pm y_1^{(1)} y_2^{(2)} \ldots y_n^{(n)} (x-a)^{-\bar{\mu}} (x-b)^{-\bar{\nu}} \ldots (x-g)^{-\bar{\rho}}$$

is a function of x, finite and single valued for all finite complex values of x, which becomes infinite of an order not exceeding $-(\bar{\mu}+\bar{\nu}+\cdots+\bar{\rho})$ for $x=\infty$. Hence it is a polynomial of degree at most $-(\bar{\mu}+\bar{\nu}+\cdots+\bar{\rho})$. This degree must be a non-negative integer if the function does not vanish identically.

The minor determinants, which are proportional to the quantities $a_0, a_1, \ldots, a_n$, thus behave like polynomials multiplied by powers of $x-a, x-b, \ldots, x-g$, whose exponents in the various minor determinants differ by integers. The quantities $a_0, a_1, \ldots, a_n$ behave like polynomials and can be replaced by these in the equations (2), and the theorem is proved.

The derivatives of the functions $y_1, y_2, \ldots, y_n$ with respect to x obviously constitute a system belonging to the same class. For the derivatives of the functions $(A)(y_1, y_2, \ldots, y_n)$, into which $(y_1, y_2, \ldots, y_n)$ are transformed when x makes a positive circuit around a, are

$$(A)\left(\frac{dy_1}{dx}, \frac{dy_2}{dx}, \ldots, \frac{dy_n}{dx}\right),$$

since the coefficients in (A) are constants. This remark yields two corollaries of the above theorem:

> *"The functions y of a system satisfy a differential equation of order n, whose coefficients are polynomials in x."*

and

> *"Every system of functions belonging to one class can be expressed as a linear combination, with rational coefficients, of these functions and their first $n-1$ derivatives."*

With the help of the latter corollary, we can give a general expression for the systems which constitute a class, from which it can immediately be seen that, as mentioned above, the number of systems is infinite. We apply this result here only to find the systems which have in common not only the same substitutions but also the same exponents. For an arbitrary system $Y_1, Y_2, \ldots, Y_n$ with the same substitutions and the same exponents as $y_1, y_2, \ldots, y_n$, it follows from the corollary that, denoting derivatives in Lagrange's manner, we obtain n linear equations of the form:

$$\begin{aligned} c_0 Y_1 &= b_0 y_1 + b_1 y_1' + \cdots + b_{n-1} y_1^{(n-1)}, \\ c_0 Y_2 &= b_0 y_2 + b_1 y_2' + \cdots + b_{n-1} y_2^{(n-1)}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ c_0 Y_n &= b_0 y_n + b_1 y_n' + \cdots + b_{n-1} y_n^{(n-1)}, \end{aligned}$$

where the coefficients are polynomials in x.

The function c_0 depends only on the functions y, and there is a finite upper bound to the degree of the polynomials b, so that they have only a finite number of coefficients. Conversely, in order that the functions $Y_1, \ldots, Y_n$ produced by these equations should have the desired properties, the coefficients must be such that, for the branch points, their exponents must not be less than those of the functions y, while for all other values of x they remain finite. These conditions give a set of simultaneous linear homogeneous equations for the coefficients of the powers of x in the polynomials b. The solution, in the case where these equations suffice to determine the coefficients, shows that the most general value of the functions (Y) is of the form const. (y). Otherwise, the general solution takes the form

$$Y_1 = ky_1 + k_1 Y_1^{(1)} + \cdots + k_m Y_1^{(m)}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$Y_n = ky_n + k_1 Y_n^{(1)} + \cdots + k_m Y_n^{(m)}$$

with arbitrary constants $k, k_1, \ldots, k_m$. These arbitrary constants can always be determined successively as functions of the others in such a way that the initial term in the development of the functions $(\alpha)^{-1}(Y), (\beta)^{-1}(Y), \ldots, (\theta)^{-1}(Y)$ vanishes, thereby ensuring that the sum of the exponents increases each time by at least 1. The final exponent sum is increased by at least m, reducing the number of arbitrary constants by the same amount. In this way, we can always derive from every system of n functions another one with greater exponents whose character is fully determined, apart from a constant factor (common to all functions) by the matrix of substitutions and the exponents. This common factor is now also determined if the coefficient of the lowest power of $x - a$ in the expansion of the first of the functions $(\alpha)^{-1}(y)$ is set equal to 1, so that the functions y are uniquely determined.

If we take care to note how the behavior of the functions varies with the position of one of the branch-points, for example a, we arrive at the theorem that the quantities y constitute an analogous system of functions of a as of x, whose branch points are $b, c, d, \ldots, x$, and whose substitutions are formed by the composition of $(A), (B), \ldots, (F)$.

In the case where it is not possible to make the functions vary with a in such a way that all the substitutions stay constant (because the number of arbitrary constants in these substitutions is smaller than the number of conditions necessary to achieve this), we can regard the system as a particular case of a system with smaller exponents in which for these special values of $a, b, \dots, g$ the coefficients of some of the initial terms in the series for $(\alpha)^{-1}(y), (\beta)^{-1}(y), \dots, (\theta)^{-1}(y)$ vanish.

It follows from this theorem that the quantities $y_1, y_2, \dots, y_n$ represent functions of the p variables $a, b, \dots, g, x$, which when all variables resume their original values, either take their original values or are transformed into linear combinations of the original values. The substitution with constant coefficients in question is a certain composition of the $p-2$ arbitrarily given systems $(A), (B), (C), \dots, (F)$.

I refrain for now from further study of these functions of several variables and the various aids afforded by this last theorem in the solution of linear differential equations. I merely remark that an integral of an algebraic function can be regarded as a special case of the functions treated here. Applying these principles to an integral of this kind leads to functions which represent general θ-series with arbitrary moduli of periodicity.

Determination of the form of the differential equation.

Our next task, in developing the theory of linear differential equations via these principles, is to find the simplest system in each class. To this end we need to determine more closely the form of the differential equation. Following Lagrange, let $y^{(1)}, y^{(2)}, \dots, y^{(n)}$ denote successive derivatives of y; then the equations (2) represent the differential equation which they satisfy.

The degree of the polynomials that can be taken as coefficients can be determined in the following manner. Each differentiation with respect to x reduces by 1 each of the exponents of the characteristic, assuming that none of these exponents is an integer. Thus

$$\sum \pm y_1 y_2^{(1)} \cdots y_n^{(n-1)} (x-a)^{-\bar{\mu}} (x-b)^{-\bar{\nu}} \cdots (x-g)^{-\bar{\rho}} = X_0$$

remains everywhere finite and single-valued. Here

$$\bar{\mu} = \sum_i \mu_i - \frac{n(n-1)}{2}, \quad \bar{\nu} = \sum_i \nu_i - \frac{n(n-1)}{2}, \dots, \bar{\rho} = \sum_i \rho_i - \frac{n(n-1)}{2}.$$

For $x = \infty$, since the functions y remain finite and single-valued, $\sum \pm y_1 y_2^{(1)} \cdots y_n^{(n-1)}$ is infinitely small of order $n(n-1)$. The degree of the polynomial X_0 is therefore

$$r = (m-2)\,\frac{n(n-1)}{2} - s,$$

where m denotes the number of branch points and s the sum of the exponents in the characteristic.

If, in the in the matrix of the $n(n+1)$ elements y, one omits column $n+1-t$ instead of the last column, the corresponding determinant has, in general, to be multiplied by powers of $x-a, x-b, \dots, x-g$ whose exponents are increased by t. The determinant becomes a polynomial of degree $r+(m-1)t$, [for $t=n$, however, this degree has the value $r+(m-2)n$].

The differential equation can therefore be put in the form

$$X_n y + \omega X_{n-1} y' + \cdots + \omega^n X_0 y^{(n)} = 0.$$

Here

$$\omega = (x-a)(x-b)\dots(x-g).$$

The X_t are polynomials of degree $r+(m-1)t$. [X_n has degree $r+(m-2)n$].

We now consider what conditions have to be satisfied by the coefficients of these polynomials if $a, b, \dots, g$ are to be the only branch points and the exponents of discontinuity have assigned values. A branch point does not occur, if and only if all the solutions of the differential equation can be expanded in a series in integral powers of the increment of x, or as long as the expansion by Maclaurin's theorem of the function y contains n arbitrary constants. This is always the case when a_n is nonzero. Only the case $a_n = 0$ needs further investigation. We write the differential equation in the form

$$b_0 y + b_1(x-a)y' + b_2(x-a)^2 y'' + \cdots + b_n(x-a)^n y^{(n)} = 0.$$

In order that the function y should have the prescribed character in the neighborhood of $x=a$, $\mu_1, \dots, \mu_n$ must all be roots of the equation

$$b_0 + b_1\mu + \cdots + b_n\mu(\mu-1)\dots(\mu-n+1) = 0.$$

This gives n conditions for the functions X. Moreover, since all the μ must be finite and distinct, b_n must be non-zero for $x=a$. The same must hold for the other roots $b, c, \dots, g$ of $\omega = 0$. Hence the equations $X_0 = 0$ and $\omega = 0$ can have no common root.

Now if, (for some root of $X_0 = 0$) $a_n = 0, a_{n-1} \neq 0$, then for this root $y, y', \dots, y^{(n-2)}$ can be assigned arbitrarily, while $y^{(n-1)}$ is determined by the differential equation

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_0 y = 0.$$

Thus while there will be $n-1$ arbitrary constants in the first $n-1$ terms of the Maclaurin expansion, the last constant appears at the earliest in the $(n+1)$-th term. Let us assume that it first appears in term $n+h$.

We now eliminate the quantities $y^{(n+h-2)}, \dots, y^{(n-1)}$ in the h-th derivative of the differential equation:

$$a_n y^{(n+h)} + (ha_n' + a_{n-1})y^{(n+h-1)} + \cdots = 0,$$

using the previous derivatives and the differential equation itself. The coefficients of $y^{(n+h-1)}, y^{(n-2)}, y^{(n-3)}, \dots, y$ must all vanish, since these functions are linearly independent. We obtain

$$ha_n' + a_{n-1} = 0,$$

and so $a_n' \neq 0$. There are $n-1$ further equations, giving n conditions for the coefficients of the functions X.

Let us now, secondly, suppose that a_n and a_{n-1} vanish simultaneously, while a_{n-2} remains finite, so that the first $n-2$ terms of the Maclaurin series contain $n-2$ arbitrary constants. Suppose that the next constant appears in term $n+h-1$ and the last in term $n+h'-1$. For $y^{(n+h-2)}$ and $y^{(n+h'-2)}$ to be independent of the values of the derivatives of lesser order, the following equations must hold:

$$\begin{gathered} a_n' = 0, \\ \frac{h(h-1)}{2} a_n'' + ha_{n-1}' + a_{n-2} = 0, \\ \frac{h'(h'-1)}{2} a_n'' + h'a_{n-1}' + a_{n-2} = 0. \end{gathered}$$

Thus a_n'' and a_{n-1}' must be different from zero, and there are $2n-3$ other equations. Consequently two linear factors of a_n are 0, and there are $2n$ conditions for the functions X.

Similarly, consider the case where a_n, a_{n-1}, a_{n-2} vanish simultaneously, but a_{n-3} remains finite, while the last three arbitrary constants first make their appearance in terms $n+h-2$, $n+h'-2$, and $n+h''-2$ respectively. We find the conditions

$$\begin{gathered} a_n' = 0, \; a_n'' = 0, \; a_{n-1}' = 0, \\ \frac{h(h-1)(h-2)}{1.2.3} a_n''' + \frac{h(h-1)}{2} a_{n-1}'' + ha_{n-2}' + a_{n-3} = 0 \end{gathered}$$

for h, h', h'' together with $3n-6$ other equations, so that a_n has three and only three equal roots, and $3n$ conditions have to be fulfilled. By an obvious generalization of this argument, each linear factor of X_0 gives rise to n conditions between the functions X.

We now suppose that one of the singular points, for example g, is at infinity, and denote by ω the polynomial of degree $m-1$,

$$\omega = (x-a)(x-b)\ldots$$

We denote the minor determinants of order n formed from the matrix with n rows

$$\begin{array}{c} y_1\ y_1' \cdots y_1^{(n)} \\ y_2\ y_2' \cdots y_2^{(n)} \\ \cdots\cdots\cdots\cdots \\ y_n\ y_n' \cdots y_n^{(n)} \end{array}$$

by $\Delta_0, \Delta_1, \ldots, \Delta_n$, so that $y_1, y_2, \ldots, y_n$ are particular solutions of the differential equation

$$y\Delta_0 + y'\Delta_1 + y''\Delta_2 + \cdots + y^{(n)}\Delta_n = 0.$$

The function

$$\Delta_k (x-a)^{-\sum\mu}(x-b)^{-\sum\nu}\ldots\omega^{-k+n(n+1)/2} = X_{n-k}$$

is then, as noted above, a polynomial in x whose degree can be deduced from a consideration of the singular point $x=\infty$. Thus, if r_t denotes the degree of X_t,

$$r_t = r + (m-2)t,$$

where

$$r = (m-2)\,\frac{n(n-1)}{2} - s$$

is the degree of X_0 and

$$s = \sum\mu + \sum\nu + \cdots + \sum\rho$$

is an integer.

The differential equation for y can now be put into the form

$$\omega^n X_0 y^{(n)} + \omega^{n-1} X_1 y^{(n-1)} + \cdots + \omega X_{n-1} y' + X_n y = 0.$$

Since there are r zeros of the function X_0 which do not belong to the set of singular points, from the preceeding considerations there must exist rn conditions between the constants in this differential equation.

Consequently there still remain in the differential equation the following number of constants at our disposal (since one of these constants may be put equal to 1):

$$\sum (r_t + 1) - 1 - rn = r + n + (m-2)\,\frac{n(n+1)}{2}.$$

Replacing r by its value, we obtain

$$-s + (m-2)n^2 + n.$$

In an arbitrary system of n particular solutions $y_1, y_2, \ldots, y_n$ which gives rise to n^2 constants of integration, there are

$$-s + (m-1)n^2 + n$$

undetermined constants.

The number of coefficients in the substitutions $(A), (B), \ldots (G)$ is mn^2 and therefore this is the number of conditions which have to be satisfied when these coefficients are arbitrarily prescribed. However the substitutions are restricted by the relation (1), so that n^2 of the conditions are identical consequences of the others. This leaves $(m-1)n^2$ conditions, and the number of constants still available is $n - s$. This number must be at least 1, because a factor common to all the y must remain arbitrary, and consequently

$$s \leq n - 1.$$

XXII.

A mathematical work that seeks to answer the question posed by the most distinguished academy of Paris.

'Determine the calorific state of an indefinite solid homogeneous body so that a system of isotherms, at a given instant, remain isotherms after an arbitrary time, in such a way that the temperature at a point can be expressed as a function of time and two other independent variables.'

Et his principiis via sternitur ad majora.

1.

We shall treat the question posed by the most distinguished academy in such a way that we answer the following more general question first:

What must be the properties of a body, which determine its conduction and distribution of heat, so that a system of lines exists that remain isotherms?

Then:

from the general solution of this problem, we select those cases in which the properties are the same at any two points, so that the body is homogeneous.

First part.

2.

In order to undertake the first question, consider the conduction of heat in an arbitrary body. Let u denote the temperature, at time t, at the point (x_1, x_2, x_3). The general equation, according to which the function u varies, takes the form

$$\text{(I)}\qquad \begin{aligned}&\frac{\partial}{\partial x_1}\left(a_{1,1}\frac{\partial u}{\partial x_1}+a_{1,2}\frac{\partial u}{\partial x_2}+a_{1,3}\frac{\partial u}{\partial x_3}\right)\\ +&\frac{\partial}{\partial x_2}\left(a_{2,1}\frac{\partial u}{\partial x_1}+a_{2,2}\frac{\partial u}{\partial x_2}+a_{2,3}\frac{\partial u}{\partial x_3}\right)\\ +&\frac{\partial}{\partial x_3}\left(a_{3,1}\frac{\partial u}{\partial x_1}+a_{3,2}\frac{\partial u}{\partial x_2}+a_{3,3}\frac{\partial u}{\partial x_3}\right)=h\frac{\partial u}{\partial t}.\end{aligned}$$

In this equation, the quantities a denote conductivities, h denotes specific heat per unit volume, that is the product of specific heat and density, and these are given functions of x_1, x_2, x_3. We restrict our investigation to the case in which conductivity is the same in opposite directions, giving the relation

$$a_{i,i'} = a_{i',i}$$

between the quantities a. Moreover, since heat moves from a warmer to a cooler place, the quadratic form

$$\begin{pmatrix} a_{1,1}, & a_{2,2}, & a_{3,3} \\ a_{2,3}, & a_{3,1}, & a_{1,2} \end{pmatrix}$$

must be positive.

3.

We now introduce in equation (I) in place of the rectangular coordinates x_1, x_2, x_3, three arbitrary new independent variables s_1, s_2, s_3.

The transformation of equation (I) can be very easily deduced, since this equation is a necessary and sufficient condition that (denoting by δu an arbitrary infinitely small variation in u) the integral

$$\text{(A)} \qquad \delta \iiint \sum_{i,i'} a_{i,i'} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_{i'}} \, dx_1 \, dx_2 \, dx_3 + \iiint 2h \frac{\partial u}{\partial t} \, \delta u \, dx_1 \, dx_2 \, dx_3$$

taken over the body, depends only on the values of the variations δu on its surface. With the introduction of the new variables, the expression (A) becomes

$$\text{(B)} \qquad \delta \iiint \sum_{i,i'} b_{i,i'} \frac{\partial u}{\partial s_i} \frac{\partial u}{\partial s_{i'}} \, ds_1 \, ds_2 \, ds_3 + \iiint 2k \frac{\partial u}{\partial t} \, \delta u \, ds_1 \, ds_2 \, ds_3.$$

Here, for brevity, we write

$$\frac{\displaystyle\sum_{i,i'} a_{i,i'} \frac{\partial s_\mu}{\partial x_i} \frac{\partial s_\nu}{\partial x_{i'}}}{\displaystyle\sum \pm \frac{\partial s_1}{\partial x_1} \frac{\partial s_2}{\partial x_2} \frac{\partial s_3}{\partial x_3}} = b_{\mu,\nu}, \qquad \frac{h}{\displaystyle\sum \pm \frac{\partial s_1}{\partial x_1} \frac{\partial s_2}{\partial x_2} \frac{\partial s_3}{\partial x_3}} = k.$$

Let the quadratic forms

$$(1)\quad \begin{pmatrix} a_{1,1}, & a_{2,2}, & a_{3,3} \\ a_{2,3}, & a_{3,1}, & a_{1,2} \end{pmatrix} \qquad (2)\quad \begin{pmatrix} b_{1,1}, & b_{2,2}, & b_{3,3} \\ b_{2,3}, & b_{3,1}, & b_{1,2} \end{pmatrix},$$

have determinants A and B, and let the adjoint forms be

$$(3)\quad \begin{pmatrix} \alpha_{1,1}, & \alpha_{2,2}, & \alpha_{3,3} \\ \alpha_{2,3}, & \alpha_{3,1}, & \alpha_{1,2} \end{pmatrix} \qquad (4)\quad \begin{pmatrix} \beta_{1,1}, & \beta_{2,2}, & \beta_{3,3} \\ \beta_{2,3}, & \beta_{3,1}, & \beta_{1,2} \end{pmatrix}.$$

It follows that

$$A = B \sum \pm \frac{\partial s_1}{\partial x_1} \frac{\partial s_2}{\partial x_2} \frac{\partial s_3}{\partial x_3}$$

and

$$\beta_{\mu,\nu} = \sum_{i,i'} \alpha_{i,i'} \frac{\partial x_i}{\partial s_\mu} \frac{\partial x_{i'}}{\partial s_\nu}.$$

Hence

$$\sum_{i,i'} \alpha_{i,i'} dx_i\, dx_{i'} = \sum_{i,i'} \beta_{i,i'} ds_i\, ds_{i'}$$

and

$$\frac{h}{A} = \frac{k}{B}.$$

From this it can easily be seen that the transformation of equation (I) reduces to the transformation of the expression $\sum_{i,i'} \alpha_{i,i'} dx_i\, dx_{i'}$.

Hence we can solve our general problem in the following way. First determine the form of functions $b_{i,i'}$ and k of s_1, s_2, s_3, so that u cannot depend on one of these variables alone. After the solution of this problem, the expression $\sum \beta_{i,i'} ds_i\, ds_{i'}$ can be formed. We then find, given the quantities $a_{i,i'}$ and h, whether u can be a function of time and only two variables; and, if this is the case, whether the expressions $\sum \beta_{i,i'} ds_i\, ds_{i'}$ can be brought to a given form. We shall see below that this question can be treated by the method used by Gauss in the theory of curved surfaces.

4.

Accordingly we first investigate the form that $b_{i,i'}$ and k must take as functions of s_1, s_2 and s_3, so that u cannot depend on a single one of these variables. For simplicity, denote the quantities s_1, s_2, s_3 by α, β, γ and the form (2) by

$$\begin{pmatrix} a, & b, & c \\ a', & b', & c' \end{pmatrix}.$$

If u is independent of γ, the differential equation takes the form

$$\text{(II)} \qquad a\,\frac{\partial^2 u}{\partial \alpha^2} + 2c'\,\frac{\partial^2 u}{\partial \alpha \partial \beta} + b\,\frac{\partial^2 u}{\partial \beta^2} + e\,\frac{\partial u}{\partial \alpha} + f\,\frac{\partial u}{\partial \beta} - k\,\frac{\partial u}{\partial t} = F = 0.$$

Here we set

$$\frac{\partial a}{\partial \alpha} + \frac{\partial c'}{\partial \beta} + \frac{\partial b'}{\partial \gamma} = e, \quad \frac{\partial b}{\partial \beta} + \frac{\partial c'}{\partial \alpha} + \frac{\partial a'}{\partial \gamma} = f.$$

When γ is given different values, we obtain from equation (II) different equations between the six partial derivatives of u whose coefficients are independent of γ. Suppose now that m of these equations are independent, say:

$$F_1 = 0, F_2 = 0, \ldots, F_m = 0,$$

and the others follow from these, so that the equation $F = 0$ follows for arbitrary γ from these m equations. Then F must be of the form

$$c_1 F_1 + c_2 F_2 + \cdots + c_m F_m;$$

in this expression only the quantities c depend on γ.

Now let us examine more closely the particular cases $m = 1, 2, 3, 4$. At the same time we simplify the equations independent of γ into which the equation $F = 0$ resolves.

First case: $m = 1$.

If $m = 1$, the ratio of the coefficients in (II) does not depend on γ. By introducing the new variables $\int k\,d\gamma$ in place of γ, we can arrange that $k = 1$ and all coefficients are independent of γ. Further, by introducing new variables in place of α and β, we can arrange that $a = b = 0$. This will indeed occur if the expression $b(d\alpha)^2 + 2c'd\alpha\,d\beta + a(d\beta)^2$ (which cannot be the square of a linear differential equation, if (2) is a positive form) takes the form $m\,d\alpha' d\beta'$ and the quantities α', β' are understood to be independent variables.

Accordingly, in this case the differential equation (II) can be reduced to the form

$$2c' \frac{\partial^2 u}{\partial\alpha\partial\beta} + e\frac{\partial u}{\partial\alpha} + f\frac{\partial u}{\partial\beta} = \frac{\partial u}{\partial t}$$

and in the form (2) a, b are then 0; a', b' are linear functions of γ; and c' is independent of γ. Thus the temperature in this case clearly remains independent of γ when the initial temperature is given as an arbitrary function of α and β.

Second case: $m = 2$.

If equation (II) can be divided into two equations independent of γ, then $\frac{\partial u}{\partial t}$ can be excluded from one or other of the equations. For brevity we write one equation as

(1) $$\Delta u = 0,$$

and the other as

(2) $$\Lambda u = \frac{\partial u}{\partial t},$$

where Δ and Λ are definite expressions in $\partial_\alpha, \partial_\beta$.

It is easy to see that, by changing the independent variables, the first equation can be brought to the form where Δ is one of

$$\partial_\alpha\partial_\beta + e\partial_\alpha + f\partial_\beta,\ \partial_\alpha^2 + e\partial_\alpha + f\partial_\beta,\ \text{or } \partial_\alpha.$$

Here we do not exclude the value 0 for e, f.

Since now

$$0 = \partial_t \Delta u = \Delta\partial_t u = \Delta\Lambda u,$$

it follows from equations (1) and (2) that

(3) $$\Delta\Lambda u = 0.$$

We now distinguish two subcases.

(α) Equation (3) follows from equation (1), that is,

$$\Delta\Lambda = \Theta\Delta$$

where Θ is a new characteristic expression;

(β) Equation (3) does not follow from (1) and represents a new equation independent of Δu.

In order to pursue the first case (α) for a particular form of Δ, suppose that

$$\Delta = \partial_\alpha \partial_\beta + e\partial_\alpha + f\partial_\beta.$$

By using the equation $\Delta u = 0$, the expression $\Delta\Lambda u$ can be reduced to a form containing only partial derivatives with respect to one of the two variables; all these coefficients must be 0. Since the term containing $\partial_\alpha\partial_\beta$ can be removed with the help of the equation $\Delta u = 0$, we write

$$\Lambda = a\partial_\alpha^2 + b\partial_\beta^2 + c\partial_\alpha + d\partial_\beta$$

and form the expression

$$\Delta\Lambda - \Lambda\Delta.$$

In this expression, since the coefficients of ∂_α^3, ∂_β^3 must vanish, we obtain $\frac{\partial a}{\partial \beta} = 0$, $\frac{\partial b}{\partial \alpha} = 0$. Excluding the special cases $a = 0$, $b = 0$, we can arrange that $a = b = 1$ by a change of the independent variable. If the coefficients of ∂_α^2, ∂_β^2 are made equal to 0 in the reduced expression $\Delta\Lambda$, we deduce that

$$\frac{\partial c}{\partial \beta} = 2\frac{\partial e}{\partial \alpha}, \ \frac{\partial d}{\partial \alpha} = 2\frac{\partial f}{\partial \beta}$$

and it follows that we may write

$$\Delta = \partial_\alpha\partial_\beta + \frac{\partial m}{\partial \beta}\partial_\alpha + \frac{\partial n}{\partial \alpha}\partial_\beta,$$
$$\Lambda = \partial_\alpha^2 + \partial_\beta^2 + 2\frac{\partial m}{\partial \alpha}\partial_\alpha + 2\frac{\partial n}{\partial \beta}\partial_\beta.$$

Here m, n denote functions of α, β that must satisfy two differential equations in order for the coefficients of $\partial_\alpha, \partial_\beta$ to vanish in the reduced expression $\Delta\Lambda$.

In a completely similar way, we can deduce the simplest forms for Δ and Λ in the other particular cases in which

$$\Delta\Lambda = \Theta\Delta$$

holds. Here we shall not pursue this discussion, which is lengthy rather than difficult.

However, it is evident that the temperature always remains independent of γ in this case, when the initial temperature is an arbitrary function of α, β and the equation $\Delta u = 0$ holds. From the equations

$$\Delta u = 0,$$
$$\Lambda u = \frac{\partial u}{\partial t},$$

it follows that

$$0 = \Theta\Delta u = \Delta\Lambda u = \Delta\partial_t u = \frac{\partial \Delta u}{\partial t}$$

and the equation $\Delta u = 0$ continues to hold, if it is valid initially and the function u varies according to the equation $\Lambda u = \frac{\partial u}{\partial t}$. Accordingly u satisfies the equation of heat conduction, $F = 0$.

5.

There remains the second subcase (β), where the equation $\Delta\Lambda u = 0$ is independent of $\Delta u = 0$. In order to include also the cases $m = 3, 4$, we shall consider the more general hypothesis that, apart from $\Delta u = 0$, there is a linear differential equation $\Theta u = 0$ at our disposal that does not contain $\partial u/\partial t$ and is independent of $\Delta u = 0$.

If Δ is of the form $\partial_\alpha\partial_\beta + e\,\partial_\alpha + f\,\partial_\beta$, then with the help of the equation $\Delta u = 0$, partial derivatives with respect to both variables can be removed from the expression Θ.

We now distinguish two cases. If all the partial derivatives with respect to one or other variable, say β, are excluded from Θ, we obtain a differential equation containing only partial derivatives with respect to α, of form

$$\sum_\nu a_\nu \frac{\partial^\nu u}{\partial \alpha^\nu} = 0. \tag{1}$$

If this is not the case, we can obtain a differential equation of form

$$\sum_\nu a_\nu \frac{\partial^\nu u}{\partial t^\nu} = 0, \tag{2}$$

that is, one containing only partial derivatives with respect to t.

For in this case the expressions $\Lambda u, \Lambda^2 u, \Lambda^3 u, \ldots$, which are equal to the partial derivatives of u with respect to t, can be transformed with the help of the equations $\Delta u = 0$, $\Theta u = 0$, so that they contain partial derivatives with respect to only one of the two variables, derivatives not of higher order than those in Θu. As there are finitely many of these, it is clear that by elimination one can reach an equation of form (2). In each of the above equations, the a_ν are functions of α and β.

We observe that one of these two equations will still hold when Δ does not take the form $\partial_\alpha \partial_\beta + e\,\partial_\alpha + f\,\partial_\beta$. The particular case $\Delta = \partial_\alpha^2 + e\,\partial_\alpha + f\,\partial_\beta$ can be reduced to either of the two cases concerned, because using the equation $\Delta u = 0$, we can remove partial derivatives with respect to β from Θu, and also from Λu, and this readily yields an equation of each of the forms concerned. The case $f = 0$, and the case $\Delta = \partial_\alpha$, can be reduced to the previous case.

We now examine the second case more precisely.

The general solution of the equation

$$\sum_\nu a_\nu \frac{\partial^\nu u}{\partial t^\nu} = 0$$

comprises terms of form $f(t)e^{\lambda t}$, where $f(t)$ is a polynomial in t, and λ is independent of t. It is easy to see that these individual terms satisfy an equation of type (I).

We shall now show that λ cannot be a function of x_1, x_2, x_3.

Let kt^n be the highest order term of $f(t)$. We distinguish two cases.

1°. If λ is either real, or of form $\mu + \nu i$ with μ, ν functions of a single variable α (which depends on x_1, x_2, x_3), then substituting $u = f(t)e^{\lambda t}$ in the left side of equation (I) yields the coefficient

$$k\left(\frac{\partial \lambda}{\partial \alpha}\right)^2 \sum_{i,i'} a_{i,i'} \frac{\partial \alpha}{\partial x_i} \frac{\partial \alpha}{\partial x_{i'}}$$

of $t^{n+2}e^{\lambda t}$.

But these quantities cannot vanish except when

$$\frac{\partial \alpha}{\partial x_1} = \frac{\partial \alpha}{\partial x_2} = \frac{\partial \alpha}{\partial x_3} = 0,$$

that is, $\alpha = \text{const.}$, because the form

$$\begin{pmatrix} a_{1,1}, & a_{2,2}, & a_{3,3} \\ a_{2,3}, & a_{3,1}, & a_{1,2} \end{pmatrix}$$

is, as already shown, positive.

2°. If λ is of form $\mu + \nu i$ with independent functions μ, ν of x_1, x_2, x_3, then the quantities $\mu + \nu i$, $\mu - \nu i$ may be taken as independent variables α and β. Now u will contain, as well as the term $f(t)e^{\alpha t}$, the complex conjugate term $\phi(t)e^{\beta t}$. If now

$$\Delta u = a\,\frac{\partial^2 u}{\partial \alpha^2} + b\,\frac{\partial^2 u}{\partial \alpha \partial \beta} + c\,\frac{\partial^2 u}{\partial \beta^2} + e\,\frac{\partial u}{\partial \alpha} + f\,\frac{\partial u}{\partial \beta},$$

we may substitute $u = f(t)e^{\alpha t}$ in the equation $\Delta u = 0$. Setting the coefficient of $t^{n+2}e^{\alpha t}$ equal to 0 gives $a = 0$. Substituting $u = \phi(t)e^{\beta t}$ likewise yields $c = 0$. Hence, by using the equation $\Delta u = 0$, the equation $\Lambda u = \frac{\partial u}{\partial t}$ can be transformed in such a way that it contains partial derivatives with respect to only one of the two variables. But by substituting

$$u = f(t)e^{\alpha t},\ u = \phi(t)e^{\beta t},$$

we find that the coefficient of each of these partial derivatives is 0. Hence all these partial derivatives can be removed from the equation $\Lambda u = \frac{\partial u}{\partial t}$. Since by hypothesis u is not constant, this yields the desired result.

In the latter case, then, the function u comprises a finite number of terms of form $f(t)e^{\lambda t}$, where λ is constant and $f(t)$ is a polynomial in t.

In the first case, when an equation of form

(1) $$\sum a_\nu \frac{\partial^\nu u}{\partial \alpha^\nu} = 0$$

holds, the function u takes the form

$$u = \sum q_\nu p_\nu.$$

Here $p_1, p_2, \ldots$ are particular solutions of (1) and $q_1, q_2, \ldots$ are arbitrary constants, that is, functions of β and t alone. When we substitute this expression in

$$\Lambda u = \frac{\partial u}{\partial t},$$

the outcome is an equation of form

$$\sum PQ = 0.$$

Here the quantities Q are partial derivatives of q, and thus functions of β and t only; on the other hand, the functions P are functions only of α and β. However, we saw above that such an equation, comprising n terms, yields μ linear equations between the functions Q and $n - \mu$ linear equations between the functions P. Here the coefficients are functions only of β; μ denotes one of $0, 1, 2, \ldots, n$. In this way we obtain expressions for the $\frac{\partial q}{\partial t}$ via partial derivatives of q with respect to β, that do not contain α.

We now investigate precisely the individual cases of our problem that belong to this case.

If $m = 2$ and Δ is of form $\partial_\alpha \partial_\beta + e\,\partial_\alpha + f\,\partial_\beta$, the reduced equation $\Delta\Lambda u = 0$, in the case that it is free of partial derivatives with respect to β, takes the form

$$\frac{\partial^3 u}{\partial \alpha^3} + r\,\frac{\partial^2 u}{\partial \alpha^2} + s\,\frac{\partial u}{\partial \alpha} = 0.$$

Hence u is of the shape

$$ap + bq + c,$$

with a, b, c functions of β and t alone, while p and q are functions only of α and β. Now we can introduce the independent variable q in place of α. This yields

$$u = ap + b\alpha + c,$$

where now p is a function of the two variables α and β. Substituting this expression in the equations

$$\Delta u = 0, \;\; \Lambda u = \frac{\partial u}{\partial t},$$

one readily finds the coefficients.

There remains the case that one of the equations, into which $F = 0$ divides, takes the shape (1) and thus has the form

$$r\,\frac{\partial^2 u}{\partial \alpha^2} + s\,\frac{\partial u}{\partial \alpha} = 0.$$

It follows that $u = ap + b$, where a and b are functions only of β and t, and p denotes a function only of α and β. Introducing in place of p the independent variable α, we find that

$$u = a\alpha + \beta, \;\; \frac{\partial^2 u}{\partial \alpha^2} = 0.$$

Thus we find that, when $m = 2$, that is when $F = 0$ splits into two equations

$$\Delta u = 0,$$
$$\Lambda u = \frac{\partial u}{\partial t},$$

that either $\Delta\Lambda = \Theta\Delta$; or the function u comprises a finite number of terms of the form $f(t)e^{\lambda t}$ with constants λ and polynomials $f(t)$; or that u takes the shape

$$\phi(\beta, t)\chi(\alpha, \beta) + \alpha\phi_1(\beta, t) + \phi_2(\beta, t).$$

Further, if $m = 3$, then either u comprises a finite number of terms $f(t)e^{\lambda t}$, or u takes the form

$$\phi(\beta, t)\alpha + \phi_1(\beta, t).$$

Finally, the case $m = 4$ can be elucidated without difficulty.

Namely, suppose that in addition to the equation $\Lambda u = \frac{\partial u}{\partial t}$, three equations hold between

$$\frac{\partial^2 u}{\partial \alpha^2}, \ \frac{\partial^2 u}{\partial \alpha \partial \beta}, \ \frac{\partial^2 u}{\partial \beta^2}, \ \frac{\partial u}{\partial \alpha}, \ \frac{\partial u}{\partial \beta}.$$

Then *either* an equation of the form

$$r\,\frac{\partial u}{\partial \alpha} + s\,\frac{\partial u}{\partial \beta} = 0$$

holds, in which case we can choose the independent variables in such a way that u is a function of only one variable. *Or*, the functions

$$\frac{\partial^2 u}{\partial \alpha^2}, \frac{\partial^2 u}{\partial \alpha \partial \beta}, \frac{\partial^2 u}{\partial \beta^2},$$

and consequently also the functions $\Lambda u, \Lambda^2 u, \Lambda^3 u$ can be expressed in terms of $\frac{\partial u}{\partial \alpha}, \frac{\partial u}{\partial \beta}$. In this case we find an equation of form

$$a\,\frac{\partial^3 u}{\partial t^3} + b\,\frac{\partial^2 u}{\partial t^2} + c\,\frac{\partial u}{\partial t} = 0,$$

whence u takes the form

$$pe^{\lambda t} + qe^{\mu t} + r \quad \text{or} \quad (p + qt)e^{\lambda t} + r,$$

where λ and μ are constants by the above discussion.

Taking p as the independent variable α, and substituting the expressions derived above into the equation $\Lambda u = \frac{\partial u}{\partial t}$, it emerges that q cannot be a function of α, at least if λ and μ are distinct. Thus p and q can take the roles of independent variables. We further get $r =$ const. from the equation

$$\Lambda u = \frac{\partial u}{\partial t}.$$

Thus in this case u is either a function of t and only one other variable, or u takes one of the two forms

$$\alpha e^{\lambda t} + \beta e^{\mu t} + \text{ const.}, \; (\alpha + \beta t) e^{\lambda t} + \text{ const.},$$

where we do not exclude the value $\mu = 0$.

Having found the forms that the function u can take, the equations $F_\nu = 0$, which for brevity we do not write out, are readily formed. Thus in every particular case, as well as the form

$$\begin{pmatrix} b_{1,1}, & b_{2,2}, & b_{3,3} \\ b_{2,3}, & b_{3,1}, & b_{1,2} \end{pmatrix},$$

we determine the adjoint form

$$\begin{pmatrix} \beta_{1,1}, & \beta_{2,2}, & \beta_{3,3} \\ \beta_{2,3}, & \beta_{3,1}, & \beta_{1,2} \end{pmatrix}.$$

If we now introduce, in the expression $\sum \beta_{t,t'} ds_t ds_{t'}$, arbitrary functions of x_1, x_2, x_3 in place of the quantities s_1, s_2, s_3, we clearly obtain all those cases in which u is a function of time and only two other variables. This solves the first question.

It remains to clarify the question of when the expression $\sum \beta_{i,i'} ds_i ds_{i'}$ can be transformed into a given form $\sum \alpha_{i,i'} dx_i dx_{i'}$.

Second Part.

On the transformation of the expression $\sum_{i,i'} b_{i,i'} ds_i ds_{i'}$ into a given form $\sum_{i,i'} a_{i,i'} dx_i dx_{i'}$.

Since the question of the most distinguished academy is restricted to homogeneous bodies in which the conductivities are constants, we first examine conditions under which the expression $\sum_{i,i'} b_{i,i'} ds_i ds_{i'}$ can be transformed into the form $\sum_{i,i'} a_{i,i'} dx_i dx_{i'}$, with constant coefficients $a_{i,i'}$, by taking the quantities s to be suitable functions of x. Accordingly we make some remarks about transformation to an arbitrary prescribed form.

It is known that an expression $\sum_{i,i'} a_{i,i'} dx_i dx_{i'}$ can always be brought to the form $\sum_i dx_i^2$ if, as we suppose here, it is a positive form in the dx. Hence, if $\sum_{i,i'} b_{i,i'} ds_i ds_{i'}$ can be transformed into $\sum_{i,i'} a_{i,i'} dx_i dx_{i'}$, it can be transformed into $\sum_i dx_i^2$, and conversely. We therefore investigate when this expression can be transformed to the form $\sum_i dx_i^2$.

Let B be the determinant $\sum \pm b_{1,1} b_{2,2} \dots b_{n,n}$, and let $\beta_{t,t'}$ be the cofactors. Then

$$\sum_i \beta_{i,i'} b_{i,i'} = B; \quad \sum_i \beta_{i,i'} b_{i,i''} = 0 \text{ if } i' \neq i''.$$

If $\sum_{i,i'} b_{i,i'} ds_i ds_{i'} = \sum_i dx_i^2$ holds for an arbitrary value of dx, then substituting $d + \delta$ for d we can also show that

$$\sum_{i,i'} b_{i,i'} ds_i \delta s_{i'} = \sum_i dx_i \delta x_i$$

for arbitrary values of dx and δx.

Accordingly, when we express the quantities ds_i via the dx_i and the quantities δx_i via the quantities δs_i, it follows that

$$\frac{\partial x_{\nu'}}{\partial s_\nu} = \sum_i b_{\nu,i} \frac{\partial s_i}{\partial x_{\nu'}}, \tag{1}$$

and so

$$\frac{\partial s_i}{\partial x_{\nu'}} = \sum_\nu \frac{\beta_{\nu,i}}{B} \frac{\partial x_{\nu'}}{\partial s_\nu}. \tag{2}$$

Since

$$\sum_\nu \frac{\partial s_i}{\partial x_\nu} \frac{\partial x_\nu}{\partial s_i} = 1 \text{ and } \sum_\nu \frac{\partial s_i}{\partial x_\nu} \frac{\partial x_\nu}{\partial s_{i'}} = 0 \ (i \neq i'),$$

we deduce further that

$$\begin{aligned} \sum_\nu \frac{\partial x_\nu}{\partial s_i} \frac{\partial x_\nu}{\partial s_{i'}} &= b_{i,i'}, \\ \sum_\nu \frac{\partial s_i}{\partial x_\nu} \frac{\partial s_{i'}}{\partial x_\nu} &= \frac{\beta_{i,i'}}{B}. \end{aligned} \tag{3}$$

Differentiation of the formula (3) yields

$$\sum_\nu \frac{\partial^2 x_\nu}{\partial s_i \partial s_{i''}} \frac{\partial x_\nu}{\partial s_{i'}} + \sum_\nu \frac{\partial^2 x_\nu}{\partial s_{i'} \partial s_{i''}} \frac{\partial x_\nu}{\partial s_i} = \frac{\partial b_{i,i'}}{\partial s_i''}. \tag{4}$$

From these expressions, we find that

$$\frac{\partial b_{i,i'}}{\partial s_{i''}}, \quad \frac{\partial b_{i,i''}}{\partial s_{i'}}, \quad \frac{\partial b_{i',i''}}{\partial s_i}$$

satisfy the relation

$$2 \sum_\nu \frac{\partial^2 x_\nu}{\partial s_{i'} \partial s_{i''}} \frac{\partial x_\nu}{\partial s_i} = \frac{\partial b_{i,i'}}{\partial s_{i''}} + \frac{\partial b_{i,i''}}{\partial s_{i'}} - \frac{\partial b_{i',i''}}{\partial s_i}. \tag{5}$$

If the right-hand side of (5) is denoted by $p_{i,i',i''}$, we obtain

$$2 \frac{\partial^2 x_\nu}{\partial s_{i'} \partial s_{i''}} = \sum_i \frac{\partial s_i}{\partial x_\nu} p_{i,i',i''}. \tag{6}$$

By further differentiation of the quantities $p_{i,i',i''}$, we obtain

$$\frac{\partial p_{i,i',i''}}{\partial s_{i'''}} - \frac{\partial p_{i,i',i'''}}{\partial s_{i''}} = 2\sum_\nu \frac{\partial^2 x_\nu}{\partial s_{i'}\partial s_{i''}}\,\frac{\partial^2 x_\nu}{\partial s_i \partial s_{i'''}} - 2\sum_\nu \frac{\partial^2 x_\nu}{\partial s_{i'}\partial s_{i'''}}\,\frac{\partial^2 x_\nu}{\partial s_i \partial s_{i''}}.$$

Finally, by substituting the values found in (6) and (4), we obtain

$$\text{(I)}\qquad \begin{aligned}&\frac{\partial^2 b_{i,i''}}{\partial s_{i'}\partial s_{i'''}} + \frac{\partial^2 b_{i',i'''}}{\partial s_i \partial s_{i''}} - \frac{\partial^2 b_{i,i'''}}{\partial s_{i'}\partial s_{i''}} - \frac{\partial^2 b_{i',i''}}{\partial s_i \partial s_{i'''}}\\ &+\frac{1}{2}\sum_{\nu,\nu'}(p_{\nu,i',i'''}p_{\nu',i,i''} - p_{\nu,i,i'''}p_{\nu',i',i''})\,\frac{\beta_{\nu,\nu'}}{B} = 0.\end{aligned}$$

Thus the functions b must satisfy equations of this form if $\sum_{i,i'} b_{i,i'}ds_i ds_{i'}$ can be transformed into the form $\sum_i dx_i^2$. We denote the left side of (I) by

$$(ii', i''i''').$$

In order to understand the structure of these equations better, we form the expression

$$\delta^2 \sum b_{i,i'}ds_i ds_{i'} - 2d\delta \sum b_{i,i'}ds_i\delta s_{i'} + d^2 \sum b_{i,i'}\delta s_i \delta s_{i'},$$

where the second order variations $d^2, d\delta, \delta^2$ are determined so that

$$\begin{aligned}&\delta' \sum b_{i,i'}ds_i ds_{i'} - \delta \sum b_{i,i'}ds_i\delta' s_{i'} - d\sum b_{i,i'}\delta s_i \delta' s_{i'} = 0,\\ &\delta' \sum b_{i,i'}ds_i ds_{i'} - 2d\sum b_{i,i'}ds_i\delta' s_{i'} = 0,\\ &\delta' \sum b_{i,i'}\delta s_i \delta s_{i'} - 2\delta\sum b_{i,i'}\delta s_i\delta' s_{i'} = 0,\end{aligned}$$

where δ' denotes an arbitrary variation. The above expression now becomes

$$\text{(II)}\qquad \sum (ii', i''i''')(ds_i\delta s_{i'} - ds_{i'}\delta s_i)(ds_{i''}\delta s_{i'''} - ds_{i'''}\delta s_{i''}).$$

From this form of the expression, it is clear that after a change of independent variable it becomes an expression that depends on the new sum corresponding to $\sum b_{i,i'}ds_i ds_{i'}$ in the same way as before. If the quantities b are constants, then all coefficients vanish in the expression (II). It follows that

(II) vanishes identically, if $\sum b_{i,i'} ds_i ds_{i'}$ can be transformed into an analogous expression with constant coefficients.

Likewise it is clear, if the expression (II) does not vanish, that the expression

$$\text{(III)} \qquad \frac{-\frac{1}{2}\sum(ii', i''i''')(ds_i\delta s_{i'} - ds_{i'}\delta s_i)(ds_{i''}\delta s_{i'''} - ds_{i'''}\delta s_{i''})}{\sum b_{i,i'} ds_i ds_{i'} \sum b_{i,i'}\delta s_i \delta s_{i'} - (\sum b_{i,i'} ds_i \ \delta s_{i'})^2}$$

is invariant under a change of independent variables, and it further remains unchanged if the variations $ds_i, \delta s_i$ are replaced by arbitrary linear combinations of form $\alpha ds_i + \beta\delta s_i, \gamma ds_i + \delta\delta s_i$. However, the maximum and minimum values of the function (III) of $ds_i, \delta s_i$ depend neither on the form of the expression $\sum b_{i,i'} ds_i ds_{i'}$ nor on the values of the variations $ds_i, \delta s_i$. Accordingly, these values can be used to determine whether two such expressions can be transformed into one another.

Our investigations can be illustrated by a geometric example, which although of unfamiliar form, is a useful supplement.

The expression $\sqrt{\sum b_{i,i'} ds_i ds_{i'}}$ can be interpreted as the line element in a general n-dimensional space that lies outside our intuition. In this space, when we draw from a point $(s_1, s_2, \ldots, s_n)$ all the shortest lines, in whose initial element the variations of s are in ratio $\alpha ds_1 + \beta\delta s_1 : \alpha ds_2 + \beta\delta s_2 : \ldots : \alpha ds_n + \beta\delta s_n$, where α and β are arbitrary, these lines form a surface that one can represent in the usual space of our perception. Now the expression (III) is the curvature of this surface at the point $(s_1, s_2, \ldots, s_n)$. [1]

If we now come back to the case $m = 3$, the expression (II) is a quadratic form in

$$ds_2\delta s_3 - ds_3\delta s_2, \ ds_3\delta s_1 - ds_1\delta s_3, \ ds_1\delta s_2 - ds_2\delta s_1$$

so that in this case we obtain six equations that the functions b must satisfy in order for $\sum b_{i,i'} ds_i ds_{i'}$ to transform into a form with constant coefficients. It is not difficult to see with the aid of familiar ideas that these six conditions are sufficient. However, we observe that only three of these are mutually independent.

Now in order to answer the question posed by the most distinguished academy, we substitute into these six equations the form of the function b found by the above method. This yields all the cases in which the temperature u in homogeneous bodies can be a function of time and only two further variables.

Considerations of time prevent us from carrying through these calculations. We must content ourselves with the exposition of particular solutions to this question that are obtained by our methods.

For brevity, consider the very simple case in which the temperature u varies according to the law

$$\text{(I)} \qquad \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} = a^2 \frac{\partial u}{\partial t}\,.$$

The other cases readily reduce to this one. The case $m = 1$ can only arise under the condition that u is constant either on parallel lines, or on circles, or on spirals. By a suitable choice of rectangular coordinates $z, r\cos\phi, r\sin\phi$, we obtain $\alpha = r, \beta = z + \phi$. const.

The case $m = 2$ occurs in case $u = f(\alpha) + \phi(\beta)$, the case $m = 3$ in case $u = \alpha e^{\lambda t} + f(\beta)$, where λ denotes a real constant. Finally the case $m = 4$, as we already showed above, occurs in case either $u = \alpha e^{\lambda t} + \beta e^{\mu t}+$ const., or $u = (\alpha + \beta t)e^{\lambda t}+$ const., or $u = f(\alpha)$.

