Quantisation as a Problem of
Proper Values (Part III)

PERTURBATION THEORY, WITH APPLICATION TO THE STARK EFFECT OF THE BALMER LINES

Introduction. Abstract

As has already been mentioned at the end of the preceding paper,[1] the available range of application of the proper value theory can by comparatively elementary methods be considerably increased beyond the "directly soluble problems"; for proper values and functions can readily be approximately determined for *such* boundary value problems as are sufficiently closely related to a directly soluble problem. In analogy with ordinary mechanics, let us call the method in question the *perturbation* method. It is based upon the important *property of continuity* possessed by proper values and functions,[2] principally, for our purpose, upon their *continuous* dependence on the *coefficients* of the differential equation, and less upon the extent of the domain and on the boundary conditions, since in our case the domain (" entire *q*-space ") and the boundary conditions (" remaining finite ") are generally the same for the unperturbed and perturbed problems.

The method is essentially the same as that used by Lord Rayleigh in investigating [3] the vibrations of a string with *small inhomogeneities* in his *Theory of Sound* (2nd edit., vol. i., pp. 115-118, London, 1894). This was a particularly simple case, as the differential equation of the unperturbed problem had *constant* coefficients, and only the perturbing terms were arbitrary functions along the string. A complete generalisation is possible not merely with regard to these points, but also for the specially important case of *several* independent variables, *i.e.* for *partial* differential equations, in which *multiple proper values* appear in the unperturbed problem, and where the addition of a

---

[1] Last two paragraphs of Part II.
[2] Courant-Hilbert, chap. vi. §§ 2, 4, p. 337.
[3] Courant-Hilbert, chap. v. § 5, 2, p. 241.

perturbing term causes the *splitting up* of such values and is of the greatest interest in well-known spectroscopic questions (Zeeman effect, Stark effect, Multiplicities). In the development of the perturbation theory in the following Section I., which really yields nothing new to the mathematician, I put less value on generalising to the *widest possible extent* than on bringing forward the very simple rudiments in the clearest possible manner. From the latter, any desired generalisation arises almost automatically when needed. In Section II., as an example, the Stark effect is discussed and, indeed, by *two* methods, of which the *first* is analogous to Epstein's method, by which he first solved[1] the problem on the basis of classical mechanics, supplemented by quantum conditions, while the *second*, which is much more general, is analogous to the method of secular perturbations.[2] The *first* method will be utilised to show that in wave mechanics also the perturbed problem can be "separated" in *parabolic* co-ordinates, and the perturbation theory will first be applied to the ordinary differential equations into which the original vibration equation is split up. The theory thus merely takes over the task which on the old theory devolved on Sommerfeld's elegant complex integration for the calculation of the quantum integrals.[3] In the *second* method, it is found that in the case of the Stark effect an exact separation co-ordinate system exists, quite by accident, for the perturbed problem also, and the perturbation theory is applied directly to the *partial* differential equation. This latter proceeding proves to be more troublesome in wave mechanics, although it is theoretically superior, being more capable of generalisation.

Also the problem of the intensity of the components in the Stark effect will be shortly discussed in Section II. Tables will be calculated, which, as a whole, agree even better with experiment than the well-known ones calculated by Kramers with the help of the correspondence principle.[4]

The application (not yet completed) to the *Zeeman effect* will naturally be of much greater interest. It seems to be indissolubly linked with a correct formulation in the language of wave mechanics of the *relativistic* problem, because in the four-dimensional formulation the vector-potential automatically ranks equally with the scalar. It was already mentioned in Part I. that the relativistic hydrogen atom may indeed be treated without further discussion, but that it leads to "half-integral" azimuthal quanta, and thus contradicts experience. Therefore "something must still be missing". Since then I have learnt *what* is lacking from the most important publications of G. E. Uhlenbeck and S. Goudsmit,[5] and then from oral and written communications from Paris (P. Langevin) and Copenhagen (W. Pauli),

[1] P. S. Epstein, *Ann. d. Phys.* 50, p. 489, 1916.
[2] N. Bohr, *Kopenhagener Akademie* (8), IV., 1, 2, p. 69 *et seq.*, 1918.
[3] A. Sommerfeld, *Atombau*, 4th ed., p. 772.
[4] H. A. Kramers, *Kopenhagener Akademie* (8), III., 3, p. 287, 1919.
[5] G. E. Uhlenbeck and S. Goudsmit, *Physica*, 1925: *Die Naturwissenschaften*, 1926; *Nature*, 20th Feb., 1926; cf. also L. H. Thomas, *Nature*, 10th April. 1926.

viz., in the language of the theory of electronic orbits, the *angular momentum* of the electron round its axis, which gives it a *magnetic moment*. The utterances of these investigators, together with two highly significant papers by Slater[1] and by Sommerfeld and Unsöld[2] dealing with the Balmer spectrum, leave no doubt that, by the introduction of the paradoxical yet happy conception of the spinning electron, the orbital theory will be able to master the disquieting difficulties which have latterly begun to accumulate (anomalous Zeeman effect; Paschen-Back effect of the Balmer lines; irregular and regular Röntgen doublets; analogy of the latter with the alkali doublets, etc.). We shall be obliged to attempt to take over the idea of Uhlenbeck and Goudsmit into wave mechanics. I believe that the latter is a very fertile soil for this idea, since in it the electron is not considered as a point charge, but as continuously flowing through space,[3] and so the unpleasing conception of a " rotating point-charge " is avoided. In the present paper, however, the taking over of the idea is not yet attempted.

To the *third section*, as " mathematical appendix ", have been relegated numerous uninteresting calculations—mainly quadratures of products of proper functions, required in the second section. *The formulae of the appendix are numbered (101), (102), etc.*

I. Perturbation Theory

§ 1. A Single Independent Variable

Let us consider a linear, homogeneous, differential expression of the second order, which we may assume to be in self-adjoint form without loss of generality, viz.

$$(1) \qquad\qquad L[y] = py'' + p'y' - qy.$$

y is the dependent function; p, p' and q are continuous functions of the independent variable x and $p \geq 0$. A dash denotes differentiation with respect to x (p' is therefore the derivative of p, which is the condition for self-adjointness).

Now let $\rho(x)$ be another continuous function of x, which never becomes negative, and also in general does not vanish. We consider the proper value problem of Sturm and Liouville,[4]

$$(2) \qquad\qquad L[y] + E\rho y = 0.$$

It is a question, first, of finding all *those* values of the constant E (" proper values ") for which the equation (2) possesses solutions $y(x)$, which are continuous and not identically vanishing within a certain domain, and which satisfy certain " boundary conditions " at the bounding points; and secondly of finding these solutions (" proper

[1] J. C. Slater, *Proc. Amer. Nat. Acad.* 11, p. 732, 1925.
[2] A. Sommerfeld and A. Unsöld, *Ztschr. f. Phys.* 36, p. 259, 1926.
[3] Cf. last two pages of previous paper.
[4] Cf. Courant-Hilbert, chap. v. § 5, 1, p. 238 *et seq.*

functions ") themselves. In the cases treated in atomic mechanics, domain and boundary conditions are always "natural". The domain, for example, reaches from 0 to ∞, when x signifies the value of the radius vector or of an intrinsically positive parabolic co-ordinate, and the boundary conditions are in these cases : *remaining finite*. Or, when x signifies an azimuth, then the domain is the interval from 0 to 2π and the condition is : Repetition of the initial values of y and y' at the end of the interval (" periodicity ").

It is only in the case of the periodic condition that *multiple*, viz. *double-valued*, proper values appear for *one* independent variable. By this we understand that to the same proper value belong *several* (in the particular case, two) linearly independent proper functions. We will now exclude this case for the sake of simplicity, as it attaches itself easily to the developments of the following paragraph. Moreover, to lighten the formulae, we will not expressly take into account in the notation the possibility that a " band spectrum " (*i.e.* a *continuum* of proper values) may be present when the domain extends to infinity.

Let now $y = u_i(x)$, $i = 1, 2, 3, \ldots$, be the series of Sturm-Liouville proper functions ; then the series of functions $u_i(x)\sqrt{\rho(x)}$, $i = 1, 2, 3, \ldots$, forms a *complete orthogonal system* for the domain ; *i.e.* in the first place, if $u_i(x)$ and $u_k(x)$ are the proper functions belonging to the values E_i and E_k, then

$$(3) \qquad \int \rho(x)u_i(x)u_k(x)dx = 0 \text{ for } i \neq k.$$

(Integrals without limits are to be taken over the domain, throughout this paper.) The expression " complete " signifies that an originally arbitrary continuous function is condemned to vanish identically, by the mere postulation that it must be orthogonal with respect to *all* the functions $u_i(x)\sqrt{\rho(x)}$. (More shortly : " There exists no further orthogonal function for the system.") We can and will always regard the proper functions $u_i(x)$ in all general discussions as " normalised ", *i.e.* we imagine the constant factor, which is still arbitrary in each of them on account of the homogeneity of (2), to be defined in *such a way* that the integral (3) takes the value unity for $i = k$. Finally we again remind the reader that the proper values of (2) are certainly all *real*.

Let now the proper values E_i and functions $u_i(x)$ be *known*. Let us, from now on, direct our attention specially to a *definite* proper value, E_k say, and the corresponding function $u_k(x)$, and ask how these alter, when we do not alter the problem in any way other than by adding to the left-hand side of (2) a small " perturbing term ", which we will initially write in the form

$$(4) \qquad -\lambda r(x)y.$$

In this λ is a small quantity (the perturbation parameter), and $r(x)$ is an arbitrary continuous function of x. It is therefore simply a matter of a slight alteration of the coefficient q in the differential expression (1). From the continuity properties of the proper quantities,

mentioned in the introduction, we now know that the altered Sturm-Liouville problem

(2') $$L[y] - \lambda r y + E \rho y = 0$$

must have, in any case for a sufficiently small λ, proper quantities in the near neighbourhood of E_k and u_k, which we may write, by way of trial, as

(5) $$E_k{}^* = E_k + \lambda \epsilon_k ; \quad u_k{}^* = u_k(x) + \lambda v_k(x).$$

On substituting in equation (2'), remembering that u_k satisfies (2), neglecting λ^2 and cutting away a factor λ we get

(6) $$L[v_k] + E_k \rho v_k = (r - \epsilon_k \rho) u_k.$$

For the defining of the perturbation v_k of the proper *function*, we thus obtain, as a comparison of (2) and (6) shows, a *non-homogeneous* equation, which belongs precisely to *that* homogeneous equation which is satisfied by our unperturbed proper function u_k (for in (6) the special proper value E_k stands in place of E). On the right-hand side of this non-homogeneous equation occurs, in addition to known quantities, the still unknown perturbation ϵ_k of the proper *value*.

This occurrence of ϵ_k serves for the calculation of this quantity *before* the calculation of v_k. It is known that the non-homogeneous equation—and this is *the starting-point of the whole perturbation theory* —for a proper *value* of the homogeneous equation possesses a solution *when*, and *only when*, its right-hand side is *orthogonal* [1] to the allied proper function (to all the allied functions, in the case of multiple proper values). (The physical interpretation of this mathematical theorem, for the vibrations of a string, is that if the force is in resonance with a proper vibration it must be distributed in a very special way over the string, namely, so that it does no work in the vibration in question; otherwise the amplitude grows beyond all limits and a stationary condition is impossible.)

The right-hand side of (6) must therefore be orthogonal to u_k, *i.e.*

(7) $$\int (r - \epsilon_k \rho) u_k{}^2 dx = 0,$$

or

(7') $$\epsilon_k = \frac{\int r u_k{}^2 dx}{\int \rho u_k{}^2 dx},$$

or, if we imagine u_i already normalised, then, more simply,

(7") $$\epsilon_k = \int r u_k{}^2 dx.$$

This simple formula expresses the perturbation of the proper value (of first order) in terms of the perturbing function $r(x)$ and the unperturbed proper function $u_k(x)$. If we consider that the proper

---

[1] Cf. Courant-Hilbert, chap. v. § 10, 2, p. 277.

value of our problem signifies mechanical energy or is analogous to it, and that the proper function u_k is comparable to "motion with energy E_k", then we see in (7") the complete parallel to the well-known theorem in the perturbation theory of classical mechanics, viz. the perturbation of the energy, to a first approximation, is equal to the perturbing function, averaged over the unperturbed motion. (It may be remarked in passing that it is as a rule sensible, or at least aesthetic, to throw into bold relief the factor $\rho(x)$ in the integrands of *all* integrals taken over the entire domain. If we do this, then, in integral (7"), we must speak of $\dfrac{r(x)}{\rho(x)}$ and not $r(x)$ as the perturbing function, and make a corresponding change in the expression (4). Since the point is quite unimportant, however, we will stick to the notation already chosen.)

We have yet to define $v_k(x)$, the perturbation of the proper *function*, from (6). We solve [1] the non-homogeneous equation by putting for v_k a series of proper functions, viz.

$$(8) \qquad v_k(x) = \sum_{i=1}^{\infty} \gamma_{ki} u_i(x),$$

and by developing the right-hand side, divided by $\rho(x)$, likewise in a series of proper functions, thus

$$(9) \qquad \left(\frac{r(x)}{\rho(x)} - \epsilon_k\right) u_k(x) = \sum_{i=1}^{\infty} c_{ki} u_i(x),$$

where

$$(10) \qquad \begin{cases} c_{ki} = \int (r - \epsilon_k \rho) u_k u_i dx \\[2mm] \quad = \int r u_k u_i dx \text{ for } i \neq k \\[2mm] \quad = 0 \qquad\quad \text{ for } i = k. \end{cases}$$

The last equality follows from (7). If we substitute from (8) and (9) in (6) we get

$$(11) \qquad \sum_{i=1}^{\infty} \gamma_{ki}(L[u_i] + E_k \rho u_i) = \sum_{i=1}^{\infty} c_{ki} \rho u_i.$$

Since now u_i satisfies equation (2) with $E = E_i$, it follows that

$$(12) \qquad \sum_{i=1}^{\infty} \gamma_{ki} \rho (E_k - E_i) u_i = \sum_{i=1}^{\infty} c_{ki} \rho u_i.$$

By equating coefficients on left and right, all the γ_{ki}'s, except γ_{kk}, are defined. Thus

$$(13) \qquad \gamma_{ki} = \frac{c_{ki}}{E_k - E_i} = \frac{\int r u_k u_i dx}{E_k - E_i} \text{ for } i \neq k,$$

while γ_{kk}, as may be understood, remains completely undefined. This indefiniteness corresponds to the fact that the postulation of

[1] Cf. Courant-Hilbert, chap. v. § 5, 1, p. 240, and § 10, p. 279.

normalisation is still available for us for the perturbed proper function. If we make use of (8) in (5) and claim for $u_k{}^*(x)$ the same normalisation as for $u_k(x)$ (quantities of the order of λ^2 being neglected), then it is evident that $\gamma_{kk}=0$. Using (13) we now obtain for the *perturbed proper function*

$$(14) \qquad u_k{}^*(x) = u_k(x) + \lambda \sum_{i=1}^{\infty}{}' \frac{u_i(x)\int r u_k u_i dx}{E_k - E_i}.$$

(The dash on the sigma denotes that the term $i=k$ has not to be taken.) And the allied perturbed proper value is, from the above,

$$(15) \qquad E_k{}^* = E_k + \lambda \int r u_k^2 dx.$$

By substituting in (2') we may convince ourselves that (14) and (15) do really satisfy the proper value problem to the proposed degree of approximation. This verification is necessary since the development, assumed in (5), in *integral* powers of the perturbation parameter is no necessary consequence of continuity.

The procedure, here explained in fair detail for the simplest case, is capable of generalisation in many ways. In the first place, we can of course consider the perturbation in a quite similar manner for the second, and then the third order in λ, etc., in each case obtaining first the next approximation to the proper value, and then the corresponding approximation for the proper function. In certain circumstances it may be advisable—just as in the perturbation theory of mechanics—to regard the perturbation function itself as a power series in λ, whose terms come into play one by one in the separate stages. These questions are discussed exhaustively by Herr E. Fues in work which is now appearing in connection with the application to the theory of *band spectra*.

In the second place, in quite similar fashion, we can consider also a perturbation of the term in y' of the differential operator (1) just as we have considered above the term $-qy$. The case is important, for the Zeeman effect leads without doubt to a perturbation of this kind—though admittedly in an equation with several independent variables. Thus the equation loses its self-adjoint form by the perturbation—not an essential matter in the case of a single variable. In a partial differential equation, however, this loss may result in the perturbed proper values no longer being real, though the perturbing term is real; and naturally also conversely, an imaginary perturbing term may have a real, physically intelligible perturbation as its consequence.

We may also go further and consider a perturbation of the term in y''. Indeed it is quite possible, in general, to add an arbitrary " infinitely small " linear [1] and homogeneous differential operator, even of higher order than the second, as the perturbing term and to calculate the perturbations in the same manner as above. In these cases,

---

[1] Even the limitation " linear " is not absolutely necessary.

however, we would use with advantage the fact that the second and higher derivatives of the proper functions may be expressed by means of the differential equation itself, in terms of the zero and first derivatives, so that this general case may be reduced, in a certain sense, to the two special cases, first considered—perturbation of the terms in y and y'.

Finally, it is obvious that the extension to equations of order higher than the second is possible.

Undoubtedly, however, the most important generalisation is that to several independent variables, *i.e.* to partial differential equations. For *this* really is the problem in the general case, and only in exceptional cases will it be possible to split up the disturbed partial differential equation, by the introduction of suitable variables, into separate differential equations, each only with one variable.

§ 2. Several Independent Variables (Partial Differential Equation)

We will represent the several independent variables in the formulae symbolically by the *one* sign x, and briefly write $\int dx$ (instead of $\int \dots \int dx_1 dx_2 \dots$) for an integral extending over the multiply-dimensioned *domain*. A notation of this type is already in use in the theory of integral equations, and has the advantage, here as there, that the structure of the formulae is not altered by the increased number of variables as such, but only by *essentially* new occurrences, which *may* be related to it.

Let therefore $L[y]$ now signify a self-adjoint *partial* linear differential expression of the second order, whose explicit form we do not require to specify ; and further let $\rho(x)$ again be a positive function of the independent variables, which does not vanish in general. The postulation " self-adjoint " is *now* no longer unimportant, as the property cannot now be generally gained by multiplication by a suitably chosen $f(x)$, as was the case with *one* variable. In the particular differential expression of wave mechanics, however, this is still the case, as it arises from a variation principle.

According to these definitions or conventions, we can regard equation (2) of § 1,

(2) $$L[y] + E\rho y = 0,$$

as the formulation of the Sturm-Liouville proper value problem in the case of several variables also. Everything said there about the proper values and functions, their orthogonality, normalisation, etc., as also *the whole perturbation theory there developed*—in short, the whole of § 1—remains *valid without change, when* all the proper values are *simple*, if we use the abbreviated symbolism just agreed upon above. And only *one* thing does *not* remain valid, namely, that they *must* be simple.

Nevertheless, from the pure mathematical standpoint, the case when the roots are all distinct is to be regarded as the *general* case for several variables also, and multiplicity regarded as a special occurrence, which, it is admitted, *is the rule in applications*, on account of the specially simple and symmetrical structure of the differential expressions $L[y]$ (and the " boundary conditions ") which appear. Multiplicity of the proper values corresponds to *degeneracy* in the theory of conditioned periodic systems and is therefore especially interesting for quantum theory.

A proper value E_k is called α-fold, when equation (2), for $E = E_k$, possesses not *one* but exactly α linearly independent solutions which satisfy the boundary conditions. We will denote these by

(16) $u_{k1}, u_{k2}, \ldots u_{k\alpha}.$

Then it is true that each of these α proper functions is *orthogonal* to each of the *other* proper functions belonging to *another* proper value (the factor $\rho(x)$ being included; cf. (3)). On the contrary, these α functions are *not* in general orthogonal *to one another*, if we merely postulate that they are α linearly independent proper functions for the proper value E_k, and nothing more. For then we can equally well replace them by α arbitrary, linearly independent, linear aggregates (with constant coefficients) of themselves. We may express this otherwise, thus. The series of functions (16) is initially *indefinite* to the extent of a linear transformation (with constant coefficients), involving a non-vanishing determinant, and such a transformation *destroys*, in general, the mutual orthogonality.

But through such a transformation this mutual orthogonality can always be *brought about*, and indeed in an infinite number of ways ; the latter property arising because *orthogonal* transformation does *not* destroy the mutual orthogonality. We are now accustomed to include this simply in *normalisation*, that orthogonality is secured for *all* proper functions, even for those which belong to the *same* proper value. We will assume that our u_{ki}'s are already *normalised* in this way, and of course for *each* proper value. Then we must have

(17) $\begin{cases} \int\int \rho(x)u_{ki}(x)u_{k'i'}(x)dx = 0 \text{ when } (k, i) \neq (k', i') \\ \qquad\qquad = 1 \text{ when } k' = k, \text{ as well as } i' = i. \end{cases}$

Each of the finite series of proper functions u_{ki}, obtained for *constant* k and *varying* i, is then only still indefinite to this extent, that it is subject to an *orthogonal* transformation.

We will now discuss, first in words, without using formulae, the consequences which follow when a perturbing term is added to the differential equation (2). The addition of the perturbing term will, in general, remove the above-mentioned symmetry of the differential equation, to which the multiplicity of the proper values (or of certain of them) is due. Since, however, the proper values and functions are *continuously* dependent on the coefficients of the differential equation, a small perturbation causes a group of α proper values, which lie close

to one another and to E_k, to enter in place of the a-fold proper value E_k. The latter is *split up*. Of course, if the symmetry is not wholly destroyed by the perturbation, it may happen that the splitting up is not complete and that several proper values (still partly multiple) of, *in summa*, equal multiplicity merely appear in the place of E_k ("*partial* removal of degeneracy").

As for the perturbed proper *functions*, those a members which belong to the a values arising from E_k must evidently also on account of continuity lie infinitely near the unperturbed functions belonging to E_k, viz. u_{ki}; $i = 1, 2, 3 \ldots a$. Yet we must remember that the last-named series of functions, as we have established above, is indefinite to the extent of an *arbitrary orthogonal transformation*. *One* of the infinitely numerous definitions, which may be applied to the series of functions, u_{ki}; $i = 1, 2, 3 \ldots a$, will lie infinitely near the series of perturbed functions; and if the value E_k is completely split up, it will be a *quite definite one*! For to the separate simple proper values, into which the value is split up, there belong proper functions which are quite uniquely defined.

This unique particular specification of the *unperturbed* proper functions (which may fittingly be designated as the " approximations of zero order " for the *perturbed* functions), which is defined by the nature of the perturbation, will naturally *not* generally coincide with that definition of the unperturbed functions which we chanced to adopt to begin with. Each group of the latter, belonging to a definite a-fold proper value E_k, will have first to be submitted to an orthogonal substitution, defined by the kind of perturbation, before it can serve as the starting-point, the " zero approximation ", for a more exact definition of the perturbed proper functions. *The defining of these orthogonal substitutions*—one for each multiple proper value—*is the only essentially new point* that arises because of the increased number of variables, or from the appearance of multiple proper values. The defining of these substitutions forms the exact counterpart to the finding of an approximate separation system for the perturbed motion in the theory of conditioned periodic systems. As we will see immediately, the definition of the substitutions can always be given in a theoretically simple way. It requires, for each a-fold proper value, merely the principal axes transformation of a quadratic form of a (and thus of a finite number of) variables.