In order to identity the form of the function u, we merely observe that the temperature u, except when it takes the form $\alpha e^{\lambda t}$, can only be a function of time and a single variable when it is constant either on parallel planes, or on cylinders with the same axis, or on concentric spheres. If u is of the form $\alpha e^{\lambda t}$, it follows from the differential equation (I) that

$$\frac{\partial^2 \alpha}{\partial x_1^2} + \frac{\partial^2 \alpha}{\partial x_2^2} + \frac{\partial^2 \alpha}{\partial x_3^2} = \lambda a^2 \alpha.$$

Then in the fourth case, by substituting the value of u into the differential equation (I), the functions α and β are easily determined, provided we note that in this case $\alpha e^{\lambda t}$ and $\beta e^{\mu t}$ can be complex conjugate quantities. [2]

Weber's remarks on XXII.

[1] (p. 380) These investigations contain the analytic execution of the results indicated in paper **XIII** 'The hypotheses on which geometry is based'. The question is that of the conditions under which a one second order differential expression can be transformed into another: in particular, one with constant coefficients. Since the first appearance of the above paper of Riemann, this question has been studied by Christoffel and Lipschitz, who obtained the same results as Riemann in different ways (*Crelle's Journal*, Vols. 70, 71, 72, 82). Later, R. Beez also took up the subject (*Schlömilch's Zeitschrift*, Vols. 20, 21, 24). In the first edition I appended some remarks based on an old (unpublished) study of R. Dedekind with a view to working out Riemann's account of the matter for the reader. The somewhat too brief exposition of these remarks may well have given rise to some doubts expressed in the above named papers. Accordingly I repeat the remarks here in somewhat greater detail.

Let the square of the line element in a space of n dimensions be

$$ds^2 = \sum_{i,i'} b_{i,i'} ds_i ds_{i'}.$$

The difference equation

$$(1) \qquad d\sum_i b_{i,\mu}\frac{ds_i}{dr} = \frac{1}{2}\,dr\sum_{i,i'}\frac{\partial b_{i,i'}}{\partial s_\mu}\,\frac{ds_i}{dr}\,\frac{ds_{i'}}{dr}$$

together with

$$\sum_{i,i'} b_{i,i'}\frac{ds_i}{dr}\,\frac{ds_{i'}}{dr} = 1$$

determines the shortest lines, where

$$r = \int\sqrt{\sum_{i,i'} b_{i,i'} ds_i ds_{i'}}$$

denotes the length of the shortest line from an arbitrary fixed point 0 to a variable point.

We take 0 as the point, in the neighborhood of which the relations of the n-dimensional space are to be investigated. We consider shortest lines drawn from this point in all directions, and introduce a system of new variables via the substitutions

$$x_1 = rc_1,\ x_2 = rc_2, \dots, x_n = rc_n,$$

where the quantities c_i are defined by

$$c_i = \left(\frac{ds_i}{dr}\right)_0 .$$

Thus these quantities are connected by the relation

$$\sum_{i,i'} b^{(0)}_{i,i'} c_i c_{i'} = 1,$$

and the c_i are constant along each shortest line issuing from the point 0. The c_i appear as constants of integration of the differential equation (1). Naturally the complete solution of this differential equation is required in order to represent the variables x_i as functions of the original variables s_i.

The characteristic property of these new variables, which we may call the *central coordinates* of a variable point m with respect to 0, is that they vanish at 0, and that their values grow in proportion to the length r as we proceed along a shortest line. These properties are invariant when $x_1, \dots, x_n$ is replaced by a system of n independent linear forms in these variables, with constant coefficients. In this way we arrive at Riemann's requirement $r^2 = \sum x_i^2$ in **XIII**, part II, §2. However, this is of no consequence, and will not be taken into account further here.

Now let the square of the line element, expressed in terms of the new variables, be

$$ds^2 = \sum_{i,i'} a_{i,i'} dx_i dx_{i'}.$$

If we proceed along a shortest line from 0, so that $ds^2 = dr^2$, we readily see that

$$\text{(2)} \qquad \sum_{i,i'} a_{i,i'} c_i c_{i'} = \sum_{i,i'} a^{(0)}_{i,i'} c_i c_{i'} = 1.$$

We now express the differential equation of shortest lines in terms of the new variables. This yields, for shortest lines issuing from the point 0,

$$d \sum_i a_{\mu,i} c_i = \frac{1}{2}\, dr \sum_{i,i'} \frac{\partial a_{i,i'}}{\partial x_\mu}\, c_i c_{i'}$$

from which it follows that

$$(3) \qquad \sum_{i,i'} p_{\mu,i,i'} x_i x_{i'} = 0.$$

Here, for brevity, we write (p. 378)

$$p_{\mu,i,i'} = \frac{\partial a_{i,\mu}}{\partial x_{i'}} + \frac{\partial a_{i',\mu}}{\partial x_i} - \frac{\partial a_{i,i'}}{\partial x_\mu}.$$

We may also write equation (3) in the form

$$(3') \qquad \sum_{i,i'} \frac{\partial a_{i,i'}}{\partial x_\mu} x_i x_{i'} = 2 \sum_{i,i'} \frac{\partial a_{i,\mu}}{\partial x_{i'}} x_i x_{i'}.$$

For brevity we now write

$$\omega_\mu = \sum_i a_{\mu,i} x_i; \;\; \frac{\partial \omega_\mu}{\partial x_\nu} = a_{\mu,\nu} + \sum_i \frac{\partial a_{\mu,i}}{\partial x_\nu} x_i.$$

This permits us to rewrite (3′) as

$$\omega_\mu + \sum_i \frac{\partial \omega_i}{\partial x_\mu} x_i = 2 \sum_i \frac{\partial \omega_i}{\partial x_i} x_i.$$

Further, let

$$2\omega = \sum_i \omega_i x_i; \;\; 2\frac{\partial \omega}{\partial x_\mu} = \omega_\mu + \sum_i \frac{\partial \omega_i}{\partial x_\mu} x_i.$$

It follows that

$$\frac{\partial \omega}{\partial x_\mu} = \sum_i \frac{\partial \omega_\mu}{\partial x_i} x_i; \;\; \frac{\partial^2 \omega}{\partial x_\mu \partial x_\nu} = \frac{\partial \omega_\mu}{\partial x_\nu} + \sum_i \frac{\partial^2 \omega_\mu}{\partial x_i \partial x_\nu} x_i,$$

and

$$\frac{\partial \omega_\mu}{\partial x_\nu} - \frac{\partial \omega_\nu}{\partial x_\mu} + \sum_i \frac{\partial}{\partial x_i} \left(\frac{\partial \omega_\mu}{\partial x_\nu} - \frac{\partial \omega_\nu}{\partial x_\mu} \right) x_i = 0.$$

We conclude that the $\frac{\partial \omega_\mu}{\partial x_\nu} - \frac{\partial \omega_\nu}{\partial x_\mu}$ are homogeneous functions of order -1. Denoting such a function by $f(x_1, \ldots, x_n)$, we have

$$f(tx_1, tx_2, \ldots, tx_n) = t^{-1} f(x_1, x_2, \ldots, x_n).$$

If we now suppose that the coefficients $a_{i,i'}$ and their derivatives at 0 have definite finite values, we conclude, on setting $t = 0$, that the function f vanishes identically; thus $\frac{\partial\omega_\mu}{\partial x_\nu} = \frac{\partial\omega_\nu}{\partial x_\mu}$. Further, then,

$$\sum_i \frac{\partial a_{\mu,i}}{\partial x_\nu} x_i = \sum_i \frac{\partial a_{\nu,i}}{\partial x_\mu} x_i.$$

With the help of (3′), this yields

$$\sum_{i,i'} \frac{\partial a_{\mu,i}}{\partial x_{i'}} x_i x_{i'} = \sum_{i,i'} \frac{\partial a_{i,i'}}{\partial x_\mu} x_i x_{i'} = 0.$$

Via integration of the differential equation of the shortest line,

$$\sum_i a_{\mu,i} c_i = \sum_i a^{(0)}_{\mu,i} c_i; \tag{4}$$

or, on multiplying by r,

$$\sum_i a_{\mu,i} x_i = \sum_i a^{(0}_{\mu,i} x_i.$$

These are all identities, that is, they hold for each choice of the system of independent variables x_i.

Now let $t_{i,i'} = t_{i',i}$ be any functions of $x_1, x_2, \ldots, x_n$, which together with their partial derivatives of order ≤ 3 have definite finite values at 0, and which satisfy the identity

$$\sum_{i,i'} t_{i,i'} x_i x_{i'} = 0.$$

Differentiating three times, and then setting $x_i = 0$, we obtain the following equation valid for the point 0:

$$t_{i,i'} = 0, \ \frac{\partial t_{i,i'}}{\partial x_{i''}} + \frac{\partial t_{i',i''}}{\partial x_i} + \frac{\partial t_{i,i''}}{\partial x_{i'}} = 0.$$

In this equation take $t_{i,i'} = p_{\mu,i,i'}$. This yields the equation, at the point 0:

$$p_{i,i',i''} = 0; \ \frac{\partial p_{i,i',i''}}{\partial x_{i'''}} + \frac{\partial p_{i,i'',i'''}}{\partial x_{i'}} + \frac{\partial p_{i,i',i'''}}{\partial x_{i''}} = 0.$$

From the first of these equations, on adding $p_{i',i,i''} = 0$, one obtains

$$\text{(5)} \qquad \frac{\partial a_{i,i'}}{\partial x_{i''}} = 0 \quad \text{at the point } 0;$$

from the second,

$$2\left(\frac{\partial^2 a_{i,i'}}{\partial x_{i''}\partial x_{i'''}} + \frac{\partial^2 a_{i,i''}}{\partial x_{i'''}\partial x_{i'}} + \frac{\partial^2 a_{i,i'''}}{\partial x_{i'}\partial x_{i''}}\right) = \frac{\partial^2 a_{i'',i'''}}{\partial x_i \partial x_{i'}} + \frac{\partial^2 a_{i''',i'}}{\partial x_i \partial x_{i''}} + \frac{\partial^2 a_{i',i''}}{\partial x_i \partial x_{i'''}}.$$

Exchanging i and i', adding, and denoting by S the sum of the six derivatives of the form $\frac{\partial^2 a_{i,i'}}{\partial x_{i''}\partial x_{i'''}}$, it follows that

$$S = 3\left(\frac{\partial^2 a_{i'',i'''}}{\partial x_i \partial x_{i'}} - \frac{\partial^2 a_{i,i'}}{\partial x_{i''}\partial x_{i'''}}\right).$$

Since S is invariant when i'', i''' is exchanged with i, i', we have

$$\text{(6)} \qquad \frac{\partial^2 a_{i'',i'''}}{\partial x_i \partial x_{i'}} = \frac{\partial^2 a_{i,i'}}{\partial x_{i''}\partial x_{i'''}},$$

$$\text{(7)} \qquad \frac{\partial^2 a_{i,i'}}{\partial x_{i''}\partial x_{i'''}} + \frac{\partial^2 a_{i,i''}}{\partial x_{i'''}\partial x_{i'}} + \frac{\partial^2 a_{i,i'''}}{\partial x_{i'}\partial x_{i''}} = \frac{\partial^2 a_{i'',i'''}}{\partial x_i \partial x_{i'}} + \frac{\partial^2 a_{i''',i'}}{\partial x_i \partial x_{i''}} + \frac{\partial^2 a_{i',i''}}{\partial x_i \partial x_{i'''}} = 0$$

at the point 0.

Now let us understand by $a_{i,i'}$, $\frac{\partial a_{i,i'}}{\partial x_{i''}}$, $\frac{\partial^2 a_{i,i'}}{\partial x_{i''}\partial x_{i'''}}$ the values of these quantities at the point 0. Under this hypothesis we have, for the line element ds_0 issuing from the point 0,

$$ds_0^2 = \sum_{i,i'} a_{i,i'} dx_i \, dx_{i'}.$$

For a line element ds issuing from a point infinitely close to 0 with (infinitely small) coordinates $x_1, x_2, \ldots, x_n$, including terms up to second order, we have

$$ds^2 = \sum_{i,i'} a_{i,i'} dx_i \, dx_{i'} + \sum_{i,i',i''} \frac{\partial a_{i,i'}}{\partial x_{i''}} x_{i''} dx_i \, dx_{i'} + \frac{1}{2} \sum_{i,i',i'',i'''} \frac{\partial^2 a_{i,i'}}{\partial x_{i''}\partial x_{i'''}} x_{i''} x_{i'''} dx_i \, dx_{i'}.$$

By (5), the second term vanishes, while the third,

$$\Theta = \frac{1}{2} \sum_{i,i',i'',i'''} \frac{\partial^2 a_{i,i'}}{\partial x_{i''} \partial x_{i'''}} x_{i''} x_{i'''} dx_i \, dx_{i'}$$

is the expression for the variation of the above n-dimensional space from flatness in the surface direction defined by x_i, dx_i. With the help of (6) and (7), Θ can be given a form from which it is apparent that it depends only on the combinations $x_i \, dx_{i'} - x_{i'} dx_i$. Namely, by interchange of indices we write Θ in the following four forms:

$$\begin{aligned}
\Theta &= \frac{1}{2} \sum \frac{\partial^2 a_{i,i'}}{\partial x_{i''} \partial x_{i'''}} x_{i''} x_{i'''} dx_i \, dx_{i'} \\
&= \frac{1}{2} \sum \frac{\partial^2 a_{i',i''}}{\partial x_i \partial x_{i'''}} x_i x_{i'''} dx_{i'} dx_{i''} \\
&= \frac{1}{2} \sum \frac{\partial^2 a_{i,i'''}}{\partial x_{i'} \partial x_{i''}} x_{i'} x_{i''} dx_i dx_{i'''} \\
&= \frac{1}{2} \sum \frac{\partial^2 a_{i'',i'''}}{\partial x_i \partial x_{i'}} x_i x_{i'} dx_{i''} dx_{i'''}.
\end{aligned}$$

We apply (6) to the fourth expression, and apply (6) and (7) to the second and third expressions. On a further exchange of indices, this yields

$$\begin{aligned}
\Theta &= \frac{1}{2} \sum \frac{\partial^2 a_{i,i'}}{\partial x_{i''} \partial x_{i'''}} x_{i''} x_{i'''} dx_i \, dx_{i'}, \\
\frac{1}{2}\Theta &= -\frac{1}{2} \sum \frac{\partial^2 a_{i,i'}}{\partial x_{i''} \partial x_{i'''}} x_i x_{i'''} dx_{i'} dx_{i''}, \\
\frac{1}{2}\Theta &= -\frac{1}{2} \sum \frac{\partial^2 a_{i,i'}}{\partial x_{i''} \partial x_{i'''}} x_{i'} x_{i''} dx_i \, dx_{i'''}, \\
\Theta &= \frac{1}{2} \sum \frac{\partial^2 a_{i,i'}}{\partial x_{i''} \partial x_{i'''}} x_i x_{i'} dx_{i''} dx_{i'''}.
\end{aligned}$$

If we add these four equations, we obtain

$$\text{(8)} \qquad \Theta = \frac{1}{6} \sum_{i,i'';i',i'''} \frac{\partial^2 a_{i,i'}}{\partial x_{i''} \partial x_{i'''}} (x_i \, dx_{i''} - x_{i''} dx_i)(x_{i'} dx_{i'''} - x_{i'''} dx_{i'}).$$

This expression for Θ is, however, derived under the hypothesis that the variables x_i signify the central coordinates. We still have the task of transforming it for arbitrary coordinates. This completes Riemann's prescription that one should express it in a form that is visibly independent of the variables used.

First of all, maintaining central coordinates, we replace the infinitely small coordinates $x_1, x_2, \ldots, x_n$ by the proportional differentials $\delta x_1, \delta x_2, \ldots, \delta x_n$. Thus

$$(9) \qquad \Theta = \frac{1}{2} \sum \frac{\partial^2 a_{i,i'}}{\partial x_{i''} \partial x_{i'''}} dx_i dx_{i'} \delta x_{i''} \delta x_{i'''}.$$

We choose the (otherwise arbitrary) differentials $dx_i, \delta x_i$ so that

$$(10) \qquad ddx_i = 0, \quad d\delta x_i = 0, \quad \delta dx_i = 0, \quad \delta\delta x_i = 0$$

which occurs for example in the case of constant $dx_i, \delta x_i$. This has the consequence that d and δ are interchangeable, that is, for an arbitrary function of position ϕ,

$$(\mathrm{I}) \qquad d\delta\phi = \delta d\phi.$$

Under this hypothesis one can derive from (5), (6), (7) the formula

$$\begin{aligned} dd \sum_{i,i'} a_{i,i'} \delta x_i \delta x_{i'} &= \delta\delta \sum_{i,i'} a_{i,i'} dx_i dx_{i'} \\ &= -2d\delta \sum_{i,i'} a_{i,i'} dx_i \delta x_{i'} \end{aligned}$$

and, with the help of this formula,

$$\begin{aligned} (\mathrm{II}) \qquad \Theta &= \frac{1}{2} dd \sum_{i,i'} a_{i,i'} \delta x_i \, \delta x_{i'} \\ &= \frac{1}{6} \Bigg\{ dd \sum_{i,i'} a_{i,i'} \delta x_i \delta x_{i'} - 2d\delta \sum_{i,i'} a_{i,i'} dx_i \delta x_{i'} \\ &\qquad + \delta\delta \sum_{i,i'} a_{i,i'} dx_i dx_{i'} \Bigg\}. \end{aligned}$$

Denote by δ' a variation, arbitrary except that it is interchangeable with d and δ. Now (5) and (10) yield the equations

$$\delta' \sum_{i,i'} a_{i,i'} dx_i \delta x_{i'} = \sum_{i,i'} a_{i,i'} d\delta' x_i \delta x_{i'} + \sum_{i,i'} a_{i,i'} dx_i \delta\delta' x_{i'},$$

$$d \sum_{i,i'} a_{i,i'} \delta' x_i \delta x_{i'} = \sum_{i,i'} a_{i,i'} d\delta' x_i \delta x_{i'},$$

$$\delta \sum_{i,i'} a_{i,i'} dx_i \delta' x_{i'} = \sum_{i,i'} a_{i,i'} dx_i \delta\delta' x_{i'}.$$

From these equations,

(III) $$\delta \sum_{i,i'} a_{i,i'} dx_i \delta x_{i'} - d \sum_{i,i'} a_{i,i'} \delta' x_i \delta x_{i'} - \delta \sum_{i,i'} a_{i,i'} dx_i \delta' x_{i'} = 0.$$

If we set $d = \delta$, then

(IV) $$\delta' \sum_{i,i'} a_{i,i'} dx_i dx_{i'} - 2d \sum_{i,i'} a_{i,i'} dx_i \delta' x_{i'} = 0,$$

(V) $$\delta' \sum_{i,i'} a_{i,i'} \delta x_i \delta x_{i'} - 2\delta \sum_{i,i'} a_{i,i'} \delta x_i \delta' x_{i'} = 0.$$

If now we introduce, in place of the variables x_i, other variables s_i that are functions of them, we obtain for entirely arbitrary differentials d, δ a transformation

$$\sum_{i,i'} a_{i,i'} dx_i \delta x_{i'} = \sum_{i,i'} b_{i,i'} ds_i \delta s_{i'}.$$

Thus we obtain the transformed expression for Θ, replacing $a_{i,i'}, x_i$ by $b_{i,i'}, s_i$ in (II); in other words the x_i need no longer be central coordinates in (II), but may be arbitrary coordinates. Admittedly the conditions (5), (6), (7), (10) will then no longer be valid. Nonetheless the conditions (I), (III), (IV), (V) will be satisfied for all systems of coordinates when they are satisfied for one system, for example the central coordinates. Thus if we make use of only the relations (I), (III), (IV), (V), in further transformation of (II), the results will be valid for arbitrary variables. The calculation is now, although somewhat long, altogether free of difficulties. On the right side of (II), due to the interchangeability of differentials, the differentials of third order cancel. We can extract the differentials of second order with the help of the following equations, consequences of (III), (IV), (V):

$$2\sum_i a_{i,i'} ddx_i = -\sum_{i,i'} p_{\nu,i,i'} dx_i dx_{i'},$$
$$2\sum_i a_{i,i'} d\delta x_i = -\sum_{i,i'} p_{\nu,i,i'} dx_i \delta x_{i'},$$
$$2\sum_i a_{i,i'} \delta\delta x_i = -\sum_{i,i'} p_{\nu,i,i'} \delta x_i \delta x_{i'}.$$

Here, as on p. 378, $p_{\nu,i,i'}$ denotes the quantity

$$p_{\nu,i,i'} = \frac{\partial a_{\nu,i}}{\partial x_{i'}} + \frac{\partial a_{\nu,i'}}{\partial x_i} - \frac{\partial a_{i,i'}}{\partial x_\nu}.$$

We obtain the expression

$$dd\sum_{i,i'} a_{i,i'}\delta x_i \delta x_{i'} - 2d\delta\sum_{i,i'} a_{i,i'} dx_i \delta x_{i'} + \delta\delta\sum_{i,i'} a_{i,i'} dx_i dx_{i'}$$
$$= \sum_{ii',i''i'''} (ii', i''i''')(dx_i\delta x_{i'} - \delta x_i dx_{i'})(dx_{i''}\delta x_{i'''} - \delta x_{i''} dx_{i'''}).$$

Here $(ii', i''i''')$ has the same significance as in Riemann's text (p. 379), and the sum is taken so that, of two pairs of indices i, i' and i', i, and likewise of two pairs i'', i''' and i''', i'', only one is employed.

From this expression we now obtain the curvature of our general space. Namely let

$$ds = \sqrt{\sum_{i,i'} a_{i,i'} dx_i dx'_i}, \quad \delta s = \sqrt{\sum_{i,i'} a_{i,i'} \delta x_i \delta x_{i'}}$$

be two line elements in the space and

$$\frac{\sum a_{i,i'} dx_i \delta x_{i'}}{ds\,\delta s} = \cos\theta$$

the cosine of the angle that they enclose.

The surface area of an infinitely small triangle formed by the line elements is then

$$\Delta = \frac{1}{2}\, ds\delta s \sin\theta.$$

This yields

$$\begin{aligned}
4\Delta^2 &= \sum_{i,i'} a_{i,i'} dx_i dx_{i'} \sum_{i,i'} a_{i,i'} \delta x_i \delta x_{i'} - \left(\sum_{i,i'} a_{i,i'} dx_i \delta x_{i'}\right)^2 \\
&= \sum_{ii',i''i'''} (a_{i,i''}a_{i',i'''} - a_{i',i''}a_{i,i'''})(dx_i\delta x_{i'} - \delta x_i dx_{i'})(dx_{i''}\delta x_{i'''} - \delta x_{i''} dx_{i'''})
\end{aligned}$$

$$-\frac{3}{8}\,\frac{dd\sum_{i,i'} a_{i,i'}\delta x_i \delta x_{i'}}{\Delta^2}$$

$$= -\frac{1}{2}\,\frac{dd\sum_{i,i'} a_{i,i'}\delta x_i\delta x_{i'} - 2d\delta\sum_{i,i'} a_{i,i'} dx_i\delta x_{i'} + \delta\delta\sum_{i,i'} a_{i,i'} dx_i dx_{i'}}{\sum_{i,i'} a_{i,i'} dx_i dx_{i'} \sum_{i,i'} a_{i,i'}\delta x_i\delta x_{i'} - \left(\sum_{i,i'} a_{i,i'} dx_i\delta x_{i'}\right)^2}$$

$$= -\frac{1}{2}\,\frac{\sum_{ii',i''i'''} (ii', i''i''')(dx_i\delta x_{i'} - \delta x_i dx_{i'})(dx_{i''}\delta x_{i'''} - \delta x_{i''} dx_{i'''})}{\sum_{ii',i''i'''} (a_{i,i''}a_{i',i'''} - a_{i,i'''}a_{i',i''})(dx_i\delta x_{i'} - \delta x_i dx_{i'})(dx_{i''}\delta x_{i'''} - \delta x_{i''} dx_{i'''})}$$

It now remains to show that this expression agrees with that which Gauss gave for the curvature of a surface, when we consider a surface formed of shortest lines in whose initial elements the variations of x are in the ratio

$$\alpha dx_1 + \beta\delta x_1 : \alpha dx_2 + \beta\delta x_2 : \ldots : \alpha dx_n + \beta\delta x_n,$$

where α and β denote arbitrary numbers.

As above, set $x_i = rc_i$, so that c_i is constant on every shortest line issuing from 0, and r is the length of this shortest line up to a variable point. As shown above,

$$\sum_{i,i'} a_{i,i'} c_i c_{i'} = \sum_{i,i'} a^{(0)}_{i,i'} c_i c_{i'} = 1.$$

We now take two fixed systems of numbers c_i, say $c_i^{(0)}$ and c'_i, as fundamental. We treat a variable system

$$c_i = \alpha c_i^{(0)} + \beta c'_i. \tag{11}$$

Then we have

$$\alpha^2 + 2\alpha\beta\cos(r^{(0)}, r') + \beta^2 = 1,$$

whence the quantities c_i become functions of a single variable, which we may take to be the angle ϕ that the initial element of r makes with the initial element of $r^{(0)}$. We obtain ϕ from the expression

$$\cos\phi = \sum_{i,i'} a_{i,i'}^{(0)} c_i c_{i'}^{(0)}.$$

Now allow variations of r, c_i by infinitely small quantities dr, dc_i, satisfying the condition

$$\sum_{i,i'} a_{i,i'}^{(0)} c_i dc_{i'} = 0.$$

With the help of equation (4), we obtain

$$\sum_{i,i'} a_{i,i'} c_i dc_{i'} = \sum_{i,i'} a_{i,i'}^{(0)} c_i dc_{i'} = 0. \tag{12}$$

We also have

$$dx_i = r dc_i + c_i dr$$

and accordingly

$$\begin{aligned} ds^2 &= \sum_{i,i'} a_{i,i'} dx_i dx_{i'} = dr^2 + r^2 \sum_{i,i'} a_{i,i'} dc_i dc_{i'} \\ &= dr^2 + r^2 \mu d\phi^2 \end{aligned}$$

with the abbreviation

$$\sum_{i,i'} a_{i,i'} dc_i dc_{i'} = \mu d\phi^2.$$

However we now have

$$\cos\phi = \sum_{i,i'} a_{i,i'}^{(0)} c_i c_{i'}^{(0)}, \quad -\sin\phi d\phi = \sum_{i,i'} a_{i,i'}^{(0)} c_i^{(0)} dc_{i'}. \tag{13}$$

From (11) follows an expression of form

$$dc_i = ac_i^{(0)} + bc_i;\ a = \beta d\alpha - \alpha d\beta,\ b = d\beta.$$

Thus from (12) and (13),

$$\begin{aligned} -\sin\phi d\phi &= a + b\cos\phi \\ 0 &= a\cos\phi + b. \end{aligned}$$

By elimination of a and b,

$$\sin\phi\, dc_i = d\phi(c_i\cos\phi - c_i^{(0)}).$$

From this we have

$$d\phi^2 = \sum_{i,i'} a_{i,i'}^{(0)} dc_i dc_{i'}$$

and also

$$\mu = \frac{\sum_{i,i'} a_{i,i'} dc_i dc_{i'}}{\sum_{i,i'} a_{i,i'}^{(0)} dc_i dc_{i'}}. \tag{14}$$

Let us write this expression as m^2/r^2, then we obtain Gauss's form of the line element on an arbitrary surface, namely

$$ds^2 = dr^2 + m^2 d\phi^2$$

(*Disquisitiones generales circa superficies curvas*, section 19). For the curvature, we obtain

$$k = -\frac{1}{m}\frac{\partial^2 m}{\partial r^2}.$$

Suppose now that the surface curves continuously at the point $r = 0$. Then at this point

$$m = 0,\ \frac{\partial m}{\partial r} = 1,\ \frac{\partial^2 m}{\partial r^2} = 0.$$

Thus at this point we have

$$k = -\frac{\partial^3 m}{\partial r^3}.$$

As for the function μ, we find that, at the same point,

$$\mu = 1, \quad \frac{\partial \mu}{\partial r} = 0, \quad k = -\frac{3}{2}\,\frac{\partial^2 \mu}{\partial r^2}.$$

The first two of these equations are satisfied in consequence of (14), (5). From the third,

$$k = -\frac{3}{2}\,\frac{\displaystyle\sum_{ii',i''i'''}\left(\frac{\partial^2 a_{i,i'}}{\partial x_{i''}\partial x_{i'''}}\right)_0 c_{i''}c_{i'''}dc_i dc_{i'}}{\displaystyle\sum_{i,i'} a^{(0)}_{i,i'}dc_i dc_{i'}},$$

which agrees with the expression found above.

We construct shortest lines issuing from the point 0, with initial directions defined via the equations (11); this gives a surface in the transcendental space. The coordinates of a point of this surface may be expressed via two independent variables. If we denote these by p, q, then there is an expression of form

$$ds^2 = Edp^2 + 2F\,dp\,dq + Gdq^2$$

for the square of the line element on the surface. Here E, F, G are functions of p and q. Take x, y, z to be a system of particular solutions of the simultaneous equations

$$\left(\frac{\partial x}{\partial p}\right)^2 + \left(\frac{\partial y}{\partial p}\right)^2 + \left(\frac{\partial z}{\partial p}\right)^2 = E,$$
$$\frac{\partial x}{\partial p}\,\frac{\partial x}{\partial q} + \frac{\partial y}{\partial p}\,\frac{\partial y}{\partial q} + \frac{\partial z}{\partial p}\,\frac{\partial z}{\partial q} = F,$$
$$\left(\frac{\partial x}{\partial q}\right)^2 + \left(\frac{\partial y}{\partial q}\right)^2 + \left(\frac{\partial z}{\partial q}\right)^2 = G,$$

then

$$ds^2 = dx^2 + dy^2 + dz^2.$$

If we view x, y, z as coordinates of a point in space, we obtain a surface upon which, by Riemann's expression, the surface in transcendental space can be developed. That is, there is a pointwise correspondence, without variation of line elements.

From these formulae one can easily derive the expression for the line element under the hypothesis of constant curvature that is given on p. Namely, let k have the constant value α; then

$$m = \frac{\sin \sqrt{\alpha}\, r}{\sqrt{\alpha}}.$$

When we replace the c_i by linear combinations of them for which

$$\sum c_i^2 = 1,$$

so that

$$d\phi^2 = \sum dc_i^2,$$

then we obtain for ds^2 the expression

$$ds^2 = dr^2 + \frac{\sin^2 \sqrt{\alpha}\, r}{\alpha} \sum dc_i^2.$$

Now let

$$x_i = \frac{2c_i}{\sqrt{\alpha}} \tan \frac{\sqrt{\alpha}\, r}{2}, \sum x_i^2 = \frac{4}{\alpha} \tan^2 \frac{\sqrt{\alpha}\, r}{2},$$

(a special case of this is the stereographic projection of the surface of the sphere on the plane). It follows that

$$\sum dx_i^2 = \frac{dr^2}{\cos^4 \frac{\sqrt{\alpha}\, r}{2}} + \frac{4}{\alpha} \tan^2 \frac{\sqrt{\alpha}\, r}{2} \sum dc_i^2$$

and

$$\begin{aligned} ds &= \cos^2 \frac{\sqrt{\alpha}\, r}{2} \sqrt{\sum dx_i^2} \\ &= \frac{1}{1 + \frac{\alpha}{4} \sum x_i^2} \sqrt{\sum dx_i^2}. \end{aligned}$$

[2] (p. 381) The complete verification of the final result stated here appears to require complicated calculations that I was only partly able to reconstruct from the very incomplete fragments available. The part that I could decipher is presented here in the hope that it could serve as the basis for a fresh investigation leading to the complete result.

We first answer the question as to the cases where the temperature depends on only one variable apart from time. In these cases the differential equation according to which heat is conducted takes the form

$$(1) \qquad a\,\frac{\partial^2 u}{\partial \alpha^2} + b\,\frac{\partial u}{\partial \alpha} = \frac{\partial u}{\partial t}.$$

If now the coefficients a, b are not functions of a single variable α, this differential equation splits into two equations as follows:

$$a'\,\frac{\partial^2 u}{\partial \alpha^2} + b'\,\frac{\partial u}{\partial \alpha} = \frac{\partial u}{\partial t}\ ;\ a''\,\frac{\partial^2 u}{\partial \alpha^2} + b''\,\frac{\partial u}{\partial \alpha} = 0,$$

where a', b', a'', b'' depend only on α.

By introducing a new variable in place of α, the second of these equations may be brought to the form $\frac{\partial^2 u}{\partial \alpha^2} = 0$, so that u takes the form $u_1\alpha + u_2$. Here u_1, u_2 are functions of time alone. The first of the above equations now assumes the form

$$(c\alpha + c_1)\frac{\partial u}{\partial \alpha} = \frac{\partial u}{\partial t},$$

where c, c_1 are constants. It follows further that

$$cu_1 = \frac{\partial u_1}{\partial t}\ ,\ 0 = \frac{\partial u_2}{\partial t},$$

so that u has the form $\alpha e^{\lambda t}$ + const.

However, if in the differential equation (1) the coefficients a, b are functions of α alone, then without loss of generality we can assume that $b = 0$ (by introduction of a new variable for α). Since the differential equation (1) must arise from a transformation of the equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \frac{\partial u}{\partial t},$$

our task reduces to the following.

We must find all functions α of coordinates x, y, z, which simultaneously satisfy the differential equations

$$\Delta = \frac{\partial^2 \alpha}{\partial x^2} + \frac{\partial^2 \alpha}{\partial y^2} + \frac{\partial^2 \alpha}{\partial z^2} = 0,\ D = \left(\frac{\partial \alpha}{\partial x}\right)^2 + \left(\frac{\partial \alpha}{\partial y}\right)^2 + \left(\frac{\partial \alpha}{\partial z}\right)^2 = f(\alpha).$$

For brevity, let

$$\frac{\partial \alpha}{\partial x} = p, \ \frac{\partial \alpha}{\partial y} = q, \ \frac{\partial \alpha}{\partial z} = r, \ p^2 + q^2 + r^2 = m.$$

We distinguish four cases.

1. If p, q, r are independent functions of the coordinates x, y, z, then α is a function of m, $\alpha = \phi(m)$, and we can introduce p, q, r as independent variables in place of x, y, z. Let

$$s = \alpha - px - qy - rz, \ ds = -xdp - ydq - zdr.$$

Then

$$x = -\frac{\partial s}{\partial p}, \ y = -\frac{\partial s}{\partial q}, \ z = -\frac{\partial s}{\partial r},$$
$$\alpha = s - p\frac{\partial s}{\partial p} - q\frac{\partial s}{\partial q} - r\frac{\partial s}{\partial r} = \phi(m).$$

We let

$$s = \psi(m) + t$$

and determine $\psi(m)$ via the differential equation

$$\psi(m) - 2m\psi'(m) = \phi(m).$$

This yields for t the first order partial differential equation

$$t - p\frac{\partial t}{\partial p} - q\frac{\partial t}{\partial q} - r\frac{\partial t}{\partial r} = 0$$

whose general solution is

$$t = p\chi\left(\frac{q}{p}, \frac{r}{p}\right) = p\chi(\beta, \gamma).$$

Here χ denotes an arbitrary function, and for brevity we write

$$\beta = \frac{q}{p}, \ \gamma = \frac{r}{p}.$$

Thus we have

$$
(2)\qquad \begin{cases} -x = \dfrac{\partial s}{\partial p} = 2p\psi'(m) + \chi - \beta\chi'(\beta) - \gamma\chi'(\gamma), \\[2ex] -y = \dfrac{\partial s}{\partial q} = 2q\psi'(m) + \chi'(\beta), \\[2ex] -z = \dfrac{\partial s}{\partial r} = 2r\psi'(m) + \chi'(\gamma). \end{cases}
$$

From the equation

$$
\Delta = \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} + \frac{\partial r}{\partial z} = 0,
$$

it follows by introducing p, q, r as independent variables that

$$
\frac{\partial y}{\partial q}\frac{\partial z}{\partial r} - \frac{\partial z}{\partial q}\frac{\partial y}{\partial r} + \frac{\partial z}{\partial r}\frac{\partial x}{\partial p} - \frac{\partial x}{\partial r}\frac{\partial z}{\partial p} + \frac{\partial x}{\partial p}\frac{\partial y}{\partial q} - \frac{\partial y}{\partial p}\frac{\partial x}{\partial q} = 0,
$$

or, by substitution from (2),

$$
\begin{aligned}
& m(12\psi'(m)^2 + 16m\psi'(m)\psi''(m)) \\
& + \sqrt{m}(4\psi'(m) + 4m\psi''(m))\sqrt{1+\beta^2+\gamma^2}\left\{(\beta^2+1)\frac{\partial^2\chi}{\partial\beta^2} + 2\beta\gamma\frac{\partial^2\chi}{\partial\beta\gamma}\right. \\
& \left. + (\gamma^2+1)\frac{\partial^2\chi}{\partial\gamma^2}\right\} \\
& + (1+\beta^2+\gamma^2)^2\left(\frac{\partial^2\chi}{\partial\beta^2}\frac{\partial^2\chi}{\partial\gamma^2} - \left(\frac{\partial^2\chi}{\partial\beta\partial\gamma}\right)^2\right) = 0.
\end{aligned}
$$

Since m, β, γ are independent variables, this equation splits into the following three:

$$
(3)\qquad \frac{\partial^2\chi}{\partial\beta^2}\frac{\partial^2\chi}{\partial\gamma^2} - \left(\frac{\partial^2\chi}{\partial\beta\partial\gamma}\right)^2 = \frac{k}{(1+\beta^2+\gamma^2)^2},
$$

$$
(4)\qquad (\beta^2+1)\frac{\partial^2\chi}{\partial\beta^2} + 2\beta\gamma\frac{\partial^2\chi}{\partial\beta\partial\gamma} + (\gamma^2+1)\frac{\partial^2\chi}{\partial\gamma^2} = \frac{k_1}{\sqrt{1+\beta^2+\gamma^2}},
$$

(5)

$$
m(12\psi'(m)^2 + 16m\psi'(m)\psi''(m)) + k_1\sqrt{m}\,(4\psi'(m) + 4m\psi''(m)) + k = 0.
$$

Here k, k_1 denote undetermined constants. We introduce in place of the function χ a new function χ_1 via the equation

$$\chi = \frac{1}{2}\, k_1 \sqrt{1+\beta^2+\gamma^2} + \chi_1 .$$

Equations (3), (4) become the following:

$$\text{(6)} \qquad \frac{\partial^2 \chi_1}{\partial \beta^2}\, \frac{\partial^2 \chi_1}{\partial \gamma^2} - \left(\frac{\partial^2 \chi_1}{\partial \beta \partial \gamma} \right)^2 = \frac{k'}{(1+\beta^2+\gamma^2)^2},$$

$$\text{(7)} \qquad (\beta^2+1)\, \frac{\partial^2 \chi_1}{\partial \beta^2} + 2\beta\gamma\, \frac{\partial^2 \chi_1}{\partial \beta \partial \gamma} + (\gamma^2+1)\, \frac{\partial^2 \chi_1}{\partial \gamma^2} = 0.$$

However, these equations can only hold simultaneously if χ_1 is a linear function of β and γ, so that $k' = 0$. We then treat

$$\chi_1 - \beta\, \frac{\partial \chi_1}{\partial \beta} - \gamma\, \frac{\partial \chi_1}{\partial \gamma},\ \frac{\partial \chi_1}{\partial \beta},\ \frac{\partial \chi_1}{\partial \gamma}$$

as rectangular coordinates. Then (6) is the differential equation of a surface with constant curvature; (7) is that of a minimal surface. These two properties are known to occur together only for planes.

Hence χ has an expression of form

$$\chi = a + b\beta + c\gamma + \frac{1}{2}\, k_1 \sqrt{1+\beta^2+\gamma^2}$$

with constants a, b, c. The equations (2) become the following:

$$x + a = -\frac{\frac{1}{2}\, k_1 + 2\sqrt{m}\, \psi'(m)}{\sqrt{1+\beta^2+\gamma^2}},$$

$$y + b = -\frac{\left(\frac{1}{2}\, k_1 + 2\sqrt{m}\, \psi'(m)\right) \beta}{\sqrt{1+\beta^2+\gamma^2}},$$

$$z + c = -\frac{\left(\frac{1}{2}\, k_1 + 2\sqrt{m}\, \psi'(m)\right) \gamma}{\sqrt{1+\beta^2+\gamma^2}},$$

$$(x+a)^2 + (y+b)^2 + (z+c)^2 = \left(\frac{1}{2}\, k_1 + 2\sqrt{m}\, \psi'(m) \right)^2 .$$

It follows that the surfaces $\alpha = \text{const.}$ or $m = \text{const.}$ are concentric spheres.

2. If the variables p, q, r are related by an equation that is free of one of the coordinates x, y, z, then we may view r as a function of p, q, and we have

$$dr = adp + bdq,$$

where we write

$$a = \frac{\partial r}{\partial p}\ , \ b = \frac{\partial r}{\partial q}\ , \ \frac{\partial a}{\partial q} = \frac{\partial b}{\partial p}.$$

It follows that

$$\frac{\partial p}{\partial z} = a\frac{\partial p}{\partial x} + b\frac{\partial p}{\partial y}\ , \frac{\partial q}{\partial z} = a\frac{\partial q}{\partial x} + b\frac{\partial q}{\partial y}\ , \ \frac{\partial r}{\partial z} = a\frac{\partial r}{\partial x} + b\frac{\partial r}{\partial y}.$$

Suppose now that

$$p^2 + q^2 + r^2 = \text{const.} \tag{8}$$

does not hold. Then α depends on the same two variables as p, q, r. Hence

$$r = ap + bq$$

and by differentiation

$$p\frac{\partial a}{\partial p} + q\frac{\partial b}{\partial p} = 0;\ p\frac{\partial a}{\partial q} + q\frac{\partial b}{\partial q} = 0,$$

$$\frac{\partial a}{\partial p}\frac{\partial b}{\partial q} - \frac{\partial a}{\partial q}\frac{\partial b}{\partial p} = 0. \tag{9}$$

Now we write as before, and also in the case where equation (8) holds,

$$s = \alpha - xp - yq - zr,$$
$$ds = -xdp - ydq - zdr = -(x + az)dp - (y + bz)dq.$$

It follows that s only depends on p and q, and

$$\frac{\partial s}{\partial p} = -(x + az)\ , \ \frac{\partial s}{\partial q} = -(y + bz). \tag{10}$$

We now introduce p, q and z as independent variables in the equation

$$\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} + \frac{\partial r}{\partial z} = 0.$$

It follows that

$$\frac{\partial x}{\partial p} + \frac{\partial y}{\partial q} - a\left(\frac{\partial y}{\partial q}\,\frac{\partial x}{\partial z} - \frac{\partial x}{\partial q}\,\frac{\partial y}{\partial z}\right) - b\left(\frac{\partial x}{\partial p}\,\frac{\partial y}{\partial z} - \frac{\partial y}{\partial q}\,\frac{\partial x}{\partial z}\right) = 0.$$

With the help of (10) we now obtain

$$z\left\{\frac{\partial a}{\partial p}\,(1+b^2) - ab\left(\frac{\partial a}{\partial q} + \frac{\partial b}{\partial p}\right) + \frac{\partial b}{\partial q}\,(1+a^2)\right\}$$
$$+\frac{\partial^2 s}{\partial p^2}\,(1+b^2) - 2ab\,\frac{\partial^2 s}{\partial p\partial q} + \frac{\partial^2 s}{\partial q^2}\,(1+a^2) = 0.$$

Since now a, b, s are independent of z, this equation splits into the following two:

$$(11)\qquad \frac{\partial^2 s}{\partial p^2}\,(1+b^2) - 2ab\,\frac{\partial^2 s}{\partial p\partial q} + \frac{\partial^2 s}{\partial q^2}\,(1+a^2) = 0,$$

$$(12)\qquad \frac{\partial a}{\partial p}\,(1+b^2) - ab\left(\frac{\partial a}{\partial q} + \frac{\partial b}{\partial p}\right) + \frac{\partial b}{\partial q}\,(1+a^2) = 0.$$

We now treat p, q, r as rectangular coordinates. Then (12) is the differential equation of a minimal surface, which by (8) or (9) must be either a sphere or a surface developable in the plane. These can only happen at once if the surface is a plane, so that a, b are constants, which can be taken to be 0 by a suitable choice of the direction of the z-axis. It now follows from (11) that

$$(13)\qquad \frac{\partial^2 s}{\partial p^2} + \frac{\partial^2 s}{\partial q^2} = 0.$$

Further, as in the first case,

$$s = \psi(m) + p\chi\left(\frac{q}{p}\right),$$
$$m = p^2 + q^2,\ r = 0,$$
$$-x = \frac{\partial s}{\partial p} = \psi'(m)2p + \chi(\beta) - \beta\chi'(\beta),$$
$$-y = \frac{\partial s}{\partial q} = \psi'(m)2q + \chi'(\beta).$$

Here we write $\beta = \frac{q}{p}$.

It now follows from (13) that

$$\sqrt{m}\,(4\psi'(m) + 4m\,\psi''(m)) + (1+\beta^2)^{3/2}\chi''(\beta) = 0.$$

This equation divides into the following two:

$$\sqrt{m}\,(4\psi'(m) + 4m\,\psi''(m)) = -k,$$
$$\chi''(\beta) = \frac{k}{(1+\beta^2)^{3/2}}$$

with a constant k. Integrating this last equation yields

$$\chi(\beta) = k\sqrt{1+\beta^2} + a + b\beta$$

with a, b arbitrary constants. Accordingly we have

$$x + a = -\frac{2\psi'(m)\sqrt{m} + k}{\sqrt{1+\beta^2}},$$

$$y + b = -\frac{(2\psi'(m)\sqrt{m} + k)\beta}{\sqrt{1+\beta^2}},$$

$$(x+a)^2 + (y+b)^2 = (2\psi'(m)\sqrt{m} + k)^2.$$

Hence in this case the isothermal surfaces are cylinders with circular cross-section and common axis.

The third case, in which p, q, r are functions of one and the same variable, cannot occur. Namely, let

$$p = \psi_1(\mu),\ q = \psi_2(\mu),\ r = \psi_3(\mu).$$

From the equations

$$\frac{\partial q}{\partial z} = \frac{\partial r}{\partial y},\ \frac{\partial r}{\partial x} = \frac{\partial p}{\partial z},\ \frac{\partial p}{\partial y} = \frac{\partial q}{\partial x},$$
$$\psi_1'(\mu) : \psi_2'(\mu) : \psi_3'(\mu) = \frac{\partial \mu}{\partial x} : \frac{\partial \mu}{\partial y} : \frac{\partial \mu}{\partial z},$$

and the equation $\Delta = 0$, it follows that

$$\psi_1'(\mu)\,\frac{\partial \mu}{\partial z} + \psi_2'(\mu)\,\frac{\partial \mu}{\partial y} + \psi_3'(\mu)\,\frac{\partial \mu}{\partial z} = 0,$$

which is evidently a contradiction.

There remains only the fourth case in which p, q, r are constant, and in this case the family of isothermal surfaces consists of parallel planes.

Concerning the more general question of when the temperature depends only on two variables apart from time, the first case, characterized in the text by $m = 1$, can be answered in the following way.

In this case we have the quadratic form

$$\begin{pmatrix} 0, & 0, & c \\ a', & b', & c' \end{pmatrix}$$

in which a', b' are linear functions of γ, while c' is independent of γ. Further, the determinant

$$\begin{vmatrix} 0 & c' & b' \\ c' & 0 & a' \\ b' & a' & c \end{vmatrix} = 2a'b'c' - cc'^2$$

is constant. The adjoint form of this one is

$$-(a'd\alpha + b'd\beta - c'd\gamma)^2 + 2(2a'b' - cc')d\alpha d\beta$$

in which $2a'b' - cc'$ is independent of γ.

Now, by introducing as a new variable a linear function of γ, in place of γ, we can transform this form into the simpler one

$$(ad\alpha + cd\gamma)^2 + 2md\alpha d\beta.$$

Here a is a linear function of γ, c; and m is independent of γ. We now need to find the cases in which this form can be transformed into another with constant coefficients, or in particular into the form $dx^2 + dy^2 + dz^2$.