When the substitution has once been accomplished, the calculation of the approximations of the *first* order runs almost word for word as in § 1. The sole difference is that the dash on the sigma in equation (14) must mean that in the summation *all* the proper functions belonging to the value E_k, *i.e. all* the terms whose denominators would vanish, must be left out. It may be remarked in passing that it is not at all necessary, in the calculation of *first* approximations, to have completed the orthogonal substitutions referred to for *all* multiple proper values, but it is sufficient to have done so for the value E_k, in whose splitting up we are interested. For the approximations

of higher order, we admittedly require them all. In all other respects, however, these higher approximations are from the beginning carried out exactly as for simple proper values.

Of course it may happen, as was mentioned above, that the value E_k, either generally or at the initial stages of the approximation, is not completely split up, and that multiplicities (" degeneracies ") still remain. This is expressed by the fact that to the substitutions already frequently mentioned there still clings a certain indefiniteness, which either always remains, or is removed step by step in the later approximations.

Let us now represent these ideas by formulae, and consider as before the perturbation caused by (4), § 1,

(4) $-\lambda r(x)y,$

i.e. we imagine the proper value problem belonging to (2) *solved*, and now consider the exactly corresponding problem (2'),

(2') $L[y] - \lambda ry + E\rho y = 0.$

We again fix our attention on a definite proper value E_k. Let (16) be a system of proper functions belonging to it, which we assume to be normalised and orthogonal to one another in the sense described above, but *not yet* fitted to the particular perturbation in the sense explained, because to find the substitution that leads *to this fitting* is precisely our chief task ! In place of (5), § 1, we must now put for the perturbed quantities the following,

(18) $E^*_{kl} = E_k + \lambda\epsilon_l; \quad u^*_{kl}(x) = \sum_{i=1}^{a} \kappa_{li} u_{ki}(x) + \lambda v_l(x)$

$$(l = 1, 2, 3 \ldots a),$$

wherein the $v_l(x)$'s are functions, and the ϵ_l's and the κ_{li}'s are systems of constants, which are still to be defined, but which we initially do not limit in any way, although we know that the system of coefficients κ_{li} must[1] form an orthogonal substitution. The index k should still be attached to the three types of quantity named, in order to indicate that the whole discussion refers to the kth proper value of the unperturbed problem. We refrain from carrying this out, in order to avoid the confusing accumulation of indices. The index k is to be assumed *fixed* in the whole of the following discussion, until the contrary is stated.

Let us select *one* of the perturbed proper functions and values by giving a definite value to the index l in (18), and let us substitute from (18) in the differential equation (2') and arrange in powers of λ. Then the terms independent of λ disappear exactly as in § 1, because the unperturbed proper quantities satisfy equation (2),

---

[1] It follows from the general theory that the perturbed system of functions $u^*_{kl}(x)$ *must* be orthogonal if the perturbation completely removes the degeneracy, and *may* be assumed orthogonal although that is not the case.

by hypothesis. Only terms containing the *first* power of λ remain, as we can strike out the others. Omitting a factor λ, we get

$$(19) \qquad L[v_l] + E_k \rho v_l = \sum_{i=1}^{a} \kappa_{li}(r - \epsilon_l \rho) u_{ki},$$

and thus obtain again for the definition of the perturbation v_l of the *functions* a *non-homogeneous* equation, to which corresponds as homogeneous equation the equation (2), with the particular value $E = E_k$, *i.e.* the equation satisfied by the set of functions u_{ki}; $i = 1, 2, \ldots a$. The form of the left side of equation (19) is independent of the index l.

On the right side occur ϵ_l and κ_{li}, the constants to be defined, and we are thus enabled to evaluate them, even *before* calculating v_l. For, in order that (19) should have a solution at all, it is necessary and sufficient that its right-hand side should be orthogonal to *all* the proper functions of the homogeneous equation (2) belonging to E_k. Therefore, we must have.

$$(20) \qquad \begin{cases} \displaystyle\sum_{i=1}^{a} \kappa_{li} \int (r - \epsilon_l \rho) u_{ki} u_{km} dx = 0 \\ (m = 1, 2, 3 \ldots a), \end{cases}$$

i.e. on account of the normalisation (17),

$$(21) \qquad \begin{cases} \kappa_{lm}\epsilon_l = \displaystyle\sum_{i=1}^{a} \kappa_{li} \int r u_{ki} u_{km} dx \\ (m = 1, 2, 3 \ldots a). \end{cases}$$

If we write, briefly, for the *symmetrical* matrix of constants, which can be evaluated by quadrature,

$$(22) \qquad \begin{cases} \displaystyle\int r u_{ki} u_{km} dx = \epsilon_{im} \\ (i, \ m = 1, 2, 3 \ldots a), \end{cases}$$

then we recognise in

$$(21') \qquad \begin{cases} \kappa_{lm}\epsilon_l = \displaystyle\sum_{i=1}^{a} \kappa_{li}\epsilon_{mi} \\ (m = 1, 2, 3 \ldots a) \end{cases}$$

a system of a linear homogeneous equations for the calculation of the a constants κ_{lm}; $m = 1, 2 \ldots a$, where the perturbation ϵ_l of the proper value still occurs in the coefficients, and is itself unknown. However, this serves for the calculation of ϵ_l before that of the κ_{lm}'s. For it is known that the linear homogeneous system (21') of equations has solutions if, and only if, its determinant vanishes. This yields the following algebraic equation of degree a for ϵ_l:

$$(23) \qquad \begin{vmatrix} \epsilon_{11} - \epsilon_l, & \epsilon_{12} & , \ldots & \epsilon_{1a} \\ \epsilon_{21} & , \epsilon_{22} - \epsilon_l, & \ldots & \epsilon_{21} \\ \cdot \cdot \cdot & \cdot \cdot \cdot & \cdot \cdot \cdot & \cdot \cdot \\ \cdot \cdot \cdot & \cdot \cdot \cdot & \cdot \cdot \cdot & \cdot \cdot \\ \epsilon_{a1} & , \epsilon_{a2} & , \ldots & \epsilon_{aa} - \epsilon_l \end{vmatrix} = 0.$$

We see that the problem is completely identical with the transformation of the quadratic form in a variables, with coefficients ϵ_{mi}, to its principal axes. The " secular equation " (23) yields a roots for ϵ_l, the " reciprocal of the squares of the principal axes ", which in general are different, and on account of the symmetry of the ϵ_{mi}'s *always real*. We thus get all the a perturbations of the proper values ($l=1, 2 \ldots a$) at the same time, and would have *inferred* the splitting up of an a-fold proper value into exactly a simple values, generally different, even had we not assumed it already, as fairly obvious. For *each* of these ϵ_l-values, equations (21') give a system of quantities κ_{li}; $i=1, 2, \ldots a$, and, as is known, *only one* (apart from a general constant factor), provided all the ϵ_l's are really different. Further, it is known that the whole system of a^2 quantities κ_{li} forms an *orthogonal* system of coefficients, defining as usual, in the principal axes problem, the *directions* of the new co-ordinate axes with reference to the old ones. We may, and will, employ the undefined factors just mentioned to normalise the κ_{li}'s completely as " direction cosines ", and this, as is easily seen, makes the perturbed proper functions $u^*{}_{ki}(x)$ turn out *normalised* again, according to (18), at least in the " zero approximation " (*i.e.* apart from the λ-terms).

If the equation (23) has multiple roots, then we have the case previously mentioned, when the perturbation does not completely remove the degeneration. The perturbed equation has then multiple proper values also and the definition of the constants κ_{li} becomes partially arbitrary. This has no consequence other than that (as is *always* the case with multiple proper values) we *must* and *may* acquiesce, even *after* the perturbation is applied, in a system of proper functions which in many respects is still arbitrary.

The main task is accomplished with this transformation to principal axes, and we will often find it sufficient in the applications in quantum theory to define the proper values to a first and the functions to zero approximation. The evaluation of the constants κ_{li} and ϵ_{li} cannot be carried out always, since it depends on the solution of an algebraic equation of degree a. At the worst there are methods [1] which give the evaluation to any desired approximation by a rational process. We may thus regard these constants as known, and will now give the calculation of the functions to the *first* approximation, for the sake of completeness. The procedure is exactly as in § 1.

We have to solve equation (19) and to that end we write v_l as a series of the *whole set* of proper functions of (2),

$$(24) \qquad v_l(x) = \sum_{(k' i')} \gamma_{l, k' i'} u_{k' i'}(x).$$

The summation is to extend with respect to k' from 0 to ∞, and, for each fixed value of k', for i' varying over the finite number of proper functions which belong to $E_{k'}$. (Now, for the first time, we take account of proper functions which do *not* belong to the a-fold value

---

[1] Courant-Hilbert, chap. i. § 3. 3, p. 14.

E_k we are fixing our attention on.)   Secondly, we develop the right-hand side of (19), divided by $\rho(x)$, in a series of the entire set of proper functions,

(25)
$$\sum_{i=1}^{a} \kappa_{li}\left(\frac{r}{\rho} - \epsilon_l\right)u_{ki} = \sum_{(k'i')} c_{l,k'i'}u_{k'i'},$$

wherein

(26)
$$\begin{cases} c_{l,k'i'} = \sum_{i=1}^{a} \kappa_{li}\int (r - \epsilon_l\rho)u_{ki}u_{k'i'}dx \\[2mm] \qquad = \sum_{i=1}^{a} \kappa_{li}\int ru_{ki}u_{k'i'}dx \text{ for } k' \neq k \\[2mm] \qquad = 0 \qquad\qquad\qquad \text{for } k' = k \end{cases}$$

(the last two equalities follow from (17) and (20) respectively).   On substituting from (24) and (25) in (19), we get

(27)
$$\sum_{(k'i')} \gamma_{l,k'i'}(L[u_{k'i'}] + E_k\rho u_{k'i'}) = \sum_{(k'i')} c_{l,k'i'}\rho u_{k'i'}.$$

Since $u_{k'i'}$ satisfies equation (2) with $E = E_{k'}$, this gives

(28)
$$\sum_{(k'i')} \gamma_{l,k'i'}\rho(E_k - E_{k'})u_{k'i'} = \sum_{(k'i')} c_{l,k'i'}\rho u_{k'i'}.$$

By equating coefficients on right and left, all the $\gamma_{l,k'i'}$'s are defined, with the exception of those in which $k' = k$.   Thus

(29)
$$\gamma_{l,k'i'} = \frac{c_{l,k'i'}}{E_k - E_{k'}} = \frac{1}{E_k - E_{k'}} \sum_{i=1}^{a} \kappa_{li}\int ru_{ki}u_{k'i'}dx \text{ (for } k' \neq k),$$

while those γ's for which $k' = k$ are of course not fixed by equation (19). This again corresponds to the fact that we have provisionally normalised the perturbed functions u^*_{kl}, of (18), only in the zero approximation (through the normalisation of the κ_{li}'s), and it is easily recognised again that we have to put the whole of the γ-quantities in question equal to zero, in order to bring about the normalisation of the u^*_{kl}'s even in the first approximation.   By substituting from (29) in (24), and then from (24) in (18), we finally obtain for the *perturbed proper functions to a first approximation*

(30)
$$u^*_{kl}(x) = \sum_{i=1}^{a} \kappa_{li}\left(u_{ki}(x) + \lambda \sum_{(k'i')}' \frac{u_{k'i'}(x)}{E_k - E_{k'}}\int ru_{ki}u_{k'i'}dx\right)$$
$$(l = 1, 2, \ldots, a).$$

The dash on the second sigma indicates that *all* the terms with $k' = k$ are to be omitted.   In the application of the formula for an arbitrary k, it is to be observed that the κ_{li}'s, as obviously also the multiplicity a of the proper value E_k, to which we have specially directed our attention, still depend on the index k, though this is not expressed in the symbols.   Let us repeat here that the κ_{li}'s are to be calculated as a system of solutions of equations (21'), normalised so that the sum of the squares is unity, where the coefficients of the equations are given by (22), while for the quantity ϵ_l in (21'), *one* of the roots of

(23) is to be taken. *This* root then gives the allied perturbed proper *value*, from

(31) $E^*_{kl} = E_k + \lambda\epsilon_l.$

Formulae (30) and (31) are the generalisations of (14) and (15) of § 1.

It need scarcely be said that the extensions and generalisations mentioned at the end of § 1 can of course take effect here also. It is hardly worth the trouble to carry out these developments generally. We succeed best in any special case if we do not use ready-made formulae, but go directly by the simple fundamental principles, which have been explained, perhaps too minutely, in the present paper. I would only like to consider briefly the possibility, already mentioned at the end of § 1, that the equation (2) perhaps may lose (and indeed in the case of several variables irreparably lose), its self-adjoint character if the perturbing terms also contain derivatives of the unknown function. From general theorems we know that then the proper values of the perturbed equation no longer need to be real. We can illustrate this further. We can easily see, by carrying out the developments of this paragraph, that the elements of determinant (23) are *no longer symmetrical*, when the perturbing term contains derivatives. It is known that in this case the roots of equation (23) no longer require to be real.

The necessity for the expansion of certain functions in a series of proper functions, in order to arrive at the first or zero approximation of the proper values or functions, can become very inconvenient, and can at least complicate the calculation considerably in cases where an extended spectrum co-exists with the point spectrum and where the point spectrum has a limiting point (point of accumulation) at a finite distance. This is just the case in the problems appearing in the quantum theory. Fortunately it is often—perhaps always—possible, for the purpose of the perturbation theory, to free oneself from the generally very troublesome extended spectrum, and to develop the perturbation theory from an equation which does *not* possess such a spectrum, and whose proper values do *not* accumulate near a finite value, but grow beyond all limits with increasing index. We will become acquainted with an example in the next paragraph. Of course, this simplification is only possible when we are not interested in a proper value of the extended spectrum.

II. APPLICATION TO THE STARK EFFECT

§ 3. Calculation of Frequencies by the Method which corresponds to that of Epstein

If we add a potential energy $+eFz$ to the wave equation (5), Part I. (p. 2), of the Kepler problem, corresponding to the influence of an electric field of strength F in the positive z-direction, on a negative

electron of charge e, then we obtain the following wave equation for the Stark effect of the hydrogen atom,

$$(32) \qquad \nabla^2\psi + \frac{8\pi^2 m}{h^2}\Big(E + \frac{e^2}{r} - eFz\Big)\psi = 0,$$

which forms the basis of the remainder of this paper. In § 5 we will apply the general perturbation theory of § 2 directly to this partial differential equation. Now, however, we will lighten our task by introducing space parabolic co-ordinates λ_1, λ_2, ϕ, by the following equations,

$$(33) \qquad \begin{cases} x = \sqrt{\lambda_1\lambda_2}\,\cos\phi \\ \quad + \\ y = \sqrt{\lambda_1\lambda_2}\,\sin\phi \\ \quad + \\ z = \tfrac{1}{2}(\lambda_1 - \lambda_2). \end{cases}$$

λ_1 and λ_2 run from 0 to infinity; the corresponding co-ordinate surfaces are the two sets of confocal paraboloids of revolution, which have the origin as focus and the positive (λ_2) or negative (λ_1) z-axis respectively as axes. ϕ runs from 0 to 2π, and the co-ordinate surfaces belonging to it are the set of half planes limited by the z-axis. The relation of the co-ordinates is *unique*. For the functional determinant we get

$$(34) \qquad \frac{\partial(x, y, z)}{\partial(\lambda_1, \lambda_2, \phi)} = \tfrac{1}{4}(\lambda_1 + \lambda_2).$$

The *space element* is thus

$$(35) \qquad dx\,dy\,dz = \tfrac{1}{4}(\lambda_1 + \lambda_2)d\lambda_1 d\lambda_2 d\phi.$$

We notice, as consequences of (33),

$$(36) \qquad x^2 + y^2 = \lambda_1\lambda_2\ ;\ \ r^2 = x^2 + y^2 + z^2 = \{\tfrac{1}{2}(\lambda_1 + \lambda_2)\}^2.$$

The expression of (32) in the chosen co-ordinates gives, if we multiply by (34) [1] (to restore the self-adjoint form),

$$(32') \qquad \begin{cases} \dfrac{\partial}{\partial\lambda_1}\Big(\lambda_1\dfrac{\partial\psi}{\partial\lambda_1}\Big) + \dfrac{\partial}{\partial\lambda_2}\Big(\lambda_2\dfrac{\partial\psi}{\partial\lambda_2}\Big) + \tfrac{1}{4}\Big(\dfrac{1}{\lambda_1} + \dfrac{1}{\lambda_2}\Big)\dfrac{\partial^2\psi}{\partial\phi^2} \\ \qquad\qquad + \dfrac{2\pi^2 m}{h^2}[E(\lambda_1 + \lambda_2) + 2e^2 - \tfrac{1}{2}eF(\lambda_1^2 - \lambda_2^2)]\psi = 0. \end{cases}$$

Here we can again take—and this is the why and wherefore of all " methods " of solving linear partial differential equations—the function ψ as the product of three functions, thus,

$$(37) \qquad \psi = \Lambda_1\Lambda_2\Phi,$$

---

[1] So far as the actual details of the analysis are concerned, the simplest way to get (32'), or, in general, to get the wave equation for any special co-ordinates, is to transform not the wave equation itself, but the corresponding variation problem (cf. Part I. p. 12), and thus to obtain the wave equation afresh as an Eulerian variation problem. We are thus spared the troublesome evaluation of the *second* derivatives. Cf. Courant-Hilbert, chap. iv. § 7, p. 193.

each of which depends on only *one* co-ordinate. For these functions
we get the ordinary differential equations

$$(38) \quad \begin{cases} \dfrac{\partial^2 \Phi}{\partial \phi^2} = -n^2 \Phi \\[2ex] \dfrac{\partial}{\partial \lambda_1}\left(\lambda_1 \dfrac{\partial \Lambda_1}{\partial \lambda_1}\right) + \dfrac{2\pi^2 m}{h^2}\left(-\tfrac{1}{2}eF\lambda_1{}^2 + E\lambda_1 + e^2 - \beta - \dfrac{n^2 h^2}{8\pi^2 m}\dfrac{1}{\lambda_1}\right)\Lambda_1 = 0, \\[3ex] \dfrac{\partial}{\partial \lambda_2}\left(\lambda_2 \dfrac{\partial \Lambda_2}{\partial \lambda_2}\right) + \dfrac{2\pi^2 m}{h^2}\left(\tfrac{1}{2}eF\lambda_2{}^2 + E\lambda_2 + e^2 + \beta - \dfrac{n^2 h^2}{8\pi^2 m}\dfrac{1}{\lambda_2}\right)\Lambda_2 = 0, \end{cases}$$

wherein n and β are two further " proper value-like " constants of in-
tegration (in addition to E), still to be defined. By the choice of
symbol for the first of these, we have taken into account the fact
that the first of equations (38) makes it take integral values, if Φ and
$\dfrac{\partial \Phi}{\partial \phi}$ are to be continuous and single-valued functions of the azimuth ϕ.
We then have

$$(39) \qquad\qquad \Phi = \genfrac{}{}{0pt}{}{\sin}{\cos} n\phi$$

and it is evidently sufficient if we do not consider negative values
of n. Thus

$$(40) \qquad\qquad n = 0, 1, 2, 3 \ldots .$$

In the symbol used for the second constant β, we follow Sommerfeld
(*Atombau*, 4th edit., p. 821) in order to make comparison easier.
(Similarly, below, with A, B, C, D.) We treat the last two equations
of (38) together, in the form

$$(41) \qquad \frac{\partial}{\partial \xi}\left(\xi \frac{\partial \Lambda}{\partial \xi}\right) + \left(D\xi^2 + A\xi + 2B + \frac{C}{\xi}\right)\Lambda = 0,$$

where

$$(42) \quad \left.\begin{matrix} D_1 \\ D_2 \end{matrix}\right\} = \mp \frac{\pi^2 m e F}{h^2}, \quad A = \frac{2\pi^2 m E}{h^2}, \quad \left.\begin{matrix} B_1 \\ B_2 \end{matrix}\right\} = \frac{\pi^2 m}{h^2}(e^2 \mp \beta), \quad C = -\frac{n^2}{4},$$

and the upper sign is valid for $\Lambda = \Lambda_1$, $\xi = \lambda_1$, and the lower one for
$\Lambda = \Lambda_2$, $\xi = \lambda_2$. (Unfortunately, we have to write ξ instead of the
more appropriate λ, to avoid confusion with the perturbation para-
meter λ of the general theory, §§ 1 and 2.)

If we omit initially in (41) the Stark effect term $D\xi^2$, which we
conceive as a perturbing term (limiting case for vanishing field), then
this equation has the same general structure as equation (7) of Part I.,
and the domain is also the same, from 0 to ∞. The discussion is almost
the same, word for word, and shows that non-vanishing solutions,
which, with their derivatives, are continuous and remain finite within
the domain, only exist if *either* $A > 0$ (extended spectrum, correspond-
ing to hyperbolic orbits) *or*

$$(43) \qquad \frac{B}{\underset{+}{\sqrt{-A}}} - \underset{+}{\sqrt{-C}} = k + \tfrac{1}{2}; \quad k = 0, 1, 2, \ldots$$

If we apply this to the last two equations of (38) and distinguish the two k-values by suffixes 1 and 2, we obtain

$$(44) \quad \begin{cases} \sqrt{-A}(k_1 + \tfrac{1}{2} + \sqrt{-C}) = B_1 \\ \phantom{\sqrt{}}{+}{+} \\ \sqrt{-A}(k_2 + \tfrac{1}{2} + \sqrt{-C}) = B_2. \\ \phantom{\sqrt{}}{+}{+} \end{cases}$$

By addition, squaring and use of (42) we find

$$(45) \qquad A = -\frac{4\pi^4 m^2 e^4}{h^4 l^2} \quad \text{and} \quad E = -\frac{2\pi^2 m e^4}{h^2 l^2}.$$

These are the well-known Balmer-Bohr elliptic levels, where as *principal quantum number* enters

$$(46) \qquad l = k_1 + k_2 + n + 1.$$

We get the *discrete* term spectrum and the allied proper functions in a way *simpler* than that indicated, if we apply results already known in mathematical literature as follows. We transform first the dependent variable Λ in (41) by putting

$$(47) \qquad \Lambda = \xi^{\frac{n}{2}} u$$

and then the independent ξ by putting

$$(48) \qquad 2\xi\sqrt{-A} = \eta.$$

We find for u as a function of η the equation

$$(41') \qquad \frac{d^2 u}{d\eta^2} + \frac{n+1}{\eta}\frac{du}{d\eta} + \left(\frac{D}{(2\sqrt{-A})^3}\eta - \tfrac{1}{4} + \frac{B}{\sqrt{-A}}\frac{1}{\eta}\right)u = 0.$$

This equation is very intimately connected with the polynomials named after Laguerre. In the mathematical appendix, it will be shown that the product of $e^{-\frac{x}{2}}$ and the nth derivative of the $(n+k)$th Laguerre polynomial satisfies the differential equation

$$(103) \qquad y'' + \frac{n+1}{x}y' + \left(-\tfrac{1}{4} + \left(k + \frac{n+1}{2}\right)\frac{1}{x}\right)y = 0,$$

and that, for a fixed n, the functions named form the complete system of proper functions of the equation just written, when k runs through all non-negative integral values. Thus it follows that, for vanishing D, equation (41') possesses the proper functions

$$(49) \qquad u_k(\eta) = e^{-\frac{\eta}{2}} L_{n+k}^n(\eta)$$

and the proper values

$$(50) \qquad \frac{B}{\sqrt{-A}} = \frac{n+1}{2} + k \quad (k = 0, 1, 2 \ldots)$$

—and no others! (See the mathematical appendix concerning the remarkable loss of the extended spectrum caused by the apparently inoffensive transformation (48); by this loss the development of the perturbation theory is made much easier.)