To this end we form the equation $(ii', i''i''') = 0$ (p. 379), which in this case takes the form

(1,1) $$m\frac{\partial^2 c}{\partial \beta^2} - \frac{\partial c}{\partial \beta}\frac{\partial m}{\partial \beta} = 0,$$

(2,2) $$mc\left(\frac{\partial^2 c}{\partial \alpha^2} - \frac{\partial^2 a}{\partial \alpha \partial \gamma}\right) + \left(\frac{\partial a}{\partial \gamma} - \frac{\partial c}{\partial \alpha}\right)\left(c\frac{\partial m}{\partial \alpha} + m\frac{\partial a}{\partial \gamma}\right) = 0,$$

(3,3) $$2mc\left(\frac{\partial^2 a}{\partial \beta^2} - 2\frac{\partial^2 m}{\partial \alpha \partial \beta}\right) + 4c\frac{\partial m}{\partial \beta}\left(\frac{\partial m}{\partial \alpha} - a\frac{\partial a}{\partial \beta}\right) - \frac{m}{c}\left(\frac{\partial (ac)}{\partial \beta}\right)^2 = 0,$$

(2,3)
$$2mc\left(\frac{\partial^2(a^2)}{\partial\beta\partial\gamma}-\frac{\partial^2(ac)}{\partial\alpha\partial\beta}\right)+4m\,\frac{\partial c}{\partial\beta}\left(a\,\frac{\partial c}{\partial\alpha}-a\,\frac{\partial a}{\partial\gamma}+c\,\frac{\partial a}{\partial\alpha}\right)$$
$$+2c\left(c\,\frac{\partial a}{\partial\beta}-a\,\frac{\partial c}{\partial\beta}\right)\left(\frac{\partial m}{\partial\alpha}-a\,\frac{\partial a}{\partial\beta}\right)-2m\,\frac{\partial c}{\partial\alpha}\,\frac{\partial(ac)}{\partial\beta}$$
$$+a\,\frac{\partial(ac)}{\partial\beta}\left(c\,\frac{\partial a}{\partial\beta}-a\,\frac{\partial c}{\partial\beta}\right)=0,$$

(3,1)
$$2mc\,\frac{\partial^2(ac)}{\partial\beta^2}-2c\,\frac{\partial(ac)}{\partial\beta}\,\frac{\partial m}{\partial\beta}-2m\,\frac{\partial c}{\partial\beta}\,\frac{\partial(ac)}{\partial\beta}=0,$$

(1,2)
$$2m\left(2c\,\frac{\partial^2 c}{\partial\alpha\partial\beta}-\frac{\partial^2(ac)}{\partial\beta\partial\gamma}\right)+\left(c\,\frac{\partial a}{\partial\beta}-a\,\frac{\partial c}{\partial\beta}\right)^2=0.$$

From (1,2) it follows that $c\,\frac{\partial a}{\partial\beta}-a\,\frac{\partial c}{\partial\beta}$, and likewise $\frac{\partial\left(\frac{a}{c}\right)}{\partial\beta}$, is independent of γ. Accordingly let $a=a_1+\gamma a_2$. It follows that a_2 is of the form $cf(\alpha)$, and $f(\alpha)$ is independent of β.

Now we have

$$(a\,d\alpha+c\,d\gamma)^2+2m\,d\alpha\,d\beta=(a_1d\alpha+c(f(\alpha)d\alpha+d\gamma))^2+2md\alpha\,d\beta.$$

If we introduce in place of γ a new variable $\gamma+\int f(\alpha)d\alpha$, the quadratic form becomes another of the same form, but in which a is independent of γ. With this hypothesis, the equation (2,2) takes the form

$$m\,\frac{\partial^2 c}{\partial\alpha^2}-\frac{\partial c}{\partial\alpha}\,\frac{\partial m}{\partial\alpha}=0.$$

In conjunction with (1,1) this leads to

$$\frac{\partial}{\partial\alpha}\left(\log\,\frac{\partial c}{\partial\alpha}\right)=\frac{\partial(\log m)}{\partial\alpha},\quad\frac{\partial}{\partial\beta}\left(\log\,\frac{\partial c}{\partial\beta}\right)=\frac{\partial(\log m)}{\partial\beta},$$

so that

$$\frac{\partial c}{\partial\alpha}=m\phi(\beta),\quad\frac{\partial c}{\partial\beta}=m\psi(\alpha).$$

We now distinguish three cases.

1) If $\phi(\beta)=\psi(\alpha)=0$, then $c=$ const., and from (1,2) it follows that $\frac{\partial a}{\partial\beta}=0$. Introducing in place of γ a new variable $c\gamma+\int a\,d\alpha$, we find in the quadratic form that $a=0, c=1$. From (3,3) it now follows that

$$\frac{\partial^2(\log m)}{\partial\alpha\partial\beta}=0,\quad 2m=\chi(\alpha)\theta(\beta).$$

If we introduce in place of α, β the variables $\int \chi(\alpha) d\alpha$, $\int \theta(\beta) d\beta$, we obtain the quadratic form

$$d\gamma^2 + d\alpha\, d\beta.$$

By substituting $\alpha = x + iy$, $\beta = x - iy$, $\gamma = z$, this becomes

$$dx^2 + dy^2 + dz^2.$$

Thus in this case the isotherms $\alpha =$ const., $\beta =$ const. are parallel straight lines.

2) If $\phi(\beta) = 0$, while $\psi(\alpha)$ is not zero, then c is independent of α. From (1,2) it follows that $\frac{a}{c}$ is independent of β. In a similar way to the above, we now find that a vanishes, and further that

$$\frac{1}{\psi(\alpha)} \frac{\partial c}{\partial \beta} = m$$

so that the equations $(1,1), \ldots, (1,2)$ are all satisfied. Introducing $\int \frac{2d\alpha}{\psi(\alpha)}, c$ as new variables in place of α, β, we obtain the quadratic form $\beta^2 d\gamma^2 + d\alpha\, d\beta$, which becomes $dx^2 + dy^2 + dz^2$ via the substitution

$$x + iy = \beta, \;\; x - iy = \alpha - \beta\gamma^2, \;\; z = \beta\gamma.$$

From this, however, one cannot obtain a real curve via the equations $\alpha =$ const., $\beta =$ const. The case $\psi(\alpha) = 0$, $\phi(\beta)$ nonzero is not essentially different from this one.

3) If neither $\psi(\alpha)$ nor $\phi(\beta)$ vanishes, we introduce new variables $\int \frac{d\alpha}{\psi(\alpha)}, \int \frac{d\beta}{\phi(\beta)}$ in place of α, β.

It follows that

$$\frac{\partial c}{\partial \alpha} = m, \;\; \frac{\partial c}{\partial \beta} = m, \;\; \frac{\partial c}{\partial \alpha} - \frac{\partial c}{\partial \beta} = 0,$$

thus $c = f(\alpha + \beta)$, $m = f'(\alpha + \beta)$.

It now follows from (1,3) that

$$\frac{\partial}{\partial \beta}\left(\log \frac{\partial}{\partial \beta}(ac)\right) = \frac{\partial}{\partial \beta}(\log cm),$$

and by integration that

$$ac = f^2 \phi(\alpha) + \psi(\alpha).$$

Introducing the variable $\gamma + \int \phi(\alpha) d\alpha$ in place of γ, we find that $\phi(\alpha) = 0$ and $ac = \psi(\alpha)$. Now it follows from (1,2) that

$$\frac{f^3 f''}{f'} = -\psi(\alpha)^2.$$

Now one side of this equation depends only on α, and the other side only on $\alpha + \beta$. Consequently both sides are equal to a constant k^2, and we obtain the second order differential equation

$$f'' - \frac{k^2 f'}{f^3} = 0$$

for f. Now equations $(1,1), \ldots, (1,2)$ are all satisfied. A single integration of this equation yields

$$2f' = k_1^2 - \frac{k^2}{f^2},$$

where k_1 denotes a new constant.

Now let $\alpha = x + iy$, $\beta = x - iy$ and introduce a new variable $\gamma - ik \int \frac{dx}{f^2}$ in place of γ. We obtain

$$\begin{aligned}(c\, d\gamma + a\, d\gamma)^2 + 2m\, d\alpha\, d\beta &= \left(f\, d\gamma + \frac{k}{f}\, dy\right)^2 + 2f'(dx^2 + dy^2) \\ &= f^2 d\gamma^2 + 2k\, d\gamma\, dy + 2f' dx^2 + k_1^2 dy^2.\end{aligned}$$

Further let

$$2f' dx^2 = \frac{df^2}{2f'} = \frac{f^2 df^2}{k_1^2 f^2 - k^2} = d\xi^2.$$

We then have

$$\xi = \frac{1}{k_1^2} \sqrt{k_1^2 f^2 - k^2},\ f^2 = k_1^2 \xi^2 + \frac{k^2}{k_1^2},$$

and our quadratic form becomes

$$\left(\frac{k}{k_1} d\gamma + k_1 dy\right)^2 + k_1^2 \xi^2 d\gamma^2 + d\xi^2.$$

We express this in terms of polar coordinates. Let

$$\xi = r,\ k_1 \gamma = \phi,\ k_1 y + \frac{k}{k_1} \gamma = z.$$

Then the quadratic form becomes

$$dr^2 + r^2 d\phi^2 + dz^2.$$

The curves $\alpha =$ const., $\beta =$ const. are now

$$r = \text{const.},\ z - \frac{k}{k_1^2}\,\phi = \text{const.}$$

where k may be 0; thus we have helices or circles.

In the particular case $k_1 = 0$, we obtain $\xi = \frac{if^2}{k}$ and the quadratic form becomes

$$-2ki\xi\, d\gamma^2 + 2k\, d\gamma\, dy + d\xi^2,$$

or, if we write α, β, γ in place of ξ, $\frac{2ky}{\sqrt{-2ki}}$, $\sqrt{-2ki}\,\gamma$,

$$\alpha\, d\gamma^2 + d\beta\, d\gamma + d\alpha^2.$$

This becomes $dx^2 + dy^2 + dz^2$ via the substitution

$$\begin{aligned} x + iy &= \beta + \alpha\gamma - \frac{1}{12}\gamma^3, \\ x - iy &= \gamma, \\ z &= \alpha - \frac{1}{4}\gamma^2. \end{aligned}$$

However the equations which this yields,

$$z + \frac{1}{4}(x - iy)^2 = \alpha = \text{const.}$$

$$x + iy - \alpha(x - iy) + \frac{1}{12}(x - iy)^3 = \beta = \text{const.}$$

do not correspond to any real curve.

In the remaining cases, I did not succeed in carrying through the calculation completely.

XXIII.

On the representation of the quotient of two hypergeometric series as an infinite continued fraction.

1.

Given an infinite continued fraction of the form

$$a + \cfrac{b_1 x}{1 + \cfrac{b_2 x}{1 + \cfrac{b_3 x}{1 + \cdots}}},$$

which converges for sufficiently small values of x and represents the function $f(x)$, it is easily seen that the m-th convergent is equal to the quotient p_m/q_m of two polynomials p_m and q_m, each of degree n when $m = 2n+1$, and of degrees n and $n-1$, if $m = 2n$. The difference between the m-th convergent and the function $f(x)$ is an infinitesimal of the m-th order when x is infinitesimal. But for this to be so, as many conditions require to be satisfied as there are arbitrary quantities in the fractional function representing the m-th convergent.

Thus the m-th convergent can be determined by the condition that the first m terms of its development in the form of a power series in x must coincide with the first m terms of the expansion of the function $f(x)$, taking account of the degrees of the polynomials representing the numerator and denominator of the m-th convergent, which are both equal to n when $m = 2n+1$, and are $n, n-1$ when $m = 2n$.

2.

This method of determining the convergents leads immediately to an expression for the convergents when it is a question of expanding the quotient of two hypergeometric series

$$P^{\alpha}\begin{pmatrix} \alpha & \beta & \gamma & \\ \alpha' & \beta' & \gamma' & x \end{pmatrix} = P \text{ and } P^{\alpha}\begin{pmatrix} \alpha & \beta+1 & \gamma & \\ \alpha'-1 & \beta' & \gamma' & x \end{pmatrix} = Q.$$

We use the characteristic property expounded in the memoir **IV**.

In fact, since for an infinitesimal x, $\frac{P}{Q} - \frac{p_m}{q_m}$ is an infinitesimal of order m and Qq_m is of order α, the expression $q_m P - p_m Q$ becomes an infinitesimal or order $m+\alpha$. It can easily be shown that this expression has all the characteristic properties of a function that can be developed in a hypergometric series, so that we have

$$
\begin{aligned}
q_{2n+1}P - p_{2n+1}Q &= P\begin{pmatrix} \alpha+2n+1 & \beta-n & \gamma & \\ \alpha'-1 & \beta'-n & \gamma' & x \end{pmatrix} \\
&= x^n P\begin{pmatrix} \alpha+n+1 & \beta & \gamma & \\ \alpha'-n-1 & \beta' & \gamma' & x \end{pmatrix} = x^n P_{n+1}, \\
q_{2n}P - p_{2n}Q &= P\begin{pmatrix} \alpha+2n & \beta+1-n & \gamma & \\ \alpha'-1 & \beta'-n & \gamma' & x \end{pmatrix} = x^n Q_n,
\end{aligned} \tag{1}
$$

where P_n, Q_n denote the values taken by P, Q when α, α' are changed into $\alpha+n$, $\alpha'-n$. Now, if the variable x and the functions of x vary continuously in such a way that the position of the complex variable x makes a circuit encircling 1, then q_m and p_m will resume their original values, whereas P, Q, P_n, Q_m, will be transformed into other branches of these functions.

Hence, if we denote by P', Q', P'_n, Q'_m the corresponding branches of these functions, we shall also have

$$
\begin{aligned}
q_{2n+1}P' - p_{2n+1}Q' &= x^n P'_{n+1}, \\
q_{2n}P' - p_{2n}Q' &= x^n Q'_n.
\end{aligned} \tag{2}
$$

From the equations (1) and (2),

$$
\frac{p_{2n+1}}{q_{2n+1}} = \frac{P\,P'_{n+1} - P'P_{n+1}}{Q\,P'_{n+1} - Q'P_{n+1}}, \quad \frac{p_{2n}}{q_{2n}} = \frac{PQ'_n - P'Q_n}{QQ'_n - Q'Q_n}.
$$

Thus to find for what values of x the convergents p_{2n}/q_{2n} and p_{2n+1}/q_{2n+1} converge to P/Q, it suffices to find out when $\frac{P_n}{P'_n}$ and $\frac{Q_n}{Q'_n}$ converge to zero as n increases without limit.

3.

For this purpose it is convenient to introduce the expressions of P_n and Q_n as definite integrals. Writing

$$
\begin{aligned}
[-\alpha' - \beta' - \gamma' &= a, \\
-\alpha' - \beta\; - \gamma &= b, \\
-\alpha - \beta' - \gamma &= c],
\end{aligned}
$$

we may express

$$P_n \text{ by } \left[x^{\alpha+n}(1-x)^{\gamma}\int_0^1 s^{a+n}(1-s)^{b+n}(1-xs)^{c-n}ds\right]$$

and

$$Q_n \text{ by } \left[x^{\alpha+n}(1-x)^{\gamma}\int_0^1 s^{a+1+n}(1-s)^{b+n}(1-xs)^{c-n}ds\right].$$

To obtain the general value of the functions P_n and Q_n, we would need to multiply the integrals by constant factors, but we can substitute in the equations (1) the integrals including the constant factors in the polynomials p_m and q_m. As regards the values of the integrands, it is immaterial which value is taken, provided that the same values are taken in each integral for s^a, $(1-s)^b$, and $(1-xs)^c$.

[Now the expressions for p_m/q_m also remain unaltered when we replace P', Q', P'_n, Q'_n by a given linear combination of these quantities, so that P, Q, P_n, Q_n become

$$AP + BP', \quad AQ + BQ', \quad AP_n + BP'_n, \quad AQ_n + BQ'_n$$

where A and B denote constants, of which B is non-zero. Corresponding functions of this kind arise when the path of the integrals in the above expression, instead of being taken from 0 to 1, is taken from any one of the four points $0, 1, 1/x, \infty$ to any other of these four points, each line integral being taken on the same contour.]

Thus we can take for P'_n and Q'_n these same integrals taken one after the other around the point $1/x$.

These integrals [by which, in virtue of the foregoing considerations, P_n, Q_n, P'_n, Q'_n can be expressed, do not change in value when the path of integration is varied in a continuous fashion between the limits indicated] because the path of integration does not pass over $1/x$, and we can choose the path of integration in such a way that the limit to which it converges as n tends to infinity can easily be found.

To this end, let $s(1-s)/(1-xs)\ldots$ [The text breaks off at this point, but a few handwritten notes and formulae suggest that Riemann proceeded somewhat as follows:

We write]

$$\frac{s(1-s)}{1-xs} = e^{f(s)}$$

[and consider the curves in the plane of the complex variable s along which the modulus of $e^{f(s)}$ has a constant value. For very small values of this modulus, these curves enclose the points 0 and 1 and closely resemble concentric circles of small radius. For very large values of the modulus, the curves enclose the point $s = 1/x$ and the point $s = \infty$. In each case, then, the curves consist of two separate portions. If we gradually increase the modulus, starting from small values, the separate portions surrounding the points 0 and 1, corresponding to the same value of the modulus, become ever closer until they form a single curve with a double point. At this double point, $f'(s)$ must be 0. Similar considerations apply when the modulus is decreased, starting from very large values.

In this way one obtains the following equations:]

$$f(s) = \log(1-s) - \log\left(\frac{1}{s} - x\right),$$

$$f'(s) = -\frac{1}{1-s} + \frac{1}{\frac{1}{s} - x}\,\frac{1}{s^2} = \frac{1-2s+xs^2}{s(1-s)(1-xs)}.$$

[Hence for $f'(s) = 0$, we have]

$$1 - 2s + xs^2 = 0,\ s(1-xs) = 1-s,\ 1-2s+s^2 = (1-x)s^2 = (1-s)^2,$$

$$\frac{1}{s} - 1 = \sqrt{1-x} = 1 - xs,$$

$$\frac{1-s}{1-xs} = s.$$

[We shall now denote by $\sqrt{1-x}$ the value of the square root whose real part is positive; the case where x is real and ≥ 1 is excluded from consideration. We shall also denote by σ and σ' the two roots of the quadratic equation

$$1 - 2s + xs^2 = 0,$$

$$\sigma = \frac{1}{1+\sqrt{1-x}},\quad \sigma' = \frac{1}{1-\sqrt{1-x}}$$

so that the modulus of σ is smaller than that of σ'.

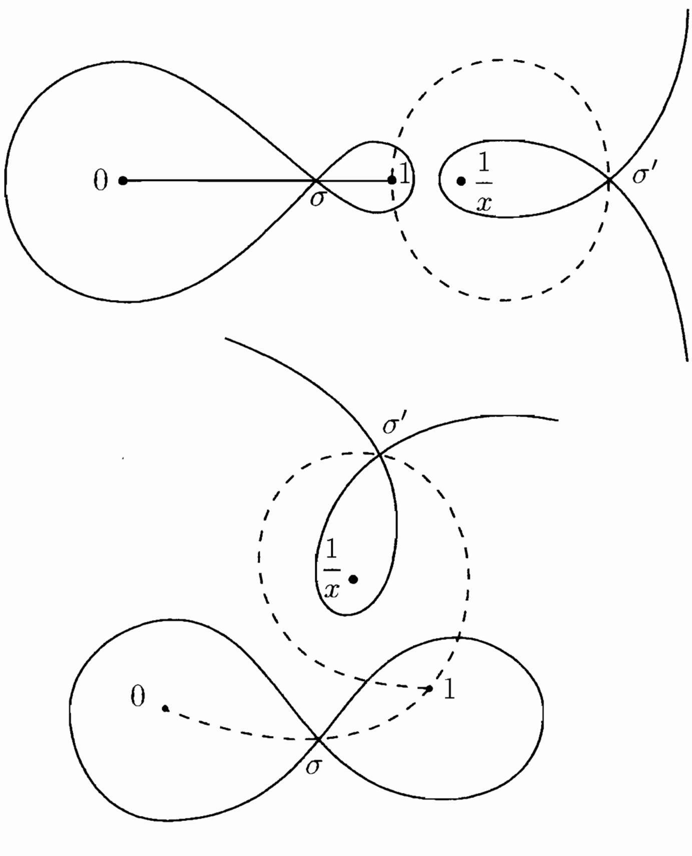
It follows that

$$e^{f(\sigma)} = \sigma^2 = \left(\frac{1}{1+\sqrt{1-x}}\right)^2,\quad e^{f(\sigma')} = \sigma'^2 = \left(\frac{1}{1-\sqrt{1-x}}\right)^2.$$

Let us now visualize a line drawn from the point $s = 0$ to the point $s = 1$, going through the point $s = \sigma$, and such that the modulus of $e^{f(s)}$ steadily increases along the path from the point $s = 0$ to the point $s = \sigma$, but continuously decreases along the path from $s = \sigma$ to $s = 1$. A line of this kind can serve as the path of integration for the integral by means of which the functions P_n and Q_n can be expressed.

On the other hand, for those integrals which represent the functions P'_n and Q'_n, one can use an integration path which starts from the point $s = 1$, goes through the point $s = \sigma'$, and then returns to the point $s = 1$, enclosing the point $s = 1/x$. This path can always be chosen so that the modulus of $e^{f(s)}$ attains its maximum value on the path only at the point $s = \sigma$.

In the following figures, based on sketches by Riemann, the paths of integration are shown by dotted lines.



Riemann's sketches of the paths of integration.

It is now a question of finding an asymptotic expression for the value of the integral

$$\int_0^1 s^{a+n}(1-s)^{b+n}(1-xs)^{c-n}ds$$

for infinitely large values of n.

Let

$$s^a(1-s)^b(1-xs)^c = \phi(s).$$

We must calculate $\int_0^1 e^{nf(s)}\phi(s)ds$ as n tends to infinity.

Those portions of the path of integration which do not lie in the neighborhood of the singularity at $s = \sigma$, make a contribution to the value of the integral which is infinitely small when n is infinitely large. Moreover, because the real part of $n(f(\sigma) - f(s))$ grows beyond all bounds, under the assumptions which have been made, this contribution is also infinitely small in relation to the contribution from the part of the path in the neighborhood of $s = \sigma$. For this reason, we need only find an asymptotic expression, valid for n tending to infinity, for the contribution of the neighborhood of $s = \sigma$. We therefore write, denoting by h a quantity whose modulus takes small values:]

$$s = \sigma + h,$$

$$f(s) = f(\sigma) + \frac{1}{2} f''(\sigma)h^2 + (h^3),$$

$$n\,f(s) = n\,f(\sigma) + \frac{n\,f''(\sigma)}{2}\,h^2 + n(h^3),$$

$$-n\,\frac{f''(\sigma)}{2}\,h^2 = z^2,$$

$$dh = \frac{dz}{\sqrt{-n\,\frac{f''(\sigma)}{2}}},$$

$$e^{nf(s)} = e^{nf(\sigma)}e^{-z^2+(z^3/\sqrt{n})},$$

$$e^{nf(s)}\phi(s)ds = e^{nf(\sigma)}\phi\left(\sigma + \frac{z}{\sqrt{-n\,\frac{f''(\sigma)}{2}}}\right) e^{-z^2}\,\frac{dz}{\sqrt{-n\,\frac{f''(\sigma)}{2}}}.$$

[If we now choose for the path, in the neighborhood of the point $s = \sigma$, a straight line through this point bisecting the right angle between the two tangents to the curve

$$|e^{f(s)}| = |e^{f(\sigma)}|,$$

then the limits of integration for the variable z tend to $-\infty$ and $+\infty$ as n tends to infinity, and consequently the contribution to the integral from the neighborhood of the point $s = \sigma$ is, for very large values of n, asymptotically equal to

$$\frac{e^{nf(\sigma)}\phi(\sigma)}{\sqrt{-n\frac{f''(\sigma)}{2}}}\int_{-\infty}^{\infty} e^{-z^2}dz = \sqrt{\frac{\pi}{-\frac{f''(\sigma)}{2}}}\,\frac{e^{nf(\sigma)}}{\sqrt{n}}\,\phi(\sigma).$$

Now

$$e^{nf(\sigma)} = \sigma^{2n} = \left(\frac{1}{1+\sqrt{1-x}}\right)^{2n},$$

$$-\frac{f''(\sigma)}{2} = \frac{1}{\sigma(1-\sigma)} = \frac{1}{\sigma^2\sqrt{1-x}},$$

$$\phi(\sigma) = \sigma^{a+b}(1-x)^{(b+c)/2}.$$

Hence the asymptotic value of $\int_0^1 e^{nf(s)}\phi(s)ds$ is equal to

$$\sqrt{\frac{\pi}{n}}\left(\frac{1}{1+\sqrt{1-x}}\right)^{2n+a+b+1}(1-x)^{\frac{b+c}{2}+\frac{1}{4}}.$$

Similarly, the asymptotic value of $\int_0^1 e^{nf(s)}\phi(s)ds$ is found to be

$$\sqrt{\frac{\pi}{n}}\left(\frac{1}{1-\sqrt{1-x}}\right)^{2n+a+b+1}(1-x)^{\frac{b+c}{2}+\frac{1}{4}}.$$

Under the given hypotheses, then, we obtain for the quotient $P_n : P'_n$ the asymptotic value:]

$$\left(\frac{1-\sqrt{1-x}}{1+\sqrt{1-x}}\right)^{2n+a+b+1}.$$

[Thus for all values of x, with the exception of those which are real and ≥ 1, the quotient $P_n : P'_n$ converges to zero as n increases without limit.

The same is obtained for the quotient $Q_n : Q'_n$, when $a+1$ is substituted for a.

We have therefore proved that the convergents to the continued fraction given in §1, in which the quotient

$$\frac{P^{\alpha}\begin{pmatrix}\alpha & \beta & \gamma & \\ \alpha' & \beta' & \gamma' & x\end{pmatrix}}{P^{\alpha}\begin{pmatrix}\alpha & \beta+1 & \gamma & \\ \alpha'-1 & \beta' & \gamma' & x\end{pmatrix}}$$

can be expanded, converge to the value of this quotient, for all values of x other than those which are real and ≥ 1.]

XXIV.

On the potential of a ring.

Suppose that we wish to determine the force, due to an arbitrary body, at a point outside the body, where the attraction or repulsion due to a piece of the body is inversely proportional to the square of the distance from that piece. It is known that we must seek a function V of the rectangular coordinates x, y, z of the point, called the potential or potential function of the acting mass. The partial derivatives $\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z}$ are the components of the accelerating force at the point x, y, z, or their opposite, according as the unit mass attracts or repels another unit mass, at unit distance, with unit force.

To determine this function, which satisfies the differential equation

$$(1) \qquad \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0,$$

it suffices if a further condition is given at each point of the surface. Frequently the problem presents itself in the form where the distribution of mass in the body is not given. Rather it suffices if the effects of the mass distribution in the surface are given. For example, V may be required to be equal to an arbitrary given function in the surface; its components in the tangential directions may be given at each points of the surface; or, at every point, the component in a particular direction may have an assigned value. It is well known that the procedure for the solution of the problem is to assemble a general expression, containing arbitrary constants $a_1, a_2, \ldots, a_n, \ldots$,

$$a_1 Q_1 + a_2 Q_2 + \cdots + a_n Q_n + \cdots = R$$

from particular solutions

$$Q_1, Q_2, \ldots, Q_n, \ldots$$

of the differential equation (1). In the general expression R, which also satisfies the differential equation (1), we must choose the constants $a_1, a_2, \ldots$ so that the boundary conditions are satisfied.

In general, the expressions R converge only for certain values of the coordinates x, y, z. For each particular expression R, the whole of infinite space is divided by a surface s into two parts. In one part, the expression R converges; in the other part, generally speaking (that is, apart from isolated points and lines), it diverges. Thus, for example, the expression

$$\sum a_n \exp\left(z\sqrt{\alpha_n^2+\beta_n^2}\right)\cos\alpha_n x\cos\beta_n y$$

ceases to converge for a definite plane perpendicular to the z-axis. If we introduce polar coordinates in place of x, y, z, and develop V in powers of the radius vector, it is known that the coefficients of the nth power are combinations of spherical functions of order n multiplied by arbitrary constants. We obtain a series that ceases to converge for a definite spherical surface with the origin at its center. It is worthy of note that a particular form of the expansion R already corresponds to a definite family of surfaces of convergence (in the first case, a family of parallel planes; in the second, a family of concentric spherical surfaces); while the values of the coefficients determine where, in this family of surfaces, the divergence begins.

Obviously the expression R must converge in the entire region where the values of the function V are to be specified. For only then can we substitute this expression into the boundary conditions in order to determine the arbitrary constants in R. On the other hand, one may readily show that an expression satisfying the differential equation (1) can only represent an arbitrary given function where the convergence ceases. Consequently, the form of the expression R has to be determined in such a way that the surface of the body is one of the family of boundary surfaces of convergence that belong to R.

We now solve this problem for a ring with circular cross-section, which might be desirable for various physical investigations.

1.

We place the z-axis in the axis of the ring and the origin of coordinates in the center of the ring. The equation of the surface of the ring takes the form

$$\left(\sqrt{x^2+y^2}\pm a\right)^2+z^2=c^2.$$

I now seek to introduce variables in place of x, y, z, one of which has a constant value on the surface of the ring, and such that the differential equation (1) takes the simplest possible form for these variables.

We introduce polar coordinates in the (x, y)-plane by writing

$$x = r\cos\phi, \quad y = r\sin\phi.$$

The differential equation (1) becomes

(1)
$$\frac{\partial^2 V}{\partial r^2} + \frac{1}{r}\frac{\partial V}{\partial r} + \frac{1}{r^2}\frac{\partial^2 V}{\partial \phi^2} + \frac{\partial^2 V}{\partial z^2} = 0.$$

The boundary equation is independent of ϕ, namely

$$(r+a)^2 + z^2 = c^2$$

and

$$(r-a)^2 + z^2 = c^2.$$

Thus in the (r, z)-plane, the boundary is formed from two circles of radius c drawn around the points $(-a, 0)$ and $(a, 0)$.

I now introduce, in place of r and z, two new variables ρ and ψ, by setting $r + zi$ equal to a function of the complex variable $\rho e^{\psi i}$:

$$r + zi = f(\rho e^{\psi i}).$$

The quantity $\rho e^{\psi i}$ is to be determined as a function of $r + zi$ in such a way that the modulus ρ is constant in each of the two boundary circles. Further, $\rho e^{\psi i}$ should remain continuous and finite everywhere outside the two circles.

These conditions are satisfied if we let

$$r + zi = \frac{\beta + \gamma\rho e^{\psi i}}{1 + \rho e^{\psi i}}$$

and

$$\beta = -\gamma = \sqrt{a^2 - c^2}.$$

For in this case,

$$a + r + zi = \frac{(a+\beta) + (a+\gamma)\rho e^{\psi i}}{1 + \rho e^{\psi i}},$$

$$(a + r + zi)(a + r - zi) = \frac{\frac{a+\beta}{(a+\gamma)\rho} + e^{\psi i}}{1 + \rho e^{\psi i}} \cdot \frac{\frac{a+\beta}{(a+\gamma)\rho} + e^{-\psi i}}{1 + \rho e^{-\psi i}} (a+\gamma)^2\rho^2.$$

This quantity will be independent of ψ, if

$$\frac{a+\beta}{(a+\gamma)\rho} = \rho.$$

Indeed, its value will be

$$(a+\gamma)^2\rho^2 = (a+\beta)(a+\gamma).$$

Likewise, the quantity

$$(-a+r+zi)(-a+r-zi)$$

will be independent of ψ, with the value

$$(-a+\beta)(-a+\gamma),$$

if

$$\rho^2 = \frac{-a+\beta}{-a+\gamma}.$$

The values

$$\rho^2 = \frac{a+\beta}{a+\gamma}\,,\ \rho^2 = \frac{-a+\beta}{-a+\gamma}$$

thus correspond to circles with centers $(-a,0),(a,0)$ and radii

$$\sqrt{(a+\beta)(a+\gamma)}\,,\ \sqrt{(-a+\beta)(-a+\gamma)}.$$

If both radii are to be c, then we must have

$$(a+\beta)(a+\gamma) - (-a+\beta)(-a+\gamma) = 2a(\beta+\gamma) = 0,$$

that is, $\gamma = -\beta$, $a^2 - \beta^2 = c^2$, and so $\beta = \sqrt{a^2-c^2}$.

2.

The transformation of the differential equation (1) can be simplified by writing $V = r^\mu U$. We obtain

$$\begin{aligned}\frac{\partial^2 V}{\partial r^2} + \frac{1}{r}\,\frac{\partial V}{\partial r} &= r^\mu\,\frac{\partial^2 U}{\partial r^2} + 2\mu\, r^{\mu-1}\,\frac{\partial U}{\partial r} \\ &\qquad + \mu(\mu-1)r^{\mu-2}U + r^{\mu-1}\,\frac{\partial U}{\partial r} + \mu r^{\mu-2}U \\ &= r^\mu\,\frac{\partial^2 U}{\partial r^2} + (2\mu+1)r^{\mu-1}\,\frac{\partial U}{\partial r} + \mu^2 r^{\mu-2}U.\end{aligned}$$

We choose μ so that the second term drops out, that is $\mu = -1/2$. Now the differential equation (I) becomes

$$r^2\left(\frac{\partial^2 U}{\partial r^2} + \frac{\partial^2 U}{\partial z^2}\right) + \frac{\partial^2 U}{\partial \phi^2} + \frac{1}{4}U = 0.$$

For brevity, we now denote the complex variables $r + zi$, $\rho e^{\psi i}$ by y, η respectively, and the conjugate quantities by y' and η'. We now have

$$r = \frac{y+y'}{2}\,,\ zi = \frac{y-y'}{2},$$
$$\frac{\partial U}{\partial y} = \frac{1}{2}\left(\frac{\partial U}{\partial r} - \frac{\partial U}{\partial z}i\right),\frac{\partial^2 U}{\partial y \partial y'} = \frac{1}{4}\left(\frac{\partial^2 U}{\partial r^2} + \frac{\partial^2 U}{\partial z^2}\right),$$

and consequently,

$$r^2\left(\frac{\partial^2 U}{\partial r^2} + \frac{\partial^2 U}{\partial z^2}\right) = (y+y')^2\,\frac{\partial^2 U}{\partial y \partial y'}.$$

Further,

$$y = \beta\,\frac{1-\eta}{1+\eta}\,,\ y' = \beta\,\frac{1-\eta'}{1+\eta'},$$
$$y + y' = 2\beta\,\frac{1-\eta\eta'}{(1+\eta)(1+\eta')},$$
$$y = \beta\left(-1 + \frac{2}{1+\eta}\right), dy = -2\beta\,\frac{d\eta}{(1+\eta)^2}\,, dy' = -2\beta\,\frac{d\eta'}{(1+\eta')^2},$$
$$(y+y')^2\,\frac{\partial^2 U}{\partial y \partial y'} = (1-\eta\eta')^2\,\frac{\partial^2 U}{\partial \eta \partial \eta'}$$
$$= \frac{(1-\eta\eta')^2}{\eta\eta'}\;\frac{\partial^2 U}{\partial \log\eta\,\partial\log\eta'}.$$

Since $\eta\eta' = \rho^2$, $\log\eta = \log\rho + \psi i$, $\log\eta' = \log\rho - \psi i$, we can write this as

$$(y+y')^2\,\frac{\partial^2 U}{\partial y \partial y'} = \left(\frac{1-\rho^2}{\rho}\right)^2\frac{1}{4}\left(\frac{\partial^2 U}{\partial(\log\rho)^2} + \frac{\partial^2 U}{\partial \psi^2}\right).$$

Thus the partial differential equation becomes

$$\left(\frac{\rho-\rho^{-1}}{2}\right)^2\left(\frac{\partial^2 U}{\partial(\log\rho)^2} + \frac{\partial^2 U}{\partial \psi^2}\right) + \frac{\partial^2 U}{\partial \phi^2} + \frac{1}{4}U = 0.$$

3.

It is now easy to expand U in a series of particular solutions of this differential equation, which simultaneously converges or diverges for all values of ϕ or ψ. To this end we need only to give these particular solutions the form

$$\cos m\psi \cos n\phi,\ \cos m\psi \sin n\phi,\ \sin m\psi \cos n\phi,\ \sin m\psi \sin n\phi$$

multiplied by a function P of ρ that satisfies the differential equation

$$\text{(II)} \qquad \left(\frac{\rho-\rho^{-1}}{2}\right)^2 \left(\frac{d^2P}{d(\log\rho)^2} - m^2 P\right) - \left(n^2 - \frac{1}{4}\right) P = 0.$$

We can then determine the arbitrary constants using Fourier series.

Let us write

$$\frac{\rho-\rho^{-1}}{2} = t,$$

then

$$\begin{aligned} \frac{dP}{d\log\rho} &= \frac{dP}{dt}\,\frac{\rho+\rho^{-1}}{2}, \\ \frac{d^2P}{d(\log\rho)^2} &= \left(\frac{\rho+\rho^{-1}}{2}\right)^2 \frac{d^2P}{dt^2} + \frac{\rho-\rho^{-1}}{2}\,\frac{dP}{dt} \\ &= (t^2+1)\,\frac{d^2P}{dt^2} + t\,\frac{dP}{dt}. \end{aligned}$$

Now the differential equation (II) becomes

$$t^2(t^2+1)\,\frac{d^2P}{dt^2} + t^3\,\frac{dP}{dt} - \left(m^2t^2 + n^2 - \frac{1}{4}\right) P = 0.$$

This differential equation contains only terms of two different degrees with respect to t. Consequently it can be solved via the procedure, known since Euler, involving hypergeometric series. The solution may be expressed in numerous ways via other hypergeometric series, namely these series whose fourth element is the value, or the reciprocal, of one of the nine quantities

$$-\left(\frac{\rho+\rho^{-1}}{2}\right)^2,\ \left(\frac{\rho+\rho^{-1}}{2}\right)^2,\ \left(\frac{1-\rho^2}{1+\rho^2}\right)^2;$$

$$\rho^2, 1-\rho^2, 1-\frac{1}{\rho^2};\ \left(\frac{1-\rho}{1+\rho}\right)^2, -\frac{(1-\rho)^2}{4\rho}, \frac{(1+\rho)^2}{4\rho}.$$

Indeed, for each of these eighteen quantities there are four distinct expansions that satisfy the differential equation, which represent the same particular solutions in pairs. In general, we expand by the smallest of these quantities. If we expand by one of these quantities that vanishes for $\rho = 1$, it is seen that of these two particular solutions, one is infinite for $\rho = 1$. Since V must remain finite, the coefficients of this particular solution must vanish in P, and P is proportional to the solution that remains finite for $\rho = 1$. Among the various expressions I shall be content to display one, denoting it by $P^{n,m}$:

$$P^{n,m} = (1-\rho^2)^{n+\frac{1}{2}}\rho^{\pm m}\, F\left(n \pm m + \frac{1}{2}, n + \frac{1}{2}, 2n+1, 1-\rho^2\right).$$

Since, in the values of $P^{n,m}$, the first three elements of the hypergeometric series differ only by integers, all of the $P^{n,m}$ can be expressed linearly in two of them, $P^{0,0}$ and $P^{0,1}$ (*Comm. Gott. rec.*, vol. II[1]). These are entire elliptic integrals of the first and second kind. Perhaps the $P^{n,m}$ can be found most conveniently by the principle of arithmetic-geometric means, that is, via repeated transformations of second order[2].

[1]Gauss, Werke vol. III, p. 131. W.

[2]All the $P^{n,m}$ may be expressed in further ways via entire elliptic integrals.

XXV.

Diffusion of heat in an ellipsoid.

According to the theory of Fourier, the problem of thermal flow in a homogeneous isotropic body is a question of solving the partial differential equation

$$\frac{\partial u}{\partial t} = a^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right), \tag{1}$$

where a^2 is a positive constant (the ratio of the conductivity to the product of the density and the specific heat). The function u is to be determined so that it satisfies: the differential equation (1) in the interior of the given body; for $t = 0, u$ becomes a given continuous function of the coordinates (the initial state); and it satisfies still another condition on the surface, for example becomes a given function.

If we are concerned with bodies bounded by an ellipsoid with semi-axes $\sqrt{\alpha}$, $\sqrt{\beta}$, $\sqrt{\gamma}$, elliptical coordinates can be introduced, in which one understands λ, μ, ν as the three roots of the cubic equation

$$\frac{x^2}{\alpha - \lambda} + \frac{y^2}{\beta - \lambda} + \frac{z^2}{\gamma - \lambda} - 1 \equiv f(\lambda) = 0, \tag{2}$$

subject to the restrictions

$$-\infty < \lambda < \gamma < \mu < \beta < \nu < \alpha,$$

so that $\lambda = 0$ at the surface of the given ellipsoid.

The transformation of equation (1) is carried out most easily by the method of Jacobi. By differentiation of (2) we obtain

$$\begin{aligned} \frac{2x}{\alpha - \lambda} + f'(\lambda) \frac{\partial \lambda}{\partial x} &= 0, \\ \frac{2y}{\beta - \lambda} + f'(\lambda) \frac{\partial \lambda}{\partial y} &= 0, \\ \frac{2z}{\gamma - \lambda} + f'(\lambda) \frac{\partial \lambda}{\partial z} &= 0, \end{aligned} \tag{3}$$

$$f'(\lambda) = \frac{x^2}{(\alpha - \lambda)^2} + \frac{y^2}{(\beta - \lambda)^2} + \frac{z^2}{(\gamma - \lambda)^2} = \frac{(\lambda - \mu)(\lambda - \nu)}{\theta(\lambda)}, \tag{4}$$

where, for brevity, we write

$$(\alpha-\lambda)(\beta-\lambda)(\gamma-\lambda)=\theta(\lambda). \tag{5}$$

Furthermore, from (3) and (4),

$$\left(\frac{\partial\lambda}{\partial x}\right)^2+\left(\frac{\partial\lambda}{\partial y}\right)^2+\left(\frac{\partial\lambda}{\partial z}\right)^2=\frac{4}{f'(\lambda)}. \tag{6}$$

From this, one obtains

$$d\tau=\frac{1}{8}\sqrt{f'(\lambda)f'(\mu)f'(\nu)}d\lambda\,d\mu\,d\nu$$

as the expression for the volume element $d\tau$ in the new coordinates, and the transformation of the integral, extended over an arbitrary region,

$$\iiint\left(\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2+\left(\frac{\partial u}{\partial z}\right)^2\right)dx\,dy\,dz$$
$$=\iiint\left(\left(\frac{\partial u}{\partial\lambda}\right)^2\frac{1}{f'(\lambda)}+\left(\frac{\partial u}{\partial\mu}\right)^2\frac{1}{f'(\mu)}+\left(\frac{\partial u}{\partial\nu}\right)^2\frac{1}{f'(\nu)}\right)$$
$$\frac{1}{2}\sqrt{f'(\lambda)f'(\mu)f'(\nu)}d\lambda\,d\mu\,d\nu.$$

By taking the first variation of this integral in both forms, one obtains the transformation required:

$$\left(\frac{\partial^2u}{\partial x^2}+\frac{\partial^2u}{\partial y^2}+\frac{\partial^2u}{\partial z^2}\right)\frac{1}{4}\sqrt{f'(\lambda)f'(\mu)f'(\nu)}$$
$$=\frac{\partial}{\partial\lambda}\sqrt{\frac{f'(\mu)f'(\nu)}{f'(\lambda)}}\frac{\partial u}{\partial\lambda}+\frac{\partial}{\partial\mu}\sqrt{\frac{f'(\nu)f'(\lambda)}{f'(\mu)}}\frac{\partial u}{\partial\mu}+\frac{\partial}{\partial\nu}\sqrt{\frac{f'(\lambda)f'(\mu)}{f'(\nu)}}\frac{\partial u}{\partial\nu}.$$

Introducing the notation from (4) and (5), the differential equation (1) becomes

$$-(\mu-\nu)(\nu-\lambda)(\lambda-\mu)\frac{\partial u}{\partial t}$$
$$=4a^2\left\{(\mu-\nu)\sqrt{\theta(\lambda)}\frac{\partial\sqrt{\theta(\lambda)}\,\frac{\partial u}{\partial\lambda}}{\partial\lambda}+(\nu-\lambda)\sqrt{\theta(\mu)}\frac{\partial\sqrt{\theta(\mu)}\,\frac{\partial u}{\partial\mu}}{\partial\mu}\right. \tag{7}$$
$$\left.+(\lambda-\mu)\sqrt{\theta(\nu)}\frac{\partial\sqrt{\theta(\nu)}\,\frac{\partial u}{\partial\nu}}{\partial\nu}\right\}.$$

In order to seek particular solutions for this equation, set

$$u = e^{-4a^2g^2t} u_\lambda u_\mu u_\nu, \tag{8}$$

where g is an arbitrary constant and u_λ is taken to be dependent only on λ, u_μ only on μ, and u_ν only on ν. Then, if we set

$$U_\lambda = \frac{\sqrt{\theta(\lambda)}}{u_\lambda} \frac{d\sqrt{\theta(\lambda)}\,\frac{du_\lambda}{d\lambda}}{d\lambda}, \tag{9}$$

so that U_λ is only dependent on λ, and U_μ, U_ν have corresponding meanings, it follows that

$$g^2(\mu-\nu)(\nu-\lambda)(\lambda-\mu) = (\mu-\nu)U_\lambda + (\nu-\lambda)U_\mu + (\lambda-\mu)U_\nu. \tag{10}$$

If this equation is differentiated twice with respect to λ, then

$$-2g^2 = \frac{d^2U_\lambda}{d\lambda^2}$$

or

$$U_\lambda = -g^2\lambda^2 - h\lambda - k,$$

and similarly,

$$\begin{aligned} U_\nu &= -g^2\nu^2 - h\nu - k \\ U_\mu &= -g^2\mu^2 - h\mu - k, \end{aligned}$$

where h and k are arbitrary constants for which (10) is satisfied. They must be the same in all three formulas. Then from (9) there is a linear second order differential equation for u_λ,

$$\sqrt{\theta(\lambda)}\,\frac{d\sqrt{\theta(\lambda)}\,\frac{du}{d\lambda}}{d\lambda} + (g^2\lambda^2 + h\lambda + k)u = 0,$$

or in rational form

$$\theta(\lambda)\,\frac{d^2u}{d\lambda^2} + \frac{1}{2}\,\theta'(\lambda)\,\frac{du}{d\lambda} + (g^2\lambda^2 + h\lambda + k)u = 0. \tag{11}$$

The same differential equation is also obtained for u_μ and u_ν, on replacing the variable λ by μ or ν.

XXVI.

Equilibrium of electricity on cylinders with circular cross-section and parallel axes. Conformal mapping of figures bounded by circles.

The problem of determining the distribution of static electricity or the steady state temperature in an infinite cylindrical conductor with parallel generators is solved under the following hypothesis. In the first case, the distributing forces are assumed constant along lines parallel to the generators. In the second, the temperature of the surface is constant along these lines. The solution requires that a solution of the following mathematical problem is found:

Determine a function u of the rectilinear coordinates x, y in a plane, one-fold, connected surface S bounded by arbitrary curves, that satisfies the differential equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

in the interior of S and assumes arbitrary prescribed values on the boundary.

This problem can first be reduced to a simpler one:

Determine a function $\zeta = \xi + \eta i$ of the complex argument $z = x + yi$, which is real on all the boundary curves of S, becomes infinite of the first order at a given point one of the boundary curves, and otherwise is finite and continuous on the whole surface S.

It can be shown easily with regard to such a function that it assumes each real value once and only once on each boundary curve. It can be shown further, that such a function assumes each complex value with positive imaginary part n times in the interior of the surface. Here n is the number of boundary curves of S (assuming that when going around one boundary curve in a positive direction, ζ goes from $-\infty$ to $+\infty$). Using this function, we obtain on the upper half of the plane which represents the complex variable ζ, an n-fold surface T, which is a conformal image of S, and which is bounded by the lines which coincide with the real axis in the n sheets. Since the surfaces S and T are both n-fold connected, T has $2n - 2$ simple branch points in its interior (cf. **VI**, section 7, p.) and our problem reduces to the following:

To find a function of the complex argument ζ, that branches as does T, whose real part u is continuous in the interior of T and has arbitrarily prescribed values on the n boundary lines.

Given a function $\tilde{\omega} = h + ig$ of ζ, that branches as does T, which is logarithmically infinite at an arbitrary point ϵ in the interior of T, whose imaginary part ig is continuous in T, except at ϵ, and vanishes on the boundary of T, then by Green's Theorem (**I**, section 10):

$$u_\epsilon = -\frac{1}{2\pi} \int u \frac{\partial g}{\partial \eta} \, d\xi,$$

where the integration is taken over the n boundary lines of T.

However, the function g can be determined in the following way. Extend the surface T over the whole plane of ζ by adjoining the mirror image of the upper half in the lower half (where ζ has negative imaginary part). We obtain a surface that covers the whole ζ-plane n times and has $4n - 4$ simple branch points. Hence g belongs to a class of algebraic functions for which the number p is $n - 1$ (**VI**, sections 7 and 12).

The function ig is now the imaginary part of an integral of the third kind, whose discontinuity points lie at ϵ and its conjugate ϵ', and whose moduli of periodicity are all real. Such a function is completely determined up to an additive constant and our problem is solved, as soon as we succeed in finding the function ζ of z.

We will deal further with this last problem under the hypothesis that the boundary of S is formed by n circles. The circles can either all lie outside of each other, so that the surface S stretches to infinity, or one circle can enclose all the others, whereby S is finite. With the help of a mapping by reciprocal radii, either case can easily be reduced to the other.

If the function ζ of z is determined on S, then ζ can be extended continuously across the boundary of S by taking the harmonic pole of each point of S with respect to each of the boundary circles, and assigning the function ζ the conjugate value there. Thus the region S is extended for the function ζ, and its boundary again consists of circles with which we can proceed similarly. This operation can be continued indefinitely, whereby the domain of the function ζ spreads out more and more over the whole z-plane.

In the following we avail ourselves of the notation

$$a = \overline{a'}$$

if a and a' are conjugate imaginary numbers. The relation so expressed for the two numbers is preserved if on both sides conjugate imaginary numbers are added, multiplied, or divided. Also the root can be extracted from both sides, if it is defined properly.