We have now to calculate the perturbation of the proper values (50) from the general theory of § 1, caused by including the D-term in (41'). The equation becomes self-adjoint if we multiply by η^{n+1}. The density function $\rho(x)$ of the general theory thus becomes η^n. As perturbation function $r(x)$ appears

$$(51) \qquad -\frac{D}{(2\sqrt{-A})^3}\eta^{n+2}.$$

(We formally put the perturbation parameter $\lambda = 1$; if we desired, we could identify D or F with it.) Now formula (7') gives, for the perturbation of the kth proper value,

$$(52) \qquad \epsilon_k = -\frac{D}{(2\sqrt{-A})^3}\frac{\int_0^\infty \eta^{n+2}e^{-\eta}[L_{n+k}^n(\eta)]^2 d\eta}{\int_0^\infty \eta^n e^{-\eta}[L_{n+k}^n(\eta)]^2 d\eta}.$$

For the integral in the denominator, which merely provides for the normalisation, formula (115) of the appendix gives the value

$$(53) \qquad \frac{[(n+k)!]^3}{k!},$$

while the integral in the numerator is evaluated in the same place, as

$$(54) \qquad \frac{[(n+k)!]^3}{k!}(n^2 + 6nk + 6k^2 + 6k + 3n + 2).$$

Consequently

$$(55) \qquad \epsilon_k = -\frac{D}{(2\sqrt{-A})^3}(n^2 + 6nk + 6k^2 + 6k + 3n + 2).$$

The condition for the kth perturbed proper value of equation (41') and therefore, naturally, also for the kth discrete proper value of the original equation (41) runs therefore

$$(56) \qquad \frac{B}{\sqrt{-A}} = \frac{n+1}{2} + k + \epsilon_k$$

(ϵ_k is retained meantime for brevity).

This result is applied twice, namely, to the last two equations of (38) by substituting the two systems (42) of values of the constants A, B, C, D; and it is to be observed that n is the *same* number in the two cases, while the two k-values are to be distinguished by the suffixes 1 and 2, as above. First we have

$$(57) \qquad \begin{cases} \dfrac{B_1}{\sqrt{-A}} = \dfrac{n+1}{2} + k_1 + \epsilon_{k1} \\[2ex] \dfrac{B_2}{\sqrt{-A}} = \dfrac{n+1}{2} + k_2 + \epsilon_{k2}, \end{cases}$$

whence comes

(58)
$$A = -\frac{(B_1 + B_2)^2}{(l + \epsilon_{k1} + \epsilon_{k2})^2}$$

(applying abbreviation (46) for the principal quantum number). In the approximation we are aiming at we may expand with respect to the small quantities ϵ_k and get

(59)
$$A = -\frac{(B_1 + B_2)^2}{l^2}\left[1 - \frac{2}{l}(\epsilon_{k1} + \epsilon_{k2})\right].$$

Further, in the calculation of these small quantities, we may use the approximate value (45) for A in (55). We thus obtain, noticing the two D values, by (42),

(60)
$$\begin{cases} \epsilon_{k1} = +\frac{Fh^4l^3}{64\pi^4m^2e^5}(n^2 + 6nk_1 + 6k_1{}^2 + 6k_1 + 3n + 2) \\[2mm] \epsilon_{k2} = -\frac{Fh^4l^3}{64\pi^4m^2e^5}(n^2 + 6nk_2 + 6k_2{}^2 + 6k_2 + 3n + 2). \end{cases}$$

Addition gives, after an easy reduction,

(61)
$$\epsilon_{k1} + \epsilon_{k2} = \frac{3Fh^4l^4(k_1 - k_2)}{32\pi^4m^2e^5}.$$

If we substitute this, and the values of A, B_1, and B_2 from (42) in (59), we get, after reduction,

(62)
$$E = -\frac{2\pi^2me^4}{h^2l^2} - \frac{3}{8}\frac{h^2Fl(k_2 - k_1)}{\pi^2me}.$$

This is our provisional *conclusion*; it is the well-known formula of Epstein for the term values in the Stark effect of the hydrogen spectrum.

k_1 and k_2 correspond fully to the parabolic quantum numbers; they are capable of taking the value zero. Also the integer n, which has evidently to do with the *equatorial* quantum number, may from (40) take the value zero. However, from (46) the sum of these three numbers must still be increased by unity in order to yield the principal quantum number. Thus $(n + 1)$ and not n corresponds to the equatorial quantum number. The value zero for the *latter* is thus *automatically* excluded by wave mechanics, just as by Heisenberg's mechanics.[1] *There is simply no proper function, i.e.* no state of vibration, which corresponds to such a meridional orbit. This important and gratifying circumstance was already brought to light in Part I. in counting the constants, and also afterwards in § 2 of Part I. in connection with the azimuthal quantum number, through the non-existence of states of vibration corresponding to *pendulum orbits*; its full meaning, however, only fully dawned on me through the remarks of the two authors just quoted.

[1] W. Pauli, jun., *Ztschr. f. Phys.* 36, p. 336, 1926; N. Bohr, *Die Naturw.* 1, 1926.

For later application, let us note the system of proper functions of equation (32) or (32') in "zero approximation", which belongs to the proper values (62). It is obtained from statement (37), from conclusions (39) and (49), and from consideration of transformations (47) and (48) and of the approximate value (45) of A. For brevity, let us call a_0 the "radius of the first hydrogen orbit". Then we get

$$(63) \qquad \frac{1}{2l\sqrt{-A}} = \frac{h^2}{4\pi^2 m e^2} = a_0.$$

The proper functions (not yet normalised!) then read

$$(64) \qquad \psi_{nk_1k_2} = \lambda_1^{\frac{n}{2}}\lambda_2^{\frac{n}{2}} e^{-\frac{\lambda_1+\lambda_2}{2la_0}} L_{n+k_1}^n\!\left(\frac{\lambda_1}{la_0}\right) L_{n+k_2}^n\!\left(\frac{\lambda_2}{la_0}\right) \frac{\sin}{\cos} n\phi.$$

They belong to the proper values (62), where l has the meaning (46). To each non-negative integral trio of values n, k_1, k_2 belong $\left(\text{on account of the double symbol } \frac{\sin}{\cos}\right)$ two proper functions or one, according as $n > 0$ or $n = 0$.

§ 4. Attempt to calculate the Intensities and Polarisations of the Stark Effect Patterns

I have lately shown[1] that from the proper functions we can calculate by differentiation and quadrature the elements of the *matrices*, which are allied in Heisenberg's mechanics to functions of the generalised position- and momentum-co-ordinates. For example, for the (rr')th element of the matrix, which according to Heisenberg belongs to the generalised co-ordinate q itself, we find

$$(65) \quad \begin{cases} q^{rr'} = \int q\rho(x)\psi_r(x)\psi_{r'}(x)dx \\ \\ \qquad\qquad \cdot \left\{ \int \rho(x)[\psi_r(x)]^2 dx \cdot \int \rho(x)[\psi_{r'}(x)]^2 dx \right\}^{-\frac{1}{2}}. \end{cases}$$

Here, for our case, the separate indices *each* deputise for a *trio* of indices n, k_1, k_2, and further, x represents the three co-ordinates r, θ, ϕ. $\rho(x)$ is the density function; in our case the quantity (34). (We may compare the self-adjoint equation (32') with the general form (2)). The "denominator" $(\ldots)^{-\frac{1}{2}}$ in (65) must be put in because our system (64) of functions is not yet normalised.

According to Heisenberg,[2] now, if q means a rectangular Cartesian co-ordinate, then the square of the matrix element (65) is to be a measure of the "probability of transition" from the rth state to the r'th, or, more accurately, a measure of the intensity of that part of the radiation, bound up with this transition, which is polarised in the q-direction. Starting from this, I have shown in the above paper that if we make

[1] Preceding paper of this collection.
[2] W. Heisenberg, *Ztschr. f. Phys.* 33, p. 879, 1925; M. Born and P. Jordan, *Ztschr. f. Phys.* 34, pp. 867, 886, 1925.

certain simple assumptions as to the electrodynamical meaning of ψ, the " mechanical field scalar ", then the matrix element in question is susceptible of a very simple physical interpretation in wave mechanics, namely, *actually*: component of the amplitude of the periodically oscillating electric moment of the atom.   The word *component* is to be taken in a double sense : (1) component in the q-direction, *i.e.* in the spatial direction in question, and (2) only the part of this spatial component which changes in a time-sinusoidal manner with exactly the frequency of the emitted light, $|E_r - E_{r'}|/h$.   (It is a question then of a kind of Fourier analysis : not in harmonic frequencies, but in the actual frequencies of emission.)   However, the idea of wave mechanics is not that of a sudden transition from one state of vibration to another, but according to it, the partial moment concerned—as I will briefly name it—arises from the *simultaneous existence* of the two proper vibrations, and lasts just as long as both are excited together.

Moreover, the above assertion that the $q^{rr'}$'s are proportional to the partial moments is more accurately phrased thus.   The ratio of, *e.g.*, $q^{rr'}$ to $q^{rr''}$ is equal to the ratio of the partial moments which arise when the proper function ψ_r and the proper functions $\psi_{r'}$ and $\psi_{r''}$ are stimulated, the first *with any strength whatever* and the last two with strengths *equal to one another*—*i.e.* corresponding to normalisation. To calculate the ratio of the *intensities*, the q-quotient must first be squared and then multiplied by the ratio of the fourth powers of the emission frequencies.   The latter, however, has no part in the intensity ratio of the Stark effect components, for there we only compare intensities of lines which have practically the same frequency.

The known *selection* and *polarisation rules* for Stark effect components can be obtained, almost without calculation, from the integrals in the numerator of (65) and from the form of the proper functions in (64).   They follow from the vanishing or non-vanishing of the integral with respect to ϕ.   We obtain the components whose electric vector vibrates *parallel* to the field, *i.e.* to the z-direction, by replacing the q in (65) by z from (33).   The expression for z, *i.e.* $\frac{1}{2}(\lambda_1 - \lambda_2)$, does *not* contain the azimuth ϕ.   Thus we see at once from (64) that a non-vanishing result after integration with respect to ϕ can only arise if we combine proper functions whose n's are *equal*, and thus whose equatorial quantum numbers are equal, being in fact equal to $n+1$.   For the components which vibrate *perpendicular* to the field, we must put q equal to x or equal to y (cf. equation (33)).   Here $\cos\phi$ or $\sin\phi$ enters, and we see almost as easily as before, that the n-values of the two combined proper functions must differ exactly by unity, if the integration with respect to ϕ is to yield a non-vanishing result.   Hence the known selection and polarisation rules are proved.   Further, it should be recalled again that we do not require to exclude any n-value after additional reflection, as was necessary in the older theory in order to agree with experience.   Our n is smaller by 1 than the equatorial quantum number, and right from the beginning cannot take negative values

(quite the same state of affairs exists, we know, in Heisenberg's theory).[1]

The numerical evaluation of the integrals with respect to λ_1 and λ_2 which appear in (65) is exceptionally tedious, especially for those of the numerator. The same apparatus for calculating comes into play as served already in the evaluation of (52), only the matter is somewhat more detailed because the two (generalised) Laguerre polynomials, whose product is to be integrated, have not the same argument. By good luck, in the *Balmer lines*, which interest us principally, one of the two polynomials L_{n+k}^n, namely that relating to the doubly quantised state, is either a constant or is a linear function of its argument. The method of calculation is described more fully in the mathematical appendix. The following tables and diagrams give the results for the first four Balmer lines, in comparison with the known measurements and estimates of intensity, made by Stark[2] for a field strength of about 100,000 volts per centimetre. The first column indicates the state of polarisation, the second gives the combination of the terms in the usual manner of description, *i.e.* in *our* symbols : of the two trios of numbers $(k_1, k_2, n+1)$ the *first* trio refers to the higher quantised state and the *second* to the doubly quantised state. The third column, with the heading Δ, gives the term decomposition in multiples of $3h^2 F/8\pi^2 me$, (see equation (62)). The next column gives the intensities observed by Stark, and 0 there signifies not observed. The question mark was put by Stark at such lines as clash either with irrelevant lines or with possible " ghosts " and thus cannot be guaranteed. On account of the unequal weakening of the two states of polarisation in the spectrograph, according to Stark his results for the $\parallel$ and for the $\perp$ components of vibration are not directly comparable with one another. Finally, the last column gives the results of our calculation in *relative numbers*, which are comparable for the collective components ($\parallel$ and $\perp$) of *one* line, *e.g.* of H_α, but not for those of H_α with H_β, etc. These relative numbers are reduced to their *smallest integral* values, *i.e.* the numbers in each of the four tables are *prime* to each other.