Now let $\zeta = \overline{\zeta'}$ and ζ, ζ' correspond to the values z, z'. Then, if r is the radius of one of the boundary circles of S and z has the value p at the center of this boundary circle:

$$\frac{z-p}{r} = \left(\frac{r}{z'-p}\right)^{-}.$$

Hence

$$z = \left(\frac{az'+b}{cz'+\delta}\right)^{-}$$

where a, b, c, δ denote constants. Now

$$\frac{dz}{d\zeta} = \left(\frac{a\delta - bc}{(cz'+\delta)^2}\,\frac{dz'}{d\zeta'}\right)^{-},$$

$$\frac{1}{\sqrt{\frac{dz}{d\zeta}}} = \left(\frac{1}{\sqrt{a\delta - bc}}\,\frac{cz'+\delta}{\sqrt{\frac{dz'}{d\zeta'}}}\right)^{-}$$

$$\frac{z}{\sqrt{\frac{dz}{d\zeta}}} = \left(\frac{1}{\sqrt{a\delta - bc}}\,\frac{az'+b}{\sqrt{\frac{dz'}{d\zeta'}}}\right)^{-}.$$

Let us write

$$\frac{1}{\sqrt{\frac{dz}{d\zeta}}} = y, \qquad \frac{z}{\sqrt{\frac{dz}{d\zeta}}} = y_1$$

and denote by y', y_1' the values that y and y_1 assume at ζ'. Then we have

$$(1)\qquad \begin{aligned} y &= \left(\frac{cy_1' + \delta y'}{\sqrt{a\delta - bc}}\right)^{-}, \\ y_1 &= \left(\frac{ay_1' + by'}{\sqrt{a\delta - bc}}\right)^{-}. \end{aligned}$$

It follows that

$$(2)\qquad \begin{aligned} \frac{d^2y}{d\zeta^2} &= \left(\frac{c\,\frac{d^2y_1'}{d\zeta'^2}+\delta\,\frac{d^2y'}{d\zeta'^2}}{\sqrt{a\delta-bc}}\right)^{-}, \\ \frac{d^2y_1}{d\zeta^2} &= \left(\frac{a\,\frac{d^2y_1'}{d\zeta'^2}+b\,\frac{d^2y'}{d\zeta'^2}}{\sqrt{a\delta-bc}}\right)^{-}. \end{aligned}$$

Now by differentiating

$$(3)\qquad z=\frac{y_1}{y}$$

it follows that

$$y\,\frac{dy_1}{d\zeta}-y_1\,\frac{dy}{d\zeta}=1,$$

$$y\,\frac{d^2y_1}{d\zeta^2}-y_1\,\frac{d^2y}{d\zeta^2}=0$$

or

$$(4)\qquad \frac{1}{y}\,\frac{d^2y}{d\zeta^2}=\frac{1}{y_1}\,\frac{d^2y_1}{d\zeta^2}.$$

Similarly,

$$(5)\qquad \frac{1}{y'}\,\frac{d^2y'}{d\zeta'^2}=\frac{1}{y_1'}\,\frac{d^2y_1'}{d\zeta'^2}.$$

From this, and from (1) and (2), it follows further that

$$(6)\qquad \frac{1}{y}\,\frac{d^2y}{d\zeta^2}=\frac{1}{y_1}\,\frac{d^2y_1}{d\zeta^2}=\left(\frac{1}{y'}\,\frac{d^2y'}{d\zeta'^2}\right)^{-}=\left(\frac{1}{y_1'}\,\frac{d^2y_1'}{d\zeta_1'^2}\right)^{-}.$$

Now let

$$(7)\qquad \frac{d^2y}{d\zeta^2}=sy.$$

Then s is a function of ζ which has conjugate imaginary values for conjugate imaginary values of ζ, and thus does not change along an arbitrary path returning to its starting point within the surface T and its symmetrical extension. Therefore s is a bounded algebraic function of ζ branched like T; y and y_1 are particular solutions of the linear differential equation (7), and z is their ratio. Conversely, take the algebraic function s in T arbitrarily so that it takes conjugate imaginary values at conjugate points, and so for real values of ζ becomes real. Take any two particular solutions of (7), then the function $z = y_1/y$ gives a conformal mapping of the surface T, which is bounded by circles. The undetermined constants that are introduced must be determined so that this image is free of singular points in its interior and thus is a one-fold surface in the z-plane, and further so that the boundary circles take given positions.

XXVII.

Examples of surfaces of least area with a given boundary.

1.

The surface of minimal area will be found whose boundary consists of three lines which intersect at two points. Thus the surface is an infinite sector having two vertices in its boundary and one in the sector passing to infinity.

Let the angles that the three lines form with one another be $\alpha\pi, \beta\pi, \gamma\pi$. The surface we are seeking maps to a spherical triangle on the sphere, whose angles are $\alpha\pi, \beta\pi$, and $\gamma\pi$. Hence $\alpha + \beta + \gamma > 1$.

The points corresponding to the two vertices and the vertex at infinity may be denoted by a, b, and c in the plane of the complex variable t (**XVII**, §13, p.). Then we have

$$u = \int \frac{\text{const. } dt}{(t-c)\sqrt{(t-a)(t-b)}}$$

or

$$u = \text{const. } \log \frac{\sqrt{\frac{t-a}{c-a}} - \sqrt{\frac{t-b}{c-b}}}{\sqrt{\frac{t-a}{c-a}} + \sqrt{\frac{t-b}{c-b}}}.$$

Assuming, as we may, that $a = 0, b = \infty$, and $c = 1$, it follows that

$$du = \text{const. } \frac{dt}{(1-t)\sqrt{t}}; \qquad u = \text{const. } \log \frac{1-\sqrt{t}}{1+\sqrt{t}}$$

and the last constant has the value $\sqrt{\frac{\gamma C}{2\pi}}$, where C is the shortest distance between the two skew lines.

If we set

$$k_1 = \sqrt{\frac{du}{d\eta}}, \quad k_2 = \eta\sqrt{\frac{du}{d\eta}},$$

as in §14 of the paper mentioned (p.), then these functions are finite at all points of the t-plane, except for $0, \infty, 1$, and are single valued. If we examine the behavior of these functions in a neighborhood of the singular points by the methods indicated on p.   , we see that k_1 and k_2 are two branches of the function

$$P\begin{Bmatrix} \frac{1}{4}-\frac{\alpha}{2} & \frac{1}{4}-\frac{\beta}{2} & -\frac{\gamma}{2} & \\ & & & t \\ \frac{1}{4}+\frac{\alpha}{2} & \frac{1}{4}+\frac{\beta}{2} & +\frac{\gamma}{2} & \end{Bmatrix},$$

and η is the quotient of two branches of this function.

2.

Let the surface of smallest area that we are seeking be bounded by two convex polygons with straight line edges lying in parallel planes. Assume the surface runs along each polygon once. In this case the surface will be twofold connected an can be changed to a simply connected surface by one transverse cut.

The image of the minimal surface on the sphere will be bounded by two systems of arcs of great circles, whose planes are perpendicular to the planes of the boundary polygons, and which therefore run together at two antipodal points on the sphere. Each of these two points corresponds to the collection of vertices of the two boundary polygons. There is a point of reversal of the normal on each of the polygon. To that point corresponds the end point of the arc of the great circle in question. Hence the image of the minimal surface will completely cover the sphere once.

If we project the surface of the sphere onto the tangent plane at one of the points in which the boundary arcs come together, then we obtain as the image of the minimal surface a piece of surface H which completely represents the plane of the complex variable η. It is bounded on the one side by a system of straight lines that run starlike from the origin to certain points $C_1, C_2, \ldots, C_n$ and on the other side by a second system of straight lines which run from certain other points $C'_1, C'_2, \ldots, C'_m$ to infinity, and hence their extensions meet at the origin. Here n and m are the number of vertices of the two given polygons.

This twofold connected surface will now be mapped to a surface T_1 in the plane of a complex variable t covering the upper half plane twice, in such a way that the two boundaries correspond to the real values of t. Since the surface is twofold connected, it contains two branch points. If we add to the surface T_1 its mirror image with respect to the real axis, we obtain a surface T that covers the whole t-plane twice, whose four branch points correspond to conjugate imaginary values of t. By introducing a new variable t' in place of t, which is related to t by a quadratic equation in both variables, we arrange for the branch points to correspond to $t' = \pm i, \pm i/k$ where k is real and < 1. In addition, any given real value of t can be made to correspond to a given real value of t' in one of the two sheets.

Therefore we must determine t as a function of the complex variable η, so that, at each point of the surface H, t has a definite value that changes continuously with position; t is real on both boundaries of H; and t is finite or infinite of first order at each point of the two boundary lines. We extend this function continuously across the boundaries by giving conjugate imaginary values to points lying symmetrically on both sides of each boundary segment. Then it is easy to see that the function $\frac{d\log\eta}{dt}$ has conjugate imaginary values for conjugate imaginary values of t. Hence it is single valued in the entire surface T, and continuous with the exception of isolated points. Therefore it must be a rational function of t, and in fact must be

$$\Delta(t) = \sqrt{(1+t^2)(1+k^2t^2)}.$$

We let $c_1, c_2, \ldots, c_n, c_1', c_2', \ldots, c_m'$ denote the real values of t which correspond to the points $C_1, C_2, \ldots, C_n,\ C_1', C_2', \ldots, C_m'$. Likewise, the real values that correspond to vertices collecting at the origin, (respectively, at infinity) of the surface H are denoted by $b_1, b_2, \ldots, b_n,\ b_1', b_2', \ldots, b_m'$. Then $\frac{d\log\eta}{dt}$ must be infinitely small of first order at

$$t = c_1, c_2, \ldots, c_n, c_1', c_2', \ldots, c_m',$$

and infinitely large of first order for

$$t = b_1, b_2, \ldots, b_n, b_1', b_2', \ldots, b_m'$$

and the branch points

$$t = \pm i,\ \pm\frac{i}{k}.$$

Therefore we may set:

$$\frac{d\log\eta}{dt} = \frac{\phi(t, \Delta(t))}{\sqrt{(1+t^2)(1+k^2t^2)}},$$

where ϕ is a rational function of t and $\Delta(t)$, which goes to zero at the points c, c', goes to infinity at the points b, b', and is determined by these conditions up to a constant real factor. However, if such a function exists there must be an equation connecting the points c, c', b, b' by virtue of which one of the points is determined by the others. (**VI**, §8, p.). Moreover, from the remarks above concerning the points c, c', b and b', one point can be chosen arbitrarily. The additive constant attached to $\log\eta$ is determined if η_0 is given as the value of η corresponding to one of the points c. We then have

$$\log\eta - \log\eta_0 = \int_c^t \frac{\phi(t,\Delta(t))dt}{\sqrt{(1+t^2)(1+k^2t^2)}}.$$

After η_0 and c are determined, there still remain $2n+2m$ undetermined constants in this expression: namely $2n+2m-2$ of the values c, c', b, b', the modulus k, and a real constant factor in ϕ.

For the determination of these constants, first, there are two conditions which say that the real part of the integral

$$\int \frac{\phi(t,\Delta(t))dt}{\sqrt{(1+t^2)(1+k^2t^2)}}$$

taken over a closed curve enclosing the two branch points i and i/k vanishes, and that the imaginary part of the same integral has the value $2\pi i$. For the remaining $2n+2m-2$ constants, we obtain the same number of conditions from the requirement that the points c and c' correspond to the given points C and C' in the η-plane.

We now consider the x-axis as being perpendicular to the plane of the two bounding polygons, and examine the image of the minimal surface in the plane of the complex variable X, after it is changed to a simply connected surface by slicing from one boundary to the other. Then the real part of X is constant on both boundaries and on each cut of the surface between the boundaries and parallel to them. The imaginary part increases steadily around such a cut, and indeed by a constant amount throughout. It follows that the image of our surface in the X-plane is bounded by a parallelogram, and is covered once. The two sides which correspond to the boundary of the surface are parallel to the imaginary axis. The other two sides, which correspond to the edges of the cut, can actually be curvilinear, but coincide with each other via a translation parallel to the imaginary axis.

This parallelogram must be mapped into the upper half T_1 of the surface T, so that both sides parallel to the imaginary axis map to the two edges of T_1, and the other two sides correspond to the two edges of a transverse cut of T_1. One such mapping is established by the function

$$X = iC\int \frac{dt}{\sqrt{(1+t^2)(1+k^2t^2)}} + C',$$

where the constant C is real, and C' can be chosen arbitrarily, if we have at our disposal the position of the origin on the x-axis. If h is the perpendicular distance between the two parallel bounding planes, it follows that

$$h = 4C \int_0^i \frac{i\,dt}{\sqrt{(1+t^2)(1+k^2t^2)}},$$

which determines the constant C.

That solves the problem, except for the determination of the constants, since by the formulae on p. one has

$$Y = \frac{1}{2} \int dX \left(\eta - \frac{1}{\eta}\right),$$

$$Z = -\frac{i}{2} \int dX \left(\eta + \frac{1}{\eta}\right)$$

whereby the coordinates of x, y, z of the minimal surface are represented as functions of two independent variables.

For the constants occurring in η, there are still two conditions. Namely, the real part of the integrals which express Y and Z are equal to 0 when taken over a closed curve in the η-plane that encloses the origin.

If we assume that h and the directions of the bounding lines are given, then in addition to the additive constants in X, Y, Z, our expression depends on $n + m - 2$ undetermined constants. For these we may take the distances of the points C and C' from the origin in the η-plane, between which there are the two relations just noted. There are just as many constants, however, which determine the mutual position of the bounding polygons. For we can retain two polygon sides for fixing the origin of the coordinate system, giving each of the remaining $n + m - 2$ a parallel shift in their plane.

The result assumes a simpler form if certain symmetries are assumed in the bounding polygons. In the following, we consider the case in which the two polygons are regular and form the top and bottom faces of a right truncated right pyramid over a regular polygonal base.

In this case, all of the reversal points of the normals lie at the midpoints of the boundary lines, and therefore fall in pairs in the same planes going through the axis of the pyramid.

If we set the y-axis perpendicular to one of the bounding lines, then a point C and a point C' will lie on the real axis of the η-plane, and have distances η_0, η_0' from the origin. The points C, respectively C', lie on two concentric circles, forming the vertices of a regular polygon on each, in such a way that each point C lies on the same radius vector as a point C'.

Since one point can be arbitrarily selected in the boundary of the surface T, we may specify that the point C lying on the real axis corresponds to $t = 0$ in one of the two sheets of T. It then follows from symmetry, that the part of the real axis in the η-plane lying between C and C' corresponds to a line in the surface T, which runs from $t = 0$ in the first sheet to the branch point $t = i$ and from there along the imaginary axis back to the point $t = 0$ in the second sheet. Hence when t has pure imaginary values, so does $\phi(t, \Delta(t))$, and the point C' corresponds to $t = 0$ in the second sheet.

Now the surface H is mapped by the substitution $\eta\eta' = \eta_0\eta_0'$ onto a congruent surface H' in such a way that the points C become the points C' and conversely (with the ordering interchanged). It follows that two points η and $\eta' = \frac{\eta_0\eta_0'}{\eta}$ lying in the surface H correspond to points that lie over each other in the two sheets of T. Since $d\log\eta + d\log\eta' = 0$, $\phi(t, \Delta(t))$ must have the same value at points that lie over each other in the two sheets. Thus $\phi(t, \Delta(t))$ can be expressed as a rational function in t, and by the remark above has the form $t\psi(t^2)$, where ψ denotes a rational function.

This occasions us to map the surface T onto a surface S using the substitution

$$\frac{1+t^2}{1+k^2t^2} = s^2,$$

whereby the upper half of the surface T corresponds to a sheet simply covering the s-plane, which is slit along the real axis between $s = 1$ and $s = 1/k$, and between $s = -1$ and $s = -1/k$. The edges of both these cuts correspond to the boundary of the surface H. From this we get the following expression for X:

$$X = \frac{h}{4K}\int \frac{ds}{\sqrt{(1-s^2)(1-k^2s^2)}},$$

where

$$K = \int_0^1 \frac{ds}{\sqrt{(1-s^2)(1-k^2s^2)}},$$

while η itself can be represented as an algebraic function of s.

If the boundary is formed from squares, then

$$\eta = c\sqrt{\frac{(1-ms)(1-m's)}{(1+ms)(1+m's)}}.$$

The points $s = 1/m$, $s = 1/m'$ on both edges of the cut correspond to the vertices of the square in one boundary. To the reversal points of the normal correspond $s = 1$, $s = 1/k$. To a point on both edges of the cut corresponds $s = 1/n$, which is to be determined by the equation $\frac{d\log\eta}{ds} = 0$. We have

$$1 > m > n > m' > k.$$[1]

When the boundary consists of equilateral triangles,

$$\eta = c\left(\frac{1-ms}{1+ms}\right)^{\frac{2}{3}}\left(\frac{1-ks}{1+ks}\right)^{\frac{1}{3}}.$$

In order to examine the possibility of determining the constants in this case, first set $s = \pm 1$, obtaining

$$\eta_0 = c\left(\frac{1-m}{1+m}\right)^{\frac{2}{3}}\left(\frac{1-k}{1+k}\right)^{\frac{1}{3}}; \quad \eta_0' = c\left(\frac{1+m}{1-m}\right)^{\frac{2}{3}}\left(\frac{1+k}{1-k}\right)^{\frac{1}{3}}.$$

Hence

$$c = \sqrt{\eta_0\eta_0'}, \quad \sqrt{\frac{\eta_0}{\eta_0'}} = \left(\frac{1-m}{1+m}\right)^{\frac{2}{3}}\left(\frac{1-k}{1+k}\right)^{\frac{1}{3}}$$

and for the special cases where the two triangles are congruent,

$$\eta_0\eta_0' = 1, \quad c = 1.$$

To the vertices of the triangle in one of the boundaries correspond the points $s = 1/m$ on both edges of the cut and the point $1/k$, so that we must have $k < m < 1$. The first reversal point of the normal occurs for $s = 1$, the other two correspond to a point $s = 1/n$ on both edges of the cut, so that

$$k < n < m.$$

[1]The preceding considerations may be extended to many cases in which the two polygons are not regular. Thus the expression for η above remains valid for a boundary consisting of two rectangles whose midpoints lie on a line perpendicular to their planes, if it is assumed that the modulus of $\eta\eta'$ has the same value for the reversal points of the normals. This occurs, for example, if both rectangles are congruent.

From the equation $\frac{d \log \eta}{dx} = 0$, we determine n:

$$n^2 = \frac{km(m+2k)}{2m+k}.$$

For each choice of k, m satisfying the condition

$$0 < k < m < 1,$$

this yields a value of n between k and m.

However, there is another equation between m, n and k which expresses the fact that $\eta^3 = \eta_0'^3$ when $s = 1/n$. This equation is

$$\left(\frac{1-m}{1+m}\right)^2 \frac{1-k}{1+k} = \left(\frac{n-m}{n+m}\right)^2 \frac{n-k}{n+k}.$$

If n is eliminated from these two equations, the following relation is obtained between k and m:

$$k \left(\frac{1+m^2+2mk}{k(1+m^2)+2m}\right)^2 = m \left(\frac{2k+m}{k+2m}\right)^3,$$

from which k is to be determined in terms of m.

For $k = 0$ the left side of the equation is zero, and the right side is $m/8$. For $k = m$ the difference between the left and right side is

$$\frac{(1-m^2)^3}{m(3+m^2)^2}.$$

This is positive for $m < 1$. Hence for each value of m less than 1 there is an odd number of values of $k < m$. Since it now follows easily that

$$\log k \frac{(1+m^2+2mk)^2(k+2m)^3}{(k(1+m^2)+2m)^2(2k+m)^3}$$

has only one maximum between $k = 0$ and $k = m$, then for each $m < 1$ one and only one value of k can be found satisfying our conditions. Accordingly there is only one corresponding value of n. For the two bounds $m = 0, 1$ we obtain $k = n = m$.

The functions X, Y, and Z are now given by the expressions

$$X = \frac{h}{4K}\int_1^s \frac{ds}{\sqrt{(1-s^2)(1-k^2s^2)}},$$
$$Y = \frac{h}{8K}\int_1^s \frac{ds}{\sqrt{(1-s^2)(1-k^2s^2)}}\left(\eta - \frac{1}{\eta}\right),$$
$$Z = -\frac{ih}{8K}\int_1^s \frac{ds}{\sqrt{(1-s^2)(1-k^2s^2)}}\left(\eta + \frac{1}{\eta}\right),$$

if we have the additive constants at our disposal.

The two remaining constants m and $\sqrt{\eta_0\eta_0'}$ can be determined from the given lengths of the sides of the triangles. Denoting these by a and b, we obtain

$$a = \frac{ih}{2K}\int_1^{1/m} \frac{ds}{\sqrt{(1-s^2)(1-k^2s^2)}}\left(\eta + \frac{1}{\eta}\right),$$
$$b = \frac{ih}{2K}\int_1^{1/m} \frac{ds}{\sqrt{(1-s^2)(1-k^2s^2)}}\left(\frac{\eta}{\eta_0\eta_0'} + \frac{\eta_0\eta_0'}{\eta}\right).$$

In the particular case $a = b$, we have $\eta_0\eta_0' = 1$, and one transcendental equation

$$\frac{a}{h} = \frac{i}{2K}\int_1^{1/m} \frac{ds}{\sqrt{(1-s^2)(1-k^2s^2)}}\left(\eta + \frac{1}{\eta}\right)$$

remains for the determination of the constant m. If we let m pass from 0 to 1 in the expression on the right, it remains positive; however, it becomes infinitely large at both limits. Thus there must be a minimum at some intermediate value of m. It follows that there is a lower bound for the ratio a/h beyond which there are no solutions of the problem, while for each value of a/h above this bound there are two values of m, hence two solutions of the problem exist. We may assume that only the smaller of the two values of m actually corresponds to the least area of the surface.

XXVIII.

Fragments concerning limiting cases of the elliptic modular functions.

1.

Addendum to §40 of Jacobi, Fundamenta nova theoriae functionum ellipticarum.

The formula contained in this section of Jacobi's work appear to merit further consideration in the case where the modulus of the variable q is 1, because they then provide examples of functions of a single variable which are discontinuous for any value of the argument.

In fact the series considered, for the most part, do not converge when the modulus of q is 1, but by integrating them convergent series can be obtained, and formulae (1)–(7) thus yield the following.

$$(48)\qquad \int_0 (\log k - \log 4\sqrt{q})\,\frac{dq}{q} = -4\log(1+q) + \frac{4}{4}\,\log(1+q^2) - \frac{4}{9}\,\log(1+q^3) + \frac{4}{16}\,\log(1+q^4) - \cdots ,$$

$$(49)\qquad \int_0 -\log k'\,\frac{dq}{q} = 4\log\left(\frac{1+q}{1-q}\right) + \frac{4}{9}\,\log\left(\frac{1+q^3}{1-q^3}\right) + \frac{4}{25}\,\log\left(\frac{1+q^5}{1-q^5}\right) + \cdots ,$$

$$(50)\qquad \int_0 \log\frac{2K}{\pi}\,\frac{dq}{q} = 4\log(1+q) + \frac{4}{9}\,\log(1+q^3) + \frac{4}{25}\,\log(1+q^5) + \cdots ,$$

$$(51)\qquad \int_0 \left(\frac{2K}{\pi} - 1\right)\frac{dq}{q} = -4\log(1-q) + \frac{4}{3}\,\log(1-q^3) - \frac{4}{5}\,\log(1-q^5) + \cdots$$
$$= 2i\,\log\left(\frac{1-qi}{1+qi}\right) + \frac{2i}{2}\,\log\left(\frac{1-q^2i}{1+q^2i}\right) + \frac{2i}{3}\,\log\left(\frac{1-q^3i}{1+q^3i}\right) + \cdots ,$$

(52)
$$\int_0 \frac{2kK}{\pi}\,\frac{dq}{q} = 4\log\left(\frac{1+q^{1/2}}{1-q^{1/2}}\right) - \frac{4}{3}\log\left(\frac{1+q^{3/2}}{1-q^{3/2}}\right) + \frac{4}{5}\log\left(\frac{1+q^{5/2}}{1-q^{5/2}}\right) + \cdots$$
$$= 4i\log\left(\frac{1+q^{1/2}i}{1-q^{1/2}i}\right) + \frac{4i}{3}\log\left(\frac{1-q^{3/2}i}{1+q^{3/2}i}\right) + \frac{4i}{5}\log\left(\frac{1-q^{5/2}i}{1+q^{5/2}i}\right) + \cdots,$$

(53)
$$\int\left(\frac{2k'K}{\pi} - 1\right)\frac{dq}{q} = -4\log(1+q) + \frac{4}{3}\log(1+q^3) - \frac{4}{5}\log(1+q^5) + \cdots$$
$$= -2i\log\left(\frac{1-qi}{1+qi}\right) + \frac{2i}{2}\log\left(\frac{1-q^2i}{1+q^2i}\right) - \frac{2i}{3}\log\left(\frac{1-q^3i}{1+q^3i}\right) + \cdots,$$

(54)
$$\int_0\left(\frac{2\sqrt{k'}K}{\pi} - 1\right)\frac{dq}{q} = -\frac{4}{2}\log(1+q^2) + \frac{4}{6}\log(1+q^6) - \frac{4}{10}\log(1+q^{10}) + \frac{4}{14}\log(1+q^{14}) - \cdots$$
$$= -\frac{2i}{2}\log\left(\frac{1-q^2i}{1+q^2i}\right) + \frac{2i}{4}\log\left(\frac{1-q^4i}{1+q^4i}\right) - \frac{2i}{6}\log\left(\frac{1-q^6i}{1+q^6i}\right) + \frac{2i}{8}\log\left(\frac{1-q^8i}{1+q^8i}\right) - \cdots$$

where the logarithms are to be interpreted in such a way that they vanish for $q = 0$.

The formulae corresponding to the functions involving the powers of q, given by Jacobi, can be represented as follows.

(55) $$\int_0 (\log k - \log 4\sqrt{q})\,\frac{dq}{q} = -4\sum \frac{\phi(p)}{p^2}\left(q^p - \frac{3}{4}\,q^{2p} - \frac{3}{16}\,q^{4p} - \frac{3}{64}\,q^{8p} - \frac{3}{256}\,q^{16p} - \cdots\right),$$

(56) $$\int_0 -\log k'\,\frac{dq}{q} = 8\sum \frac{\phi(p)}{p^2}\,q^p,$$

(57) $$\int_0 \log\frac{2K}{\pi}\,\frac{dq}{q} = 4\sum \frac{\phi(p)}{p^2}\left(q^p - \frac{1}{2}\,q^{2p} - \frac{1}{4}\,q^{4p} - \frac{1}{8}\,q^{8p} - \frac{1}{16}\,q^{16p} - \cdots\right),$$

(58) $$\int_0 \left(\frac{2K}{\pi} - 1\right)\frac{dq}{q} = 4\sum \frac{\psi(n)q^{2^\ell(4m-1)^2 n}}{2^\ell(4m-1)^2 n},$$

(59) $$\int_0 \frac{2kK}{\pi}\,\frac{dq}{q} = 8\sum \frac{\psi(n)q^{(4m-1)^2 n/2}}{(4m-1)^2 n},$$

(60) $$\int_0 \left(\frac{2k'K}{\pi} - 1\right)\frac{dq}{q} = -4\sum \frac{\psi(n)q^{(4m-1)^2 n}}{(4m-1)^2 n} + 4\sum \frac{\psi(n)q^{2^{\ell+1}(4m-1)^2 n}}{2^{\ell+1}(4m-1)^2 n},$$

(61) $$\int_0 \left(\frac{2\sqrt{k'}K}{\pi} - 1\right)\frac{dq}{q} = -4\sum \frac{\psi(n)q^{2(4m-1)^2 n}}{2(4m-1)^2 n} + 4\sum \frac{\psi(n)q^{2^{\ell+2}(4m-1)^2 n}}{2^{\ell+2}(4m-1)^2 n}.$$

To discuss more precisely the nature of these functions, we shall begin by enunciating the following general theorem as a lemma.

If the series

$$a_0 + a_1 + a_2 + \cdots$$

whose terms are summed in the order in which we have written them, has a convergent sum, then the function of the variable r which is the sum of the series

$$a_0 + a_1 r + a_2 r^2 + \cdots$$

tends to the sum of the series when r tends to 1.

From the lemma it is easily deduced that if $f(q)$ is a function of the complex variable q defined by the series

$$a_0 + a_1 q + a_2 q^2 + \cdots$$

when the modulus of q is less than 1, then the value of the sum of the series (if it has a sum when q has the value q_0 of modulus 1) is given by the value of the function $f(q)$ obtained when q converges towards q_0 in such a way that its modulus changes by the same amount; that is, with the usual geometrical representation, when the point which represents q approaches q_0 along a line perpendicular to the boundary of the domain of definition of the function.

By these means, the various formulae giving the values of the functions listed can be justified, although (48)–(54) are obvious enough.

For brevity, we denote by (x) the distance of the real number x from the nearest integer, except when x differs from an integer by $1/2$, in which case $(x) = 0$. Moreover, we define $E(x)$ to be the greatest integer not exceeding x. Then we obtain from (48), by assigning to q the value $q_0 = e^{xi}$:

$$(62)\quad \int_0^{e^{xi}} (\log k - \log 4\sqrt{q})\,\frac{dq}{q}$$

$$= -2\log\left(4\cos\frac{x^2}{2}\right) + \frac{2}{4}\log\left(4\cos\frac{2x^2}{2}\right) - \frac{2}{9}\log\left(4\cos\frac{3x^2}{2}\right)$$

$$+\frac{2}{16}\log\left(4\cos\frac{4x^2}{2}\right) - \cdots$$

$$-4\pi i\left(\frac{x}{2\pi}\right) + \frac{4\pi i}{4}\left(\frac{2x}{2\pi}\right) - \frac{4\pi i}{9}\left(\frac{3x}{2\pi}\right) + \frac{4\pi i}{16}\left(\frac{4x}{2\pi}\right) - \cdots$$

$$= 2\sum\frac{(-1)^n\log\left(4\cos\frac{nx^2}{2}\right)}{n^2}\left[+4\pi i\sum\frac{(-1)^n}{n^2}\left(\frac{nx}{2\pi}\right)\right].$$

The imaginary part of this series converges for all x, but the real part does not converge when $\frac{x}{2\pi}$ is a surd. However, when $\frac{x}{2\pi} = \frac{m}{n}$ and m/n are relatively prime integers, the series takes the form:

1°: if n is odd,

$$\frac{\pi^2}{n^2}\sum_{s=1}^{n-1}\frac{(-1)^s\cos\frac{\pi s}{n}}{\sin\frac{\pi s^2}{n}}\log\left(4\cos\frac{s\pi m^2}{n}\right)-\frac{\pi^2}{6n^2}\log 4;$$

2°: if n is even, and p denotes an odd integer,

$$\frac{\pi^2}{n^2}\sum_{s=1}^{(n/2)-1}\frac{2(-1)^s\log\left(4\cos\frac{s\pi m^2}{n}\right)}{\sin\frac{\pi s^2}{n}}+\frac{\pi^2}{3n^2}\log 4$$
$$+\frac{2\pi^2}{n^2}(-1)^{n/2}\left(\log\frac{q_0-q}{q_0+q}+\log n+\frac{8}{\pi^2}\sum\frac{\log p}{p^2}\right).$$

Here it is clear that the formula must be understood as meaning that, after subtracting the function

$$\frac{2\pi^2}{n^2}(-1)^{n/2}\log\frac{q_0-q}{q_0+q}$$

whose value is determined by making q tend to the limit q_0 in the manner specified earlier, the remainder converges to the finite number given by the remaining terms of the formula.

Similarly, we obtain

(63)
$$\int_0^{e^{xi}}-\log k'\,\frac{dq}{q}=-2\log\tan\frac{x^2}{2}-\frac{2}{9}\log\tan\frac{3x^2}{2}-\frac{2}{25}\log\tan\frac{5x^2}{2}-\cdots$$
$$+4\pi i\left(\left(\frac{x}{2\pi}\right)-\left(\frac{x}{2\pi}+\frac{1}{2}\right)\right)$$
$$+\frac{4\pi i}{9}\left(\left(\frac{3x}{2\pi}\right)-\left(\frac{3x}{2\pi}+\frac{1}{2}\right)\right)$$
$$+\frac{4\pi i}{25}\left(\left(\frac{5x}{2\pi}\right)-\left(\frac{5x}{2\pi}+\frac{1}{2}\right)\right)+\cdots$$
$$=-\sum_{-\infty}^{\infty}\frac{\log\tan\frac{px^2}{2}}{p^2}\left[+4\pi i\sum_{1}^{\infty}\frac{1}{p^2}\left(\left(\frac{px}{2\pi}\right)-\left(\frac{px}{2\pi}+\frac{1}{2}\right)\right)\right],$$

$$\text{(64)} \qquad \int_0^{e^{xi}} \log \frac{2K}{\pi} \frac{dq}{q} = 2\log\left(4\cos\frac{x^2}{2}\right) + \frac{2}{9}\log\left(4\cos\frac{3x^2}{2}\right) + \cdots$$

$$+ 4\pi i\left(\frac{x}{2\pi}\right) + \frac{4\pi i}{9}\left(\frac{3x}{2\pi}\right) + \frac{4\pi i}{25}\left(\frac{5x}{2\pi}\right) + \cdots$$

$$= \sum_{-\infty}^{\infty} \frac{\log\left(4\cos\frac{px^2}{2}\right)}{p^2}\left[+4\pi i \sum_{1}^{\infty} \frac{1}{p^2}\left(\frac{px}{2\pi}\right)\right],$$

$$\text{(65)} \quad \int_0^{e^{xi}} \left(\frac{2K}{\pi} - 1\right)\frac{dq}{q} = -2\log\left(4\sin\frac{x^2}{2}\right) + \frac{2}{3}\log\left(4\sin\frac{3x^2}{2}\right)$$

$$- \frac{2}{5}\log\left(4\sin\frac{5x^2}{2}\right) + \cdots$$

$$- 4\pi i\left(\frac{x}{2\pi} + \frac{1}{2}\right) + \frac{4\pi i}{3}\left(\frac{3x}{2\pi} + \frac{1}{2}\right) - \cdots$$

$$= i\log\tan\left(\frac{2x+\pi}{4}\right)^2 + \frac{i}{2}\log\tan\left(\frac{4x+\pi}{4}\right)^2$$

$$+ \frac{i}{3}\log\tan\left(\frac{6x+\pi}{4}\right)^2 + \cdots$$

$$+ 2\pi\left(\left(\frac{x}{2\pi} + \frac{1}{4}\right) - \left(\frac{x}{2\pi} + \frac{3}{4}\right)\right)$$

$$+ \frac{2\pi}{2}\left(\left(\frac{2x}{2\pi} + \frac{1}{4}\right) - \left(\frac{2x}{2\pi} + \frac{3}{4}\right)\right)$$

$$+ \frac{2\pi}{3}\left(\left(\frac{3x}{2\pi} + \frac{1}{4}\right) - \left(\frac{3x}{2\pi} + \frac{3}{4}\right)\right) + \cdots,$$

$$(66)\quad \int_0^{e^{xi}} \frac{2kK}{\pi}\,\frac{dq}{q} = -2\log\tan\frac{x^2}{4} + \frac{2}{3}\log\tan\frac{3x^2}{4} - \frac{2}{5}\log\tan\frac{5x^2}{4} + \cdots$$

$$+ 4\pi i\left(\left(\frac{x\pi}{4}\right) - \left(\frac{x\pi}{4} + \frac{1}{2}\right)\right)$$

$$- \frac{4\pi i}{3}\left(\left(\frac{3x}{4\pi}\right) - \left(\frac{3x}{4\pi} + \frac{1}{2}\right)\right) + \cdots$$

$$= 2i\log\tan\left(\frac{x+\pi}{4}\right)^2 + \frac{2i}{3}\log\tan\left(\frac{3x+\pi}{4}\right)^2$$

$$+ \frac{2i}{5}\log\tan\left(\frac{5x+\pi}{4}\right)^2 + \cdots$$

$$+ 4\pi\left(\left(\frac{x}{4\pi} + \frac{1}{4}\right) - \left(\frac{x}{4\pi} + \frac{3}{4}\right)\right)$$

$$+ \frac{4\pi}{3}\left(\left(\frac{3x}{4\pi} + \frac{1}{4}\right) - \left(\frac{3x}{4\pi} + \frac{3}{4}\right)\right) + \cdots,$$

$$(67)\quad \int_0^{e^{xi}} \left(\frac{2k'K}{\pi} - 1\right) = -2\log\left(4\cos\frac{x^2}{2}\right) + \frac{2}{3}\log\left(4\cos\frac{3x^2}{2}\right)$$

$$- \frac{2}{5}\log\left(4\cos\frac{5x^2}{2}\right) + \cdots$$

$$- 4\pi i\left(\frac{x}{2\pi}\right) + \frac{4\pi i}{3}\left(\frac{3x}{2\pi}\right) - \frac{4\pi i}{5}\left(\frac{5x}{2\pi}\right) + \cdots$$

$$= -i\log\tan\left(\frac{2x+\pi}{4}\right)^2 + \frac{i}{2}\log\tan\left(\frac{4x+\pi}{4}\right)^2$$

$$- \frac{i}{3}\log\tan\left(\frac{6x+\pi}{4}\right)^2 + \cdots$$

$$- 2\pi\left(\left(\frac{x}{2\pi} + \frac{1}{4}\right) - \left(\frac{x}{2\pi} + \frac{3}{4}\right)\right)$$

$$+ \frac{2\pi}{2}\left(\left(\frac{2x}{2\pi} + \frac{1}{4}\right) - \left(\frac{2x}{2\pi} + \frac{3}{4}\right)\right) - \cdots,$$

(68)

$$\begin{aligned}\int_0^{e^{xi}} \left(\frac{2\sqrt{k}'K}{\pi} - 1\right)\frac{dq}{q} &= -\log(4\cos x^2) + \frac{1}{3}\log(4\cos 3x^2) \\ &\quad - \frac{1}{5}\log(4\cos 5x^2) + \cdots \\ &\quad - 2\pi i\left(\frac{x}{\pi}\right) + \frac{2\pi i}{3}\left(\frac{x}{\pi}\right) - \frac{2\pi i}{5}\left(\frac{5x}{\pi}\right) + \cdots \\ &= -\frac{i}{2}\log\tan\left(x + \frac{\pi}{4}\right)^2 + \frac{i}{4}\log\tan\left(2x + \frac{\pi}{4}\right)^2 \\ &\quad - \frac{i}{6}\log\tan\left(3x + \frac{\pi}{4}\right)^2 + \cdots \\ &\quad - \pi\left(\left(\frac{x}{\pi} + \frac{1}{4}\right) - \left(\frac{x}{\pi} + \frac{3}{4}\right)\right) \\ &\quad + \frac{\pi}{2}\left(\left(\frac{2x}{\pi} + \frac{1}{4}\right) - \left(\frac{2x}{\pi} + \frac{3}{4}\right)\right) \\ &\quad - \frac{\pi}{3}\left(\left(\frac{3x}{\pi} + \frac{1}{4}\right) - \left(\frac{3x}{\pi} + \frac{3}{4}\right)\right) + \cdots\end{aligned}$$

The imaginary part of the formula (65), after substituting $\frac{2\pi m}{n}$ for x, becomes:

1°: when n is an even integer,

$$\sum_{s=0}^{\infty} -4\pi i \sum_{1\le p\le n-1} \frac{(-1)^{(p-1)/2}}{p+ns}\left(\frac{pm}{n} + \frac{1}{2}\right)(-1)^{ns/2};$$

2°: when n is an odd integer,

$$\sum_{s=0}^{\infty} -4\pi i \sum_{1\le p\le 2n-1} \frac{(-1)^{(p-1)/2}}{p+2ns}\left(\frac{pm}{n} + \frac{1}{2}\right)(-1)^{ns},$$

which clearly has a finite value unless $n \equiv 0 \pmod 4$.

The convergence of the series

$$a_0 + a_1 + a_2 + a_3 + \cdots$$

implies that, given any quantity ϵ no matter how small, it is always possible to assign a term a_n, beyond which the sum of the subsequent terms up to and including a_m shall always be less in absolute value than ϵ. For brevity, let

$$\begin{aligned}\epsilon_{n+1} &= a_{n+1},\\ \epsilon_{n+2} &= a_{n+1} + a_{n+2},\\ \epsilon_{n+3} &= a_{n+1} + a_{n+2} + a_{n+3}, \dots\end{aligned}$$

Then the function

$$f(r) = a_0 + a_1 r + a_2 r^2 + \cdots$$

can readily be expressed in the form

$$\begin{aligned}&a_0 + a_1 r + a_2 r^2 + \cdots + a_n r^n + \epsilon_{n+1} r^{n+1} + (\epsilon_{n+2} - \epsilon_{n+1}) r^{n+2}\\ &\qquad + (\epsilon_{n+3} - \epsilon_{n+2}) r^{n+3} + \cdots\\ &= a_0 + a_1 r + a_2 r^2 + \cdots + a_n r^n + \epsilon_{n+1}(r^{n+1} - r^{n+2})\\ &\qquad + \epsilon_{n+2}(r^{n+2} - r^{n+3}) + \cdots\end{aligned}$$

Hence it is evident that, as r tends to the limit 1, the function $f(r)$ differs from the value of the series $a_0 + a_1 + a_2 + \cdots$ by an arbitrarily small amount. As for the sums of the terms whose indices exceed n, $\epsilon_{n+1}, \epsilon_{n+2}, \dots$ are all, by hypothesis, less than ϵ in absolute value, and since the differences $r^{n+1} - r^{n+2}$ are all positive, it is clear that the sum of these terms is

$$\begin{aligned}&< \epsilon(r^{n+1} - r^{n+2}) + \epsilon(r^{n+2} - r^{n+3}) + \cdots\\ &< \epsilon r^{n+1}.\end{aligned}$$

On the other hand, the sum of the terms whose indices do not exceed n is an algebraic function of r whose value can be made to approach as closely as we wish to

$$a_0 + a_1 + \cdots + a_n$$

by choosing r sufficiently close to 1. It therefore follows that the function $f(r)$ differs from the value of the series

$$a_0 + a_1 + a_2 + \cdots$$

by an amount which can be indefinitely reduced as r tends to 1.

From this theorem, to which Dirichlet, who attributed it to Abel, has recently (September 14, 1852) drawn attention, in a paper which came to my attention after the preceding pages were written, it can easily be deduced that...

2.

$$\log k = \log 4\sqrt{q} + \sum (-1)^n \frac{4}{n}\ \frac{q^n}{1+q^n},\ q = e^{xi}.$$

1) $x = \frac{2m}{n}\pi$, n odd. Let $\alpha = e^{2\pi i/2n}$. Then

$$\begin{aligned}
\log k &= i\left(\frac{x}{2} + \sum (-1)^s \frac{2}{s}\ \tan \frac{sx}{2}\right) \\
&= i\left(\frac{x}{2} + \sum_{t=0}^{\infty}\sum_{s=1}^{2n} (-1)^s \frac{2}{2nt+s}\ \tan \frac{sm\pi}{n}\right) \\
&= i\,\frac{x}{2} + 2i\int_0^1 \sum_{s=1}^{2n} (-1)^s \tan \frac{sm\pi}{n}\ \frac{x^{s-1}}{1-x^{2n}}\, dx \\
&= i\,\frac{x}{2} + 2\int_0^1 \sum_{s=1}^{2n} (-1)^s \frac{\alpha^{2sm}-1}{\alpha^{2sm}+1}\ \frac{1}{2n}\sum_{t=1}^{2n} \frac{\alpha^{-ts}\alpha^t}{1-\alpha^t x}\, dx \\
&= i\,\frac{x}{2} + \frac{1}{2n}\int_0^1 \sum_{t=1}^{2n} \frac{\alpha^t dx}{1-\alpha^t x}\, 2\sum_{\sigma=1}^{n-1}\ \sum_{s=1}^{2n} (-1)^{s+\sigma-1}\alpha^{s(2m\sigma-t)}
\end{aligned}$$

$$\left(\text{since } \frac{1}{1+r\alpha^{2sm}} = \sum \frac{(-1)^\sigma \alpha^{2s\sigma m} r^\sigma}{1-r^{2n}} = -\frac{1}{2n}\sum_{0}^{2n-1}(-1)^\sigma \sigma \alpha^{2s\sigma m} \right.$$

$$\left. = \frac{1}{2}\sum_{0}^{n-1}(-1)^\sigma \alpha^{2s\sigma m}\right)$$

$$\begin{aligned}
&= \frac{ix}{2} + 2\sum_{1}^{n-1} \log(1-\alpha^{n+2m\sigma})(-1)^\sigma \\
&= \frac{ix}{2} + \sum_{1}^{n-1} \log \alpha^{2m\sigma}(-1)^\sigma \\
&= \frac{ix}{2} + 2\pi i\left(\sum_{1}^{n-1} \frac{2m\sigma}{2n}(-1)^\sigma - \sum_{1}^{n-1}(-1)^\sigma E\left(\frac{2m\sigma}{2n}+\frac{1}{2}\right)\right).
\end{aligned}$$

2) $x = \frac{m\pi}{n}$, m, n odd. Let $\alpha = e^{2\pi i/4n}$. Then

$$\log k = -\frac{q+q_0}{q-q_0}\,\frac{3}{2n^2}\sum_1^\infty \frac{1}{s^2} - \frac{1}{n}\log\frac{1+q^n}{1-q^n}$$

$$+\frac{x}{2}i + 2\int_0^1 \sum_{s=1}^{4n-1} (-1)^s \frac{x^{s-1}dx}{1-x^{4n}}\,\frac{\alpha^{2ms}-1}{\alpha^{2ms}+1}$$

$$= A + \frac{x}{2}i + 2\int_0^1 \sum_{t=1}^{4n} \frac{\alpha^t dx}{1-\alpha^t x}\,\frac{1}{4n}\cdot -\frac{1}{2n}\sum_{s=1}^{4n-1}\sum_{\sigma=0}^{2n-1}(-1)^{s+\sigma}\sigma\,\alpha^{2\sigma sm}(\alpha^{2ms}-1)\alpha^-$$

$$= A + \frac{x}{2}i + 2.2\pi i \sum_1^{n-1} \frac{s}{n}(-1)^s\left(\frac{ms-n}{2n} - E\left(\frac{ms}{2n}\right)\right)$$

$$= A + \pi i\left(\frac{m-\mu}{2} + \frac{\mu}{2n} + 2\sum_{s=1}^{n-1} E\left(\frac{\mu s}{2n}\right)(-1)^s - 2\sum_{s=1}^{n-1} E\left(\frac{ms}{2n}\right)(-1)^s\right)$$

(where $m\mu \equiv 1 \pmod{2n}$).

3) $x = \frac{m}{2n}\pi$, m odd. Let $\alpha = e^{2\pi i/8n}$. Then

$$\log k = \frac{q+q_0}{q-q_0}\,\frac{3}{8n^2}\sum\frac{1}{s^2} + \frac{1}{2n}\log\left(\frac{1+q^{2n}}{1-q^{2n}}\right)$$

$$+\frac{xi}{2} + i\sum_t \sum_{s=1}^{8n-1} (-1)^2 \frac{2}{8nt+s}\tan\frac{sm\pi}{4n}$$

$$= A + \frac{x}{2}i + 2\int_0^1 \sum_{s=1}^{8n-1} \frac{x^{s-1}dx}{1-8x^n}\,\frac{\alpha^{2ms}-1}{\alpha^{2ms}+1}(-1)^s$$

$$= A + \frac{x}{2}i + 2\int_0^1 \sum_{t=1}^{8n} \frac{\alpha^t dx}{1-\alpha^t x}\,\frac{1}{8n}\cdot -\frac{1}{4n}\sum_{s=1}^{8n-1}\sum_{\sigma=0}^{4n-1}(-1)^{s+\sigma}\sigma\alpha^{2s\sigma m}(\alpha^{2ms}-1)\alpha^{-st}$$

$$= A + \frac{x}{2}i + 2\sum_{r=1}^{4n-1}\log(1-\alpha^{4n+2rm})\,\frac{1}{8n}\cdot\frac{1}{4n}(8n((-1)^{r-1}(r-1) - (-1)^r r)$$

$$+ 8n(-1)^r(4n-1))$$

(substituting $t \equiv 2rm + 4n \pmod{8n}$)

$$= A + \frac{x}{2} i + 2 \sum_{s=-2n+1}^{2n-1} \log(1 - \alpha^{2sm}) \cdot \frac{-s}{2n} (-1)^s$$

$$= A + \frac{x}{2} i - 4 \sum_{s=0}^{2n-1} \log(1 - \alpha^{2sm}) \frac{s}{4n} (-1)^s$$

$$= A + \frac{x}{2} i - 4 \sum_{s=0}^{2n-1} \left(\frac{sm}{4n} + \frac{1}{2}\right) \left(\frac{s}{4n}\right) (-1)^s 2\pi i$$

where (x) is as above.