[1] W. Pauli, jun., *Ztschr. f. Physik*, 36, p. 336, 1926.
[2] J. Stark, *Ann. d. Phys.* 48, p. 193, 1915.

[TABLES

INTENSITIES IN THE STARK EFFECT OF THE BALMER LINES

TABLE 1

H_α

Polarisation.	Combination.	Δ	Observed Intensity.	Calculated Intensity.
‖	(111) (011)	2	1	729
	(102) (002)	3	1·1	2304
	(201) (101)	4	1·2	1681
	(201) (011)	8	0	1
				Sum : 4715
⊥	(003) (002)	0	} 2·6 {	4608
	(111) (002)	0		882
	(102) (101)	1	1	1936
	(102) (011)	5	0	16
	(201) (002)	6	0	18
				Sum * : 4715

* Undisplaced components halved.

TABLE 2

H_β

Polarisation.	Combination.	Δ	Observed Intensity.	Calculated Intensity.
‖	(112) (002)	0	1·4	0
	(211) (101)	2	1·2	9
	—	(4)	1	0
	(211) (011)	6	4·8	81
	(202) (002)	8	9·1	384
	(301) (101)	10	11·5	361
	—	(12)	1	0
	(301) (011)	14	0	1
				Sum : 836
⊥	—	(0)	1·4	0
	(112) (011)	2	3·3	72
	(103) (002)	4	} 12·6 {	384
	(211) (002)	4		72
	(202) (101)	6	9·7	294
	—	(8)	1·3	0
	(202) (011)	10	1·1 ?	6
	(301) (002)	12	1 ?	8
				Sum : 836

INTENSITIES IN THE STARK EFFECT OF THE BALMER LINES

TABLE 3

H_γ

Polarisation.	Combination.	Δ	Observed Intensity.	Calculated Intensity.
‖	(221) (011)	2	1·6	15 625
	(212) (002)	5	1·5	19 200
	(311) (101)	8	1	1 521
	(311) (011)	12	2·0	16 641
	(302) (002)	15	7·2	115 200
	(401) (101)	18	10·8	131 769
	(401) (011)	22	1 ?	729
				Sum : 300 685
⊥	(113) (002)	0	} 7·2 {	115 200
	(221) (002)	0		26 450
	(212) (101)	3	3·2	46 128
	(212) (011)	7	1·2	5 808
	(203) (002)	10	} 4·3 {	76 800
	(311) (002)	10		11 250
	(302) (101)	13	6·1	83 232
	(302) (011)	17	1·1	2 592
	(401) (002)	20	1	4 050
				Sum : * 300 685

* Undisplaced components halved.

TABLE 4

H_δ

Polarisation.	Combination.	Δ	Observed Intensity.	Calculated Intensity.
‖	(222) (002)	0	0	0
	(321) (101)	4	1	8
	(321) (011)	8	1·2	32
	(312) (002)	12	1·5	72
	(411) (101)	16	1·2	18
	(411) (011)	20	1·1	18
	(402) (002)	24	2·8	180
	(501) (101)	28	7·2	242
	(501) (011)	32	1 ?	2
				Sum: 572
⊥	(222) (011)	2	1·3	36
	(213) (002)	6	} 3·2 {	162
	(321) (002)	6		36
	(312) (101)	10	2·1	98
	(312) (011)	14	1	2
	(303) (002)	18	} 2·0 {	90
	(411) (002)	18		9
	(402) (101)	22	2·4	125
	(402) (011)	26	1·3	5
	(501) (002)	30	1 ?	9
				Sum : 572

In the *diagrams* it is to be noticed that, on account of the huge differences in the theoretical intensities, some theoretical intensities
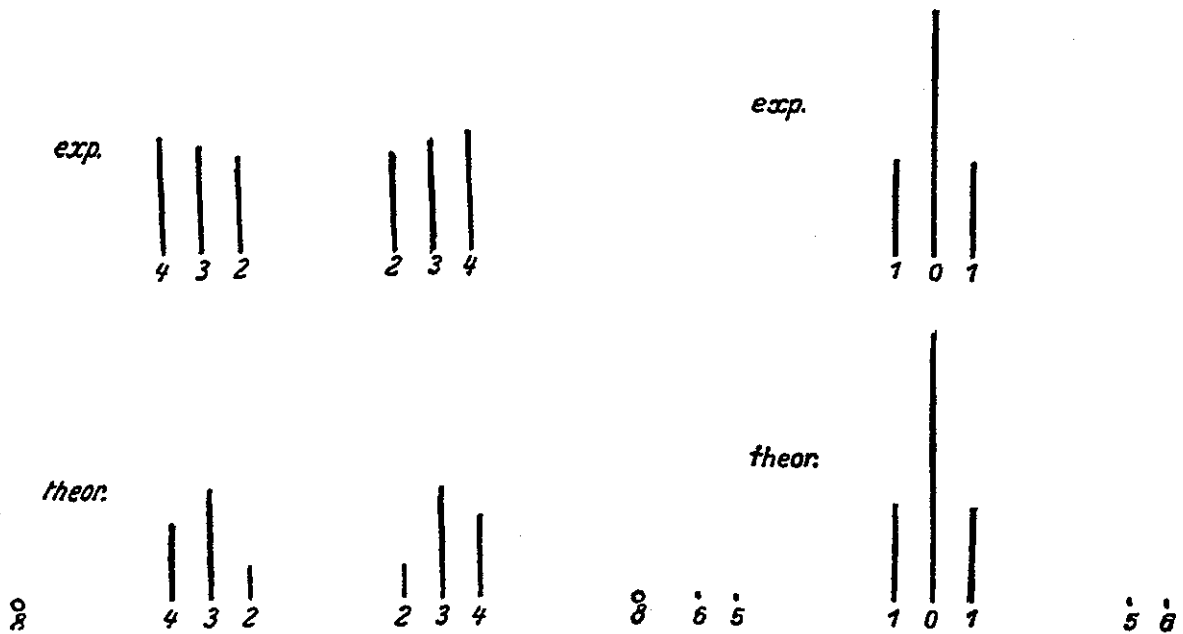
exp.

theor.

FIG. 1.—H_α ∥-components.
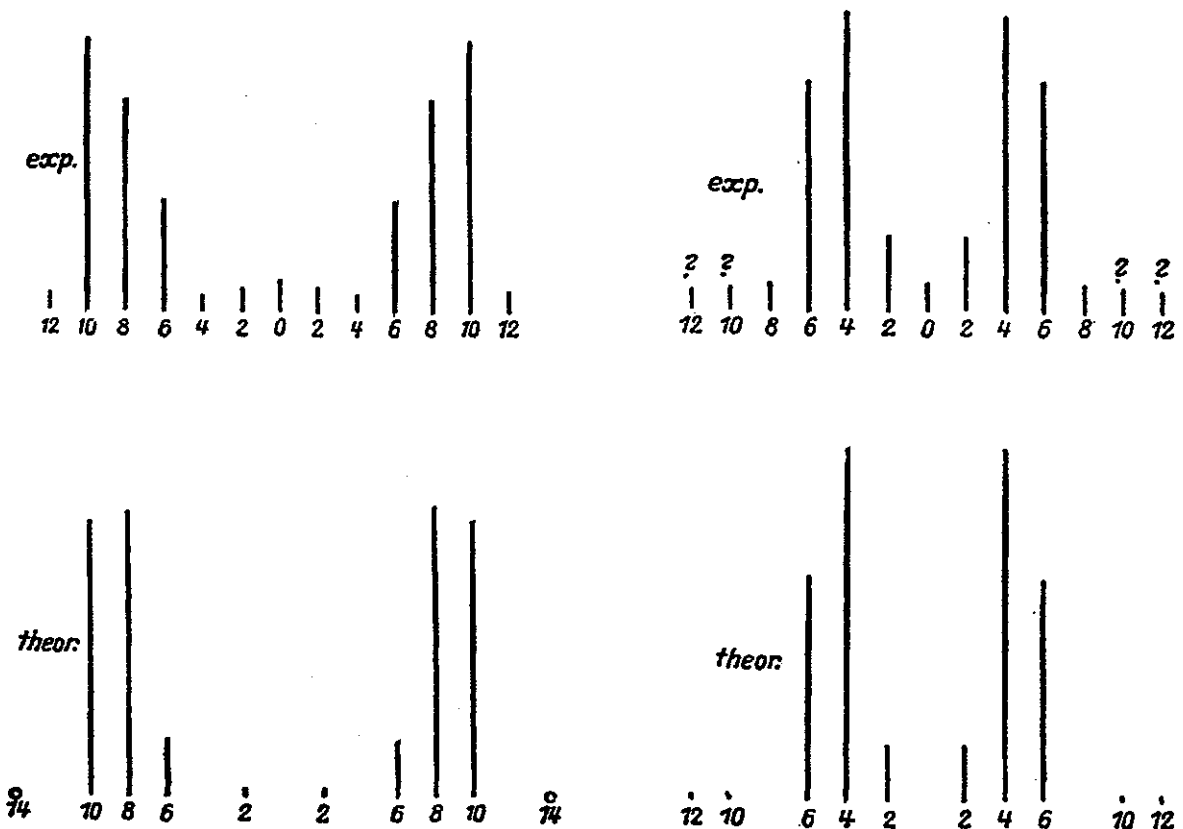
exp.

theor.

FIG. 2.—H_α ⊥-components

exp.

theor.

FIG. 3.—$H\beta$ ∥-components.

exp.

theor.

FIG. 4.—$H\beta$ ⊥-components.

cannot be truly represented to scale, as they are much too small. These are indicated by *small circles*.