Now

$$-\log k' = 8 \sum \frac{1}{t} \frac{q^t}{1 - q^t} = 4i \sum \frac{1}{t \sin t\, x}, \quad q = e^{xi}.$$

1) $x = \frac{m}{2n}\pi$, m odd. Let $\alpha = e^{2\pi i/4n}$. Then

$$-\log k' = 4i \sum_{t=0}^{\infty} \sum_{s=1}^{4n-1} \frac{1}{4nt + s} \frac{1}{\sin \frac{sm\pi}{2n}}$$

$$= 8 \int_0^1 \sum_s \frac{x^{s-1}}{1 - x^{4n}} \frac{\alpha^{sm}}{1 - \alpha^{2ms}}$$

$$= 8 \int_0^1 \sum_{t=1}^{4n} \frac{\alpha^t dx}{1 - \alpha^t x} \frac{1}{4n} \cdot -\frac{1}{2n} \sum_{s=1}^{4n-1} \sum_{\sigma=0}^{2n-1} \sigma \alpha^{ms(2\sigma+1)} \alpha^{-ts}$$

$$= \sum_0^{n-1} \left[\log(1 + \alpha^{m(2r+1)}) - \log(1 + \alpha^{-m(2r+1)})\right]$$

$$\left(\text{using } \frac{1}{1 - r\alpha^{2ms}} = \sum_{\sigma=0}^{2n-1} \frac{r^\sigma \alpha^{2ms\sigma}}{1 - r^{2n}}, \right.$$

$$\left. \frac{1}{1 - \alpha^{2ms}} = -\frac{1}{2n} \sum_{\sigma=0}^{2n-1} \sigma \alpha^{2ms\sigma} = \frac{1}{2} \sum_{\sigma=0}^{n-1} \alpha^{2ms\sigma} \right)$$

$$= -\pi i \left((m-2)n - 4 \sum_{s=0}^{n-1} E\left(\frac{m(2s+1)}{4n}\right) \right).$$

2) $x = \frac{m\pi}{n}$, n odd. Let $\alpha = e^{2\pi i/2n}$. Then

$$-\log k' = -\frac{q+q_0}{q-q_0}\,\frac{\pi^2}{4n^2}\,q_0^{-n} + 8\int_0^1 \sum_{s=1}^{2n-1} \frac{x^{s-1}dx}{1-x^{2n}}\,\frac{\alpha^{ms}}{1-\alpha^{2ms}}$$

$$= A + 8\int_0^1 \sum_{t=1}^{2n} \frac{\alpha^t dx}{1-\alpha^t x}\cdot -\frac{1}{2n}\sum_{s=1}^{2n-1}\sum_{\sigma=0}^{n-1}\left(\frac{\sigma-\left(\frac{n-1}{2}\right)}{n}\right)\alpha^{ms(2\sigma+1)}\alpha^{-ts}$$

$$= A + 8\sum_0^{n-1}\log(1-\alpha^{m(2r+1)})\,\frac{1}{2n}\left(\frac{r-\frac{n-1}{2}}{n}\right)n$$

$$-8\sum\log(1-\alpha^{m(2r+1)+n})\,\frac{1}{2n}\left(\frac{r-\frac{n-1}{2}}{n}\right)n$$

(using 1) $t \equiv m(2r+1) \pmod{2n}$, 2)$t \equiv m(2r+1)+n \pmod{2n}$)

$$= A + 8\sum_1^{(n-1)/2}\frac{1}{2}\left(\frac{s}{n}\right)\left(\log(1-\alpha^{2ms+mn}) - \log(1-\alpha^{-2ms+mn})\right)$$

$$-4\sum\left(\frac{s}{n}\right)\left(\log(1-\alpha^{2ms+(m+1)n}) - \log(1-\alpha^{-2ms+(m+1)n})\right)$$

$$= A + 8\pi i\sum_1^{(n-1)/2}\left(\frac{s}{n}\right)\left(\left(\frac{2ms+(m+1)n}{2n}\right) - \left(\frac{2ms+mn}{2n}\right)\right)$$

$$= A + 4\pi i\sum\left(\frac{s}{n}\right)(\cdots)$$

$$= A + 4\pi i\sum\left(\frac{\mu s}{n}\right)\left(\left(\frac{2s+(m+1)n}{2n}\right) - \left(\frac{2s+mn}{2n}\right)\right)$$

(where $m\mu \equiv 1 \pmod{n}$)

$$= A + 4\pi i(-1)^{m+1}\sum_1^{(n-1)/2}\left(\frac{\mu s}{n}\right)$$

$$= (-1)^{m+1}\left[\frac{\pi^2}{4n^2}\,\frac{q+q_0}{q-q_0} + \pi i\left(\frac{n^2-1}{2n}\mu - 4\sum_1^{(n-1)/2}E\left(\frac{\mu s}{n}+\frac{1}{2}\right)\right)\right].$$

Finally,

$$\log\frac{2K}{\pi}=4\sum\frac{q^t}{t(1+q^t)}=\log\left(\frac{q_0+q}{q_0-q}\right)+4\sum\frac{1}{t}\left(\frac{q^t}{1+q^t}-\frac{1}{2}\,\frac{q^t}{q_0^t}\right)$$
$$=\log\left(\frac{q_0+q}{q_0-q}\right)+2i\sum\frac{1}{t}\tan\frac{tx}{2}.$$

1) $x=\frac{2m}{n}\pi$, n odd. Let $\alpha=e^{2\pi i/2n}$. Then using

$$\frac{1}{1+r\alpha^{2sm}}=\sum_{\sigma=0}^{n-1}\frac{(-1)^\sigma r^\sigma\alpha^{2s\sigma m}}{1+r^n},$$

we have

$$\log\frac{2K}{\pi}=\log\left(\frac{q_0+q}{q_0-q}\right)+2\sum\sum_{s=1}^{2n-1}\frac{1}{2nt+s}\,\frac{\alpha^{2ms}-1}{\alpha^{2ms}+1}$$
$$=\log\left(\frac{q_0+q}{q_0-q}\right)+2\int_0^1\sum_{t=1}^{2n}\frac{\alpha^t dx}{1-\alpha^t x}\cdot-\frac{1}{2n}\sum_s\alpha^{-ts}\sum_{\sigma=1}^{n-1}(-1)^\sigma\alpha^{2s\sigma m}$$
$$=\log\left(\frac{q_0+q}{q_0-q}\right)+2\sum_1^{n-1}\log(1-\alpha^{2rm})(-1)^r\,\frac{1}{2n}\,n$$
$$-2\sum_1^{n-1}\log(1-\alpha^{2rm+n})(-1)^r\,\frac{1}{2n}\,n$$

$$=A+\frac{1}{2}\sum\left(\frac{rm}{n}+\frac{1}{2}\right)(-1)^r2\pi i-\frac{1}{2}\sum\left(\frac{rm}{n}\right)(-1)^r2\pi i$$

$$=\log\left(\frac{q_0+q}{q_0-q}\right)+2\pi i\sum_{s=1}^{(n-1)/2}\left(\left(s\,\frac{2m}{n}+\frac{1}{2}\right)-\left(s\,\frac{2m}{n}\right)\right).$$

2) $x = \frac{m}{n}\pi$, n odd, m odd. Let $\alpha = e^{2\pi i/4n}$. Then

$$\log \frac{2K}{\pi} = \frac{q+q_0}{q-q_0}\,\frac{\pi^2}{4n^2} + \log\left(\frac{q_0+q}{q_0-q}\right) + 2\sum\sum_{s=1}^{4n-1} \frac{1}{4nt+s}\,\frac{\alpha^{2ms}-1}{\alpha^{2ms}+1}$$

$$= A + 2\int_0^1 \sum_{t=1}^{4n} \frac{\alpha^t dx}{1-\alpha^t x}\,\frac{1}{4n} \cdot -\frac{1}{2n}\sum_{s=1}^{4n-1}\sum_{\sigma=0}^{2n-1}(-1)^\sigma \sigma \alpha^{2s\sigma m}(\alpha^{2ms}-1)\alpha^{-ts}$$

$$= A + 2\int_0^1 \sum_{t=1}^{4n} \frac{\alpha^t dx}{1-\alpha^t x}\,\frac{1}{4n}\, 2\sum_{s=1}^{4n-1}\sum_{\sigma=1}^{2n-1}(-1)^\sigma \left(\frac{\sigma-n}{2n}\right)\alpha^{2ms\sigma}\alpha^{-ts}$$

$$= A - 2\sum_1^{2n-1} \log(1-\alpha^{2mr})\,\frac{1}{4n}\,(-1)^r\left(\frac{r-n}{sn}\right)4n$$

$$+ 2\sum \log(1-\alpha^{2mr+2n})(-1)^r\left(\frac{r-n}{2n}\right)$$

(using 1) $t \equiv 2mr \pmod{4n}$, 2) $t \equiv 2mr + n \pmod{4n}$)

$$= A - 2\pi i \sum_1^{2n-1}(-1)^r\left(\left(\frac{mr+n}{2n}\right) - \left(\frac{mr}{2n}\right)\right)\left(\frac{r-n}{2n}\right)$$

$$= A - 2\pi i \sum_1^{2n-1}(-1)^r\left(\left(\frac{r+n}{2n}\right) - \left(\frac{r}{2n}\right)\right)\left(\frac{\mu r-n}{2n}\right)$$

(where $m\mu \equiv 1 \pmod{2n}$)

$$= A + 2\pi i \sum_1^{n-1}(-1)^r\left(\frac{\mu r - n}{2n}\right).$$

3) $x = \frac{m}{2n}\pi$, m odd. Let $\alpha = e^{2\pi i/4n}$. Then

$$\log \frac{2K}{\pi} = \log \frac{q_0 + q}{q_0 - q} + 2 \sum \sum_{s=1}^{4n-1} \frac{1}{4nt + s} \, \frac{\alpha^{ms} - 1}{\alpha^{ms} + 1}$$

$$= A + 2 \int_0^1 \sum_{t=1}^{2n} \frac{\alpha^t dx}{1 - \alpha^t x} \, \frac{1}{4n} \, 2 \sum_{s=1}^{4n-1} \sum_{\sigma=1}^{4n-1} (-1)^\sigma \left(\frac{\sigma - 2n}{4n} \right) \alpha^{ms\sigma} \alpha^{-ts}$$

$$= A + 2\pi i \sum_{1}^{2n-1} (-1)^r \left(\frac{\mu r - 2n}{4n} \right)$$

where $m\mu \equiv 1 \pmod{4n}$.

Commentary on XXVIII.

R. Dedekind

The time of writing of the first of the two fragments (September 1852) makes it likely that Riemann started from here in finding examples of functions with infinitely many discontinuities in each interval for his work on trigonometric series (**XII**). Perhaps the second investigation, which occurs on the barely legible sheet, has the same object. However, the method used by Riemann here to determine the behavior of the modular functions appearing in the theory of elliptic functions, in the case that the complex period ratio

$$\omega = \frac{K'i}{K} = \frac{\log q}{\pi i} \tag{1}$$

tends to a rational value, permits a very interesting application. The application, in the so-called theory of the infinitely many forms of the theta-functions, concerns the determination of the constants appearing via transformations of first degree, which as is known, were reduced by Jacobi and Hermite to Gauss sums, and thus to the theory of quadratic residues. The following commentary illustrates these relationships.

The central point, in some sense, of the theory of these modular functions can also be put forward almost independently of elliptic functions, and has been the object of numerous investigations since the first edition of Riemann's works appeared. This central point is the function

$$\eta(\omega) = 1^{\omega/24} \prod (1 - 1^{\omega\nu}) = q^{1/12} \prod (1 - q^{2\nu}) \tag{2}$$

where for brevity we set

$$e^{2\pi i z} = 1^z, \text{ thus } q = 1^{\omega/2}. \tag{3}$$

The product is taken over all natural numbers ν. Now this function of the complex variable $\omega = x + yi$, whose ordinate y is always positive, is finite and non-zero in the simply connected region bounded in this fashion. Hence all powers of $\eta(\omega)$ with arbitrary exponents, and likewise $\log \eta(\omega)$, are single valued functions of ω once their values have been fixed at a certain point. We specify $\log \eta(\omega)$ by the condition that when y tends to infinity, and thus q vanishes, then

$$\log \eta(\omega) - \frac{\omega \pi i}{12} = 0. \tag{4}$$

The conjugate of $\log\eta(\omega)$ is then $\log\eta(-\omega')$. Here and below, ω' denotes the conjugate of ω. We know (Fund. nova §36) that

$$\eta(2\omega)\eta\left(\frac{\omega}{2}\right)\eta\left(\frac{1+\omega}{2}\right) = 1^{1/48}\eta(\omega)^3,$$

$$\sqrt[4]{k} = 1^{1/48}\sqrt{2}\,\frac{\eta(2\omega)}{\eta\left(\frac{1+\omega}{2}\right)},$$

$$\sqrt[4]{k'} = 1^{1/48}\,\frac{\eta(\omega/2)}{\eta\left(\frac{1+\omega}{2}\right)},$$

$$\sqrt{\frac{2K}{\pi}} = 1^{-1/24}\,\frac{\eta\left(\frac{1+\omega}{2}\right)^2}{\eta(\omega)}.$$

Thus, by the rule given above,

$$(5)\qquad \begin{cases} \log\eta(2\omega) + \log\eta\left(\frac{\omega}{2}\right) + \log\eta\left(\frac{1+\omega}{2}\right) = \frac{\pi i}{24} + 3\log\eta(\omega), \\ \log k = \log 4 + \frac{\pi i}{6} + 4\log\eta(2\omega) - 4\log\eta\left(\frac{1+\omega}{2}\right), \\ \log k' = \frac{\pi i}{6} + 4\log\eta\left(\frac{\omega}{2}\right) - 4\log\eta\left(\frac{1+\omega}{2}\right), \\ \log\frac{2K}{\pi} = -\frac{\pi i}{6} + 4\log\eta\left(\frac{1+\omega}{2}\right) - 2\log\eta(\omega). \end{cases}$$

As in Fund. nova, §40, the logarithms on the left side are single valued functions of ω, defined so that the three quantities

$$\log k - \log 4 - \frac{\omega\pi i}{2} = \log k - \log 4\sqrt{q},\ \log k',\ \log\frac{2K}{\pi}$$

tend to 0 with q.

From this behavior of the functions, we now obtain, with the help of linear fractional transformations of the theta-functions, the behavior investigated by Riemann when ω becomes close to a rational value. Equivalently, q tends to a definite root of unity q_0. Let

$$\theta_1(z,\omega) = \sum 1^{\left(s+\frac{1}{2}\right)^2\frac{\omega}{2}+\left(s+\frac{1}{2}\right)\left(z-\frac{1}{2}\right)}$$
$$= 2\eta(\omega)1^{\omega/12}\sin z\pi\prod(1-1^{\omega\nu+z})(1-1^{\omega\nu-z})$$

where the summation extends over all integers s. Using an accent to denote differentiation with respect to z, we have

$$\theta_1'(0,\omega) = 2\pi\eta(\omega)^3.$$

Now let $\alpha, \beta, \gamma, \delta$ be four integers satisfying the condition

$$\alpha\delta - \beta\gamma = 1. \tag{6}$$

We know that

$$\theta_1\left(z, \frac{\gamma+\delta\omega}{\alpha+\beta\omega}\right) = c\sqrt{\alpha+\beta\omega}\,1^{\frac{1}{2}\beta(\alpha+\beta\omega)z^2}\theta_1((\alpha+\beta\omega)z, \omega).$$

Here c is an eighth root of unity, depending on $\alpha, \beta, \gamma, \delta$ and the choice of the square root, whose determination was reduced to Gauss sums by Hermite (Liouville's Journal, Ser. 2, vol 3, 1858). For $z = 0$, this yields

$$\theta_1'\left(0, \frac{\gamma+\delta\omega}{\alpha+\beta\omega}\right) = c(\alpha+\beta\omega)^{3/2}\theta_1'(0,\omega),$$

thus

$$\eta\left(\frac{\gamma+\delta\omega}{\alpha+\beta\omega}\right) = c^{1/3}(\alpha+\beta\omega)^{1/2}\eta(\omega). \tag{7}$$

From this transformation of $\eta(\omega)$, the transformation of $\log\,\eta(\omega)$ may be deduced.

The case $\beta = 0$ is dealt with directly via the definitions (2), (4) of $\eta(\omega)$, $\log\,\eta(\omega)$, and we have

$$\log\eta(1+\omega) = \log\eta(\omega) + \frac{\pi i}{12}, \tag{8}$$

or more generally, for any integer n,

$$\log\eta(n+\omega) = \log\eta(\omega) + \frac{n\pi i}{12}. \tag{9}$$

However, if $\beta \neq 0$, the quantity $\mu = -(\alpha+\beta\omega)^2$ cannot be negative. Consequently one can define $\log\,\mu$ uniquely in such a way that its imaginary part always remains between $-\pi i$ and πi. Consequently conjugate values of $\log\,\mu$ correspond to conjugate values of μ. Now, as a consequence of (7),

$$\log\eta\left(\frac{\gamma+\delta\omega}{\alpha+\beta\omega}\right) = \log\eta(\omega) + \frac{1}{4}\log\{-(\alpha+\beta\omega)^2\} + (\alpha,\beta,\gamma,\delta)\frac{\pi i}{12}, \tag{10}$$

where $(\alpha, \beta, \gamma, \delta)$, denotes an integer, completely determined by $\alpha, \beta, \gamma, \delta$, that remains unchanged when these four numbers are multiplied by -1.

Clearly it is even more desirable to determine $(\alpha, \beta, \gamma, \delta)$ than the above root of unity c, and this determination is the actual objective of the following investigation.

We begin by reducing $(\alpha, \beta, \gamma, \delta)$ to a number that depends only on α, β. Namely, let γ', δ', be numbers that likewise satisfy $\alpha\delta' - \beta\gamma' = 1$, then we know that $\gamma' = \gamma + n\alpha$, $\delta' = \delta + n\beta$, where n denotes a natural number. From (9), then,

$$\log \eta\left(\frac{\gamma' + \delta'\omega}{\alpha + \beta\omega}\right) = \log \eta\left(n + \frac{\gamma + \delta\omega}{\alpha + \beta\omega}\right)$$
$$= \log \eta\left(\frac{\gamma + \delta\omega}{\alpha + \beta\omega}\right) + \frac{n\pi i}{12}.$$

Now, according to (10), it follows that the quantity

$$(\alpha, \beta, \gamma', \delta') - \frac{\delta'}{\beta} = (\alpha, \beta, \gamma, \delta) - \frac{\delta}{\beta}$$

depends only on the pair of numbers α, β. Hence we may write

$$(11) \qquad \beta(\alpha, \beta, \gamma, \delta) = \alpha + \delta - 2(\alpha, \beta).$$

Thus

$$(12) \quad \log \eta\left(\frac{\gamma + \delta\omega}{\alpha + \beta\omega}\right) = \log \eta(\omega) + \frac{1}{4}\log\{-(\alpha + \beta\omega)^2\} + \frac{\alpha + \delta - 2(\alpha, \beta)}{12\beta}\pi i.$$

The quantity $2(\alpha, \beta)$ and (as we shall see later) (α, β) itself, is an integer depending only on the pair of relatively prime integers α, β: and we also have

$$(13) \qquad (-\alpha, -\beta) = -(\alpha, \beta).$$

If we replace each term in equation (12) by the corresponding conjugate quantity, then according to the above remarks we obtain

$$\log \eta\left(-\frac{\gamma + \delta\omega'}{\alpha + \beta\omega'}\right) = \log \eta(-\omega') + \frac{1}{4}\log\{-(\alpha + \beta\omega')^2\} - \frac{\alpha + \delta - 2(\alpha, \beta)}{12\beta}\pi i.$$

Since the left side can, by (12), also be written in the form

$$\log \eta\left(\frac{-\gamma + \delta(-\omega')}{\alpha - \beta(-\omega')}\right) = \log \eta(-\omega') + \frac{1}{4}\log\{-(\alpha + \beta\omega')^2\} + \frac{\alpha + \delta - 2(\alpha, -\beta)}{12(-\beta)}\pi i$$

we obtain

$$(-\alpha,-\beta)=(\alpha,\beta) \tag{14}$$

and, in consequence of (13), also

$$(-\alpha,\beta)=-(\alpha,\beta). \tag{15}$$

If the result (12) is also to hold for the case $\beta=0$, $\alpha=\delta=\pm1$, then we should complete the definition of the symbol (α,β) via the rule

$$(\pm1,0)=\pm1. \tag{16}$$

This is consistent with (13), (14), (15).

From (15), we have $(0,\pm1)=0$. Thus if we let $\alpha=0$, $\beta=1$, $\gamma=-1$, $\delta=0$, the result (12) becomes the special case of the complimentary transformation

$$\log\eta\left(\frac{-1}{\omega}\right)=\log\eta(\omega)+\frac{1}{4}\log(-\omega^2). \tag{17}$$

In (12), we now replace ω by $1+\omega$ and by $-\frac{1}{\omega}$, and we express the quantities

$$\log\eta\left(\frac{\gamma+\delta+\delta\omega}{\alpha+\beta+\beta\omega}\right),\quad \log\eta\left(\frac{\delta-\gamma\omega}{\beta-\alpha\omega}\right)$$

in terms of log $\eta(\omega)$ via (12). Recalling (8), (17), we readily obtain the following two results valid for every relatively prime pair α,β:

$$(\alpha+\beta,\beta)=(\alpha,\beta), \tag{18}$$

$$2\alpha(\alpha,\beta)+2\beta(\beta,\alpha)=1+\alpha^2+\beta^2-3|\alpha\beta|. \tag{19}$$

With the aid of the last result, which is closely related to the reciprocity theorem in the theory of quadratic residues, we can also give equation (11) the form

$$(\alpha,\beta,\gamma,\delta)=2\gamma(\alpha,\beta)+2\delta(\beta,\alpha)-(\alpha\gamma+\beta\delta)\pm3\alpha\delta. \tag{20}$$

The sign $\pm$ is to be chosen so that $\pm\alpha\beta=|\alpha\beta|$. In this way the number $(\alpha,\beta,\gamma,\delta)$, which first occurs in (10), again appears in the form of an integer.

It is now plain that the two results (18), (19) not only include the earlier properties (13)–(16), but also suffice to determine the value of (α,β) in every case via a continued fraction expansion, in fact giving an integer value for (α, β). This already follows from the rule

$$(21) \qquad (\alpha, \alpha+\beta) = (\alpha, \beta) - (\beta, \alpha) - (\beta, \alpha) + \beta - \alpha, \text{ if } \alpha\beta \geq 0,$$

which is an easy consequence of (18) and (19). Conversely it is clear that (21) in conjunction with (18), that is, with the rule

$$(22) \qquad (\alpha', \beta) = (\alpha, \beta) \quad \text{if } \alpha' \equiv \alpha \pmod{\beta}$$

likewise provides for the complete determination of the symbol (α, β) and furnishes a very convenient calculation of a table. Finally, it is very convenient to give the symbol (α, β) a definite meaning when the pair of integers α, β are not relatively prime but have an arbitrary (positive) greatest common divisor p. In this case, let

$$(23) \qquad (\alpha, \beta) = p\left(\frac{\alpha}{p}, \frac{\beta}{p}\right).$$

Obviously the two rules (21), (2.) remain valid without change. Admittedly, the first term 1 on the right side in the rule (19) must be replaced by p^2. However, the two rules (21), (2.) now yield the complete determination of (α, β) without the aid of (23). Moreover, these rules are also valid for the case $\alpha = \beta = 0$ if we let

$$(24) \qquad (0,0) = 0.$$

This extension of the symbol (α, β) often permits results to be given a unified statement, which would otherwise have to be decomposed into various cases. (Compare the results contained in (28), (34).)

Although the symbol (α, β) is now completely defined for every pair of rational integers α, β by the properties (21), (2.), it would be difficult to derive a general expression for (α, β) from these properties. With the help of the method applied by Riemann in the two fragments, we can set down such an expression in the form of a finite sum. This method consists of the investigation of the behavior of the modular functions when $\omega = x + yi$ approximates a rational number, $-\alpha/\beta$ in lowest terms. If this approximation takes place in such a way that $\alpha + \beta x$ is infinitely small of higher order than $\sqrt{y}$, then the ordinate in the quantity

$$\omega_1 = \frac{\gamma + \delta\omega}{\alpha + \beta\omega} = \frac{\delta}{\beta} - \frac{1}{\beta(\alpha + \beta\omega)},$$

appearing in (12), becomes infinitely large and positive. Now according to (4),

$$\log \eta(\omega_1) - \frac{\omega_1 \pi i}{12} = 0.$$

Thus

$$\log \eta(\omega) + \frac{\pi i}{12\beta(\alpha + \beta\omega)} + \frac{1}{4}\log\{-(\alpha + \beta\omega)^2\} = \frac{2(\alpha, \beta) - \alpha}{12\beta}\,\pi i$$

In order to get close to Riemann's notation, we replace α, β by $-m, n$. Now we may state this result as follows. If the variable $\omega = x + yi$ approaches the irreducible fraction m/n in such a way that $nx - m$ is infinitely small of higher order than $\sqrt{y}$, then ultimately

$$\text{(25)} \qquad \log \eta(\omega) + \frac{\pi i}{12n(n\omega - m)} + \frac{1}{4}\log\{-(n\omega - m)^2\} = \frac{m - 2(m, n)}{12n}\,\pi i.$$

If we subject the approximation to the sharper condition that $nx - m$ is infinitely small of higher order than y^2, then the imaginary parts of the second and third terms on the left vanish. On subtracting the conjugate quantity we get the limiting result

$$\text{(26)} \qquad \log \eta(\omega) - \log \eta(-\omega') = \frac{m - 2(m, n)}{6n}\,\pi i.$$

As a consequence of the above extension of the symbol (m, n), (26) also holds if the pair of integers m, n has arbitrary greatest common divisor.

Before we make use of this relation to solve our problem, we note some consequences. Let a, ∂ be positive integers and c an arbitrary integer. Suppose that the approach of ω to its limiting value obeys the latter, sharper condition. Clearly the same is true for the approach of

$$\frac{c + \partial\omega}{a} \text{ to the value } \frac{cn + \partial m}{an}.$$

Consequently, at the same time as (26), we have the limiting result

$$\log \eta\left(\frac{c + \partial\omega}{a}\right) - \log \eta\left(-\left(\frac{c + \partial\omega'}{a}\right)\right) = \frac{cn + \partial m - 2(cn + \partial m, an)}{6an}\,\pi i.$$

We have the following easily deduced rule for any prime number p, which derives from the transformation of order p or from (2):

$$\log \eta(p\omega) + \sum_{s=0}^{p-1} \log \eta\left(\frac{s+\omega}{p}\right) = \frac{(p-1)\pi i}{24} + (p+1)\log \eta(\omega). \tag{27}$$

We subtract from (27) the equation obtained from (27) by passing to conjugate quantities. On taking the limiting value we obtain the rule

$$p(pm, n) + \sum_{s}(m + ns, np) = p(p+1)(m, n). \tag{28}$$

Here s runs over an arbitrary complete residue system (mod p). From (27) we can derive, in various ways, more general results valid for arbitrary composite integers p. Arising from these results there is an analogous result for the symbol (m, n). However, we shall not enter here into these properties of the function log $\eta(\omega)$ and the symbol (m, n), which are of great independent interest.

Turning now to our problem, we employ the representation

$$\log \eta(\omega) = \frac{\omega\pi i}{12} + \sum_{\nu} \log(1 - 1^{\omega\nu}) \tag{29}$$

that follows from (2) and (4). Here ν runs over the natural numbers, and the logarithm on the right side vanishes together with 1^{ω}. Hence

$$\log(1 - 1^{\omega\nu}) = -\sum_{\mu=1}^{\infty} \frac{1^{\omega\nu\mu}}{\mu}.$$

Carrying out the summation with respect to ν, we obtain the transformation

$$\log \eta(\omega) = \frac{\omega\pi i}{12} - \sum \frac{1}{\mu}\frac{1^{\omega\mu}}{1 - 1^{\omega\mu}}, \tag{30}$$

given by Jacobi (Fund. nova, §39). Hence

$$\log \eta(\omega) - \log \eta(-\omega') = \frac{(\omega + \omega')\pi i}{12} - \sum_{\mu=1}^{\infty} \frac{a_\mu}{\mu}.$$

Here, for brevity, we let

$$a_\mu = \frac{1}{1 - 1^{\omega\mu}} - \frac{1}{1 - 1^{-\omega'\mu}}.$$

Now we let the positive ordinate y of $\omega = x + yi$ tend to zero, and the abscissa x, as above, tend to the rational constant m/n, subject to the above sharper condition on the approximation. The integers m, n may have an arbitrary common divisor below: however, we take the denominator n to be positive. For brevity, let

$$1^x = 1^{m/n} = e^{2\pi i m/n} = \theta; \; 1^{yi} = e^{-2\pi y} = r.$$

The constant θ satisfies the condition $\theta^n = 1$, and r denotes a positive variable less than 1, increasing to the limit 1. Now we have

$$a_\mu = \frac{1}{1 - \theta^\mu r^\mu} - \frac{1}{1 - \theta^{-\mu} r^\mu},$$

and we are concerned with the determination of the limiting value of

$$\log \eta(\omega) - \log \eta(-\omega') = \frac{m\pi i}{6n} - \sum \frac{a_\mu}{\mu}.$$

By combining each pair of numerators a_μ corresponding to the numbers $\mu = sn + \nu$, $\mu = (s+1)n - \nu$ $(0 < \nu < n/2)$, we readily see that the absolute value of the sum

$$A_\mu = a_1 + a_2 + \cdots + a_\mu$$

remains below a finite constant independent of ν and μ, for all r, including $r = 1$. By a general theorem[1], it follows that the series

$$\sum \frac{a_\mu}{\mu} = \sum A_\mu \left(\frac{1}{\mu} - \frac{1}{\mu + 1} \right),$$

with the terms in order of increasing μ, is also convergent for $r = 1$ and is continuous at this point. Recalling (26), it follows that

$$\frac{(m,n)\pi i}{3n} = \sum \frac{b_\mu}{\mu},$$

[1]Dirichlet, Vorlesungen über Zahlentheorie, vol. 2, §143.

where $b_\mu = \lim a_\mu$ is 0 or

$$\frac{1}{1-\theta^\mu} - \frac{1}{1-\theta^{-\mu}}$$

according as $\theta^\mu = 1$ or not. Applying the transformation

$$\frac{1}{1-\theta^\mu} = -\frac{1}{n}\sum_{\sigma=1}^{n-1} \sigma\theta^{\mu\sigma},$$

we obtain a representation

$$b_\mu = \frac{1}{n}\sum_{\sigma=1}^{n-1} \sigma(\theta^{-\mu\sigma} - \theta^{\mu\sigma})$$

that is valid for all μ. This representation yields the value of our infinite series without the use of definite integrals.

For a real z, we use the notation $((z))$, rather than (z), for the difference of z from the nearest integer, a value between $-1/2$ and $1/2$, for the sake of clarity. Following Riemann (pp, and) if z is halfway between two integers, the value of the periodic function $((z))$ at this discontinuity is taken to be 0. The value 0 at such a z is the arithmetic mean of the two limiting values $((z+0)) = -1/2, ((z-0)) = 1/2$. By a well-known result from the theory of trigonometric series, which also follows directly from the logarithmic series, the representation

$$2\pi i((z)) = \sum_{\mu=1}^{\infty} \frac{(-1)^\mu(1^{-z\mu} - 1^{z\mu})}{\mu}$$

is always valid. Thus

$$(31) \qquad 2\pi i((z-1/2)) = \sum \frac{1^{-z\mu} - 1^{z\mu}}{\mu}.$$

It follows that

$$\sum \frac{\theta^{-\mu\sigma} - \theta^{\mu\sigma}}{\mu} = 2\pi i\left(\left(\frac{\sigma m}{n} - \frac{1}{2}\right)\right),$$

hence

$$\frac{(m,n)}{6n} = \sum \frac{\sigma}{n}\left(\left(\frac{\sigma m}{n} - \frac{1}{2}\right)\right).$$

However, replacing σ by $n - \sigma$ yields

$$\frac{1}{2}\sum\left(\left(\frac{\sigma m}{n} - \frac{1}{2}\right)\right) = 0.$$

Subtracting this readily gives the expression

$$(m,n) = 6n\sum\left(\left(\frac{s}{n} - \frac{1}{2}\right)\right)\left(\left(\frac{ms}{n} - \frac{1}{2}\right)\right), \tag{32}$$

where n is assumed to be positive and s runs over a complete residue system (mod n). This expression for the symbol (m,n) in the form of a finite sum permits various transformations and simplifications, which we examine more closely below. The result (32) also holds when m, n have an arbitrary (positive) common divisor p. We obtain this on recalling (23), supplemented by the formula

$$\sum\left(\left(\frac{x+p'}{p} - \frac{1}{2}\right)\right) = \left(\left(x - \frac{1}{2}\right)\right) \tag{33}$$

(x an arbitrary real number; p' runs over a complete residue system (mod p)), which is important in itself.

Now let us assume that m, n are relatively prime. For brevity, let

$$B = \frac{\pi i}{24n(n\omega - m)}, \qquad C = \frac{1}{4}\log\{-(n\omega - m)^2\},$$

$$\mu = \frac{1-(-1)^m}{2}, \qquad \nu = \frac{1-(-1)^n}{2}.$$

Then $(1-\mu)(1-\nu) = 0$, $m \equiv \mu, n \equiv \nu \pmod 2$. From the limiting result (25),

$$\log \eta(\omega) = \frac{m - 2(m,n)}{12n}\pi i - 2B - C.$$

At the same time this yields

$$\log\eta(2\omega) = \frac{m-(2m,n)}{6n}\pi i - (4-3\nu)B - C + \frac{\nu}{2}\log 2,$$

$$\log\eta\left(\frac{\omega}{2}\right) = \frac{m-2(m,2n)}{24n}\pi i - (4-3\mu)B - C + \frac{1-\mu}{2}\log 2,$$

$$\log\eta\left(\frac{1+\omega}{2}\right) = \frac{m+n-2(m+n,2n)}{24n}\pi i + (2-3\mu-3\nu)B - C$$
$$+\frac{\mu+\nu-1}{2}\log 2.$$

The symbols that appear here are, in consequence of (28), always connected by the relation

(34) $$2(2m,n)+(m,2n)+(m+n,2n) = 6(m,n).$$

At the same time we obtain, in consequence of (5), the limiting results

(35) $$\begin{cases} \log k = \frac{3m+2(m+n,2n)-4(2m,n)}{6n}\pi i + (\mu+2\nu-2)(12B-2\log 2), \\ \log k' = \frac{(m+n,2n)-(m,2n)}{3n}\pi i + (2\mu+\nu-2)(12B-2\log 2), \\ \log\frac{2K}{\pi} = \frac{(m,n)-(m+n,2n)}{3n}\pi i + (1-\mu-\nu)(12B-2\log 2) - 2C. \end{cases}$$

Comparison of these results with the eight formulae in the second fragment reveals that Riemann attached less value to the determination of the infinitely large real parts contained in the terms with B, C; they are partly not precisely given, and partly dropped. Moreover, in the imaginary parts (in the third, fourth and fifth formulae) there were some small slips. These slips were already freely corrected in the first edition; the real part is reproduced without change. It cannot always be recognized at a first glance that Riemann's formulae coincide in their imaginary parts with the above results (35), and it would take us too far afield to give complete proof of this here. However, since this is of some importance, we append the following remarks for clarification.

By the *denominator* of a rational number x, we understand the smallest positive integer n for which nx is an integer m; we call m the *numerator* of x. Now there are always infinitely many numbers x' having the same denominator n, whose numerators m' satisfy the congruence $mm' \equiv 1$ (mod

n). Each such number x' will be called an associate (socius) of x (compare Section 77 of Disqu. Arithm.). The real numbers x, y are said to be *simply congruent* if their difference is an integer, and we denote this by $x \equiv y$. To each class of congruent numbers x corresponds exactly one class of numbers x', and if p is an integer relatively prime to n, then $p(px')' \equiv x$. Now, for brevity, let us write

$$D(x) = \frac{(m,n)}{n} = 6\sum\left(\left(\frac{s}{n}-\frac{1}{2}\right)\right)\left(\left(\frac{ms}{n}-\frac{1}{2}\right)\right). \tag{36}$$

Then this function, as we easily see from the above expression or alternatively from (18), (15), (12), (34), has the properties

$$\begin{cases} D(x) = D(x+1) = -D(-x) = D(x'), \\ D(2x) + D\left(\frac{x}{2}\right) + D\left(\frac{x+1}{2}\right) = 3D(x). \end{cases} \tag{37}$$

For the function $E(x)$, sometimes used in Riemann's formulae, namely the greatest integer $\leq x$, we substitute the expression

$$E(x) = x - \frac{1}{2} - \left(\left(x - \frac{1}{2}\right)\right). \tag{38}$$

Here, for integer x we are to replace $E(x)$ by the arithmetic mean $x - \frac{1}{2}$ of $E(x+0)$ and $E(x-0)$. Then in most of these formulae, only functions of the form

$$R(x) = \sum((\nu x)), \quad S(x) = \sum\left(\left(\nu x - \frac{1}{2}\right)\right) \tag{39}$$

will appear. Here the summations are taken over non-negative integers ν less than half the denominator of x. These functions have the properties

$$\begin{cases} R(x) = R(x+1) = -R(-x), \\ S(x) = S(x+1) = -S(-x), \\ R(x) - S(x) = R(x') - S(x') = \frac{1}{2}h. \end{cases} \tag{40}$$

Here h denotes the excess of the number of positive terms $((\nu x))$ over the number of negative terms. We have the following relations of these functions to the function $D(x)$. In general, according to (36),

$$6S(x') = D(2x) - 2D(x). \tag{41}$$

If x has even denominator n, then

$$\begin{cases} R(x) = -S(x) = \frac{1}{4}\,h = \frac{1}{3}\,D(x) - \frac{1}{6}\,D(2x), \\ R\left(\frac{x}{2}\right) + R\left(\frac{x+1}{2}\right) = 2R(x). \end{cases} \tag{42}$$

However, if x has odd denominator n, then the numbers y that satisfy the condition $2y \equiv x$, that is, $y \equiv \frac{x}{2}$ and $y \equiv \frac{1}{2}\,(x+1)$, fall into two classes. We denote those having the same denominator n by x_1, and the other numbers y by x_2. The latter numbers have the denominator $2n$. Now

$$R(x_2) = R(x) - S(x) = 2R(x) - S(2x) \tag{43}$$

and

$$\begin{cases} D(x) &= 6R(x_2) - 4R(x) - 4R(x'), \\ D(2x) &= 6R(x_2) - 8R(x) - 2R(x'), \\ D(x_1) &= 6R(x_2) - 2R(x) - 8R(x'), \\ D(x_2) &= 6R(x_2) - 2R(x) - 2R(x'). \end{cases} \tag{44}$$

From these relations, we recover the above condition

$$D(2x) + D(x_1) + D(x_2) = 3D(x). \tag{45}$$

The validity of the first three representations in (44) follows from the earlier properties of $R(x)$ on taking into account the relations

$$x_1 \equiv x_2 + \frac{1}{2} \equiv (2x')'; \ \left(x + \frac{1}{2}\right)' \equiv (4x)' + \frac{1}{2}; \ (x_2)' \equiv (x')_2.$$

We have the inverse relations

$$\begin{cases} 6R(x) &= 3D(x) - 2D(2x) - D(x_1) = D(x_2) - D(2x), \\ 6R(x') &= 3D(x) - D(2x) - 2D(x_1) = D(x_2) - D(x_1), \\ 6R(x_2) &= 5D(x) - 2D(2x) - 2D(x_1) = 2D(x_2) - D(x). \end{cases} \tag{46}$$

We must postpone the derivation of these and numerous other relations, all of which are closely related to the theory of quadratic residues, to another occasion.

XXIX.

Fragment on Analysis Situs.

Two 1-solids will be placed in the same or different classes depending on whether or not it is possible to continuously change one into the other.

Every two 1-solids which are bounded by the same pair of points, together form a connected 1-solid without boundary. This can form the whole boundary of a 2-solid, or not, depending on whether they belong to the same or different classes.

An inner, connected, unbounded 1-solid, if taken once, can either be the whole boundary of an inner 2-solid, or not.

Let $a_1, a_2, \ldots, a_m$ be m inner connected n-solids without boundary which taken once, neither individually nor in combination can completely bound an inner $(n+1)$-solid. Let $b_1, b_2, \ldots, b_m$ be n-solids with the same properties, where each b taken together with one or more of the a's form the complete boundary of an inner $(n+1)$-solid. Thus each inner connected n-solid which with the a's forms the complete boundary of an inner $(n+1)$-solid, likewise does so with the b's and conversely.

If any unbounded inner n-solid when taken together with the a's forms the entire boundary of an inner $(n+1)$-solid, then as a consequence of the hypotheses the a's can successively be eliminated and replaced by the b's.

An n-solid A is said to be changeable into another n-solid B if an inner $(n+1)$-solid can be completely bounded by A and by pieces of B.

If in the interior of a continuous manifold, each unbounded n-solid is bounded with the help of m fixed parts of n-solids which do not bound it by themselves, then the manifold is $m+1$ times connected in dimension n.

A continuous connected manifold is said to be simply connected, if it is simply connected in each dimension.

Let A be a bounded continuous manifold. A connected manifold B of lower dimension lying inside A whose entire boundary falls in the boundary of A is called a transverse cut of A.

The connectedness of an n-solid will be changed by a transverse cut with a simply connected $(n-m)$-solid by a decrease of 1 in the m-th dimension, or by an increase of 1 in dimension $m-1$.

The connectivity of the μ-th dimension can only be changed either by changing unbounded nonbounding μ-solids into bounded, or bounding into non-bounding. The first case changes to the extent of bounding a μ-solid, the last to the extent that new parts are added to the boundary of a $(\mu+1)$-solid.

The dependence of the connectivity of the boundary *B* of a continuous manifold *A* on the connectivity *A*.

The solids without boundary which are non-bounding within B fall into those which are not bounding within A and those which are bounding within A. We examine first, how the connectivity of B can be changed by a simply connected transverse cut of A.

Let A be n-dimensional, q an m-dimensional transverse cut, a the $(n-1-m)$-dimensional envelope of a point of q which q does not disconnect, and p the boundary of q.

The connectivity of A will be increased by 1 in the $(n-1-m)$-th dimension, if a is not bounding within A', and decreased by 1 in the $(n-m)$-th dimension if a is bounding within A',

$$A' - A = \begin{pmatrix} m+1 \\ +1 \end{pmatrix} \text{ if } a \text{ is not bounding within } A' \quad (\alpha)$$

$$= \begin{pmatrix} m \\ -1 \end{pmatrix} \text{ if } a \text{ is bounding within } A' \quad (\beta)$$

......[1]

	Change of A	Change of B
I. a not bounding within A', a not bounding within B', consequently p bounding within B,	$\begin{pmatrix} m+1 \\ +1 \end{pmatrix}$	$\begin{pmatrix} n-m-1 & m \\ +1 & +1 \end{pmatrix}$
II. a bounding within A', a not bounding within B', consequently p bounding within B,	$\begin{pmatrix} m \\ -1 \end{pmatrix}$	$\begin{pmatrix} n-m-1 & m \\ +1 & +1 \end{pmatrix}$
III. a bounding within A', a bounding within B', consequently p not bounding within B.	$\begin{pmatrix} m \\ -1 \end{pmatrix}$	$\begin{pmatrix} n-m-1 & m-1 \\ -1 & -1 \end{pmatrix}$

[1] There are a few symbols here in the manuscript whose significance and relevance I could not decipher. W.

Two parts of solids (space parts) are said to be connecting or to belong to one piece, if a line can be drawn from an inner point of one, through the interior of the solids (space) to an interior point of the other.

Theorems from the Theoria Situs.

(1) A solid of dimension less than $n-1$ cannot separate one part of an n-solid from another. A connected n-solid may, or may not, have the property of being decomposed into pieces by each transverse cut by an $(n-1)$-solid. We designate the first kind by a.

If an n-solid belonging to a is changed into another by a transverse cut by an $(n-2)$-solid, then this is connected and either belongs to a or not.

Those n-solids of a, which are changed into non-a by every transverse cut by an $(n-2)$-solid, we designate by a_1.

(2) If a solid A is changed into another, A', by transverse cut by a μ-solid, then each transverse cut of A of dimension more than $\mu+1$ forms a transverse cut of A' and conversely.

If one of the a_1 n-solids is changed by a transverse cut by an $(n-3)$-solid into another a_1, then this belongs to the $a(2)$. It is possible for it to belong, or not belong, to a_1.

The a_1's, which are changed by every transverse cut by an $(n-3)$-solid into a non-a_1, we denote by a_2.

If one continues in this way, then at last one obtains a category a_{n-2} of n-solids, which consists of those a_{n-3}'s which, with each transverse cut by a 1-solid (linear transverse cut), are changed to non-a_{n-3}'s. We call these a_{n-2} n-solids simply connected. The a_μ n-solids are hence simply connected apart from transverse cuts of dimension $n-\mu-2$ or less, and will be called simply connected up to dimension $n-\mu-2$.

An n-solid that is not simply connected up to dimension $n-1$ can be decomposed by a transverse cut by an $(n-1)$-solid, without becoming disconnected. The n-solid arising, if it is not simply connected up to dimension $n-1$, can be decomposed by a similar transverse cut. Clearly the procedure can be continued so long as one does not arrive at a solid connected up to dimension $n-1$. The number of transverse cuts which brings about such a decomposition of the n-solid into one that is simply connected up to the first dimension can differ according to the choices of the transverse cuts. Clearly there must be a type of decomposition for which the number of cuts is smallest.

XXX.

Convergence of p-fold infinite theta-series.

The investigation of the convergence of an infinite series with positive terms can always be reduced to the investigation of a definite integral using the following theorem:

Let

$$a_1 + a_2 + a_3 + \cdots$$

be a series with positive decreasing terms. Further, let $f(x)$ be a function decreasing with increasing x; then

$$f(\alpha) > \int_\alpha^{\alpha+1} f(x)dx > f(\alpha+1).$$

Hence

$$f(0) + f(1) + \cdots + f(n) > \int_0^{n+1} f(x)\,dx > f(1) + f(2) + \cdots + f(n+1).$$

Accordingly, the series

$$f(0) + f(1) + f(2) + \cdots$$

converges and diverges together with the integral

$$\int_0^\infty f(x)dx.$$

If now $f(n)$ is positive and $a_n < f(n)$, then the series

$$a_1 + a_2 + a_3 + \cdots$$

will likewise converge when this integral converges. This gives the theorem:

If $a_n < f(x)$ for $n \geq x$, then the series $\sum_{n=1}^{\infty} a_n$ converges provided that the integral $\int_0^\infty f(x)dx$ converges.

Now let $x = \phi(y), f(x) = f(\phi(y)) = F(y)$. We have

$$\int_0^\infty f(x)dx = \int F(y)\phi'(y)dy.$$

If the two variables x, y increase and decrease together (and tend to infinity together) then, according to the above hypotheses, $F(y)$ decreases and $\phi(y)$ increases with increasing y. Thus the above condition for convergence becomes the following:

The series $\sum a_n$ converges if for $n \geq \phi(y)$ we have $a_n < F(y)$, or equivalently, if for $a_n \geq F(y)$ we have $n < \phi(y)$, and the integral

$$\int_b^\infty F(y)\phi'(y)dy$$

converges.

If now $a_n > F(y)$, then the same holds for $a_1, a_2, \ldots, a_{n-1}$. Thus if $a_{n+1} < F(y)$, we find that n is the number of terms of the series that exceed $F(y)$. Accordingly, we can also express the result as follows:

Let $F(y), \phi(y)$ be two functions, the first decreasing with positive y, the second increasing to infinity with y. Suppose that the number of terms of a series with positive terms that are $\geq F(y)$ is $< \phi(y)$. Then the series converges if the integral $\int_b^\infty F(y)\phi'(y)dy$ converges.

We now look for such functions for the p-fold infinite series

$$\sum_{m_1,\ldots,m_p=-\infty}^{\infty} \exp\left(\sum_{i=1}^{p}\sum_{i'=1}^{p} a_{i,i'}m_i m_{i'} + 2\sum_{i=1}^{p} m_i v_i\right).$$

We may assume without loss of generality that the quantities $a_{i,i'}$ and v_i are real.

The general term of the series

$$\exp\left(\sum_{i=1}^{p}\sum_{i'=1}^{p} a_{i,i'}m_i m_{i'} + 2\sum_{i=1}^{p} m_i v_i\right)$$

is greater than e^{-h^2} if

$$-\sum_{i=1}^{p}\sum_{i'=1}^{p} a_{i,i'}m_i m_{i'} - 2\sum_{i=1}^{p} m_i v_i < h^2.$$

For our purpose, then, we wish to determine how many combinations of the natural numbers $m_1, m_2, \ldots m_p$ satisfy this inequality.

To this end we first treat the multiple definite integral

$$A = \iint \cdots \int dx_1 dx_2 \ldots dx_p$$

taken over the region defined by the inequaltiy

$$-\sum_{i=1}^{p}\sum_{i'=1}^{p} a_{i,i'}x_i x_{i'} < 1.$$

The integral will have a finite value if and only if the quadratic form

$$-\sum_{i=1}^{p}\sum_{i'=1}^{p} a_{i,i'}x_i x_{i'}$$

can be written as a sum of p positive squares. For if

$$-\sum\sum a_{i,i'}x_i x_{i'} = t_1^2 + t_2^2 + \cdots + t_p^2,$$

then the domain of integration is defined via the inequality

$$t_1^2 + t_2^2 + \cdots + t_p^2 < 1,$$

and the integral A becomes

$$A = \iint \cdots \int \left(\sum \pm \frac{\partial x_1}{\partial t_1}\,\frac{\partial x_2}{\partial t_2}\cdots\frac{\partial x_p}{\partial t_p}\right) dt_1 dt_2 \ldots dt_p.$$

The functional determinant is a finite constant and none of the variables t can have absolute value > 1.

On the other hand, if the t^2 were not all positive, or some of them were absent in the transformed quadratic form, then infinite values of t would occur in the integral A, and A itself would be infinite.

This result would not change in any way if, in place of the above domain of integration for A, we took the following:

$$-\sum_{i}\sum_{i'} a_{i,i'}x_i x_{i'} - 2\sum_{i} \alpha_i x_i < 1$$

where the α_i are arbitrary real numbers. Now let us consider the inequality

$$-\sum_{i}\sum_{i'} a_{i,i'}m_i m_{i'} - 2\sum_{i} v_i m_i < h^2$$

or, writing $\frac{m_i}{h} = x_i$,

$$-\sum_{i}\sum_{i'} a_{i,i'}x_i x_{i'} - 2\sum_{i} \frac{v_i}{h} x_i < 1.$$

It follows at once that, for every finite h, a finite number of combinations of the integers $m_1, m_2, \ldots, m_p$ satisfy this inequality, since the x_i must all remain between definite finite bounds, and within these bounds there are only finitely many rational numbers with a given denominator h.

Let Z_h be the number of possible combinations of the number m.

We now consider the sum

$$\sum_{m_1, m_2, \ldots, m_p} \int_{\frac{m_1}{h}}^{\frac{m_1+1}{h}} dx_1 \int_{\frac{m_2}{h}}^{\frac{m_2+1}{h}} dx_2 \ldots \int_{\frac{m_p}{h}}^{\frac{m_p+1}{h}} dx_p = \frac{Z_h}{h^p}.$$

This quantity is finite for finite h, and as h tends to infinity has the limit A. We have already established that A is finite if the function $-\sum_i \sum_{i'} a_{i,i'} x_i x_{i'}$ can be represented as the sum of p positive squares. If we write the above sum as $A + k$, then k is a finite quantity tending to 0 as h tends to infinity. Thus

$$Z_h = (A + k) h^p$$

is the number n of terms of the theta-series that are $> e^{-h^2}$. Accordingly

$$n < (A + K) h^p$$

where K is a constant, which can be taken as small as we wish if we consider only sufficiently large h. The functions $F(y), \phi(y)$ can now be taken as follows:

$$F(y) = e^{-y^2}, \quad \phi(y) = (A + K) y^p.$$

Since the integral

$$\int_b^\infty e^{-y^2} (A + K) p y^{p-1} dy$$

converges, the same holds for the θ-series under the given hypotheses. We conclude that *the p-fold theta-series converges for all values of the variables $v_1, v_2, \ldots, v_p$, provided that the real part of the quadratic form in the exponent is essentially negative.*

XXXI.

On the theory of Abelian functions.

Let $(e_1, e_2, \ldots, e_p)$ be a system of numbers with the property

$$\theta(e_1, e_2, \ldots, e_p) = 0.$$

By **VI**, §23 (p. 126), under these hypotheses we may satisfy the congruence

$$(e_1, e_2, \ldots, e_p) \equiv \left(\sum_1^{p-1} \alpha_1^{(\nu)}, \ldots, \sum_1^{p} \alpha_p^{(\nu)} \right) \equiv \left(-\sum_p^{2p-2} \alpha_p^{(\nu)}, \ldots, -\sum_p^{2p-2} \alpha_p^{(\nu)} \right)$$

with certain points $\eta_1, \eta_2, \ldots, \eta_{2p-2}$ that are linked via an equation $\phi = 0$. Accordingly, let u_μ and u'_μ be the values taken by the integral u_μ of first kind for two indefinite systems of values s, z and s_1, z_1. Then the function

$$\theta(u_1 - u'_1 - e_1, \ldots, u_p - u'_p - e_p)$$

vanishes as a function of s, z for $(s, z) = (s_1, z_1)$ and at the $p - 1$ points $\eta_1, \eta_2, \ldots, \eta_{p-1}$. Considered as a function of s_1, z_1, it vanishes for $(s_1, z_1) = (s, z)$, and at the points $\eta_p, \ldots, \eta_{2p-2}$. Thus if $(f_1, f_2, \ldots, f_p)$ is a system of numbers with the same property as $(e_1, e_2, \ldots, e_p)$, the function

$$(1) \qquad \frac{\theta(u_1 - u'_1 - e_1, \ldots)\theta(u_1 - u'_1 + e_1, \ldots)}{\theta(u_1 - u'_1 - f_1, \ldots)\theta(u_1 - u'_1 + f_1, \ldots)}$$

which is rational with respect both to (s, z) and (s, z_1), become infinitely large, and infinitely small, of first order at each of the systems of points linked by the equation $\phi = 0$. Accordingly, the functions (1) can be represented in the form

$$(2) \qquad \frac{\sum_1^p c_\nu \phi_\nu(s, z) \sum_1^p c_\nu \phi_\nu(s_1, z_1)}{\sum_1^p b_\nu \phi_\nu(s, z) \sum_1^p b_\nu \phi_\nu(s_1, z_1)},$$

with coefficients b, c independent of s, z and s_1, z_1.

Now if the system of numbers e, f have the property that

$$\begin{cases} (e_1, e_2, \ldots, e_p) \equiv (-e_1, -e_2, \ldots, -e_p), \\ (f_1, f_2, \ldots, f_p) \equiv (-f_1, -f_2, \ldots, -f_p), \end{cases} \tag{3}$$

then the points, where the function (1) or (2) respectively becomes zero or infinite, coincide in pairs and we obtain a function that is infinitely large, and infinitely small, of second order at only $p-1$ points. Hence the function

$$\sqrt{\frac{\sum_1^p c_\nu \phi_\nu(s,z) \sum_1^p c_\nu \phi_\nu(s_1,z_1)}{\sum_1^p b_\nu \phi_\nu(s,z) \sum_1^p b_\nu \phi_\nu(s_1,z_1)}}$$

branches like the surface T' and on crossing the transverse cuts acquires factors ± 1. The functions

$$\sqrt{\sum_1^p c_\nu \phi_\nu(s,z)}$$

determined in this way, which become infinitely small of first order at $p-1$ points, are called Abelian functions. They arise from the functions ϕ via pairwise coincidence of the zeros and extraction of roots. In general, the number of such functions is finite.

That is to say, the congruence (3) requires that the systems of numbers e, f take the form

$$\left(\epsilon_1' \frac{\pi i}{2} + \frac{1}{2}\epsilon_1 a_{1,1} + \cdots + \frac{1}{2}\epsilon_p a_{p,1}, \ldots, \epsilon_p' \frac{\pi i}{2} + \frac{1}{2}\epsilon_1 a_{1,p} + \cdots + \frac{1}{2}\epsilon_p a_{p,p}\right).$$

where the ϵ, ϵ' denote integers which may be reduced to their smallest remainders (modulo 2). In general, the condition $\theta(e_1, e_2, \ldots, e_p) = 0$ will be fulfilled by such a system of numbers only when

$$\epsilon_1\epsilon_1' + \epsilon_2\epsilon_2' + \cdots + \epsilon_p\epsilon_p' \equiv 1 \pmod 2. \tag{4}$$

There are $2^{p-1}(2^p - 1)$ such systems of numbers ϵ, ϵ' and accordingly, in general, the same number of Abelian functions. The matrix

$$\begin{pmatrix} \epsilon & \epsilon_2 & \cdots & \epsilon_p \\ \epsilon_1' & \epsilon_2' & \cdots & \epsilon_p' \end{pmatrix}$$

is called the *characteristic* of the function

$$\sqrt{\sum_1^p c_\nu \phi_\nu(s,z)}$$

and will be denoted by

$$\left(\sqrt{\sum_1^p c_\nu \phi_\nu(s,z)}\right).$$

The characteristic is said to be *odd* when the congruence (4) is satisfied; otherwise it is *even*. The number of even characteristics amounts to $2^{p-1}(2^p + 1)$ and in general there are no Abelian functions corresponding to these.

By the sum of two characteristics, we understand the characteristic obtained by adding corresponding elements and reducing them to 0 or 1 modulo 2. The sum and difference of two characteristics are accordingly identical.

We first bring the equation $F(s,z) = 0$ into a symmetric form by introducing new variables. If $p \geq 3$, there exist at least three linearly independent functions ϕ, and accordingly we can transform the equation $F(s,z) = 0$ by introducing the variables

$$\xi = \frac{\phi_1}{\phi_3}\ , \ \eta = \frac{\phi_2}{\phi_3}$$

(provided that no identical equation holds between ξ and η, a condition which is fulfilled in general).

If the function ϕ_1, ϕ_2, ϕ_3 does not satisfy special conditions, then to each value of ξ belong $2p-2$ values of η, and conversely, since each of the functions

$$\phi_1 - \xi\phi_3\ , \ \phi_2 - \eta\phi_3$$

vanish at $2p-2$ points for a constant ξ, respectively η. The resulting equation $F(\xi, \eta) = 0$ is thus of degree $2p - 2$ with respect to each variable. Since, moreover, this degree must still occur under any linear substitution for ξ, η, no term in the equation can have total degree in ξ, η exceeding $2p-2$. The remaining functions ϕ, expressed in terms of ξ, η, become functions in which no term can have degree more than $2p-5$, as we perceive from the fact that

$$\int \frac{\phi}{\frac{\partial F}{\partial \xi}}\, d\eta$$

must remain finite for infinite values of ξ and η.

The number of constants that occur in such a function of degree $2p-5$ is $(p-2)(2p-3)$. We determine r of these constants in such a way that the functions ϕ vanish for the r pairs of values (γ, δ) where $\frac{\partial F}{\partial \xi}, \frac{\partial F}{\partial \eta}$ are simultaneously zero. Then p constants remain, since there are p linearly independent integrals of the first kind. Accordingly,

$$(p-2)(2p-3) = p + r,$$

and consequently.

$$r = 2(p-1)(p-3).$$

We may reach the same result in the following way: the function $\frac{\partial F}{\partial \xi}$ becomes infinitely small of first order at $(2p-2)(2p-3)$ points, and this number is $w + 2r$, where w is the number of simple branch points. On the other hand (**VI**, Section 7, p. 105),

$$w = 2(n+p-1)\ ,\ n = 2p-2,$$
$$w = 2(3p-3),$$

hence

$$r = (p-1)(2p-3) - \frac{1}{2}\, w = 2(p-1)(p-3).$$

If we now express all the functions ϕ in terms of ξ, η, then the pair of equations

$$\xi = \frac{\phi_1}{\phi_3}\ , \eta = \frac{\phi_2}{\phi_3}$$

must become identities. Thus

$$\phi_1 = \xi\phi_3\ ,\ \phi_2 = \eta\phi_3.$$

Hence there must be a function ϕ_3 that is of degree $2p-6$ with respect to ξ, η. Thus this function ϕ will vanish for $(2p-2)(2p-6) = 2r$ pairs of values of ξ, η satisfying the equation $F = 0$, and accordingly is only zero at the r pairs (γ, δ).

Finally, introducing the new variables $\xi = \frac{x}{z}$, $\eta = \frac{y}{z}$, and multiplying by z^{2p-2}, the equation $F = 0$ becomes a homogeneous equation of degree $2p-2$ in the three variables x, y, z:

$$\overset{2p-2}{F(x, y, z)} = 0.$$

As we have seen, one among the functions ϕ is of degree $2p-6$ with respect to ξ, η; denote this function by ψ, then $\frac{\phi}{\psi}$ is a function, always finite for finite ξ, η, that becomes infinite of first order for infinite ξ and η. Conversely, each function that has these properties may be represented in the form $\frac{\phi}{\psi}$ (**VI**, Section 10, p. 109).

Functions that remain for finite values of ξ and η, and become infinite of second order for infinite ξ, η, can be represented in the form

$$\frac{f(\xi, \eta)}{\psi},$$

where $f(\xi, \eta)$ is a polynomial of degree $2p-4$ in ξ, η, which must vanish for the r pairs of values γ, δ. The function $f(\xi, \eta)$ contains

$$(p-1)(2p-3) - r = 3p - 3$$

constants, and accordingly (**VI**, Section 5, p. 99) can represent every function with these properties. Apart from the r pairs of values γ, δ, the function $f(\xi, \eta)$ becomes infinitely small of first order at $4p-4$ points.

Every polynomial of degree two in the $p-1$ variables $\frac{\phi}{\psi}$ belongs to this set of functions; and such a polynomial contains $p(p+1)/2$ constants. Since the general function f/ψ only contains $3p-3$ constants, there must be

$$p(p+1)/2 - 3p + 3 = (p-2)(p-3)/2$$

equations of second degree between the $p-1$ variables ϕ/ψ. That is to say, there are $(p-2)(p-3)/2$ homogeneous equations of second degree between the p functions ϕ.

In the case $p = 3$, the equation $F(\xi, \eta) = 0$, or $F(x, y, z) = 0$, is of degree 4; we have $r = 0$, and the function ψ reduces to a constant. None of the functions ϕ can have degree more than 1, and the general expression for these functions is

$$\phi = c\xi + c'\eta + c'',$$

or, if only the ratio of such functions is important,

$$\phi = cx + c'y + c''z,$$

where c, c', c'' are constants. Each function ϕ becomes infinitely small of first order at 4 points, and there are 28 such functions, whose zeros coincide in pairs. The square roots of these are the Abelian functions, and we must investigate how the characteristics of these 28 functions are associated.

As variables x, y, z we introduce three such functions ϕ that twice become infinitely small of first order, so that $\sqrt{x}, \sqrt{y}, \sqrt{z}$ are Abelian functions. The equation $F(x, y, z) = 0$ that emerges has the property that it becomes a perfect square when one of x, y or z is set equal to 0. Accordingly,

$$\begin{aligned} \text{for } x = 0 : F &= (y - \alpha z)^2 (y - \alpha' z)^2; \\ \text{for } y = 0 : F &= (z - \beta x)^2 (z - \beta' x)^2; \\ \text{for } z = 0 : F &= (x - \gamma y)^2 (x - \gamma' y)^2. \end{aligned}$$

Now let a, b, c be the coefficients of x^4, y^4, z^4 in $F(x, y, z)$; then

$$\alpha\alpha' = \pm\sqrt{\frac{c}{b}}\,, \quad \beta\beta' = \pm\sqrt{\frac{a}{c}}\,, \quad \gamma\gamma' = \pm\sqrt{\frac{b}{a}}.$$

It follows that

$$\alpha\alpha'\ \beta\beta'\ \gamma\gamma' = \pm 1. \tag{5}$$

Thus if the quantities $\alpha, \alpha', \beta, \beta', \gamma, \gamma'$ are known, all terms of the function $F(x, y, z)$ that do not contain the product xyz can be formed; in addition, F contains a term $xyzt$, where t is a linear form in x, y, z.

Now if the upper sign holds in (5), one can always represent the first part of F as the square of a quadratic form f in x, y, z. Let us write

$$f = a_{1,1}x^2 + a_{2,2}y^2 + a_{3,3}z^2 + 2a_{2,3}yz + 2a_{3,1}zx + 2a_{1,2}xy.$$

For determination of the coefficients $a_{i,k}$ we have the equations

$$\begin{aligned} \alpha\alpha' &= \frac{a_{3,3}}{a_{2,2}}\,, & \alpha + \alpha' &= -2\,\frac{a_{2,3}}{a_{2,2}}, \\ \beta\beta' &= \frac{a_{1,1}}{a_{3,3}} & \beta + \beta' &= -2\,\frac{a_{3,1}}{a_{3,3}} \\ \gamma\gamma' &= \frac{a_{2,2}}{a_{1,1}}\,, & \gamma + \gamma' &= -2\,\frac{a_{1,2}}{a_{1,1}}, \end{aligned}$$

which can always be satisfied if $\alpha\alpha'\beta\beta'\gamma\gamma' = 1$. Under this condition, the equation $F = 0$ becomes

$$f^2 - xyzt = 0. \tag{6}$$

If we let $t = 0$, then from $f^2 = 0$ we obtain two further pairs of equal roots. Accordingly, $\sqrt{t}$ is an Abelian function, and has the property that $\sqrt{xyzt}$ is a rational function of x, y, z. If we write $(a), (b), (c), (d)$ for the characteristics of $\sqrt{x}, \sqrt{y}, \sqrt{z}, \sqrt{t}$, we must have

$$(a+b+c+d) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

or

$$(d) = (a+b+c).$$

Hence the sum of the characteristics of the three functions $\sqrt{x}, \sqrt{y}, \sqrt{z}$ is an odd characteristic.

Conversely, if the latter condition is fulfilled, and $\sqrt{t}$ is the Abelian function belonging to the characteristic $(a+b+c)$, then $\sqrt{xyzt}$ is a function that varies continuously across the transverse cuts and hence can be expressed rationally in x, y, z. However, the degree of this function cannot exceed 2 and accordingly, under this hypothesis, we always have an equation of the form (6). The equation cannot hold identically if $\sqrt{x}, \sqrt{y}, \sqrt{z}, \sqrt{t}$ are distinct Abelian functions.

Since there are 28 Abelian functions, the equation $F = 0$ can be reduced in several ways to the form (6). We investigate first whether the pair of Abelian functions $\sqrt{z}, \sqrt{t}$ can be replaced by another pair $\sqrt{p}, \sqrt{q}$.

Suppose $F = 0$ may also be brought, by introducing x, y, p, q, into the form

$$\psi^2 - xypq = 0.$$

With the introduction of an appropriate constant factor, we obtain the identical equation

$$f^2 - xyzt = \psi^2 - xypq,$$

or

$$(f-\psi)(f+\psi) = xy(zt-pq).$$

Accordingly, $f - \psi$ or $f + \psi$ must be divisible by xy and, since both are of degree 2, can only differ from xy by a constant factor. Now if

$$\begin{aligned} \psi - f &= \alpha xy, \\ \alpha(\psi + f) &= -zt + pq, \end{aligned} \tag{7}$$

we have

$$(8) \qquad \begin{aligned} \psi &= \alpha xy + f, \\ 2\alpha f + \alpha^2 xy + zt &= pq. \end{aligned}$$

The left side of the last equation therefore decomposes into two linear factors. Consider the expansion of this function in the form

$$a_{1,1}x^2 + a_{2,2}y^2 + a_{3,3}z^2 + 2a_{2,3}yz + 2a_{3,1}zx + 2a_{1,2}xy.$$

Then the coefficients $a_{i,k}$ are second degree polynomials in α. As the determinant

$$\sum \pm a_{1,1}a_{2,2}a_{3,3}$$

must vanish, we obtain an equation of degree 6 for α, which we readily see has the roots $\alpha = 0$, $\alpha = \infty$ corresponding to the two decompositions zt and xy.

Thus an equation of degree 4 is left, whose roots produce four pairs of functions p, q having the required property.

From the second equation (8), combined with (6), it follows that

$$pqzt = z^2t^2 + 2\alpha f zt + \alpha^2 f^2 = (zt + \alpha f)^2,$$

so that we can also establish the desired form of the equation $F = 0$ in terms of the functions p, q, z, t. If we start from two arbitrary Abelian functions $\sqrt{x}, \sqrt{y}$, we obtain 6 pairs of such functions:

$$\sqrt{xy}, \sqrt{zt}, \sqrt{p_1q_1}, \sqrt{p_2q_2}, \sqrt{p_3q_3}, \sqrt{p_4q_4},$$

having the property that, taking any pair of them, the equation $F = 0$ may be brought to the form

$$f^2 - xyzt = 0.$$

These 6 functions must acquire the same factors in crossing transverse cuts, for otherwise the product of two of them could not be rational. Such sets of 6 products of pairs of Abelian functions are said to *belong to a group.* Since the system of factors at the transverse cuts for products of Abelian functions are determined by the sums of the characteristics, it follows that the characteristics of all pairs belonging to a group must yield the same sum, which we call the *characteristic of the group.*

From the equations (8) and (6), we have further

$$2f = \frac{pq - zt}{\alpha} - \alpha xy = 2\sqrt{xy}\sqrt{zt},$$

whence

$$pq = \alpha^2 xy + 2\alpha\sqrt{xy}\sqrt{zt} + zt$$

or:

$$\sqrt{pq} = \sqrt{zt} + \alpha\sqrt{xy}. \tag{9}$$

We conclude that each product in a group can be expressed linearly in two products from the same group.

If we arrange all 28 Abelian functions into pairs, we obtain $\frac{28.27}{2} = 6.63$ pairs, which decompose in sixes into 63 groups. Every one of the 63 characteristics except $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ can be the characteristic of a group.

To obtain the characteristics of the 6 pairs in a group, we must decompose the group characteristic in question into two odd characteristics in 6 ways. The group whose characteristic is $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ serves as an example:

$$\begin{aligned}
\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}.
\end{aligned}$$

When three pairs of Abelian functions are known, we can obtain the other pairs from the same group by solution of a cubic equation, and with the help of this equation we can determine all the other Abelian functions and their characteristics.

To carry this out, let us suppose that $\sqrt{x\xi}$, $\sqrt{y\eta}$, $\sqrt{z\zeta}$ are three pairs from a group, so that ξ, η, ζ are given as linear forms in x, y, z.

By appropriate choice of constant factors, we can bring equation (9) into the form

$$\sqrt{x\xi} + \sqrt{y\eta} + \sqrt{z\zeta} = 0, \tag{10}$$

which yields

$$z\zeta = x\xi + y\eta + 2\sqrt{x\xi y\eta},$$

or

(11) $$4x\xi\, y\eta = (z\zeta - x\xi - y\eta)^2,$$

so that

(12) $$f = z\zeta - x\xi - y\eta.$$

In order to find all the pairs belonging to the group $\sqrt{x\xi}$, $\sqrt{y\eta}$, we must, by the above, solve a quartic equation, one of whose roots, corresponding to the pair $\sqrt{z\zeta}$, is already known. The calculation will be more symmetric if we first look for the pairs of the group $\sqrt{x\eta}$, in which the pair $\sqrt{y\xi}$ also belongs.

If $\sqrt{pq}$ is a further undetermined pair from this group, then along with the equation (11) we have another equation that is identical with it,

(13) $$4y\xi\, pq = \phi^2,$$

with (according to (8))

$$\phi = f + 2\lambda\, y\xi.$$

Here λ denotes an as yet undetermined constant. From this we obtain, by means of (11) and (12),

$$\phi^2 = 4\lambda\, y\xi\left(x\xi + y\eta - z\zeta + \frac{x\eta}{\lambda} + \lambda y\xi\right).$$

Thus, apart from the factor λ, we have

$$\begin{aligned} pq &= x\xi + y\eta - z\zeta + \frac{x\eta}{\lambda} + \lambda y\xi \\ &= \left(\xi + \frac{\eta}{\lambda}\right)(x + \lambda y) - z\zeta. \end{aligned}$$

For $x + \lambda y = 0$ and $z = 0$, one of the pair of functions p, q must vanish, say p. Thus, denoting by μ a further unknown coefficient, we find that

(14) $$\begin{aligned} p &= x + \lambda y + \mu z, \\ pq &= p\left(\xi + \frac{\eta}{\lambda}\right) - \mu z\left(\xi + \frac{\eta}{\lambda} + \frac{\zeta}{\mu}\right). \end{aligned}$$

Moreover, since p and z are not identical,

$$\xi + \frac{\eta}{\lambda} + \frac{\zeta}{\mu} = -a^2 p. \tag{15}$$

Thus, with the help of (13), we obtain

$$ax + a\lambda y + a\mu z + \frac{\xi}{a} + \frac{\eta}{\lambda a} + \frac{\zeta}{\mu a} = 0.$$

Replacing $\lambda a, \mu a$ by b, c,

$$ax + by + cz + \frac{\xi}{a} + \frac{\eta}{b} + \frac{\zeta}{c} = 0. \tag{16}$$

Now, since p and q are determined only up to a constant factor, we obtain

$$p = ax + by + cz = -\left(\frac{\xi}{z} + \frac{\eta}{b} + \frac{\zeta}{c}\right),$$

$$q = \frac{\xi}{a} + \frac{\eta}{b} + cz = -\left(ax + by + \frac{\zeta}{c}\right).$$

Since there are four pairs p, q, it must be possible to determine four systems a, b, c.

To obtain these, we recall that 3 homogeneous linear equations hold between the 6 functions $x, y, z, \xi, \eta, \zeta$; we denote these equations by $u_1 = 0$, $u_2 = 0$, $u_3 = 0$. Now we form a linear combination of these equations with undetermined coefficients ℓ_1, ℓ_2, ℓ_3:

$$\ell_1 u_1 + \ell_2 u_2 + \ell_3 u_3 = \alpha x + \beta y + \gamma z + \alpha' \xi + \beta' \eta + \gamma' \zeta = 0.$$

Here $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$ are linear forms in ℓ_1, ℓ_2, ℓ_3. This relation will take the form (16) if the conditions

$$\alpha\alpha' = \beta\beta' = \gamma\gamma'$$

are fulfilled. From this one obtains four sets of values for the ratios $\ell_1 : \ell_2 : \ell_3$.

We obtain our objective in the most elegant way if we consider the functions ξ, η, ζ to be given by three equations of the form

$$\begin{cases} x + y + z + \xi + \eta + \zeta = 0, \\ \alpha x + \beta y + \gamma z + \dfrac{\xi}{\alpha} + \dfrac{\eta}{\beta} + \dfrac{\zeta}{\alpha} = 0, \\ \alpha' x + \beta' y + \gamma' z + \dfrac{\xi}{\alpha'} + \dfrac{\eta}{\beta'} + \dfrac{\zeta}{\gamma'} = 0. \end{cases} \tag{17}$$

One can arrange for the coefficients in the first of these equations to have the values 1 by attaching constant factors to $x, y, z, \xi, \eta, \zeta$, in such a way that (10) does not change in form.

The equations (17) must yield, as an identity, a fourth equation of the same form:

$$\alpha'' x + \beta'' y + \gamma'' z + \frac{\xi}{\alpha''} + \frac{\eta}{\beta''} + \frac{\zeta}{\gamma''} = 0. \tag{18}$$

Thus to obtain $\alpha'', \beta'', \gamma''$, we must determine the coefficients $\lambda, \lambda', \lambda''$ from the following equations:

$$\begin{cases} \lambda''\alpha'' = \lambda'\alpha' + \lambda\alpha + 1, & \dfrac{\lambda''}{\alpha''} = \dfrac{\lambda'}{\alpha'} + \dfrac{\lambda}{\alpha} + 1, \\ \lambda''\beta'' = \lambda'\beta' + \lambda\beta + 1, & \dfrac{\lambda''}{\beta''} = \dfrac{\lambda'}{\beta'} + \dfrac{\lambda}{\beta} + 1, \\ \lambda''\gamma'' = \lambda'\gamma' + \lambda\gamma + 1, & \dfrac{\lambda''}{\gamma''} = \dfrac{\lambda'}{\gamma'} + \dfrac{\lambda}{\gamma} + 1. \end{cases} \tag{19}$$

By multiplying corresponding pairs of these equations, we obtain

$$\begin{cases} \lambda''^2 = \lambda'^2 + \lambda^2 + \lambda\lambda'\left(\dfrac{\alpha}{\alpha'} + \dfrac{\alpha'}{\alpha}\right) + \lambda\left(\alpha + \dfrac{1}{\alpha}\right) + \lambda'\left(\alpha' + \dfrac{1}{\alpha'}\right) + 1, \\ \lambda''^2 = \lambda'^2 + \lambda^2 + \lambda\lambda'\left(\dfrac{\beta}{\beta'} + \dfrac{\beta'}{\beta}\right) + \lambda\left(\beta + \dfrac{1}{\beta}\right) + \lambda'\left(\beta' + \dfrac{1}{\beta'}\right) + 1, \\ \lambda''^2 = \lambda'^2 + \lambda^2 + \lambda\lambda'\left(\dfrac{\gamma}{\gamma'} + \dfrac{\gamma'}{\gamma}\right) + \lambda\left(\gamma + \dfrac{1}{\gamma}\right) + \lambda'\left(\gamma' + \dfrac{1}{\gamma'}\right) + 1. \end{cases} \tag{20}$$

If we eliminate λ'' from each pair of equations, we obtain the following pair of linear equations for $\frac{1}{\lambda}$ and $\frac{1}{\lambda'}$:

$$0 = \frac{1}{\lambda'}\left(\alpha + \frac{1}{\alpha} - \beta - \frac{1}{\beta}\right) + \frac{1}{\lambda}\left(\alpha' + \frac{1}{\alpha'} - \beta' - \frac{1}{\beta'}\right) + \left(\frac{\alpha'}{\alpha} + \frac{\alpha}{\alpha'} - \frac{\beta'}{\beta} - \frac{\beta}{\beta'}\right),$$

$$0 = \frac{1}{\lambda'}\left(\alpha + \frac{1}{\alpha} - \gamma - \frac{1}{\gamma}\right) + \frac{1}{\lambda}\left(\alpha' + \frac{1}{\alpha'} - \gamma' - \frac{1}{\gamma'}\right) + \left(\frac{\alpha'}{\alpha} + \frac{\alpha}{\alpha'} - \frac{\gamma'}{\gamma} - \frac{\gamma}{\gamma'}\right).$$

From these equations λ, λ' can be calculated uniquely.

Now, from one of the equations (20), we obtain λ'' apart from sign, and finally from (19) we likewise obtain $\alpha'', \beta'', \gamma''$ up to the set of common signs which the nature of the matter leaves undetermined.[1]

Once we have found $\alpha'', \beta'', \gamma''$ in this way, we obtain in the group $\sqrt{x\eta}$, $\sqrt{y\xi}$ the following four pairs of Abelian functions:

$$\begin{aligned}
&\sqrt{x+y+z} \quad , \quad \sqrt{\xi+\eta+z};\\
&\sqrt{\alpha x+\beta y+\gamma z} \quad , \quad \sqrt{\frac{\xi}{\alpha}+\frac{\eta}{\beta}+\gamma z};\\
&\sqrt{\alpha' x+\beta' y+\gamma' z} \quad , \quad \sqrt{\frac{\xi}{\alpha'}+\frac{\eta}{\beta'}+\gamma' z};\\
&\sqrt{\alpha'' x+\beta'' y+\gamma'' z} \quad , \quad \sqrt{\frac{\xi}{\alpha''}+\frac{\eta}{\beta''}+\gamma'' z}.
\end{aligned}$$

[1]For brevity, let

$$\begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' \end{vmatrix} = (\alpha,\beta,\gamma), \quad , \quad \begin{vmatrix} 1 & 1 & 1 \\ \frac{1}{\alpha} & \beta & \gamma \\ \frac{1}{\alpha'} & \beta' & \gamma' \end{vmatrix} = \left(\frac{1}{\alpha},\beta,\gamma\right),$$

and so on. Then one can determine $\alpha'', \beta'', \gamma''$ from

$$\alpha\alpha'\alpha'' : \beta\beta'\beta'' = (\alpha,\beta,\gamma)\left(\alpha,\beta,\frac{1}{\gamma}\right) : \left(\frac{1}{\alpha},\frac{1}{\beta},\gamma\right)\left(\frac{1}{\alpha},\frac{1}{\beta},\frac{1}{\gamma}\right),$$

$$\alpha\alpha'\alpha'' : \beta\beta'\beta'' = \left(\alpha,\frac{1}{\beta},\gamma\right)\left(\alpha,\frac{1}{\beta},\frac{1}{\gamma}\right) : \left(\frac{1}{\alpha},\beta,\alpha\right)\left(\frac{1}{\alpha},\beta,\frac{1}{\gamma}\right)$$

and the analogous equations.

In the same way we obtain the pairs

$$\sqrt{x+y+z} \quad , \quad \sqrt{\xi+y+\zeta};$$
$$\sqrt{\alpha x+\beta y+\gamma z} \quad , \quad \sqrt{\frac{\xi}{\alpha}+\beta y+\frac{\zeta}{\gamma}};$$
$$\sqrt{\alpha' x+\beta' y+\gamma' z} \quad , \quad \sqrt{\frac{\xi}{\alpha'}+\beta' y+\frac{\zeta}{\gamma'}};$$
$$\sqrt{\alpha'' x+\beta'' y+\gamma'' z} \quad , \quad \sqrt{\frac{\xi}{\alpha''}+\beta'' y+\frac{\zeta}{\gamma''}}$$

in the group $\sqrt{x\zeta}$, $\sqrt{z\zeta}$, and

$$\sqrt{x+y+z} \quad , \quad \sqrt{x+\eta+\zeta};$$
$$\sqrt{\alpha x+\beta y+\gamma z} \quad , \quad \sqrt{\alpha x+\frac{\eta}{\beta}+\frac{\zeta}{\gamma}};$$
$$\sqrt{\alpha' x+\beta' y+\gamma' z} \quad , \quad \sqrt{\alpha' x+\frac{\eta}{\beta'}+\frac{\zeta}{\gamma'}};$$
$$\sqrt{\alpha'' x+\beta'' y+\gamma'' z} \quad , \quad \sqrt{\alpha'' x+\frac{\eta}{\beta''}+\frac{\zeta}{\gamma''}}$$

in the group $\sqrt{y\zeta}$, $\sqrt{z\eta}$. Thus from the given 6 Abelian functions, we determine 16 more. To obtain their characteristics, we need only note that the three groups considered here contain four Abelian functions in common. Thus if we form the characteristics of the groups, these four characteristics must also be common to them. These four characteristics must be associated to the functions

$$\sqrt{x+y+z},\ \sqrt{\alpha x+\beta y+\gamma z},\ \sqrt{\alpha' x+\gamma' y+\gamma' z},\ \sqrt{\alpha'' x+\beta'' y+\gamma'' z}$$

in an arbitrary fashion. The characteristics of the remaining Abelian functions are fully determined by this, since they appear paired with these characteristics in three groups in the same way as the corresponding Abelian functions. These characteristics may be represented symmetrically in the following way.

Let the characteristics of the groups $\sqrt{y\zeta}$, $\sqrt{z\xi}$, $\sqrt{x\eta}$ respectively be denoted by $(p), (q), (r)$. Further denote by $(d), (e), (f), (g)$ the characteristics of the four functions

$$\sqrt{x+y+z},\ \sqrt{\alpha x+\beta y+\gamma z},\ \sqrt{\alpha' x+\beta' y+\gamma' z},\ \sqrt{\alpha'' x+\beta'' y+\gamma'' z}$$

and that of $\sqrt{x}$ by $(n+p)$. From this we obtain the following expressions for characteristics:

$$(21)\quad \begin{cases} (\sqrt{x}) = (n+p)\ ,\ (\sqrt{y}) = (n+q)\ ,\ (\sqrt{z}) = (n+r), \\ (\sqrt{\xi}) = (n+q+r)\ ,\ (\sqrt{\eta}) = (n+r+p)\ ,\ (\sqrt{\zeta}) = (n+p+q), \\ (\sqrt{x+y+z}) = (d)\ ,\ (\sqrt{x+\eta+\zeta}) = (p+d), \\ (\sqrt{\alpha x+\beta y+\gamma z}) = (e)\ ,\ \left(\sqrt{\alpha x+\frac{\eta}{\beta}+\frac{\zeta}{\gamma}}\right) = (p+e), \\ (\sqrt{\alpha' x+\beta' y+\gamma' z}) = (f)\ ,\ \left(\sqrt{\alpha' x+\frac{\eta}{\beta'}+\frac{\zeta}{\eta'}}\right) = (p+f), \\ (\sqrt{\alpha'' x+\beta'' y+\gamma'' z}) = (g)\ ,\ \left(\sqrt{\alpha'' x+\frac{\eta}{\beta''}+\frac{\zeta}{\gamma''}}\right) = (p+g), \\ (\sqrt{\xi+y+\zeta}) = (q+d)\ ,\ (\sqrt{\xi+\eta+z}) = (r+d), \\ \left(\sqrt{\frac{\xi}{\alpha}+\beta y+\frac{\zeta}{\gamma}}\right) = (q+e)\ ,\ \left(\sqrt{\frac{\xi}{\alpha}+\frac{\eta}{\beta}+\gamma z}\right) = (r+e), \\ \left(\sqrt{\frac{\xi}{\alpha'}+\beta' y+\frac{\zeta}{\gamma'}}\right) = (q+f)\ ,\ \left(\sqrt{\frac{\xi}{\alpha'}+\frac{\eta}{\beta'}+\gamma' z}\right) = (r+f), \\ \left(\sqrt{\frac{\xi}{\alpha''}+\beta'' y+\frac{\zeta}{\gamma''}}\right) = (q+g)\ ,\ \left(\sqrt{\frac{\xi}{\alpha''}+\frac{\eta}{\beta''}+\gamma'' z}\right) = (r+g). \end{cases}$$

As an example, suppose that

$$(\sqrt{x}) = \begin{pmatrix}1 & 0 & 1\\ 1 & 0 & 0\end{pmatrix}\ ,\ (\sqrt{y}) = \begin{pmatrix}1 & 1 & 1\\ 1 & 0 & 0\end{pmatrix}\ ,\ (\sqrt{z}) = \begin{pmatrix}1 & 0 & 1\\ 1 & 1 & 0\end{pmatrix}$$
$$(\sqrt{\xi}) = \begin{pmatrix}1 & 0 & 0\\ 1 & 0 & 0\end{pmatrix}\ ,\ (\sqrt{\eta}) = \begin{pmatrix}1 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}\ ,\ (\sqrt{\zeta}) = \begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 0\end{pmatrix}$$

which is admissible, since from this $\sqrt{x\xi}$, $\sqrt{y\eta}$, $\sqrt{z\zeta}$ belong in the same group

$\begin{pmatrix}0&0&1\\0&0&0\end{pmatrix}$. It follows that

$$(p)=\begin{pmatrix}0&1&1\\0&1&0\end{pmatrix}, \ (q)=\begin{pmatrix}0&0&1\\0&1&0\end{pmatrix}, \ (r)=\begin{pmatrix}0&1&1\\0&0&0\end{pmatrix}, \ (n)=\begin{pmatrix}1&1&0\\1&1&0\end{pmatrix}.$$

The complete groups $(p), (q)$ are

$$\begin{aligned}\begin{pmatrix}0&1&1\\0&1&0\end{pmatrix}&=\begin{pmatrix}1&0&0\\1&1&0\end{pmatrix}+\begin{pmatrix}1&1&1\\1&0&0\end{pmatrix}=\begin{pmatrix}1&0&1\\1&1&0\end{pmatrix}+\begin{pmatrix}1&1&0\\1&0&0\end{pmatrix}\\&=\begin{pmatrix}0&1&0\\0&1&1\end{pmatrix}+\begin{pmatrix}0&0&1\\0&0&1\end{pmatrix}\\&=\begin{pmatrix}1&1&0\\0&1&1\end{pmatrix}+\begin{pmatrix}1&0&1\\0&0&1\end{pmatrix}=\begin{pmatrix}1&1&1\\1&1&1\end{pmatrix}+\begin{pmatrix}1&0&0\\1&0&1\end{pmatrix}\\&=\begin{pmatrix}0&1&0\\1&1&1\end{pmatrix}+\begin{pmatrix}0&0&1\\1&0&1\end{pmatrix},\end{aligned}$$

$$\begin{aligned}\begin{pmatrix}0&0&1\\0&1&0\end{pmatrix}&=\begin{pmatrix}1&0&0\\1&1&0\end{pmatrix}+\begin{pmatrix}1&0&1\\1&0&0\end{pmatrix}=\begin{pmatrix}1&0&1\\1&1&0\end{pmatrix}+\begin{pmatrix}1&0&0\\1&0&0\end{pmatrix}\\&=\begin{pmatrix}0&1&0\\0&1&1\end{pmatrix}+\begin{pmatrix}0&1&1\\0&0&1\end{pmatrix}\\&=\begin{pmatrix}1&1&0\\0&1&1\end{pmatrix}+\begin{pmatrix}1&1&1\\0&0&1\end{pmatrix}=\begin{pmatrix}1&1&1\\1&1&1\end{pmatrix}+\begin{pmatrix}1&1&0\\1&0&1\end{pmatrix}\\&=\begin{pmatrix}0&1&0\\1&1&1\end{pmatrix}+\begin{pmatrix}0&1&1\\1&0&1\end{pmatrix},\end{aligned}$$

whence we obtain

$$(d)=\begin{pmatrix}0&1&0\\0&1&1\end{pmatrix}, \ (e)=\begin{pmatrix}1&1&0\\0&1&1\end{pmatrix}, \ (f)=\begin{pmatrix}1&1&1\\1&1&1\end{pmatrix}, \ (g)=\begin{pmatrix}0&1&0\\1&1&1\end{pmatrix},$$

and the characteristics of the functions collected in (21) are, in the same order,

$$\begin{pmatrix}1&0&1\\1&0&0\end{pmatrix}, \ \begin{pmatrix}1&1&1\\1&0&0\end{pmatrix}, \ \begin{pmatrix}1&0&1\\1&1&0\end{pmatrix},$$

$$\begin{pmatrix}1&0&0\\1&0&0\end{pmatrix}, \ \begin{pmatrix}1&1&0\\1&0&0\end{pmatrix}, \ \begin{pmatrix}1&0&0\\1&1&0\end{pmatrix},$$

$$\begin{pmatrix}0&1&0\\0&1&1\end{pmatrix},\ \begin{pmatrix}0&0&1\\0&0&1\end{pmatrix},\ \begin{pmatrix}0&1&1\\0&0&0\end{pmatrix},\ \begin{pmatrix}0&0&1\\0&1&1\end{pmatrix},$$
$$\begin{pmatrix}1&1&0\\0&1&1\end{pmatrix},\ \begin{pmatrix}1&0&1\\0&0&1\end{pmatrix},\ \begin{pmatrix}1&1&1\\0&0&1\end{pmatrix},\ \begin{pmatrix}1&0&1\\0&1&1\end{pmatrix},$$
$$\begin{pmatrix}1&1&1\\1&1&1\end{pmatrix},\ \begin{pmatrix}1&0&0\\1&0&1\end{pmatrix},\ \begin{pmatrix}1&1&0\\1&0&1\end{pmatrix},\ \begin{pmatrix}1&0&0\\1&1&1\end{pmatrix},$$
$$\begin{pmatrix}0&1&0\\1&1&1\end{pmatrix},\ \begin{pmatrix}0&0&1\\1&0&1\end{pmatrix},\ \begin{pmatrix}0&1&1\\1&0&1\end{pmatrix},\ \begin{pmatrix}0&0&1\\1&1&1\end{pmatrix}.$$

For three Abelian functions of a group, two of which belong to one pair, we have the result that the sum of their characteristics is always an even characteristic. Consider, for example, the three functions $\sqrt{x}$, $\sqrt{y}$, $\sqrt{z}$ and express ξ, η, ζ linearly in x, y, z. Then the equation (10) can be written in the form

$$\sqrt{x(ax+by+cz)} + \sqrt{y(a'x+b'y+c'z)} + \sqrt{z(a''x+b''y+c''z)} = 0.$$

If we successively set $x = 0$, $y = 0$, $z = 0$ in the equation, then we find the following values for the products of the roots of the quadratic equations that emerge for the ratio of the other two variables:

$$-\frac{c''}{b'},\ -\frac{a}{c''},\ -\frac{b'}{a}$$

with product -1. However, according to pp. 488–9, this is the criterion for the sum of the characteristics of the functions $\sqrt{x}$, $\sqrt{y}$, $\sqrt{z}$ to be an even characteristic.

Based on this result, one can prove that the 16 Abelian functions that we determined above are distinct from the 12 functions occurring in the group $\sqrt{x\xi}$. For if $\sqrt{pq}$ is a pair belonging in the group $\sqrt{x\xi}$, then the characteristics

$$(\sqrt{x}) + (\sqrt{\xi}) + (\sqrt{p}),\quad (\sqrt{y}) + (\sqrt{\eta}) + (\sqrt{p}),\quad (\sqrt{z}) + (\sqrt{\zeta}) + (\sqrt{p})$$

are odd, and by the result just proved, $\sqrt{p}$ cannot occur in any of the 3 groups

$$(\sqrt{x\eta}) = (\sqrt{y\xi}),\ (\sqrt{x\zeta}) = (\sqrt{z\xi}),\ (\sqrt{y\zeta}) = (\sqrt{z\eta}).$$

Hence the 16 functions determined above run over all Abelian functions not contained in the group $\sqrt{x\xi}$, and once we have found the 6 absent functions of this group, the 28 Abelian functions will be determined.

To attain this, we write

$$t = x + y + z, \quad u = \xi + \eta + z$$

and start from the equation

$$\sqrt{tu} = \sqrt{x\eta} + \sqrt{y\xi}, \tag{22}$$

which follows easily from (10) and (17). In the above treatment we insert the functions

$$t, x, y, u, \eta, \xi$$

in place of

$$x, y, z, \xi, \eta, \zeta.$$

We obtain at once the following equation between the variables:

$$t - x - y - u + \eta + \xi = 0, \tag{23}$$

alongside which there must be 3 more equations of the form

$$at + bx + cy + a'u + b'\eta + c'\xi = 0 \tag{24}$$

with the condition

$$aa' = bb' = cc'.$$

In place of the group $(p+q+r)$, (p), (q), (r), the following now appear:

$$\begin{cases} (\sqrt{tu}) = (\sqrt{x\eta}) = (\sqrt{y\xi}) = (r), \\ (\sqrt{x\xi}) = (\sqrt{y\eta}) = (\sqrt{z\zeta}) = (p+q+r), \\ (\sqrt{t\xi}) = (\sqrt{uy}) = (n+d+q+r), \\ (\sqrt{t\eta}) = (\sqrt{ux}) = (n+d+p+r). \end{cases} \tag{25}$$

In the first of these groups (r), the following pairs of characteristics occur:

$$\begin{aligned}(r) &= (n+p) + (n+r+p) = (n+q) + (n+r+q) \\ &= (d) + (r+d) = (e) + (r+e) = (f) + (r+f) = (g) + (r+g),\end{aligned}$$

and from the equation (23) we obtain the following Abelian functions:

$$\sqrt{t-x-y}=\sqrt{z},\ \sqrt{t+\eta+\xi}=\sqrt{-\zeta},$$
$$\sqrt{-u-x+\xi}=\sqrt{\xi+y+\zeta},\ \sqrt{-u+\eta-y}=\sqrt{x+\eta+\zeta},$$

whose characteristics are

$$(n+r),\ (n+p+q),\ (q+d),\ (p+d)$$

which are distributed in the last 3 groups (25) in the following way:

$$\begin{aligned}(p+q+r)&=(n+r)+(n+p+q),\\(n+d+q+r)&=(n+r)+(q+d),\\(n+d+p+r)&=(n+r)+(p+d).\end{aligned}$$

The characteristics of the Abelian functions that have not yet been determined must now, as shown above, be contained in the group $(p+q+r)$. If we denote these characteristics by $(k_1), (k_1'), (k_1''), (k_2), (k_2'), (k_2'')$, then we must have

$$(p+q+r)=(k_1+k_2)=(k_1'+k_2')=(k_1''+k_2''),$$

and these characteristics do not occur in the group (r).

However, comparison of the group (25) with the groups $(p+q+r)$, (p), (q), (r), now shows that in these groups all odd characteristics must actually occur, and further, that the three remaining pairs of the groups $(p+q+r)$, $(n+d+q+r)$, $(n+d+p+r)$ must all have one characteristic in common.

Now the characteristic $(q+e)$ occurs neither in the group (r) nor in $(p+q+r)$. It follows that (k_1) can be chosen so that either

$$(k_1+q+e)=(n+d+q+r)$$

or

$$(k_1+q+e)=(n+d+p+r).$$

From the first assumption, it would follow that

$$(k_1) = (n + r + d + e).$$

However, this is not possible, since in the group (p) we have the pairs

$$\begin{gathered}(n + r), (n + r + p),\\ (d), (d + p),\\ (e), (e + p)\end{gathered}$$

and accordingly, by the result proved on p. 499,

$$(n + r + d + e)$$

is even. Now we have

$$(k_1) = (n + d + e + p + q + r),$$

and it follows that

$$(k_2) = (n + d + e).$$

Likewise we conclude that

$$\begin{gathered}(k_1') = (n + d + f + p + q + r),\ (k_2') = (n + d + f),\\ (k_1'') = (n + d + g + p + q + r),\ (k_2'') = (n + d + g),\end{gathered}$$

and the group $(n + d + p + r)$ contains the pairs

$$(k_1), (q + e);\ (k_1'), (q + f);\ (k_1''), (q + g).$$

Thus we obtain the following pairs for the group $(n + d + q + r)$:

$$(k_1), (p + e);\ (k_1'), (p + f);\ (k_1''), (p + g).$$

From the results of the earlier treatment, an equation of the form (24) yields the 4 Abelian functions:

$$\begin{aligned}\sqrt{at + bx + cy} &= \sqrt{-(a'u + b'\eta + c'\xi)},\\ \sqrt{a'u + bx + cy} &= \sqrt{-(at + b'\eta + c'\xi)},\\ \sqrt{at + b'\eta + cy} &= \sqrt{-(a'u + bx + c'\xi)},\\ \sqrt{at + bx + c'\xi} &= \sqrt{-(a'u + b'\eta + cy)}\end{aligned}$$

with respective characteristics

$$(k_1), (k_2), (p+e), (q+e).$$

Accordingly our problem will be solved once we have determined the coefficients a, b, c, a', b', c'.

However, the function whose characteristic is $(p+e)$ is already determined above. It is

$$\sqrt{\alpha x + \frac{\eta}{\beta} + \frac{\zeta}{\gamma}}\,.$$

If we write

$$v = \alpha x + \frac{\eta}{\beta} + \frac{\zeta}{\gamma} = -\left(\frac{\xi}{\alpha} + \beta y + \gamma z\right),$$

we may now determine the coefficients a, b, c, a', b', c' by writing v in the following two ways:

$$v = at + b'\eta + cy = -a'u - bx - c'\xi.$$

We proceed in the following way: using

$$u = \xi + \eta + z = -x - y - \zeta,$$

we eliminate the variables z and ζ from the two expressions for v. We obtain

$$v + \frac{u}{\gamma} = x\left(\alpha - \frac{1}{\gamma}\right) - \frac{\eta}{\beta} - \frac{y}{\gamma},$$

$$v + \gamma u = -\xi\left(\frac{1}{\alpha} - \gamma\right) + \gamma\eta - \beta y.$$

Eliminating η and y from these equations, we have

$$v = u\,\frac{\beta-\gamma}{1-\beta\gamma} + x\,\frac{\beta(1-\alpha\gamma)}{1-\beta\gamma} - \frac{\xi}{\alpha}\,\frac{1-\alpha\gamma}{1-\beta\gamma}.$$

In the same way, we have

$$v = t\,\frac{1-\alpha\gamma}{\alpha-\gamma} + \frac{\eta}{\beta}\,\frac{\beta-\gamma}{\alpha-\gamma} - y\,\frac{\alpha(\beta-\gamma)}{\alpha-\gamma}.$$

We now obtain

$$a = \frac{1-\alpha\gamma}{\alpha-\gamma}, \; a' = -\frac{\beta-\gamma}{1-\beta\gamma},$$
$$b = -\frac{\beta(1-\alpha\gamma)}{1-\beta\gamma}, \; b' = \frac{1}{\beta}\,\frac{\beta-\gamma}{\alpha-\gamma},$$
$$c = -\frac{\alpha(\beta-\gamma)}{\alpha-\gamma}, \; c' = \frac{1}{\alpha}\,\frac{1-\alpha\gamma}{1-\beta\gamma}.$$

From this we may form the pair of Abelian functions

$$\sqrt{at+bx+cy}\;,\;\sqrt{a'u+bx+cy}.$$

In these functions, we replace t and u by their expressions in $x, y, z, \xi, \eta, \zeta$. Suppressing constant factors, we obtain the pair of expressions for the functions that belong to the characteristic (k_1):

$$\sqrt{\frac{x}{1-\beta\gamma}+\frac{y}{1-\gamma\alpha}+\frac{z}{1-\alpha\beta}}\;,\sqrt{\frac{\xi}{\alpha(\gamma-\beta)}+\frac{\eta}{\beta(\gamma-\alpha)}+\frac{z}{1-\alpha\beta}},$$

and for the functions belonging to the characteristic (k_2):

$$\sqrt{\frac{\xi}{\alpha(1-\beta\gamma)}+\frac{y}{\beta(1-\gamma\alpha)}+\frac{\zeta}{\gamma(1-\alpha\beta)}},\;\sqrt{\frac{x}{\gamma-\beta}+\frac{y}{\gamma-\alpha}+\frac{\zeta}{\gamma(1-\alpha\beta)}}.$$

We obtain at once the functions belonging to the characteristics $(k_1'), (k_2')$; $(k_1''), (k_2'')$, by replacing α, β, γ by α', β', γ' and $\alpha'', \beta'', \gamma''$ respectively. This determines all Abelian functions and their characteristics. In the example chosen above, the characteristics $(k_1), (k_2), (k_1'), (k_2'), (k_1''), (k_2'')$ would be arranged in the following way:

$$(k_1) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \; (k_1') = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \; (k_1'') = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$
$$(k_2) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \; (k_2') = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \; (k_2'') = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

Now as shown above, $\alpha'', \beta'', \gamma''$ can be expressed in terms of $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$. Hence all Abelian functions, with all their algebraic relations can be expressed via $3p-3=6$ constants, which one can regard as the moduli of the class for the case $p=3$.