A consideration of the diagrams shows that the agreement is tolerably good for almost all the strong components, and taken all over it is somewhat better than for the values deduced from correspondence considerations.[1] Thus, for example, is removed one of the most serious contradictions which arose, in that the correspondence principle gave the ratio of the intensities of the two strong $\perp$-components of H_β, for $\Delta = 4$ and 6, inversely and indeed very much out, in fact
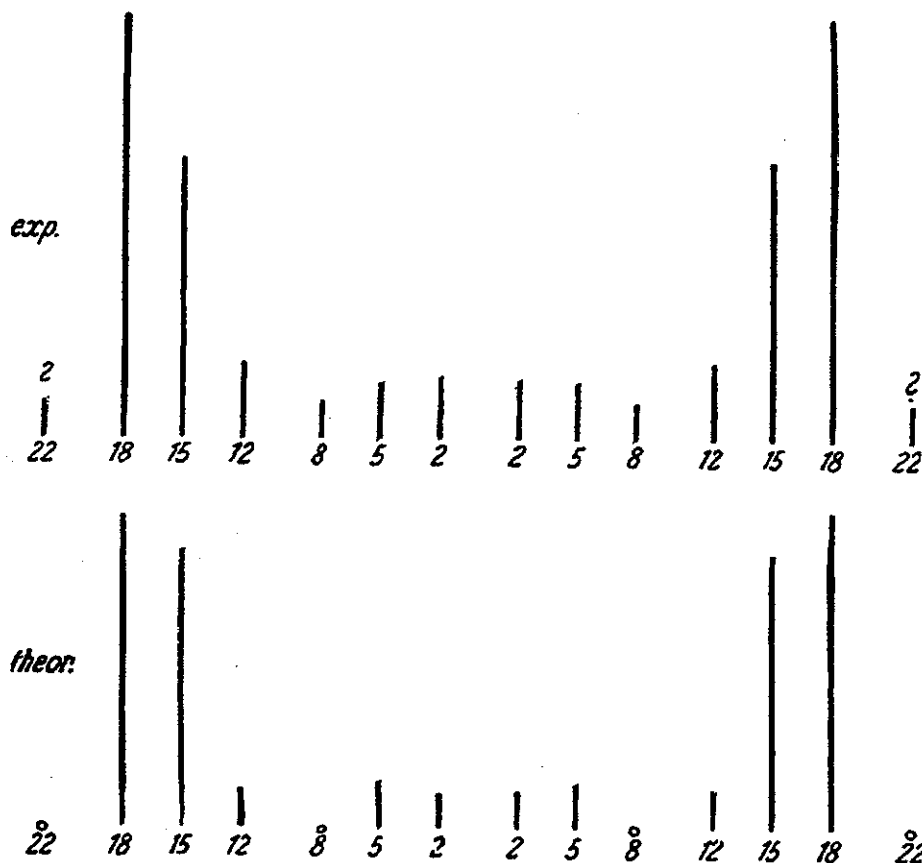
Fig. 5.—H_γ ||-components.

as almost $1 : 2$, while experiment requires about $5 : 4$. A similar thing occurs with the mean $(\Delta = 0)$ $\perp$-components of H_γ, which decidedly preponderate experimentally, but are given as *far* too weak by the correspondence principle. In *our* diagrams also, it is admitted that such "reciprocities" between the intensity ratios of intense components demanded by theory and by experiment are not entirely wanting. The theoretically most intense ||-component $(\Delta = 3)$ of H_α is furthest out; by experiment, it should lie *between* its neighbours in intensity. And the two strongest ||-components of H_β and two $\perp$-components $(\Delta = 10, 13)$ of H_γ are given "reciprocally" by the

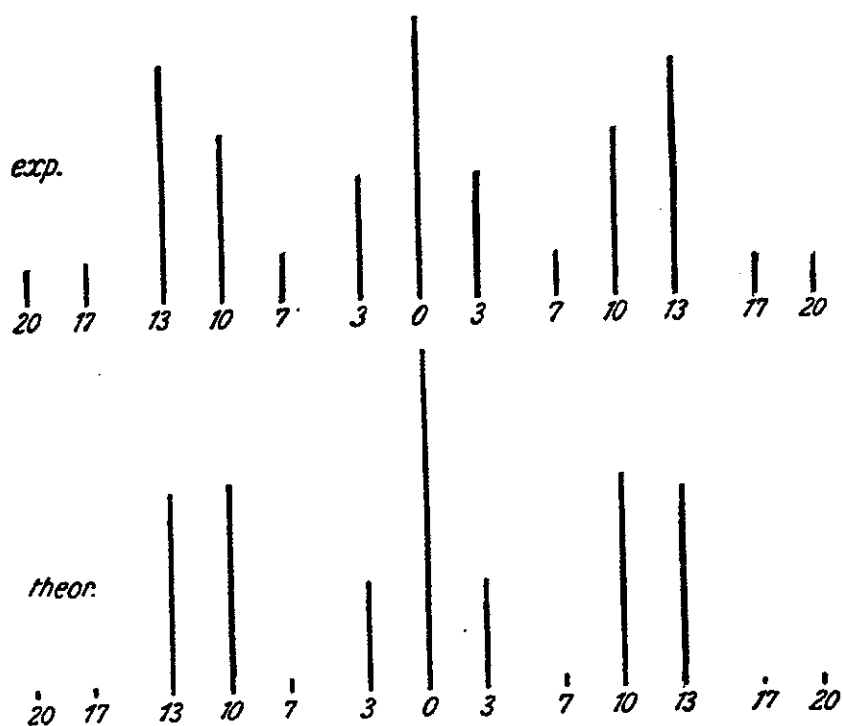[1] H. A. Kramers, *Dänische Akademie* (8), iii. 3, p. 333 *et seq.*, 1919.
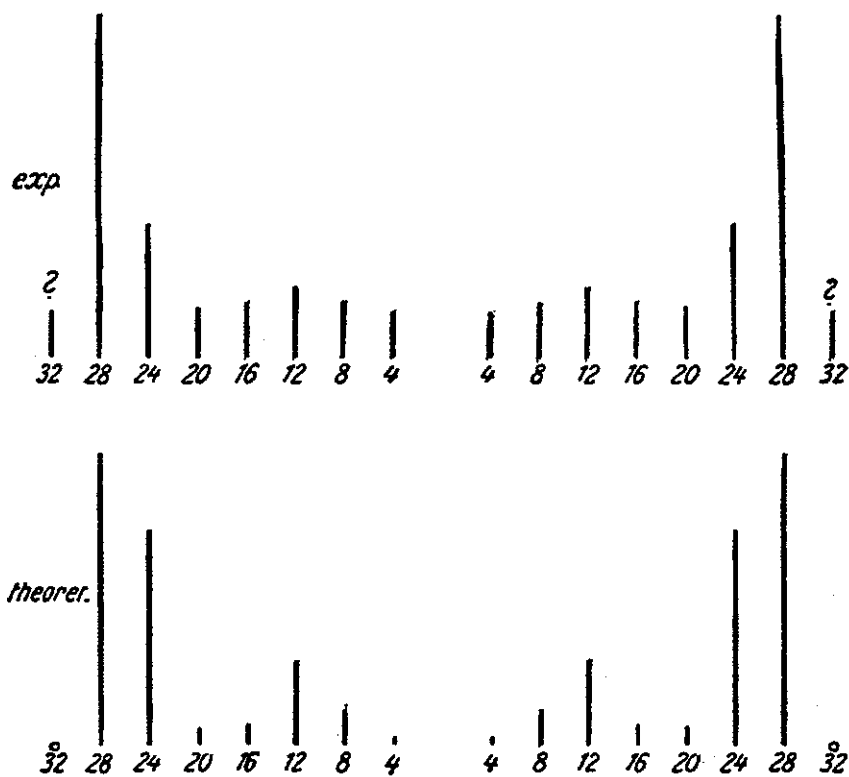
Fig. 6.—$H_\gamma \perp$-components.

Fig. 7.—$H_\delta \parallel$-components.

theory. Of course, in both cases the intensity ratios, both experimentally and theoretically, are pretty near unity.

Passing now to the weaker components, we notice first that the contradiction which exists for some weak observed components of H_β to the selection and polarisation rules, of course still remains in the new theory, since the latter gives these rules in conformity with the older theory. However, components which are extremely weak theoretically are for the most part *un*observed, or the observations are *questionable*. The strength *ratios* of weaker components to one another or to stronger ones are *almost never* given even approximately correctly; cf. especially H_γ and H_δ. Such serious mistakes in the experimental determination of the blackening are of course out of the question.

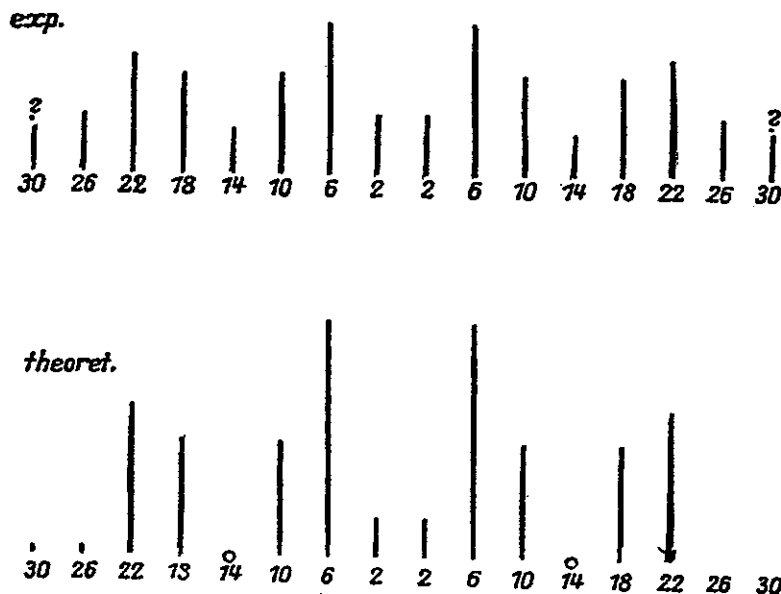Considering all this, we might feel inclined to be very sceptical of

exp.

theoret.

FIG. 8.—H_δ ⊥-components.

the thesis that the integrals (65) or their squares are measures of intensity. I am far from wishing to represent this thesis as irrefutable. There are still many alterations conceivable, and these may, perhaps, be necessitated by internal reasons when the theory is further extended. Yet the following should be remembered. The whole calculation has been performed with the *unperturbed* proper functions, or more precisely, with the *zero* approximation to the perturbed ones (cf. above § 2). It, therefore, represents an approximation for a *vanishing* field strength! However, just for the weak or almost vanishing components we should expect theoretically a fairly powerful growth with increasing field strength, for the following reason. According to the view of wave mechanics, as explained at the beginning of this section, the integrals (65) represent the amplitudes of the electrical partial moments, which are produced by the distribution of charges which flow round about the nucleus within the atom's domain. When for

a line component we get as a zero approximation very weak or even vanishing intensity, this is not caused in any way by the fact that to the simultaneous existence of the two proper vibrations corresponds only an insignificant motion of electricity, or even none at all. The vibrating mass of electricity—if this vague expression is allowed— may be represented as the same in all components, on the ground of *normalisation*. Rather is the reason for the low line intensity to be found in a high degree of *symmetry* in the motion of the electricity, through which only a small, or even no, dipole moment arises (on the contrary, *e.g.*, only a four-pole moment). Therefore it is to be expected that the *vanishing* of a line component in presence of perturbations of any kind is a relatively *unstable* condition, since the symmetry is probably destroyed by the perturbation. And thus it may be expected that weak or vanishing components gain quickly in intensity with increasing field strength.

This has now actually been observed, and the intensity ratios, indeed, alter quite considerably with field strength, for strengths of about 10,000 gauss and upwards; and, if I understand aright, in the way [1] shown by the present general discussion. Certain information on the question whether this really explains these discrepancies could of course only be got from a continuation of the calculation to the next approximation, but this is *very* troublesome and complicated.

The present considerations are of course nothing but the " translation " into the language of the new theory of very well-known considerations which Bohr [2] has brought forward in connection with calculation of line intensities by means of the principle of correspondence.

The theoretical intensities given in the tables satisfy a fundamental requirement, which is set up not only by intuition but also by experiment,[3] viz., the sum of the intensities of the $\parallel$-components is equal to that of the $\perp$-components. (Before adding, *undisplaced* components must be *halved*—as a compensation for the *duplication* of all the others, which occur on both sides.) This makes a very welcome " control " for the arithmetic.