Natural Philosophy.

1. Molecular mechanics.

Consider a system of point masses $m_1, m_2, \ldots$ with rectangular coordinates $x_1, y_1, z_1;\ x_2, y_2, z_2; \ldots$, with forces $X_1, Y_1, Z_1;\ X_2, Y_2, Z_2, \ldots$ acting on them parallel to the three axes. The free motion of the system occurs according to the equations

$$(1) \qquad m_i \frac{d^2 x_i}{dt^2} = X_i \ , \ m_i \frac{d^2 y_i}{dt^2} = Y_i, \ , \ m_i \frac{d^2 z_i}{dt^2} = Z_i.$$

This law may also be stated as follows: the accelerations are determined in such a way that

$$\sum m_i \left(\left(\frac{d^2 x_i}{dt^2} - \frac{X_i}{m_i} \right)^2 + \left(\frac{d^2 y_i}{dt^2} - \frac{Y_i}{m_i} \right)^2 + \left(\frac{d^2 z_i}{dt^2} - \frac{Z_i}{m_i} \right)^2 \right)$$

is a minimum. For this function of the accelerations takes its smallest value 0 when all accelerations obey the equations (1), that is, all the quantities $\frac{d^2x_i}{dt^2} - \frac{X_i}{m_i}$ vanish; and the function only takes a minimum value in this case. For if one of these quantities, for example $\frac{d^2x_i}{dt^2} - \frac{X_i}{m_i}$, were nonzero, one could always vary $\frac{d^2x_i}{dt^2}$ continuously in such a way that the absolute value of this quantity, and thus its square, decreases. The function would then become smaller if all other accelerations were left unchanged.

The above function of the accelerations differs from

$$\sum m_i \left(\left(\frac{d^2 x_i}{dt^2} \right)^2 + \left(\frac{d^2 y_i}{dt^2} \right)^2 + \left(\frac{d^2 z_i}{dt^2} \right)^2 \right)$$
$$-2 \sum \left(X_i \frac{d^2 x_i}{dt^2} + Y_i \frac{d^2 y_i}{dt^2} + Z_i \frac{d^2 z_i}{dt^2} \right)$$

only by a constant, that is, a quantity independent of the accelerations.

Suppose that the forces derive only from attractions and repulsions between the points that are functions of distance, and the i-th and i'-th points at distance r repel with force $f_{i,i'}(r)$ or attract with force $-f_{i,i'}(r)$. It is known that the components of the forces may be expressed via the partial derivatives of a function of the coordinates of all points,

$$P = \sum_{i,i'} F_{i,i'}(r_{i,i'}).$$

Here $F_{i,i'}(r)$ denotes a function whose derivative is $f_{i,i'}(r)$, and i, i' run over all pairs of distinct indices.

Let us substitute the values

$$X_i = \frac{\partial P}{\partial x_i}\ ,\ Y_i = \frac{\partial P}{\partial y_i}\ ,\ Z_i = \frac{\partial P}{\partial z_i}$$

into the above function of the accelerations, and multiply by $\frac{dt^2}{4}$, which leaves unchanged the position of its maxima and minima. We obtain an expression that differs from

$$\frac{1}{4}\sum\left(\left(d\,\frac{dx_i}{dt}\right)^2 + \left(d\,\frac{dy_i}{dt}\right)^2 + \left(d\,\frac{dz_i}{dt}\right)^2\right) - P_{(t+dt)}$$

only by a quantity that is independent of the accelerations. If the locations and velocities of the points are given at time t, the locations will be determined at time $t + dt$ in such a way that the above quantity is as small as possible. Accordingly a tendency exists to make this quantity as small as possible.

We can now interpret this law from actions that strive to make the individual terms of the expression as small as possible, if we suppose that conflicting forces balance out in such a way that the sum of the quantities, which the individual actions strive to minimize, gives a minimum.

Suppose that the point masses $m_1, m_2, \ldots, m_n$ are in the same ratios as the integers $k_1, k_2, \ldots, k_n$, so that $m_i = k_i\mu$. The expression to be minimized consists of the sum of quantities

$$\frac{\mu}{4}\left(\left(d\,\frac{dx_i}{dt}\right)^2 + \left(d\,\frac{dy_i}{dt}\right)^2 + \left(d\,\frac{dz_i}{dt}\right)^2\right)$$

over a system of particles μ, together with $-P_{t+dt}$. Following Gauss, we consider the quantity

$$\left(d\,\frac{dx_i}{dt}\right)^2 + \left(d\,\frac{dy_i}{dt}\right)^2 + \left(d\,\frac{dz_i}{dt}\right)^2$$

as a measure of the departure of the state of motion of the mass μ, at time $t + dt$, from the state at time t. Then the decomposition of the total action, with respect to each mass, gives an action that strives to minimize the departure of the state of motion at time $t+dt$ from that at time t, or a tendency toward maintaining the state of motion, together with an action striving to minimize $-P$.

The latter action may be decomposed into the tendency to maintain the individual terms of the sum $\sum_{i,i'} F_{i,i'}(r_{i,i'})$ as small as possible; that is, we consider the decomposition of the sum into attractions and repulsions between each pair of points. This would lead back to the customary explanation of the law of motion from the law of inertia and attraction and repulsion. However, all the natural forces with which we are familiar may also be reduced to forces acting between neighboring elements of space, as discussed in the section on gravitation below.

2. New mathematical principles of natural philosophy.[1]

Although the title of this section will scarcely awaken a favorable prejudice in most readers, it seems the best way to express my intention. Its object is to penetrate the inner nature of the foundations of astronomy and physics laid down by Galileo and Newton. For astronomy, this speculation admittedly has no direct practical use. However, I hope that this circumstance will not impair the interest of these pages in the eyes of the reader....

The basis of the general laws of motion of ponderable bodies assembled in the preamble to Newton's *Principia* lies in the intrinsic state of the bodies. Let us try to draw an analogy from our own inner perception. Mental images continually appear to us, which quickly vanish from our consciousness. We observe a continuous activity of our minds. Every mental act has something lasting underlying it, which announces itself as such on particular occasions (via the memory) without exerting a lasting influence on appearances. Thus continually (with each mental act), something lasting enters our minds, which however exerts no lasting influence in the world of phenomena. Each mental act is thus founded on something of a lasting nature, which enters our mind with this act, but at the same moment completely disappears from the world of phenomena.

Guided by these data, I make the hypothesis that space is filled with a substance which continually flows into ponderable atoms, and vanishes there from the world of phenomena, the corporeal world.

Both hypotheses may be replaced by a single one, that in all ponderable

[1]Found on March 1, 1853.

atoms, a substance perpetually appears from the corporeal world into the mental world. The cause for the vanishing of the substance there is to be looked for in the formation immediately beforehand of a mental substance there. Accordingly, ponderable bodies are the site where the mental world engages the physical world.[2]

The influence of general gravitation, which will now be interpreted in terms of this hypothesis, is known to be fully determined in each part of space if the potential function P of the system of all ponderable masses is given for this part of space. Equivalently, a function P of location should be given such that the ponderable mass in the interior of a closed surface S is $\frac{1}{4\pi}\int \frac{\partial P}{\partial p}\, dS$.

We now assume that the space-filling substance is a homogeneous incompressible fluid without inertia, and that in equal times, equal quantities of the substance flow into each ponderable atom, the quantity being proportional to its mass. Obviously the pressure undergone by the ponderable atom (is proportional to the velocity of motion of the substance at the location of the atom (?)).

Thus the influence of general gravitation on a ponderable atom can be expressed in terms of the pressure of the space-filling substance in the immediate neighborhood, and may be thought of as depending on this pressure.

From our hypothesis, it necessarily follows that the space-filling substance must transmit the waves which we perceive as light and heat.

Consider a simply polarized ray and denote by x the distance of an undetermined point of the ray from a fixed initial point. Let y denote the amplitude of the ray at time t. Since the velocity of propagation of the waves must, in all circumstances, be very nearly constant (say α) in space free of matter, the equation

$$y = f(x + \alpha t) + \phi(x - \alpha t)$$

must be fulfilled, or at least very nearly so.

If the equation holds exactly, then we must have

$$\frac{\partial y}{\partial t} = \alpha^2 \int^t \frac{\partial^2 y}{\partial x^2}\, d\tau.$$

[2]In each ponderable atom, a definite quantity of the substance enters in each moment, proportional to the force of gravitation, and vanishes there.

It is a consequence of the psychology based on Herbart's principles that it is not the mind, but the individual ideas that we form, that possess substance.

However, our empirical knowledge can also be satisfied via the equation

$$\frac{\partial y}{\partial t} = \alpha^2 \int^t \frac{\partial^2 y}{\partial x^2} \phi(t-\tau) d\tau,$$

with $\phi(t-\tau)$ not identically 1 for all positive values of $t-\tau$ (but decreasing with $t-\tau$ increasing to infinity). Here ϕ would only be significantly less than 1 for a sufficiently large time interval....

We express the location of points of the substance at a particular time t via a rectangular coordinate system. Let the coordinates of an undetermined point be $O : x, y, z$. Similarly let the coordinates of the point likewise with respect to a rectangular coordinate system, be $O' : x', y', z'$. Then x', y', z' are functions of x, y, z. Then x', y', z' are functions of x, y, z and $ds'^2 = dx'^2 + dy'^2 + dy'^2$ will be a quadratic form in dx, dy, dz. By a known theorem, we can uniquely determine linear forms in dx, dy, dz,

$$\begin{aligned} \alpha_1 dx + \beta_1 dy + \gamma_1 dz &= ds_1, \\ \alpha_2 dx + \beta_2 dy + \gamma_2 dz &= ds_2, \\ \alpha_3 dx + \beta_3 dy + \gamma_3 dz &= ds_3, \end{aligned}$$

so that

$$dx'^2 + dy'^2 + dz'^2 = G_1^2 ds_1^2 + G_2^2 ds_2^2 + G_3^2 ds_3^2,$$

while

$$ds^2 = dx^2 + dy^2 + dz^2 = ds_1^2 + ds_2^2 + ds_3^2.$$

The quantities $G_1 - 1$, $G_2 - 1$, $G_3 - 1$ are called the principal dilatations of the particle of substance at O in passing from the former to the latter form. I denote these quantities by $\lambda_1, \lambda_2, \lambda_3$.

I now assume that the difference between the earlier form of the particles of the substance, from the form at time t gives rise to a force that strives to alter this form. I assume that, other things being equal, the influence of an earlier form becomes weaker at further times from t, and the influence from earlier than a certain limit can be neglected. I further assume those earlier states that still exercise a noticeable influence differ little enough from those at time t that the dilatations may be treated as infinitely small. The forces that strive to reduce $\lambda_1, \lambda_2, \lambda_3$ can then be viewed as linear functions in $\lambda_1, \lambda_2, \lambda_3$. Now, on account of the homogeneity of the ether, the force which strives to reduce λ_1 must be a function of $\lambda_1, \lambda_2, \lambda_3$ invariant under interchange of λ_2, λ_3, and the other forces must arise from this one when λ_2, respectively λ_3, is exchanged with λ_1. We deduce the following expression for the total moment of this force:

$$\delta\lambda_1(a\lambda_1 + b\lambda_2 + b\lambda_3) + \delta\lambda_2(b\lambda_1 + a\lambda_2 + b\lambda_3) + \delta\lambda_3(b\lambda_1 + b\lambda_2 + a\lambda_3)$$

or, with a somewhat different notation for the constants,

$$\begin{aligned}\delta\lambda_1(a(\lambda_1 + \lambda_2 + \lambda_3) + b\lambda_1) &+ \delta\lambda_2(a(\lambda_1 + \lambda_2 + \lambda_3) + b\lambda_2)\\ &+ \delta\lambda_3(a(\lambda_1 + \lambda_2 + \lambda_3) + b\lambda_3)\\ &= \frac{1}{2}\,\delta(a(\lambda_1 + \lambda_2 + \lambda_3)^2 + b(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)).\end{aligned}$$

Now we can consider the moment of the force, that strives to vary the form of the infinitely small particle of substance at O, as the result of forces that strive to vary the lengths of line elements ending at O. We reach the following law of influence:

Let dV be the volume of an infinitely small particle of substance at O at time t and let dV' be the volume of the same particle of substance at time t'. Then the force striving to lengthen ds, arising from the difference in states of the substance, can be written

$$a\,\frac{dV - dV'}{dV} + b\,\frac{ds - ds'}{ds}.$$

The first part of this expression arises from the force with which a particle of substance resists alteration in volume without changing shape. The second arises from the force with which a physical line element resists a variation in length.

There are no grounds to assume that the two causative influences both vary with time according to the same law. If we take together the effects of all earlier forms of a particle of substance on the variation of the line element ds at time t, then the value of $\delta\,\frac{ds}{dt}$ that they seek to bring about is

$$\int_{-\infty}^{t} \frac{dV' - dV}{dV}\,\psi(t - t')\delta t' + \int_{-\infty}^{t} \frac{ds' - ds}{ds}\,\phi(t - t')\delta t'.$$

Now what must be the nature of the functions ψ and ϕ, in order that gravitation, light and radiant heat are conveyed through the spatial substance?

The influences of ponderable matter on ponderable matter are:

1) Attracting and repelling forces inversely proportional to the square of the distance.

2) Light and radiant heat.

Both classes of phenomena may be explained, if we suppose that the whole of infinite space is filled with a uniform substance, and each particle of substance acts only on its immediate neighborhood.

The mathematical law according to which this occurs can be considered as divided into

1) the resistance of a particle of substance to alteration in volume;

2) the resistance of a physical line element to alteration in length.

Gravitation and electrostatic attraction and repulsion are founded on the first part; propagation of light and heat, and electrodynamic or magnetic attraction and repulsion on the second.

3. Gravitation and light.

Newton's explanation of falling bodies and the motion of heavenly bodies consists in the assumption of the following cause:

1. There exists an infinite space with the properties ascribed to it by geometry, and ponderable bodies whose location in space may only change in continuous fashion.

2. In every ponderable body at each moment exists a cause determined in magnitude and direction, according to which the motion is determined. (Matter in a determinate state of motion.) The measure of this cause is the velocity.[3]

The phenomena to be explained here do not yet lead to the hypothesis of different masses of ponderable bodies.

3. At each point of space, at each moment, exists a cause determined in magnitude and direction (acceleration force) which transmits to any ponderable point found there a definite motion, the same for all such points. This motion combines geometrically with the existing motion of the body.

[3]Every material body, if considered to be alone in space, would either remain in the same location, or move through space with constant velocity in a straight line.

This law of motion cannot be explained from the principle of sufficient reason. For the body to continue in motion there must be a cause, which can only be sought in the inner state of matter.

4. At each point of a ponderable body exists a cause determined in magnitude (absolute force of gravity), giving an accelerative force at each point of space, inversely proportional to the square of the distance from this ponderable body and directly proportional to its gravitational force. This accelerative force combines geometrically with the other accelerative forces existing at each point.[4]

I seek the forces determined in magnitude and direction (accelerative forces) which are found at each point of space according to 3), in the form of motion of a substance, continuously spread throughout the whole of infinite space. Indeed, I assume that the direction of motion is the same as the direction of the force explained by this substance, and the magnitude of the force is proportional to its velocity. This substance may thus be represented as a physical space whose points move within it geometrically.

According to this hypothesis, all the effects, produced by each ponderable body on others across empty space, must be transmitted via this substance. Thus the forms of motion in which the light and heat occur, which heavenly bodies transmit to one another, are forms of motion of this substance. However, these two phenomena, gravitation and the motion of light through empty space, are the only phenomena that must simply be explained from the motion of this substance.

I now assume that the actual motion of the substance in empty space is a combination of the motion required to explain gravitation and the motion required to explain light.

The further development of this hypothesis falls into two parts, in so far as we seek

1. The laws of motion of the substance that must be assumed in order to explain phenomena.

2. The causes which could explain these motions.

The first task is a mathematical one; the second, a metaphysical one. Regarding the latter I remark in anticipation that the explanation of the cause

[4]The same ponderable point would undergo changes in motion at different points, whose direction coincides with the direction of the forces, of magnitudes proportional to the forces.

Hence the force, divided by the change in motion, yields the same quotient for the same ponderable point. This quotient, different for different ponderable points, is called the *mass* of the point.

striving to alter the distance between two particles of substance is not to be considered an objective. This method of explanation via attractive and repulsive forces owes its general application in physics, not to direct evidence (particular logical argument), nor—apart from electricity and gravity—to a particular ease of application. Rather, we should look to the circumstance that Newton's law of attraction has operated so long on the notions of researchers that they seek no further for explanations.[5]

I. Laws of motion of the substance, which on our hypotheses give rise to the phenomena of gravitation and light.

I express the location of a point of space in rectangular coordinates x_1, x_2, x_3. Let u_1, u_2, u_3 denote the components in these directions of velocity of the motion at time t that account for the phenomena of gravitation; w_1, w_2, w_3 for the corresponding components for the motion that account for the phenomena of light. Let v_1, v_2, v_3 denote the corresponding components of the actual motion, so that $v = u + w$. It will emerge from the laws of motion that the substance, which has uniform density everywhere at any given moment, has a density that is always the same. I assume that this density is 1 at time t.

a. The motion that gives rise to gravitational phenomena only.

The gravitational force is determined at each point via the potential function V, whose partial derivatives $\frac{\partial V}{\partial x_1}, \frac{\partial V}{\partial x_2}, \frac{\partial V}{\partial x_3}$ are the components of the gravitational force. This function V is further determined by the following condition (disregarding an additive constant):

1. $dx_1\, dx_2\, dx_3 \left(\frac{\partial^2 V}{\partial x_1^2} + \frac{\partial^2 V}{\partial x_2^2} + \frac{\partial^2 V}{\partial x_3^2} \right)$ vanishes outside the attracting body and has a constant value for every ponderable element of the body. This is the product of -4π and the absolute value of the attractive force that the attraction theory attributes to the element, and will be denoted by dm.

2. If all attractive forces are found within a finite space, then, at infinite

[5]Newton says "That gravity should be innate, inherent, and essential to matter, so that one body may act upon another at a distance through a vacuum, without the mediation of anything else, by and through which their action and force may be conveyed from one to another, is to me so great an absurdity, that I believe no man who has in philosophical matters a competent faculty of thinking can ever fall into it." See the third letter to Bentley.

distance r from a point of this finite space, $r\frac{\partial V}{\partial x_1}, r\frac{\partial V}{\partial x_2}, r\frac{\partial V}{\partial x_3}$ are infinitely small.

According to our hypotheses we now have $\frac{\partial V}{\partial x} = u$. Consequently

$$dV = u_1 dx_1 + u_2 dx_2 + u_3 dx_3.$$

This contains the conditions:

$$(1) \quad \frac{\partial u_2}{\partial x_3} - \frac{\partial u_3}{\partial x_2} = 0, \quad \frac{\partial u_3}{\partial x_1} - \frac{\partial u_1}{\partial x_3} = 0, \quad \frac{\partial u_1}{\partial x_2} - \frac{\partial u_2}{\partial x_1} = 0,$$

$$(2) \quad \left(\frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3}\right) dx_1 dx_2 dx_3 = -4\pi\, dm,$$

$$(3) \quad ru_1 = 0,\ ru_2 = 0,\ ru_3 = 0 \text{ for } r = \infty.$$

Conversely, if the quantities u satisfy these conditions, they are the components of the gravitational force. For the conditions (1) contain the existence of a function U with differential $dU = u_1 dx_1 + u_2 dx_2 + u_3 dx_3$, and thus the partial derivatives satisfy $\frac{\partial U}{\partial x} = u$. The other conditions then yield $U = V + \text{const.}$[6]

[6]This function U is thus obtained from empirical reality (from relative motion) by means of the general laws of motion, but only on disregarding a linear function of the coordinates, since we can only observe relative motion.

The determination of this function is based on the following mathematical theorem: a function V of location is determined up to a constant in a finite space if it is not discontinuous throughout any surface, and for all elements of this space,

$$\left(\frac{\partial^2 V}{\partial x_1^2} + \frac{\partial^2 V}{\partial x_2^2} + \frac{\partial^2 V}{\partial x_3^2}\right) dx_1\, dx_2\, dx_3$$

is given, while either V or its directional derivative along an inward normal is given on the boundary. In connection with this we note:

1. If the directional derivative at the boundary element ds is denoted by $\frac{\partial V}{\partial p}$, then in the latter case $\int \sum \frac{\partial^2 V}{\partial x^2} dx_1 dx_2 dx_3$, taken over the finite space, is equal to $-\int \frac{\partial V}{\partial p} ds$ taken over the boundary. Apart from this, the choice is arbitrary throughout the boundary in both cases. Accordingly, this choice is necessary to the determination.

2. For a space element where $\sum \frac{\partial^2 V}{\partial x^2}$ becomes infinitely large, we replace the product by $-\int \frac{\partial V}{\partial p} ds$ taken over the boundary of this element.

3. If $\sum \frac{\partial^2 V}{\partial x^2}$ only differs from 0 within a finite space, then the boundary condition may be replaced by the condition that, at an infinite distance R from a point of this space, $R\frac{\partial V}{\partial x}$ is infinitely small.

b. The motion that gives rise to light phenomena only.

The motion that must be assumed in empty space to explain light phenomena may be considered (in consequence of a theorem) as a combination of plane waves, that is motion such that along each plane of a family of parallel planes (wave planes) the form of motion is constant. Each of these wave systems then (according to empirical reality) consists of motions parallel to the wave plane, which propagate for all forms of motion (type of light) with the same constant velocity c perpendicular to the wave plane.

For such a wave system, let ξ_1, ξ_2, ξ_3 be rectangular coordinates of a point of space, the first perpendicular and the others parallel to the wave plane. Let $\omega_1, \omega_2, \omega_3$ be the velocity components in the same directions at this point, at time t. Then we have

$$\frac{\partial \omega}{\partial \xi_2} = 0, \quad \frac{\partial \omega}{\partial \xi_3} = 0.$$

Empirically, we have firstly

$$\omega_1 = 0,$$

and secondly, the motion is the combination of a motion advancing on the positive side, and one on the negative side, of the wave plane, with velocity c. let ω' be the velocity components of the former, and ω'' those of the latter. Then ω' is invariant when t increases by dt and ξ_1 increases by cdt; ω'' is invariant when t increases by dt and ξ_1 by $-cdt$; and we have $\omega = \omega' + \omega''$. It follows that

$$\left(\frac{\partial \omega'}{\partial t} + c\frac{\partial \omega'}{\partial \xi_1}\right) dt = 0, \quad \left(\frac{\partial \omega''}{\partial t} - c\frac{\partial \omega''}{\partial \xi_1}\right) dt = 0,$$

$$\frac{\partial^2 \omega'}{\partial t^2} = -c\frac{\partial^2 \omega'}{\partial \xi_1 \partial t} = c^2\frac{\partial^2 \omega'}{\partial \xi_1^2}, \quad \frac{\partial^2 \omega''}{\partial t^2} = c\frac{\partial^2 \omega''}{\partial \xi_1 \partial t} = c^2\frac{\partial^2 \omega''}{\partial \xi_1^2},$$

and thus

$$\frac{\partial^2 \omega}{\partial t^2} = c^2\frac{\partial^2 \omega}{\partial \xi_1^2}.$$

These equations yield the following symmetric equations:

$$\frac{\partial \omega_1}{\partial \xi_1} + \frac{\partial \omega_2}{\partial \xi_2} + \frac{\partial \omega_3}{\partial \xi_3} = 0,$$

$$\frac{\partial^2 \omega}{\partial t^2} = c^2\left(\frac{\partial^2 \omega}{\partial \xi_1^2} + \frac{\partial^2 \omega}{\partial \xi_2^2} + \frac{\partial^2 \omega}{\partial \xi_3^2}\right).$$

Expressed in terms of the original coordinates, these become equations of the same form:

$$\frac{\partial w_1}{\partial x_1}+\frac{\partial w_2}{\partial x_2}+\frac{\partial w_3}{\partial x_3}=0, \tag{1}$$

$$\frac{\partial^2 w}{\partial t^2}=c^2\left(\frac{\partial^2 w}{\partial x_1^2}+\frac{\partial^2 w}{\partial x_2^2}+\frac{\partial^2 w}{\partial x_3^2}\right). \tag{2}$$

These equations are valid for every plane wave passing through the point (x_1, x_2, x_3) at time t, and consequently for the motion composed of all these waves.

c. The motion that gives rise to both types of phenomena. From the conditions found for u and w flow the following conditions for v, or laws of motion of the substance in empty space:

$$\frac{\partial v_1}{\partial x_1}+\frac{\partial v_2}{\partial x_2}+\frac{\partial v_3}{\partial x_3}=0, \tag{I}$$

$$\begin{cases}(\partial_t^2-c^2(\partial_{x_1}^2+\partial_{x_2}^2+\partial_{x_3}^2))\left(\dfrac{\partial v_2}{\partial x_3}-\dfrac{\partial v_3}{\partial x_2}\right)=0,\\[2ex] (\partial_t^2-c^2(\partial_{x_1}^2+\partial_{x_2}^2+\partial_{x_3}^2))\left(\dfrac{\partial v_3}{\partial x_1}-\dfrac{\partial v_1}{\partial x_3}\right)=0,\\[2ex] (\partial_t^2-c^2(\partial_{x_1}^2+\partial_{x_2}^2+\partial_{x_3}^2))\left(\dfrac{\partial v_1}{\partial x_2}-\dfrac{\partial v_2}{\partial x_1}\right)=0.\end{cases} \tag{II}$$

This is easily obtained on carrying out the differentiations.

These equations show that the motion of a point of the substance depends only on the motions in adjacent parts of space and time, and their (complete) causation may be sought in neighboring influences.

Equation (I) confirms our earlier claim that density is invariant under the motion of the substance. For

$$\left(\frac{\partial v_1}{\partial x_1}+\frac{\partial v_2}{\partial x_2}+\frac{\partial v_3}{\partial x_3}\right)dx_1\,dx_2\,dx_3\,dt,$$

which according to this equation is 0, expresses the quantity of the substance flowing into the space element $dx_1\,dx_2\,dx_3$ in the time element dt. Accordingly the quantity of the substance in that element remains constant.

The conditions (II) are identical with the conditions that

$$(\partial_t^2 - c^2(\partial_{x_1}^2 + \partial_{x_2}^2 + \partial_{x_3}^2))(v_1 dx_1 + v_2 dx_2 + v_3 dx_3)$$

is a complete differential dW. Now

$$(\partial_t^2 - c^2(\partial_{x_1}^2 + \partial_{x_2}^2 + \partial_{x_3}^2))(w_1 dx_1 + w_2 dx_2 + w_3 dx_3) = 0$$

and consequently

$$\begin{aligned} dW &= (\partial_t^2 - c^2(\partial_{x_1}^2 + \partial_{x_2}^2 + \partial_{x_3}^2))(u_1 dx_1 + u_2 dx_2 + u_3 dx_3) \\ &= (\partial_t^2 - c^2(\partial_{x_1}^2 + \partial_{x_2}^2 + \partial_{x_3}^2))dV \\ &= d\,\frac{\partial^2 V}{\partial t^2} \end{aligned}$$

since $(\partial_{x_1}^2 + \partial_{x_2}^2 + \partial_{x_3}^2)dV = 0$.

d. Common expression for the laws of motion of the substance and the influence of gravitational force on the motion of ponderable bodies.

The laws of these phenomena may be combined in the condition that the variation of the integral

$$\begin{aligned} \frac{1}{2}\int\Bigg\{\sum\left(\frac{\partial\eta_i}{\partial t}\right)^2 - c^2\Bigg[\left(\frac{\partial\eta_2}{\partial x_3} - \frac{\partial\eta_3}{\partial x_2}\right)^2 + \left(\frac{\partial\eta_3}{\partial x_1} - \frac{\partial\eta_1}{\partial x_3}\right)^2 \\ + \left(\frac{\partial\eta_1}{\partial x_2} - \frac{\partial\eta_2}{\partial x_2}\right)^2\Bigg]\Bigg\} dx_1\, dx_2\, dx_3\, dt \\ + \int V\left(\sum \frac{\partial\eta_i}{\partial x_i \partial t}\, dx_1\, dx_2\, dx_3 + 4\pi dm\right) dt + 2\pi\int dm \sum\left(\frac{\partial x_i}{\partial t}\right)^2 dt \end{aligned}$$

becomes 0 under suitable boundary conditions.

In this expression, the first two integrals are taken over the whole geometrical space, the last over all elements of ponderable bodies. However, the coordinates of each ponderable element must be determined as functions of time, and $\eta_1, \eta_2, \eta_3, V$ as functions of x_1, x_2, x_3 and t, in such a way that a variation of these functions satisfying the boundary conditions generates only a second order variation of the integral.

In this case, the quantities $\frac{\partial\eta}{\partial t}\,(= v)$ are the velocity components of the substance, and V is the potential at the point (x_1, x_2, x_3) at time t.

The life of Bernhard Riemann.

R. Dedekind

The following account of Bernhard Riemann's life is not in any way intended to assess the significance of his contribution to science or to shed any light on its relation to earlier mathematics, or the present state of mathematics. Its purpose is rather to provide for those readers who would like to know something about the upbringing, character, external circumstances, and fate of this great mathematician, whose complete works are now for the first time being published in a collected edition.

Georg Friedrich Bernhard Riemann was born on 17th September 1826 in Breselenz, a village in the kingdom of Hanover near Dannenberg, close to the river Elbe. His father, Friedrich Bernhard Riemann, born in Boitzenburg on the Elbe in Mecklenburg, who had taken part as a lieutenant under Wallmoden in the war of liberation, was the local pastor and was married to Charlotte, the daughter of Privy Councillor Ebell of Hanover. He later moved, with his family, to the parish of Quickborn, at that time a three-hour journey away. Bernhard was the second of six children. His eagerness to learn was awakened very early by his father who taught him, almost single-handed, until Bernhard's departure to the Gymnasium. As a five-year-old he was very interested in history, in the stories of antiquity, and more especially in the sad fate of Poland, which his father had to relate to him over and over again. Very soon afterwards, however, all this faded into the background and his remarkable talent for calculation broke out; there was nothing he liked better than to discover for himself hard problems and then to set them as exercises to his brothers and sisters. Later on, when Bernhard was ten years old his father arranged for a tutor, Schulz, to help him with his son's education. Schulz gave Bernhard a good grounding in arithmetic and geometry, but he soon found that he was hard put to it to follow his pupil's solutions to the problems which he set, as they were often better and faster than his own.

At the age of thirteen and a half, Bernhard was confirmed by his father and left the parental home, in which a serious and pious domestic atmosphere prevailed. The parents regarded the proper education of their children as their main duty; and the heartfelt love which bound Riemann to his family was maintained throughout the remainder of his life. It shows itself in the letters he wrote from afar to those who were so dear to him. He displays a vivid interest in everything to do with the parental home, even in the smallest incidents, and lets the members of his family truly share his joys and tribulations.

At Easter 1840 Riemann came to Hanover, where his grandmother lived and where, until her death some two years later, he attended the third form of the local Lyceum. Initially, as was to be expected from his previous education, he had to overcome all kinds of difficulties, but it was not long before he was being praised for his progress in the various subjects taught. He was always a diligent and obedient pupil. Numerous letters from Riemann to his beloved parents and to his brothers and sisters, dating from this period, have survived, in which he reports incidents from his school life, often with an amusing sense of humor. But the predominant theme is a yearning for home; as the holiday-time approaches he constantly asks for permission to spend it at home in Quickborn, and long beforehand he tries to devise means for the journey to be arranged at the least possible cost; he makes small purchases of birthday presents for his parents and brothers and sisters, and is eager to make them a real surprise. In his thoughts he is still living at home with his family. At times, however, one can sense in his letters that he finds conversation with strangers very difficult, and he wistfully laments his shyness. This shyness, a natural consequence of his sheltered upbringing, which to his sorrow even caused him to appear in a false light to his teachers, never completely left him. Indeed, this often drove him to take refuge in solitude and in the world of his imagination, where he displayed the greatest boldness and openmindedness.

After the death of his grandmother Riemann, apparently at his own wish, was taken at Easter 1842, by his father, to the Johanneum in Luneburg, where he remained for two years in the second form and two years in the top form until his departure for the university. The great fire of Hamburg occurred during the first part of this period and it made a profound impression on Riemann, who reported on it in great detail to his parents. The greater nearness to home and the possibility of spending his holidays in Quickborn with his family contributed to making this further time at school a happy one. Admittedly the journeys to and fro, for the most part on foot, required considerable effort, as his bodily growth left much to be desired; even at this time, in the beautiful letters to his mother—whose loss he would soon, alas, have to mourn—there are references to anxieties over his health, and to her repeated kindly warnings to avoid undue physical exertions. Later on, he lived with the schoolmaster Seffer, who took an active interest in the lad, and who, as is clear from Riemann's letters, became a fatherly friend and protector. He had good school reports in other subjects, but in mathematics they were always brilliant, and on leaving school he was the top pupil. His outstanding talent for this science was recognised by the excellent headmaster Schmalfuss, who lent him mathematical works for private study, and was often surprised or rather absolutely astonished when Riemann returned the books to him after a few days and then showed in conversation how completely he had worked through and mastered their content. These studies, far outside the school curriculum, must have led him into the realms of higher mathematics. His knowledge of advanced analysis, as far as one can tell, was obtained by studying the works of Euler; but he must also have read Legendre's *Théorie des Nombres* at this time.

At the age of nineteen and a half, at Easter 1846, Riemann entered the University of Göttingen. His father, as a minister of religion, had naturally cherished the hope that his son would devote himself to theology, and in fact Riemann on the 25th April matriculated as a student of philology and theology. The decision to do so, which seems at complete variance with his own inclinations and his gift for mathematics, must above all have been attributable to his concern for his family, and to his hope that he might be able to help his impoverished father with his numerous motherless children, by an early appointment to some living in the church. However, in addition to attending the lectures on philology and theology, he also heard the mathematical ones, and in fact already in the summer semester he attended those given by Stern, on the numerical solution of equations, and those on terrestrial magnetism by Goldschmidt; and then, in the winter semester of 1846–1847, those of Gauss on the method of least squares, and those of Stern on definite integrals. With this continued preoccupation with mathematics, he soon saw that his attraction to the subject was too strong within him to be denied, and he obtained his father's permission to devote his time entirely to his favourite study.

Although Gauss had now, for nearly half a century, held the rank of the greatest mathematician alive, he confined his lecturing activities—which were admittedly highly stimulating as far as they went—to a narrow field, belonging more to applied mathematics. Riemann, with his existing knowledge of mathematics, could not therefore expect any great enrichment of his store of knowledge in Göttingen, or to learn anything which might bear fruit in the form of new ideas. He therefore at Easter 1847 transferred to the University of Berlin, where Jacobi, Lejeune Dirichlet and Steiner were attracting large numbers of students by the brilliance of their discoveries, which formed the subject-matter of their lectures. He stayed there for two years, until Easter 1849. Among the lectures which he attended were those of Dirichlet, on the theory of numbers, on the theory of definite integrals, and on partial differential equations, and those of Jacobi on analytical mechanics and on higher algebra. Unfortunately very few of his letters from this time have survived; in one of them (dated 29th November 1847) he writes of his great joy on learning that Jacobi had changed his mind and had decided, after all, to lecture on Mechanics. He came into close contact with Eisenstein, whose lectures on elliptic functions he attended during his first year. Riemann later recounted that they had discussed between them the question of how complex numbers should be introduced into the theory of functions, and that they held radically different opinions on the basic principles. Eisenstein remained firmly attached to the formal algorithmic approach, whereas he had felt that the true definition of a function of a complex variable was to be found in the partial differential equation. It is probable that it was in the autumn vacation of 1847 that Riemann first worked out the basic principles of these ideas, which were to prove so important during the remainder of his life.

Very little is to be gleaned from his letters about other aspects of his life during his two-year stay in Berlin. He was deeply stirred by the great political events of 1848; he was an eyewitness of the March revolution and had, as a member of the student corps, been on guard duty at the Royal Palace from 9 a.m. on the 24th March until 1 p.m. on the following day.

At Easter 1849, after he had experienced the arrival in Berlin of the Kaiser's deputation from Frankfurt, he returned to Göttingen. During the three following semesters he attended some courses of lectures on scientific subjects and on philosophy, among others a brilliant course on experimental physics given by Wilhelm Weber which he followed with great interest. He later became very attached to Weber, who became a true friend and counsellor until his death. It must have been at this time, when he was simultaneously occupied in studying philosophy—especially the teachings of Herbart—that his ideas on natural philosophy first began to germinate. At least, as far as his striving to obtain an unified conception of Nature is concerned, this seems to be so, to judge from a passage in an essay which he wrote in November 1850. This memoir, drawn up in his capacity as a member of a pedagogical seminar, is entitled "On the scope, arrangement and methods of teaching scientific subjects in schools." In it he says "Thus, for example, it is possible to present a mathematical theory complete in itself, proceeding from the elementary laws valid for individual points to the events in the actual space or continuum, without having to distinguish between whether we are concerned with gravitation, electricity, magnetism, or heat and its equilibrium." In autumn 1850 he joined the recently formed seminar on mathematical physics led by Professors Weber, Ulrich, Stern, and Listing, and took part in the experimental work, even though it often took him away from his primary objective, the preparation of his doctoral dissertation. It was partly for this reason, and partly also because of the scrupulous attention which he paid to the preparation of any work intended for publication—a scrupulousness which was to have the effect of seriously restricting the publication of his work later on—that it was not until November of the following year, 1851, that he was able to submit his thesis "Foundations for a general theory of functions of a complex variable." This doctoral dissertation received a very favourable judgment from Gauss, who informed Riemann at his interview that he had been for several years preparing a paper dealing with the same subject, but not restricted to it. The examination took place on Wednesday, 3rd December, and the public disputation and ceremony of promotion to the degree of Doctor on Tuesday, 16th December. To his father he writes "Now that I have completed my dissertation, I think that my prospects have significantly improved. I also hope that with time I shall learn to write more fluently and more rapidly, particularly if I try to meet more people socially, and also I have the opportunity for the first time to give lectures; I am now in good spirits." At the same time he blames himself for the additional costs which he has imposed on his father by not showing more eagerness in applying for the vacancy at the Observatory[7] occasioned by the death of Goldschmidt. He goes on to say that nothing stands in the way of his habilitation as a Privatdocent once his habilitation thesis has been completed. It appears that he had from early on intended to choose as the subject of this thesis the theory of trigonometric series, but two and a half years were to elapse before his habilitation. In the autumn vacation of 1852, Lejeune Dirichlet, whom Riemann knew well from Berlin, came to spend

[7]From information given to me by W. Weber, it seems that Gauss himself did not wish Riemann to be given this appointment. It goes without saying that he did not doubt that Riemann had the necessary theoretical knowledge and practical aptitude for the post; but he already at that time held such a high opinion of Riemann's scientific work that he feared that the time-consuming duties associated with the position, and to some extent the subordinate official business, would divert him far too much from his true work.

some time in Göttingen. Riemann, who had just returned from Quickborn, was lucky enough to see him almost daily. On his first visit to the Crown, where Dirichlet was staying, and next day at a noon meeting at the house of Sartorius von Waltershausen at which Professor Dove, from Berlin, and Professor Listing were present, he asked Dirichlet, whom he recognised to be, after Gauss, the greatest mathematician then alive, for advice about his work. "The other morning"—Riemann writes to his father—"Dirichlet was with me for about two hours; he gave me some notes which I needed for my habilitation thesis, which were so comprehensive that it will significantly lighten my work; otherwise I might have had to spend a lot of time in the library searching for all kinds of things. He also went through my dissertation with me and, in general, was extremely friendly towards me, which I scarcely expected because of the great gap in standing. I have hopes that he will not forget me later." A few days later Wilhelm Weber returned to Göttingen from the Wiesbaden Association for the Advancement of Science. There was a very rewarding excursion, in a large group, to the Hohen Hagen a few hours off, and on the following day Dirichlet and Riemann once again met at Weber's house. Personal encounters of this kind were highly beneficial to Riemann. He wrote about this to his father "You see that I have not been spending all my time at home, but even so I am hard at work in the morning and I find that I am making better progress than if I had spent the whole day at my books."

In those days he also wrote about his habilitation and the commencement of his lecturing as though they were something in the immediate future, and he would undoubtedly have progressed more rapidly in his career if more of such outside stimuli had come his way. It is clear that at the beginning of 1853 he was almost exclusively occupied with natural philosophy; his new ideas take on a fixed shape to which he keeps coming back after all interruptions. Finally, however, the habilitation thesis gets finished, and he writes to his younger brother Wilhelm on 28th December 1853: "My work is now in a reasonable state; I handed in my habilitation thesis[8] at the beginning of December, and I must now propose three subjects for the trial lecture, one of which is then chosen by the faculty. I had already prepared the first two, and I had hoped that one of these would be chosen, but Gauss chose the third[9], and so I am again in something of a tight spot, as I now have

[8] On the representation of a function by a trigonometric series.

[9] On the hypotheses which underlie geometry.

to do some more work on it. I had resumed my other investigations into the connection between electricity, galvanism, light and gravity immediately after finishing my habilitation thesis, and have now made sufficient progress in them, that I can publish my conclusions in this form without hesitation. At the same time I have become more and more convinced that Gauss has been working on this question for several years, and has told a few of his friends—among others Weber—under the seal of secrecy. Of course I can say this to you, without any risk of appearing presumptuous; anyhow I hope that I am not too late, and that it will be recognised that my results were entirely independent work."

At this time, Riemann was acting as assistant to Wilhelm Weber in the seminar on mathematical physics, and as such had to supervise the work of the new entrants, as well as to give a few lectures. On 26th June 1854, he writes from Quickborn to his brother about the further progress of his work: "Around Christmas, if I remember rightly, I wrote to you from Göttingen, telling you that I had completed my habilitation thesis at the beginning of December, and handed it in to the Dean, and that immediately afterwards I had continued with my researches into the interrelations between the fundamental laws of physics and had become so engrossed in this study that when the subject was set for the trial lecture at the colloquium, I could not tear myself away from it. Soon after this I became ill, partly as a result of too much brooding, and partly because of sitting in front of a stove during this foul weather; my old trouble started again with a vengeance, which was not without adverse effect on my work. It was not until several weeks later, when the weather improved and I was able to get about a bit, that my health got better. I have now rented a summer-house for the summer and since then, thank God, have had nothing to complain of as regards my health. For about a fortnight after Easter I had to deal with other work which I could not very well avoid. That done, I went ahead enthusiastically with the preparation of my trial lecture and had completed it by Whitsun. Although it was quite an effort, I needed to have it done so that I could have my examination immediately, and would not have to visit Quickborn without having achieved my object. In fact, Gauss's state of health has recently become so bad, that everyone now fears that he will die this year, and he himself felt that he was too weak to examine me. He now asked me to wait—since in any case I could not start lecturing until the next semester—at least until August, when he might be feeling better. I had already resolved to bow to the inevitable. Then suddenly at midday on the Friday after Whitsun he decided to give in to my repeated pleas to get rid of "the halter round my neck" by agreeing to hold the examination at half past ten on the following day, and so by one o'clock on the Saturday it was all happily ended. Let me now tell you, in all haste, what this other business which took up my time at Easter was all about. During the Easter vacation Kohlrausch—a son of the Inspector of Schools, and a cousin and brother-in-law of Schmalfuss—who is now a professor in Marburg, was on a fortnight's visit at Weber's, to carry out some electrical experiments with him. Weber had been responsible for part of the research, and Kohlrausch on the other hand had done the preliminary work, and had designed and constructed the apparatus. I had participated in the experimental side and had got to know Kohlrausch on this occasion. Now Kohlrausch had some time previously made and published some very precise measurements in connection with an electrical phenomenon not hitherto investigated (the residual charge in a Leyden jar), and, in the course of my general research into the connection between electricity, light, and magnetism, I had found the explanation. I spoke to Kohlrausch about this, and this gave me the opportunity to work out the details of the theory of this phenomenon, which I sent to him. Kohlrausch replied in a most friendly manner, and offered to send my work to Poggendorf, the editor of the Annals of Physics and Chemistry in Berlin, for publication in that journal. He also invited me to visit him during the next autumn vacation, to pursue the matter. This is very important to me because it is the first time I could apply my theories to a hitherto unknown phenomenon, and I hope that the publication of this small work will contribute to my larger work getting a favourable reception. Here in Quickborn I may well have to spend part of my time on preparing it for printing, as the page proofs will probably be sent to me, and part on working out a course of lectures for next semester."

As regards the first part of this letter, it should be noted that Riemann, in composing his trial lecture on the hypotheses of geometry had made his task very much harder by striving to ensure that it would be as easily understandable as possible even to members of the faculty who had had no mathematical training. By doing so, however, he created a most admirable masterpiece of exposition, inasmuch as while suppressing all the detailed mathematical analysis, he nevertheless succeeds in conveying so accurately the train of his thoughts, that it can be completely reconstructed from the indications which he gave. Gauss, contrary to the usual practice, had selected from the three proposed themes not the first but the third, because he was keen to know how such a difficult subject would be handled by such a young man. The lecture, which exceeded all his expectations, greatly astonished him, and on his way back from the meeting of the faculty he spoke to Wilhem Weber enthusiastically, and with quite uncharacteristic excitement, about the profundity of the ideas which Riemann had put forward.

After a fairly long stay in Quickborn, Riemann returned to Göttingen in September to take part in a meeting of the Association for Scientific Research. At the invitation of Weber and Stern he decided to give a lecture on the distribution of electricity in non-conductors to the section on mathematics, physics and astronomy. He writes to his father about this: "My turn to give a lecture came on Thursday, and as nothing else had been announced for this session of our section, I worked on it the previous evening so as to fill up to a certain extent the usual period of the session. I had initially wanted only to lecture on the physical law which I wished to tell them about, but I now applied it to several different phenomena and showed how it agreed with experience. My presentation in this latter part was admittedly somewhat disjointed, but I still think that the impression of the whole gained by the addition;I spoke for about an hour and a quarter. The fact that I had spoken in public to this particular gathering before gave me a bit more courage to give my lecture, but I now see that there is a vast difference between, on the one hand, having thought about a subject for a long time and having sorted everything out, and on the other having only just prepared the subject-matter immediately before the lecture. I hope that in six months time I shall be able to think more calmly about my lectures, and that the thought of them will not again haunt my stay in Quickborn, letting it spoil the pleasure of being with you, as it did last time." He met Kohlrausch again in Göttingen, but after a further exchange of letters Riemann decided not to go ahead with the publication of his paper on the residual charge in Leyden jars, presumably because of his unwillingness to agree to certain proposed alterations. Instead there appeared in *Poggendorf's Annalen* Riemann's memoir on the theory of Nobili's color rings, concerning which he writes to his elder sister Ida: "This matter is of particular importance because it lends itself to very accurate measurements from which the laws governing the flow of electricity can be verified with great precision."

In this same letter of 9th October 1854 he writes with great delight of his achievement in attracting, against all his expectations, a considerable audience to his first course of lectures, as many as eight students having enrolled. The subject of the lectures was the theory of partial differential equations and its applications to problems of physics; they were modeled, in the main, on the course of lectures which Dirichlet had given, under the same title, in Berlin. On 18th November 1854 he writes to his father: "My life here has gradually settled down into a fairly regular routine. I have so far managed to hold my lectures regularly, and my initial awkwardness and diffidence have almost disappeared. I have trained myself to think more about my listeners and less about myself, and I have learnt to read from their faces whether I ought to keep going or explain the subject in more detail." There can be little doubt that communication by word of mouth always caused him great difficulties in the first years of his academic activities. His dazzling power of thought and his fantastic premonitions often led him, especially in conversations on scientific subjects, to take huge steps which others could not easily follow, and if one asked him for more detailed explanations of some of the intermediate steps leading to his conclusions, he would be taken aback. With some difficulty, he would adjust his style to the slower thought processes of others and rapidly dispel their doubts. His sensitive nature was often upset, when during his lectures he could see from the looks on the faces of his audience—to which he refers in the letter quoted above—that he would be obliged to give a special proof of some point which seemed to him to be quite self-evident. After long practice he eventually got over this, and the comparatively large number of his pupils bears witness not only to his fame as the author of the most profound work, but to the attraction of his lectures, which were prepared with the greatest care. He succeeded in guiding his listeners through great difficulties in order to penetrate the new principles he had created.

On 23rd February 1855, Gauss died and soon afterwards Lejeune Dirichlet was called from Berlin to Göttingen. On this occasion attempts were made from several sides to get Riemann appointed as an extraordinary professor, but they came to naught. However, it was arranged that he would receive an annual stipend of 200 thalers from the government, and although the amount was so small it came as a considerable relief to Riemann, who was then, and for some time to come, often gloomy about the future. It began a series of sad and somber years, in which fate dealt him one heavy blow after another. In 1855 he lost his father and one of his sisters, Clara. His old home in Quickborn that he loved so much had to be abandoned, and his three sisters went to live with his brother Wilhelm in Bremen, where he was Secretary for Postal Services. From then on, Wilhelm took over the responsibility for the support of the family.

Riemann had now returned with renewed eagerness to his researches into the theory of Abelian functions, which he had begun in the years 1851 and 1852 and for the first time, from Michaelmas 1855 until Michaelmas 1856, took as the subject of his lectures, in which three listeners, Schering, Bjerknes, and his colleague Dedekind took part. In summer 1856, he was appointed assessor of the mathematical section of the Göttingen Scientific Society; and as such he presented to the society, on 2nd November, a paper on Gauss's hypergeometric series. On the same day he wrote to his brother: "I also hope that my work will bear fruit. My paper is, as I wrote to you before, now ready to be printed, and perhaps the Society will allow it to be published in their transactions; a great honor, because during the last 50 years the only mathematical papers which they have contained have been those by Gauss. The mathematical section of the society, consisting of Weber, Ulrich, and Dirichlet, will, in Weber's opinion, at any rate, make an application for my paper to be printed. I am fairly happy with my lectures, that is to say with the attendance, especially in view of the small number of new student arrivals. There are hardly any mathematicians among them, and that is probably why Dedekind and Westphal have been unable to get the number of pupils they wanted for private tuition. The number attending my lectures on the four days I lectured was three to begin with, then four, and on the last two occasions five, but one of these was just sitting in on the course. I am particularly gratified that this time I also have a few first-year students and not simply, as in the past, students from the sixth and later semesters, because it is a sign that my lectures have become easier to understand. For all that, I cannot yet pretend that my lectures are all they should be, because no-one has yet signed up, and it is still possible that my listeners may leave me in the lurch. I shall devote my free time to my work on Abelian functions, which I have told you about. Shortly before my return here to Göttingen, Dr. Borchardt of Berlin, the principal editor of the mathematical journal, was here and has passed on through Dirichlet and Dedekind a request to send him, as soon as possible, an account of my researches into Abelian functions, in as rough and ready a state as I like. Weierstrass is now busily publishing, but the latest volume of the journal, which Scherk has told me about, contains only the preliminary indications of his theory."

In fact Riemann now devoted all his strength to working out the details of this work, so that he was able to send the manuscript of the first three smaller articles to Berlin on 18th May, and the larger fourth on 2nd July 1857. The excessive effort involved severely affected his health and he reached by the end of the summer semester a state of mental exhaustion very much darkened by his temperament. To refresh his spirits and improve his health, he took a few weeks leave in Harzburg, accompanied for a few days by his friend Ritter (then a teacher at the Polytechnic in Hanover, and now a Professor at Aachen). His colleague Dedekind followed him thither, and together they went for numerous walks and longer excursions into the Harz region. These walks raised his spirits, and began to restore his trust in others and his self-confidence; his harmless jests and highly entertaining conversations on scientific subjects made him one of the most kindly and stimulating companions. At this time his thoughts kept returning to natural philosophy, and one evening, after coming back from a strenuous walk, he took hold of Brewster's Life of Newton, and spoke at length with admiration about the letter to Bentley in which Newton himself states the impossibility of direct action at a distance.

Soon after his return to Göttingen, he was on 9th November promoted to extraordinary professor, and his stipend increased from 200 to 300 thalers a year. Almost at the same time his happiness was shattered by the death of his beloved brother Wilhelm; he now assumed responsibility for the upkeep of his three remaining sisters and insisted that they should come to Göttingen and stay with him during that winter; this happened again at the beginning of March 1858, but shortly afterwards his youngest sister, Marie, was taken away from him by death. After so many cruel strokes of fate, the company of his two surviving sisters consoled him in the depths of his depression, and the recognition which his work was, albeit slowly, gaining in wider circles, gradually restored his sunken self-confidence and gave him fresh courage to undertake new work. Earlier, he had written an essay entitled "A contribution to electrodynamics" which was later much talked about, and had written to his sister Ida: "I have submitted to the Royal Society [of Göttingen] my discovery on the connection between electricity and light. From various remarks which I have picked up on this subject, I can only conclude that Gauss has put forward a theory on this matter which is different from mine, and has communicated it to his closest friends. But I am convinced that mine is the right one, and that in a few years time it will be recognised as such." It is known that he soon afterwards withdrew his paper, and never resubmitted it , probably because he was no longer satisfied with the deductions it contained.

In the autumn vacation of 1858 he got to know the Italian mathematicians Brioschi, Betti and Casorati who were on a tour of Germany, and were spending a few days in Göttingen; this acquaintance was to be renewed later in Italy.

At this time, Dirichlet fell ill and succumbed on 5th May 1859 after a long period of suffering. He had always taken the liveliest personal interest in Riemann and had taken every opportunity to do what he could to improve his external relationships. Meanwhile Riemann's scientific reputation had become so generally recognized that, after Dirichlet's death, the authorities saw no need to look for a mathematician from outside. At Easter 1859, living quarters were found for him in the Observatory; on 30th July he was appointed ordinary professor, and in December unanimously elected as an ordinary member of the Royal Society of Sciences in Göttingen. Earlier, on 11th August, the Berlin Academy of Sciences had elected him as a corresponding member in the physics and mathematics section. This gave him the opportunity to travel to Berlin in September, accompanied by Dedekind, where he met the leading mathematicians Kummer, Borchardt, Kronecker, and Weierstrass, who all welcomed him with honor and great cordiality. One of the consequences of his nomination and visit (which was followed by his election as a foreign member[10] in March 1866) was that he submitted to the Academy in October his paper on the frequency of prime numbers, and sent Weierstrass a letter on multiply periodic functions, which was published posthumously.

One month later he submitted to the Göttingen Society of Sciences his memoir on the propagation of planar air-waves of finite amplitude.

In the Easter vacation of 1860 he went on a trip to Paris, where he stayed for a month from 26th March; unfortunately the weather was raw and unfriendly and in the last week of his visit there was one day after another of snow and hail which made it almost impossible to see the sights. However, he was delighted with the friendly reception which he received from the Parisian scholars Serret, Bertrand, Hermite, Puiseux and Briot, with whom he spent a pleasant day in the country at Chatenay, along with Bouquet.

In the same year, he completed his paper on the motion of a fluid ellipsoid and turned his attention to a prize problem proposed by the Paris Academy. This concerned the conduction of heat, and Riemann had already done the groundwork in his investigations into the hypotheses of geometry. In June

[10]Concerning the honors conferred on Riemann, I note here that he was elected, on 28th November 1859,a corresponding member of the Bavarian Academy of Sciences, then an ordinary member on 28th November 1863. He became a corresponding member of the Paris Academy on 19th March 1866. On 14th June 1866, shortly before his death, he was elected a foreign member of the Royal Society of London.

1861, he sent in his solution, which was written in Latin, with the motto "Et his principiis via sternitur ad majora." It did not gain him the prize, because he had not had time to fill in all the details of the necessary calculations.

The untroubled happy life of these last years, with which Riemann could be so gratified, reached its highest point when, on 3rd June 1862, he married Fräulein Elise Koch from Körchow in Mecklenburg-Schwerin, a friend of his sister. It was her lot to share the troubles of the remaining years of his life, and to ameliorate them through her untiring love. Already in July of this same year he had an attack of pleurisy, from which he had apparently quickly recovered, but this was the origin of a lung infection which was to lead to his early demise. When the doctors advised a lengthy sojourn in a southern climate, Wilhelm Weber and Sartorius von Waltershausen pressed the authorities to accord Riemann, not only the necessary leave, but sufficient financial support to enable him to undertake a visit to Italy, which began in November 1862. At the warm recommendation of Sartorius von Waltershausen, Riemann found the friendliest of welcomes at the family of Jäger, the consul in Messina, at whose villa in the suburb of Gazzi he spent the winter. His state of health rapidly improved, and he was able to go on excursions to Taormina, Catania, and Syracuse. On the return journey, which he began on 19th March 1863, he visited Palermo, Naples, Rome, Leghorn, Pisa, Florence, Bologna, and Milan. During the lengthy stays which he made in these cities, whose art treasures and antiquities aroused his greatest interest, he was able at the same time to get to know the leading Italian scholars. In particular he was able to renew his warm friendship with Professor Enrico Betti of Pisa, whom he had already met in Göttingen in 1858. On the whole, the years spent in Italy, despite the sad way in which they ended, represented the highlight of his life, not only because of the unending pleasure which this enchanting land with its natural beauty and its artistic treasures gave him, but also because it was here that he felt himself free. He was free from the petty restraints in his relations with other people which he thought that he had to observe at every step in Göttingen. The beneficial effect of the climate on his health also meant that he was often quite cheerful and was able to have many happy days.

It was with high hopes that he left his beloved Italy, but as he crossed over the Splügen pass, where he unexpectedly had to take a long journey on foot through the snow, he caught a severe cold, and when he arrived in Göttingen on 17th June, his condition was so bad that he soon had to decide on a second voyage to Italy, which began on 21st August 1863. He first went via Merano, Venice, and Florence to Pisa, where on 22nd December 1863 his wife gave birth to a daughter, who was christened Ida, in memory of his older sister. Unfortunately the winter was so cold that the Arno froze. In May 1864 he took a villa near Pisa, and here at the end of August he lost his younger sister Helene; he himself had an attack of jaundice, which resulted in a worsening of his chest complaint. He turned down an invitation to an appointment at the university of Pisa, formerly held by Professor Mosotti, which had already been extended to him in 1863, through the intermediary of Betti, partly on the advice of his friends in Göttingen, but mainly, no doubt, on the grounds that he feared that he would be unable, in the precarious state of his health, to fulfil the duties associated with the position. He felt therefore that he was not in a position in which he could honourably assume the responsibilities involved. This same sense of duty aroused a strong desire to return to Göttingen to resume his teaching role, and it was only the earnest representations of his friends and medical advisers which induced him to come back for another winter in Italy, which he passed in the congenial company of his learned friends in Pisa, Betti, Felici, Novi, Villari, Tassinari, and Beltrami. At that time he was also working on his memoir on the vanishing of the theta-functions. A deterioration in his condition made him spend May and June 1865 in Leghorn, July and August at Lake Maggiore, and September in Pegli, near Genoa, where a gastric fever brought about a significant worsening of his physical state.

In these circumstances Riemann could no longer withstand his desire to go back to Göttingen; he arrived back on 3rd October, and over the winter had tolerable health, which usually allowed him to work a few hours every day. He completed the memoir on the vanishing of the theta-functions and entrusted his former pupil Hattendorf with completing the paper on minimal surfaces. He often expressed the wish to speak before the end, to Dedekind, about some of his unfinished works, but he felt too weak and exhausted to be able to invite him to Göttingen. In the last months he worked on a paper dealing with the mechanism of the ear, which was unfortunately never completed, but was published only in fragmentary form after his death, by Henle and Schering.

The completion of this memoir, and of some other works, was very close to his heart, and he hoped that he might assemble the necessary strength for this task after a few months at Lake Maggiore, in the country which he had come to love so much, and which he longed to visit once more. He therefore resolved, on 15th June 1866, in the first days of the war, to make his third journey to Italy. It was interrupted at Cassel, because the railway there had been destroyed, but fortunately he succeeded in making his way to Giessen by carriage, and from there progress was unimpeded.

On 28th June he arrived at Lake Maggiore, where he lived in the Villa Pisoni in Selasca near Intra. His strength soon failed him and he himself clearly felt that his end was drawing near, but even on the day before his death he was working quietly under a fig-tree, filled with a great joy looking out over the beautiful landscape, on his last, and alas unfinished work. His life ended very peacefully, without any struggle or fear of death; it seemed as though he were following with interest the separation of the soul from the body. His wife had to give him bread and wine; he gave his blessing to those he loved at home and said to her: kiss our child. She recited the Lord's Prayer, but he could no longer speak; at the words "forgive us our trespasses" his eyes looked upwards, she felt his hand in hers grow colder, and after a few breaths his pure and noble heart ceased to beat. The religious sense which had been implanted in his paternal home remained with him throughout his life, and he served, though in a different way, the same God. With the greatest piety he avoided disturbing others in their beliefs; a daily self-examination in the sight of God was, in his expressed opinion, an important religious principle.

He lies buried in the churchyard of Biganzolo, which belongs to the same parish as Selasca. His gravestone bears the inscription:

Hier ruhet in Gott
GEORG FRIEDRICH BERNHARD RIEMANN, Prof. zu Göttingen,
geb. in Breselenz 17 Sept. 1826, gest. in Selasca 20. Juli 1866.
Denen die Gott lieben müssen alle Dinge zum Besten dienen.[11]

Here lies in the peace of God
GEORG FRIEDRICH BERNHARD RIEMANN, Professor at Göttingen,
born in Breselenz 17th September 1826, died in Selasca 20th July 1866.
Those who love God must serve him in all things to the best of their ability.

[11] The gravestone, which was donated by his Italian friends and colleagues, has now been removed as the result of the cemetery being transferred elsewhere.

Notes on the Individual Papers.

I.

Riemann's doctoral thesis, **I**, represents a major advance from the conceptions of complex function theory of Cauchy and the French school. Apart from the identification of the Cauchy-Riemann equations as fundamental, and the concept of an analytic function as a conformal mapping, there is the splendid invention of Riemann surfaces. This marks the first systematic use of topological ideas as a tool in mathematics. Riemann follows Gauss in referring to a conformal mapping as a mapping that 'preserves similarity in the smallest parts'.

Even Riemann's faulty use of the "Dirichlet principle" to prove the existence of certain functions was an invaluable stimulus to research. This complicated story is related in Gray (1994) and Monna (1975). Much was accomplished in conformal mapping by H. A. Schwarz and C. Neumann. Osgood gave the first completely general proof of the Riemann mapping theorem in 1900, and soon afterwards Hilbert gave a satisfactory account of the Dirichlet principle. The landmark book of Weyl (1913) was the first rigorous account of Riemann surfaces.

Further reading: Ahlfors (1953), Archibald (1996), Farkas and Kra (1991), Gray (1994), Kellogg (1929), Koch (1991), Laugwitz (1999), Monna (1975), Neuenschwander (1981), Remmert (1991, 1998), Springer (1957), Weber (1902, 46–48), Weyl (1913).

II.

This brief report of Riemann's first venture into mathematical physics is based on his 1854 lecture at a scientific meeting in Göttingen; see Dedekind's biography, p. 526. The final version **XX** of this work was not published during Riemann's lifetime. Weber (1902, p. 367) suggests that this was because Riemann did not care for an alteration to the paper that he had been asked to make, presumably by Kohlrausch. See also the Notes to **XX**.

III.

This is an early demonstration of Riemann's power in the modeling of problems in physics. There is a notable use of complex function theory. Riemann appears to have drawn some of the ideas from an 1854 paper of Kirchhoff on induced magnetization. Riemann's use of Bessel functions and 'half-convergent series' (asymptotic expansions) is historically notable; see Watson (1922).

Beetz and von Guebhard conducted further experiments on Nobili's rings later in the nineteenth century. By the end of the century it appeared that forces not taken into account by Riemann, in particular the electromotive force of polarization, could not be neglected. See Weber (1902), p. 62.

Further reading: Archibald (1991), Jungnickel and McCormach (1986, Vol. I), Riemann-Weber (1900), Watson (1922), Weber (1902, 62–66).

IV.

In this notable paper, Riemann applied his new ideas in complex function theory to the hypergeometric function. He provided an entirely new approach to problems treated by Gauss in 1812 and Kummer in 1837. He specified geometrically the P-functions to be studied and, in effect, used groups of matrices to specify the behavior of P-functions under circuits of the branch points. This is an early version of the theory of 'monodromy groups' (the term was first used by Jordan in 1870). Fuchs wrote extensively on related topics in the 1860s. Gray (2000) links the works of Riemann and Fuchs with a great deal of important work culminating with papers of Poincaré in the 1880s. Riemann's lectures (supplement to Weber (1902), 67–94) contain further ideas beyond **IV** and **XXI**, but these had been rediscovered by others when the supplement appeared.

The function $\Pi(s)$ (Gauss's notation) is $\Gamma(s+1)$ in modern notation.

Further reading: Ahlfors (1978), 318–321, Bühler (1985), Gray (2000), Laugwitz (1999,§1.3.1), Weber (1902, 86–87 and supplement, 67–94).

V.

This announcement concerning **IV** tells us what Riemann thought most important in the memoir. We note his high regard for applications of special functions, his wide acquaintance with the literature including Gauss's posthumous papers, and his telling comment that his approach 'yields all the earlier results, almost without calculation'. The hypergeometric functions turn up in a number of Riemann's later papers on physical problems and minimal surfaces.

VI.

The underlying object of study in this crucial paper is the integral of a rational function

$$\int_0^z R(x,y)dx$$

on an algebraic curve $F(x,y)=0$. Jacobi had shown in 1834 that if F is nonsingular of degree greater than 3, there will be more than 2 periods to such an integral. In the simplest case $F(x,y)=y^2-f(x)$, $\deg f=5$ or 6, his idea of taking pairs of integrals and pairs of endpoints together had been carried out by Göpel in 1847, and Rosenhain in 1851, independently. Weierstrass gained his reputation in extending this to $F(x,y)=y^2-f(x)$, $\deg f=2n+1$, in 1853 and 1856. Riemann's methods were entirely different, requiring Riemann surfaces. He relied on the Dirichlet principle for existence results. He introduced theta functions of several variables to solve the Jacobi inversion problem (his predecessors had restricted themselves to one or two variables). In 1913 Weyl carried out Riemann's ideas in **I** and **VI** rigorously. The 'Riemann-Roch' theorem, concerning the number of linearly independent meromorphic functions on a surface of genus p having m prescribed simple poles, is taken as fundamental in modern accounts (e.g. Farkas and Kra (1991)). The contribution of Riemann's student Gustav Roch (1839–1866) was published in Crelle's Journal in 1865.

Riemann's paper became important to algebraic geometers (relationship between the geometry of an algebraic curve and its Jacobian variety; birational geometry of plane curves).

In a Crelle paper of 1882, Dedekind and Weber gave Riemann's subject matter an alternative, purely algebraic, treatment. For a modern presentation, see Deuring (1973).

A lecture course of Riemann, partially reconstructed in the supplement to Weber (1902), goes beyond **VI** and its continuation **XI** in some respects.

Further reading: Deuring (1793), Farkas and Kra (1991), Dieudonné (1985), Gray (1989, 1998), Griffiths and Harris (1978), Koch (1991), Laugwitz (1999), Markushevich (1991), Mumford (1983, 1984), Siegel (1973), Weber (1902, 143–144 and supplement, 1–66), Weyl (1913).

VII.

This paper is a crucial step towards modern analytic number theory. Riemann shows that the zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \quad (\operatorname{Re} s > 1)$$

(a function studied by Euler) can be extended meromorphically over $\mathbb{C}$ and proves its functional equation. He then turns to the distribution of the zeros of $\zeta(s)$ in the 'critical strip' $0 \leq \operatorname{Re} s \leq 1$. Giving heuristic arguments, Riemann states correctly the asymptotic formula for $N(T)$, the number of these zeros with imaginary part in $[0, T]$, and correctly writes down the product formula for

$$\frac{1}{2}\, s(s-1)\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s)$$

in terms of these zeros. He provides an explicit formula for $\pi(x)$, the number of primes x, in terms of these zeros. The main term confirms the prime number theorem conjectured by Legendre and Gauss. Riemann proposes the (still open) 'Riemann hypothesis' that the zeros in the critical strip have real part 1/2. His other claims, including the prime number theorem, were proved between 1893 and 1896 by Hadamard, de la Vallée Poussin, and von Mangoldt.

By careful study of the Nachlass, Siegel (1932) was able to reconstruct more of Riemann's work on the zeta function, including an important asymptotic expansion in the critical strip. It is difficult to separate the contributions of Riemann and Siegel.

Further reading: Conrey (2003), Davenport (1980), Edwards (1974), Ivic (2003), Laugwitz (1999, §1.4), Motohashi (1997), Siegel (1932), Titchmarsh (1986), Weber (1902, 154–155).

VIII.

In this paper, probably his best work in mathematical physics, Riemann applies nonlinear partial differential equations to study the propagation of waves in a compressible gas. The concepts that he introduced here—Riemann invariants, the Riemann initial value problem, jump conditions for nonlinear equations, and the Riemann function for linear equations—are still basic today.

Riemann's account of shock waves does not allow for changes in entropy. As Rayleigh pointed out in 1878, energy is not conserved. The complete set of jump relations was supplied by Rankine in 1870, and in a posthumous publication of Hugoniot in 1887. Christoffel, in 1877, applied Riemann's results to a more general problem. Hadamard's monograph gives the state of the art in 1903.

Further reading: Hadamard (1903), Hölder (1981), Lax (1990), Smoller (1982), Weber (1902, 179–181).

IX.

This announcement concerning **VIII** is interesting for Riemann's view of the subject and of the significance of **VIII**. He remarks that his work clears up questions raised by Challis, Airy and Stokes. This is discussed by Friedrichs and Courant, §51.

Further reading: Friedrichs and Courant (1948).

X.

Newton first showed that the departure of the figure of the earth from a sphere is due to its rotation. Jacobi showed in 1834 that gravitational equilibrium of a rotating spheroid is consistent with three distinct axes if the angular momentum exceeds a critical value. Dirichlet had posed and partially analyzed the conditions for a configuration which is an ellipsoid varying with time, such that the motion, in an inertial frame, is linear in the coordinates. His results were edited posthumously by Dedekind in 1860. Riemann took up this problem of Dirichlet in **X**.

The paper contains a wealth of new results, and also some lapses and erroneous conclusions concerning the conditions for stability. The principal conclusions and defects of the paper, and its relation to work on stability in modern theoretical astrophysics, are discussed in Chandrasekhar (1969) and Chandrasekhar and Lebovitz (1990).

Further reading: Chandrasekhar (1969), Chandrasekhar and Lebovitz (1990), Dirichlet (1860).

XI.

This important paper is a continuation of **VI** required because 'insufficient account has been taken of the possibility that the theta function may vanish identically...when for its variables are substituted integrals of algebraic functions of a single variable' (p. 203). Riemann remedies this defect. His result is called the Riemann vanishing theorem. The modern version, in terms of a compact Riemann surface M, the Abel-Jacobi map from M to the Jacobian variety $J(M)$, and the theta divisor, is Theorem VI.3.5 of Farkas and Kra.

Further reading: Farkas and Kra (1991), Gunning (1976), Narasimhan (1990, 1–20).

XII.

This is Riemann's habilitation thesis. See Dedekind's biography, p. 523. Dedekind published **XII** posthumously from a complete manuscript in the Nachlass.

Riemann was concerned in **XII** with necessary conditions for the representation of a function by a trigonometric series. The careful historical introduction owes something to discussions with Dirichlet. The 'Riemann integral' is a great advance, given the function concepts of the period. Riemann's necessary and sufficient condition for integrability can be rewritten as 'discontinuities form a null set'. See e.g. Jones (1993, Ch. 7). Hawkins (1970) discusses the evolution of the Lebesgue integral from the solid starting point provided by **XII**.

Riemann's idea of using the function F obtained by two integrations of the trigonometric series was entirely new and remains crucial (see e.g. Zygmund (1959, Ch. 9). Other highlights include the Riemann-Lebesgue lemma in §10, and essentially the earliest use of the principle of stationary phase in §13. Compare Watson (1922), §8.2. For once, Watson misses a source—he overlooks **XII**.

Georg Cantor (1845–1918) investigated several questions suggested by **XII**, and his results form the basis of the modern theory of uniqueness (Zygmund (1959, Ch. 9), Meyer (1972)). These investigations eventually led Cantor to his famous results in set theory (transfinite cardinal and ordinal numbers).

Further reading: Dauben (1979), Dirichlet (1829), Hawkins (1970), Jones (1993), Laugwitz (1999), Meyer (1972), Monna (1975), Watson (1922), Weber (1902, 266–271), Zygmund (1959).

XIII.

This is the 'trial lecture' for Riemann's habilitation (see Dedekind's biography, p. 523). The colloquium of the philosophical faculty was held on June 10th, 1854. Riemann creates a new object, an n-dimensional differentiable manifold, and discusses different metrical scenarios and relations to physical reality. He was inspired by the great paper of Gauss (1828) on curved surfaces in 3-dimensional space. The trial lecture was published in 1868 from a manuscript in the Nachlass.

Riemann suppressed all the calculations in **XIII** for an audience more philosophically than mathematically inclined. (See, however, the notes to **XXII**.) Weyl (1919) is the best discussion of the mathematics underlying **XIII**. Despite its format, **XIII** proved a tremendous stimulus to differential geometers for the remainder of the century. After 1868, the writings of Christoffel, Helmholtz and Lipschitz were sharply influenced by **XIII**.

One can trace the early development of Riemannian geometry in the bibliography in Schouten (1954). Ricci's numerous contributions began in 1884; he developed the concept of tensor in the years 1887 to 1896. His 1901 paper with Levi-Civita later provided the tools for general relativity. Marcel Grossmann and Einstein developed the mathematical aspects of general relativity from 1912 onwards.

An interesting question is whether Riemann was aware in 1854 of the pioneering works of Bolyai and Lobachewski on non-Euclidean geometry. See Laugwitz (1999), Gray (1979).

Further reading: Butzer (1984), Gauss (1828), Gray (1979), Howard and Stachel (1986), Jungnickel and McCormach (1986, vol. II), Koch (1991), Laugwitz (1999), Portnoy (1982), Reich (1973), Schouten (1954), Struik (1989), Tazzioli (1989), Weyl (1919), Zund (1983).

XIV.

Riemann's paper **XIV** was submitted to the *Königl. Gesellschaft der Wissenschaften* in 1858 and then withdrawn. At this time the principal theories of electrodynamics were those of Wilhelm Weber (1804–1891) and Franz Neumann (1798–1895), dating from the 1840s. James Clerk Maxwell's electromagnetic field equations appeared in 1865.

After the posthumous publication of **XIV**, it was criticized by Clausius in Poggendorff's Annalen (vol. 135, p. 606) on the grounds that, by Riemann's hypotheses in the paper, the expression

$$P = -\int_0^t \sum\sum \epsilon\epsilon' F\left(\tau - \frac{r}{\alpha}, \tau\right) d\tau$$

has a vanishingly small value. This probably explains why Riemann withdrew **XIV**. However, this paper stimulated the work of Carl Neumann on electrodynamic theory (1868). The key assertion in **XIV** is that the electrodynamic actions of galvanic currents can be explained if one assumes that the action of one electrical mass on others propagates with the speed of light. Riemann's differential equation

$$\frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} = \Delta U + 4\pi\rho$$

was later deduced from Maxwell's theory.

Further reading: Archibald (1989), Jungnickel and McCormach (1986, vol. I), Laugwitz (1999, §3.2.4), Weber (1902, p. 293).

XV.

For the friendly contacts between Riemann and Weierstrass in Berlin that led to the writing of this letter, see Dedekind's biography, p. 530. For the modern version of **XV**, see Siegel (1973), Chapter 5: 'if a meromorphic function $f(z)$ of n variables cannot be transformed by a nonsingular linear mapping into a meromorphic function of fewer than n variables, then its period group is a lattice'. Note the use of Dirichlet's pigeonhole principle on p. 280 of **XV**.

Further reading: Siegel (1973), Markushevich (1991).

XVI.

This letter was written (in Italian) while Riemann was staying in Pisa. Enrico Betti (1823–1892) was one of the leading Italian mathematicians of the period. He was praised by Klein for his work of 1853 on Galois theory. He became friendly with Riemann in 1858. See Dedekind's biography, p. 529. Later Riemann spend a good deal of time in Pisa and conversed almost daily with Betti on the foundations of real and complex analysis, on mathematical physics, and—most significantly—on topological ideas. See the Notes to **XXIX**.

Further reading: Bottazini (1977, 1986, 1991), Klein (1928).

XVII.

Riemann worked on this problem in 1860 and 1861. In April 1866, he passed a manuscript containing only formulae to his student K. Hattendorff and asked him to write it up. The resulting paper, **XVII**, appeared in 1867. (Some further revision was carried out by Hattendorff for the 1876 version.) Meanwhile, Weierstrass in 1866 had given formulae which associate to every analytic function $F(t)$, $t = t_1 + it_2$, a minimal surface $x(t_1, t_2)$. See Kreyszig (1968, p. 83). Riemann's representation (p. 294) can be reduced to that of Weierstrass (Kreyszig (1992), p. 152).

Weierstrass's work was more influential than **XVII**. The reader can follow the progress of research on minimal surfaces from 1865 to 1885 in Volume 1 of the collected papers of H.A. Schwarz (1890). In the twentieth century, interest has centered on existence theorems for minimal surfaces rather than closed formulae.

Further reading: Garnier (1928), Kreyszig (1968, 1992), Laugwitz (1999, §1.3.5), Nitsche (1992), Schwarz (1890), Weber (1902, 334–337).

XIX.

Papers **XIX–XXI** (except **XXV**) first appeared in Weber (1876). Weber had searched through the 4,000 page Nachlass (now held at Göttingen University Library) for publishable material. The manuscript for **XIX** was dated 14 January 1847. Riemann might well have objected to the inclusion of this essay from his student years. There are some points of interest, noted by Laugwitz (1999), and the paper attracted the attention of Cayley in 1880.

References: Cayley (1880), Laugwitz (1999, §0.4.2).

XX.

This paper was written in 1854 (see Note to **II**). Weber (1902, p. 371) observes that in the manuscript, the whole of §3 was crossed out, and suggests that the reason for this was that contemporary physicists would have been shocked by these speculations.

The experimental work for the paper had been done by R. H. A. Kohlrausch (1809–1857). Kohlrausch was visiting Wilhelm Weber to work on the determination of the constant c in Weber's law at the time of his discussions with Riemann which led to **XX**.

Further reading: Archibald (1988), Jungnickel and McCormach (1986, vol. I).

XXI.

Riemann had claimed in **IV** that the ideas there should be applicable to linear differential equations with algebraic functions as coefficients. He deals with this question in **XXI**, which is dated 20 February 1857, the period when **IV** was written. Riemann considered the n-th order case only when the monodromy matrices can be diagonalized, but gave an essentially complete account in the 2×2 case. Later in the manuscript Riemann expressed doubts about his reasoning; see Weber (1902, p. 385). The last section was written by Weber from fragments in the Nachlass, and the enumeration of constants had to be corrected in the 1892 edition following objections by Hilbert.

Klein commented in 1894 that Riemann had not tackled the existence problem corresponding to given branch points and monodromy matrices. Hilbert pose a variant of this as the 21st of his 23 Paris problems in 1900. For the twists and turns of the subsequent story, see Gray (2000).

Further reading: Gray (2000), Laugwitz (1999, §1.3.1), Weber (1902, supplement, 67–94).

XXII.

This paper was an answer to the prize question on heat distribution posed by the Paris Academy. Written in Latin, it was submitted on 1 July 1861. The motto, which we did not translate in the text, could be rendered as 'These principles pave the way to higher things.' The submitted manuscript was made available to Weber for the collected works. Riemann had it in mind to revise **XXII**, but was diverted from the project by his ill health. The main interest of **XXII** today is in a computation that Riemann must have worked out earlier in connection with **XIII**, a necessary condition for flatness of a manifold with element of distance

$$ds^2 = \sum g_{ij}(\boldsymbol{x})dx_i dx_j$$

in terms of the 'Riemann curvature tensor'. By the time that **XXII** appeared, Christoffel and Lipschitz, in 1869–70, had shown that the condition is necessary and sufficient.

Riemann left out many details in **XXII**, and did not receive the prize (in fact, it was not awarded). The extensive remarks provided by Weber, highly praised in Weyl (1919), draw on an unpublished paper of Dedekind from about 1867. Dedekind's paper can be found in Sinaceur (1990).

Further reading: Butzer (1981), Laugwitz (1999, §3.1.5), Portnoy (1982), Sinaceur (1990), Tazzioli (1989), Weyl (1919), Zund (1983).

XXIII.

This fragment, written in Italian, dates from October 1863, when Riemann was traveling in northern Italy (see Dedekind's biography, p. 532). The paper was edited by H. A. Schwarz, and §3 had to be reconstructed by Schwarz from a few lines of text and formulae. Watson (§8.2) cites **XXIII** for a very early use of the principle of stationary phase (but see Notes to **XII**).

Further reading: Watson (1922).

XXIV.

In the list of Riemann's lecture courses in the supplement to Weber (1902), we find a number of courses from 1858 to 1864 on gravity, electricity and magnetism. The influential textbooks Riemann–Hattendorff (1869, 1876) are based on Riemann's lecture courses, but have been criticized for the changes made by Hattendorff. Riemann-Weber (1900) is a complete reworking by Weber of the material in Riemann-Hattendorff (1869).

Probably the computations in **XXIV** and **XXV** arose from illustrative examples in Riemann's lectures. For more on the extant lecture courses of Riemann, student notes of which can be found in a number of libraries in Germany, see Neuenschwander (1988).

Further reading: Jungnickel and McCormach (1986, vol. I), Laugwitz (1999, §§3.2.3, 3.2.4), Neuenschwander (1988), Riemann–Hattendorff (1876), Weber (1902, supplement, 114–116).

XXV.

Weber (1902, p. 439) dates this fragment between Easter and Whitsun 1856, and makes some comments on the differential equation (11), which reduces to Lamé's equation when $g = 0$. This paper, not included in the first edition of Riemann's works, was extracted when Weber made a further scan of the Nachlass for the second edition.

XXVI.

According to Weber (1902, p. 440), **XXVI–XXIX** consist of reconstructions from sheets in the Nachlass containing formulae and a few remarks. Weber considered the use of conformal mapping in **XXVI** of particular interest. In Weber (1902), we find the awful notation $z_1 \mp z_2$ for a conjugate pair; we have substituted $z_1 = \bar{z}_2$.

XXVII.

Here Weber took considerable liberties. Riemann had given the results for the first example briefly on a single sheet in the Nachlass. For the second example, he remarked that the problem could be solved. Weber, drawing on special cases in an 1871 paper of Schwarz, provided all the mathematics for the second example.

Further reading: Weber (1902, p. 445), Schwarz (1890).

XXVIII.

Presumably because the work of Jacobi to which he was adding was in Latin, Riemann wrote the first section of this fragment in Latin. The second section consists of formulae. Dedekind's explanatory essay, rederiving the formulae in Riemann's fragment, is of independent value and was included in his own collected works. The Dedekind sum in (32), now denoted by

$$s(m,n) = \sum_{t=1}^{n-1} \left(\left(\frac{t}{n} - \frac{1}{2}\right)\right) \left(\left(\frac{mt}{n} - \frac{1}{2}\right)\right)$$

appears in Dedekind's essay for the first time. The function $\eta(\omega)$ is now named after Dedekind. Both objects have been much studied. See particularly the notes on Chapter VIII of Chandrasekharan's book.

Further reading: Note by R. Fricke, p. 173 (vol. I) of Dedekind (1920–1932), Chandrasekharan (1985), Rademacher (1973).

XXIX.

This fragment consists of remarks (expanded by Weber) on topological concepts that Riemann had noted down. In the 1892 edition, Weber notes the relationship of these notions with those in an 1871 paper of Betti that he had overlooked in 1876. We are fortunate to have Betti's enthusiastic letter of 1863 to Tardy about his conversations with Riemann on simply connected spaces. See Weil (1979, 1979/1980), Pont (1974, 76–80). Dieudonné (1994) describes these ideas as "remaining vague and confused" in Betti's work, and gives Poincaré the credit for developing a coherent theory after studying Betti. Poincaré introduced the term 'Betti numbers' in 1895.

Further reading: Bottazini (1977), Dieudonné (1994), Pont (1974), Weber (1902, p. 482), Weil (1979, 1979/1980).

XXX.

This paper and **XXXI** were extracted by Weber from notes of Riemann's lectures taken in 1861 and 1862 by Roch. Weber made these exceptions to the rule in Weber (1892) that Riemann's lectures were excluded, because he thought the topics were highly interesting and somewhat self-contained. Both papers are often referred to in Weber (1902, supplement, 1–66).

XXXI.

Another source for Riemann's ideas on theta functions with characteristic is Riemann–Stahl (1899). Stahl based this account on lecture notes taken by Hattendorff and Schering, but made many changes. Rauch and Lebowitz (1973) organize their textbook in a similar way to Riemann–Stahl, although the exposition is quite different.

Further reading: Laugwitz (1999, §1.3.6), Rauch and Lebowitz (1973), Riemann–Stahl (1899), Weber (1902, supplement, 1–66).

Natural Philosophy.

This fragment is included in Weber (1876 and subsequent editions) as an appendix. There is no doubt that Riemann thought of this as important work—see Weber (1902, 507–508)—and at one time intended to publish some of the material. Altogether the Nachlass contains 203 pages classified as 'Fragments on natural philosophy'. In this particular piece, it is possible to discern an early version of the twentieth century goal of a unified field theory of gravitation and electromagnetism.

Further reading: Laugwitz (1999, §3.3.3), Schering (1909), Scholz (1982).

The life of Bernhard Riemann.

This essay for the first edition of the collected works is rounded out by Laugwitz (1999, Chapter 1) using Dedekind's correspondence. The footnotes are Dedekind's. There are some further points of interest in Klein (1928) and Monastyrsky (1999). Jungnickel and McCormach (1986, vol. I) skillfully reconstruct the contemporary scientific atmosphere in Germany.

A photograph of Riemann's gravestone can be found in Narasimhan (1990).

Bibliography

L. V. Ahlfors (1953), *Development of the theory of conformal mapping and Riemann surfaces through a century*, Contributions to the theory of Riemann surfaces, 3–13, Princeton University Press.

L. V. Ahlfors (1978) *Complex Analysis*, 3rd edn, McGraw-Hill.

T. Archibald (1989), *Physics as a constraint on mathematical research: the case of potential theory and electrodynamics*, Rowe and McCleary (1989), Vol. II, 29–75.

T. Archibald (1991), *Riemann and the theory of electrical phenomena: Nobili's rings*, Centaurus **34**, 247–271.

T. Archibald (1996), *From attraction theory to existence proofs: the evolution of potential-theoretic methods in the study of boundary-value problems*, 1860–1890, Rev. Histoire math. **2**, 67–93.

R. Bentley (1966) The works of Richard Bentley, Vol. 3, AMS Press.

U. Bottazini (1977), *Riemann's Einfluss auf Betti and F. Casorati*, Arch. History Exact Sci. **18**, 27–37.

U. Bottazini (1986), *The higher calculus: a history of real and complex analysis from Euler to Weierstrass*, Springer.

U. Bottazini (1991), *Riemann in Italia*, Conferenza internazionale nel 125° anniversario della norte di G.F.B. Riemann, Instituto Santa Maria.

W. K. Bühler (1985), *The hypergeometric function—a biographical sketch*, Mathematical Intelligencer **7**, 35–40.

P. L. Butzer, *An outline of the life and work of E. B. Christoffel*, in E. B. Christoffel, 2–29, P. L. Butzer and F. Feher (eds.), Birkhäuser 1981.

A. Cayley (1880), *Note on Riemann's paper 'Versuch einer allgemeinen Auffassung der Integration und Differentiation'*, Collected Mathematical Papers, Vol. II, 235–236.

S. Chandrasekhar, *Ellipsoidal figures of equilibrium*, Yale University Press, 1969.

S. Chandrasekhar and N. Lebowitz (1990), *On Riemann's paper: Ein Beitrag zu den Untersuchungen über die Bewegung eines flüssigen gleichartigen Ellipsoides*, Narasimhan (1990), 811–820.

K. Chandrasekharan, *Elliptic functions*, Springer 1985.

J. B. Conrey (2003), *The Riemann hypothesis*, Notices Amer. Math. Soc. **50**, 341-353.

J. W. Dauben (1979), *Georg Cantor. His mathematics and philosophy of the infinite.* Harvard University Press.

H. Davenport (1980), *Multiplicative number theory*, 2nd edn. Springer.

R. Dedekind (1920–1932), *Gesammelte mathematische Werke*, Vols. I–III, reprinted by Chelsea in two volumes (1968).

M. Deuring (1973) *Lectures on the theory of algebraic functions of one variable*, Lecture notes in mathematics, no. 314, Springer.

J. Dieudonné (1985), *History of algebraic geometry*, Wadsworth.

J. Dieudonné (1994), *Une breve histoire de la topologie*, in J-P. Pier, 35-153.

P. G. L. Dirichlet, (1829), *Sur la convergence des séries trigonometriques qui servent à répresenter un fonction arbitraire entre les limites données*, Mathematische Werke, Vol. I, 283–306, reprinted by Chelsea (1969).

P. G. L. Dirichlet (1860), *Untersuchungen über ein problem der hydrodynamik*, Mathematische Werke Vol. II, 263–301, reprented by Chelsea (1969).

H. M. Edwards (1974), *Riemann's zeta function*, Academic Press.

H. M. Farkas and I. Kra (1991), *Riemann surfaces*, 2nd edn., Springer.

K. Friedrichs and R. Courant (1948), *Supersonic flow and shock waves*, Interscience.

R. Garnier (1928), *Le problème de Plateau*, Ann. de l'école Norm. Sup. **45**, 53–144.

J. J. Gray (1979), *Ideas of space. Euclidean, non-Euclidean and relativistic.* Clarendon Press.

J. J. Gray (1989), *Algebraic geometry in the late nineteenth century*, Rowe and McLeary (1989), vol. I, 361–385.

J. J. Gray (1994), *On the history of the Riemann mapping theorem*, Supplemento ai Rendiconti del Circolo Matematico di Palermo, ser. II, no. 34, 47–94.

J. J. Gray (1998), *The Riemann-Roch theorem and geometry*, 1854-1914, Proceedings of the International Congress of Mathematicians Vol. III, 811–822, Documenta Mathematica.

J. J. Gray (2000), *Linear differential equations and group theory from Riemann to Poincaré*, 2nd. edn., Birkhäuser.

C. F. Gauss (1828), *Disquisitiones generales circa superficies curvas, Werke*, Vol. IV, 217–258. English translation by A. Hiltebeitel and J. Morehead, Astérisque **62**, 1–100.

P. J. Griffiths and J. Harris (1978), *Principles of algebraic geometry*, Wiley.

R. C. Gunning (1976), *Riemann surfaces and generalized theta functions*, Springer.

J. Hadamard (1903), *Lecons sur la propagation des ondes et les equations de l'hydrodynamique*. Hermann, Paris.

T. Hawkins (1970), *Lebesgue's theory of integration. Its origins and development*, University of Wisconsin Press.

E. Hölder (1981), *Historisches überblick zur mathematischen Theorie von Unstetigkeitswellen seit Riemann und Christoffel*, E. B. Christoffel, 412-434, ed. P. L. Butzer and F. Feher, Birkhäuser.

D. Howard and J. Stachel (eds.) (1986), *Einstein and the history of general relativity*, Birkhäuser.

A. Ivic (2003), *The Riemann zeta-function*, Dover.

F. Jones (1993), *Lebesgue integration on Euclidean space*, Jones and Bartlett.

C. Jungnickel and R. McCormach (1986), *Intellectual mastery of nature*, Vols. I and II, University of Chicago Press.

O. D. Kellogg (1929), *Foundations of potential theory*, Springer.

F. Klein (1928), *Development of mathematics in the nineteenth century*, translated by M. Ackerman from the 1928 original, Math. Sci. Press.

E. Knobloch and D. E Rowe (eds.)(1989), *The history of modern mathematics*, Vol. III, Academic Press.

H. Koch (1991), *Introduction to classical mathematics I*, Kluwer.

E. Kreyszig (1968), *Introduction to differential geometry and Riemannian geometry*, University of Toronto Press.

E. Kreyszig (1992), *On the theory of minimal surfaces*, in The problem of Plateau, 138–164, ed. T. M. Rassias, World Scientific.

D. Laugwitz (1999), *Bernhard Riemann, 1826–1866.* Birkhäuser.

P. Lax (1990), *On Riemann's paper: über die Fortplanzung ebener Luftwellen von endlicher Schwingungsweite*, Narasimhan (1990), 807–810.

A. I. Markushevich (1991), *Introduction to the classical theory of Abelian functions*, American Mathematical Society.

Y. Meyer (1972), *Algebraic numbers and harmonic analysis*, North-Holland.

M. Monastyrsky (1999), *Riemann, topology and physics*, 2nd. edn., Springer.

A. F. Monna (1975), *Dirichlet's principle, a mathematical comedy of errors and its influence on the development of analysis*, Oosthoek, Scheltema and Holkema.

Y. Motohashi (1997), *Spectral theory of the Riemann zeta-function*, Cambridge.

D. Mumford (1983, 1984), *Tata lectures on theta*, Vols. I, II., Birkhäuswer.

R. Narasimhan (1990), *Bernhard Riemann, Collected Papers*, reprint of Weber (1902) with additional papers by H. Weyl et. al., Springer.

E. Neuenschwander (1981), *Studies in the history of complex function theory II: Interactions among the French school, Riemann and Weierstrass*, Bulletin Amer. Math. Soc. N.S. **5**, 87–105.

E. Neuenschwander (1988), *A brief report on a number of recently discovered sets of notes of Riemann's lectures and on the transmission of the Riemann Nachlass*, Historia Math. **15**, 101–113, reprinted in Narasimhan (1990), 855–867.

J. Nitsche (1989), *Lectures on minimal surfaces, Vol. I*, Cambridge University Press.

H. L. S. Orde (1978), *Dirichlet's class number formula*, J. London Math. Soc. (2) **18**, 409–420.

J-P. Pier (ed.) (1994), *Development of mathematics 1900–1950*, Birkhäuser.

J-C. Pont (1974), *La topologie algebrique des origines à Poincaré*, P. U. F.

E. Portnoy (1982), *Riemann's contribution to differential geometry*, Historia Math. **9**, 1–18.

H. Rademacher (1973), *Topics in analytic number theory*, Springer.

H. E. Rauch and A. Lebowitz (1973), *Elliptic functions, theta functions and Riemann surfaces*, Williams and Wilkins.

K. Reich (1973), *Die Geschichte der Differentialgeometrie von Gauss bis Riemann (1828-1868)*, Arch. Hist. Exact Sciences **11**, 273–282.

R. Remmert (1991), *Theory of complex functions*, Springer.

R. Remmert (1998), *Classical topics in complex function theory*, Springer.

B. Riemann and K. Hattendorff (1869), *Partielle Differentialgleichungen und deren Anwendung auf physikalische Fragen*, Vieweg.

B. Riemann and K. Hattendorff (1876), *Schwere, Elektricität und Magnetismus*, Rümpler.

B. Riemann and H. Stahl (1899), *Elliptische Functionen*, Teubner.

B. Riemann and H. Weber (1900), *Die partiellen Differentialgleichungen der Mathematischen Physik*, Vieweg.

D. E. Rowe and J. McCleary (eds.) (1989), *The history of modern mathematics*, Vols. I and II, Academic Press.

E. Schering (1909), *Zum Gedächtniss an B. Riemann*, Gesammelte Mathematische Werke, Vol. II, 367–383, Meyer and Muller; reprinted in Narasimhan (1990), 828–844.

E. Scholz (1982), *Herbart's influence on Bernhard Riemann*, Historia Math. **9**, 413–440.

J. A. Schouten (1954), *Ricci-Calculus*, Springer.

H. A. Schwarz (1980), Gesammelte mathematische Abhandlungen, Vols. I, II, reprinted by Chelsea (1972).

C. L. Siegel (1932), *Über Riemanns Nachlass zur analytischen Zahlenheorie*, Quellen and Studien zur Geschichte der Mathematik, Astronomie und Physik, B: Studien, **2**, 45–80; reprinted in *Gesammelte Abhandlungen* Vol. I, 275–310 and Narasimhan (1990), 770–805.

C. L. Siegel (1973), *Topics in complex function theory*, Vol. III, Wiley.

M. A. Sinaceur (1990), *Dedekind et le programme de Riemann*, Revue d'Histoire des Sciences **43**, 221–296.

J. Smoller (1982), *Shock waves and reaction-diffusion equations*, Springer.

G. Springer (1957), *Introduction to Riemann surfaces*, Addison-Wesley.

D. J. Struik (1989), *Schouten, Levi-Civita and the emergence of tensor calculus*, Rowe and McLeary, vol. II, 99–105.

R. Tazzioli (1989), *Rudolf Lipschitz's work on differential geometry and mechanics*, Knobloch and Rowe (1989), 113–138.

E. C. Titchmarsh (1986), *The theory of the Riemann zeta-function*, 2nd. edn., Oxford University Press.

G. N. Watson (1922), *A treatise on the theory of Bessel functions*, Cambridge University Press.

H. Weber (ed.) (1892), *Bernhard Riemann's gesammelte mathematische Werke und wissenschaftlicher Nachlass*, 2nd. edn., Teubner.

H. Weber (ed.) (1902), Reprint of Weber (1892), with supplement edited by M. Noether and W. Wirtinger; reprinted by Dover (1953).

A. Weil (1979), *Riemann, Betti and the birth of topology*, Arch. Hist. Exact Sci. **20**, 91–96.

A. Weil (1980), *A postscript to my article "Riemann, Betti and the birth of topology"*, Arch. Hist. Exact Sci. **21**, 387.

H. Weyl (1913), *The concept of a Riemann surface*, translation of 3rd edn., Addison-Wesley (1953).

H. Weyl (1919), *Vorwort und Erläuterungen zu: Riemann, Über die Hypothesen, welche der Geometrie zu Grunde liegen*, Das Kontinuum und Andere Monographien, H. Weyl, E. Landau, B. Riemann, Chelsea (1960), III–V, 23–48; reprinted in Narasimhan (1990), 739–768.

J. D. Zund (1983), *Some comments on Riemann's contributions to differential geometry*, Historia Math. **10** (1983), 84–89.

A. Zygmund (1959), *Trigonometric series*, Vols. I and II, Cambridge University Press.