It is also of interest to compare the *total intensities* of the four lines by using the four "sums" given in the tables. For this purpose I take back from my numerical calculations the four factors, which were omitted in order to represent the intensity ratios within each of the four line groups by the smallest integers possible, and multiply by them. Further, I multiply each of these four products by the *fourth* power of the appropriate emission frequency. Thus I obtain the following four numbers :

$$\text{for } H_\alpha \ldots \quad \frac{2^6 \cdot 23 \cdot 41}{3^2 \cdot 5^9} \quad = 0\cdot003433 \ldots$$

$$\text{for } H_\beta \ldots \quad \frac{4 \cdot 11 \cdot 19}{3^{12}} \quad = 0\cdot001573 \ldots$$

[1] J. Stark, *Ann. d. Phys.* 43, p. 1001 *et seq.*, 1914.
[2] N. Bohr, *Dänische Akademie* (8), iv. 1. 1, p. 35, 1918.
[3] J. Stark, *Ann. d. Phys.* 43, p. 1004, 1914.

$$\text{for } H_\gamma \ldots \frac{2^6 \cdot 3^6 \cdot 11^2 \cdot 71}{5 \cdot 7^{13}} = 0 \cdot 0008312 \ldots$$

$$\text{for } H_\delta \ldots \frac{11 \cdot 13}{2^{15} \cdot 3^2} = 0 \cdot 0004849 \ldots$$

I give these numbers *with still greater reserve* than the former ones because I am not sure, theoretically, about the *fourth* power of the frequency. Investigations[1] which I have lately published seem to call, perhaps, for the *sixth*. The above method of calculation corresponds exactly to the assumptions of Born, Jordan, and Heisenberg.[2] Fig. 9 represents the results diagrammatically.
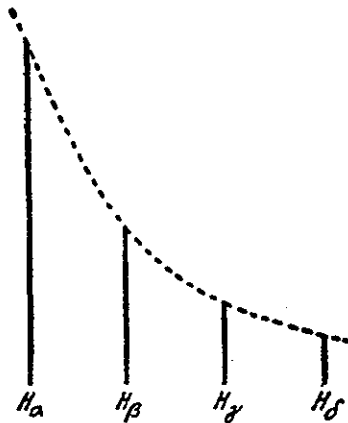

FIG. 9.—Total Intensities.

Actual measured intensities of emission lines, which are known to depend greatly on the conditions of excitation, naturally cannot here be used in a comparison with experience. From his researches[3] on dispersion and magneto-rotation in the neighbourhood of H_α and H_β, R. Ladenburg has, with F. Reiche,[4] calculated the value 4·5 (limits 3 and 6) for the ratio of the so-called " electronic numbers " of these two lines. If I assume that the above numbers may be taken as proportional to Ladenburg's [5] expression,

$$\Sigma \frac{g_k}{g_i} a_{ki} \nu_0,$$

then they may be reduced to (relative) " electronic numbers " by division by $\nu_0{}^3$, *i.e.* by

$$\left(\frac{5}{36}\right)^3, \quad \left(\frac{3}{16}\right)^3, \quad \left(\frac{21}{100}\right)^3, \quad \text{and} \quad \left(\frac{2}{9}\right)^3 \text{respectively.}$$

Hence we obtain the four numbers,

$$1 \cdot 281, \quad 0 \cdot 2386, \quad 0 \cdot 08975, \quad 0 \cdot 04418.$$

The ratio of the first to the second is 5·37, which agrees sufficiently with Ladenburg's value.

---

[1] Equation (38) at end of previous paper of this collection. The *fourth* allows for the fact that for the radiation it is a question of the square of the *acceleration* and not of the electric moment itself. In this equation (38) occurs explicitly another factor $(E_k - E_m)/h$. This is occasioned by the appearance of $\partial/\partial t$ in statement (36). *Addition at proof correction :* Now I recognise this $\partial/\partial t$ to be incorrect, though I hoped it would make the later relativistic generalisation easier. Statement (36), *loc. cit.*, is to be replaced by $\psi\bar{\psi}$. The above doubts about the *fourth* power are therefore dissolved.

[2] Cf. M. Born and P. Jordan, *Ztschr. f. Phys.* 34, p. 887, 1925.

[3] R. Ladenburg, *Ann. d. Phys.* (4), 38, p. 249, 1912.

[4] R. Ladenburg and F. Reiche, *Die Naturwissenschaften*, 1923, p. 584.

[5] Cf. Ladenburg-Reiche, *loc. cit.*, the first formula in the second column, p. 584. The factor ν_0 in the *above* expression comes from the fact that the " transition probability " a_{ki} is still to be multiplied by the " energy quantum " to give the intensity of the radiation.

§ 5. Treatment of the Stark Effect by the Method which corresponds to that of Bohr

Mainly to give an *example* of the general theory of § 2, I wish to outline *that* treatment of the proper value problem of equation (32), which must have been adopted, if we had *not* noticed that the perturbed equation is also exactly " separable " in parabolic co-ordinates. We therefore now keep to the polar co-ordinates r, θ, ϕ, and thus replace z by $r \cos \theta$. We also introduce a new variable η for r by the transformation

$$(66) \qquad 2r\sqrt{-\frac{8\pi^2 m E}{h^2}} = \eta,$$

(which is closely akin to transformation (48) for the parabolic co-ordinate ξ). For one of the unperturbed proper values (45), we get from (66)

$$(66') \qquad \eta = \frac{2r}{l a_0},$$

where a_0 is the same constant as in (63). (" Radius of the innermost hydrogen orbit.") If we introduce this and the unperturbed value (45) into the equation (32), which is to be treated, then we obtain

$$(67) \qquad \nabla'^2 \psi + \left(-\frac{1}{4} - g\eta \cos \theta + \frac{l}{\eta} \right)\psi = 0,$$

where for brevity

$$(68) \qquad g = \frac{a_0^2 F l^3}{4e}.$$

The dash on the Laplacian operator is merely to signify that in it the letter η is to be written for the radius vector.

In equation (67) we *conceive l to be the proper value*, and the term in g to be the perturbing term. The fact that the perturbing term *contains* the proper value need not trouble us in the first approximation. If we neglect the perturbing term, the equation has as proper values the natural numbers

$$(69) \qquad l = 1, 2, 3, 4 \ldots$$

and no others. (The extended spectrum is again cut out by the artifice (66), which would be valuable for closer approximations.) The allied *proper functions* (not yet normalised) are

$$(70) \qquad \psi_{lnm} = P_n^m (\cos \theta) \frac{\cos}{\sin} (m\phi) . \eta^n e^{-\frac{\eta}{2}} L_{n+l}^{2n+1}(\eta).$$

Here P_n^m signifies the mth " associated " Legendre function of the nth order, and L_{n+l}^{2n+1} is the $(2n+1)$th derivative of the $(n+l)$th Laguerre polynomial.[1] So we must have

$$n < l,$$

[1] I lately gave the proper functions (70) (see Part I.), but without noticing their connection with the Laguerre polynomials. For the proof of the above representation, see the Mathematical Appendix, section 1.

otherwise L_{n+l}^{2n+1} would vanish, because the number of differentiations would be greater than the degree. With reference to this, the numbering of the spherical surface harmonics shows that l is an l^2-fold proper value of the unperturbed equation. We now investigate the *splitting up* of a *definite* value of l, supposed fixed in what follows, due to the addition of the perturbing term.

To do this we have, *in the first place*, to normalise our proper functions (70), according to § 2. From an uninteresting calculation, which is easily performed with the aid of the formulae in the appendix,[1] we get as the normalising factor

$$(71) \qquad \frac{1}{\sqrt{\pi}} \sqrt{\frac{2n+1}{2}} \sqrt{\frac{(n-m)!}{(n+m)!}} \sqrt{\frac{(l-n-1)!}{[(n+l)!]^3}},$$

if $m \neq 0$, but, for $m = 0$, $\dfrac{1}{\sqrt{2}}$ times this value. *Secondly*, we have to calculate the symmetrical matrix of constants ϵ_{im}, according to (22). The r there is to be identified[2] with our perturbing function $-g\eta^3 \cos\theta \sin\theta$, and the proper functions, there called u_{ki}, are to be identified with our functions (70). The fixed suffix k, which characterises the proper value, corresponds to the *first* suffix l of ψ_{lnm}, and the *other* suffix i of u_{ki} corresponds now to the *pair* of suffixes n, m in ψ_{lnm}. The matrix (22) of constants forms in our case a square of l^2 rows and l^2 columns. The quadratures are easily carried out by the formulae of the appendix and yield the following results. Only those elements of the matrix are different from zero, for which the two proper functions ψ_{lnm}, $\psi_{ln'm'}$, to be combined, satisfy the following conditions simultaneously :

1. The *upper indices* of the " associated Legendre functions " must agree, *i.e.* $m = m'$.

2. The *orders* of the two Legendre functions must differ exactly by unity, *i.e.* $|n - n'| = 1$.

3. To each trio of indices lnm, if $m \neq 0$, there belong, according to (70), *two* Legendre functions, and thus also *two* proper functions ψ_{lnm}, which only differ from each other in that one contains a factor $\cos m\phi$ and the other $\sin m\phi$. The third condition reads : we may only combine sine with sine, or cosine with cosine, and not sine with cosine.

The remaining non-vanishing elements of the desired matrix would have to be characterised from the beginning by *two* index-*pairs* (n, m) and $(n+1, m)$. (We renounce any idea of showing the fixed index l explicitly.) Since the matrix is symmetrical, *one* index pair (n, m) is sufficient, if we stipulate that the first index, *i.e.* n, shall mean the *greater* of the two orders n, n', in every case.

---

[1] It is to be noticed that the *density function*, generally denoted by $\rho(x)$, reads as $\eta \sin\theta$ in equation (67), because the equation must be multiplied by $\eta^2 \sin\theta$, in order to acquire self-adjoint form.

Then the calculation gives

(72)
$$\epsilon_{nm} = -6lg \sqrt{\frac{(l^2-n^2)(n^2-m^2)}{4n^2-1}}.$$

We have now to form the determinant (22) out of these elements. It is advantageous to *arrange* its rows *as well as* its columns on the following principle. (To fix our ideas, let us speak of the columns, and therefore of the index-pair characterising the *first* of the two Legendre functions.) Thus : first come all terms with $m=0$, then all with $m=1$, then all with $m=2$, etc., and finally, all terms with $m=l-1$, which last is the greatest value that m (like n) can take. *Inside* each of these groups, let us arrange the terms thus : first, all terms with $\cos m\phi$, and then all with $\sin m\phi$. Within these " half groups " let us arrange them in order of increasing n, which runs through the values $m, m+1, m+2 \ldots l-1$, *i.e.* $(l-m)$ values in all.

If we carry this out, we find that the non-vanishing elements (72) are exclusively confined to the two secondary diagonals, which lie immediately alongside the principal diagonal. On the latter are the proper value perturbations which are to be found, but taken negatively, while everywhere else are zeros. Further, the two secondary diagonals are interrupted by zeros at *those* places, where they break through the *boundaries* between the so-called " half-groups ", in very convenient fashion. Hence the whole determinant *breaks up* into a product of just so many smaller determinants as there are " half-groups " present, viz. $(2l-1)$. It will be sufficient if we consider one of them. We write it here, denoting the desired perturbation of the proper value by ϵ (without suffix) :

(73)
$$\begin{vmatrix} -\epsilon & \epsilon_{m+1,m} & 0 & 0 & \ldots & 0 \\ \epsilon_{m+1,m} & -\epsilon & \epsilon_{m+2,m} & 0 & \ldots & 0 \\ 0 & \epsilon_{m+2,m} & -\epsilon & \epsilon_{m+3,m} & \ldots & 0 \\ 0 & 0 & \epsilon_{m+3,m} & -\epsilon & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \epsilon_{l-1,m} & -\epsilon \end{vmatrix}$$

If we divide each term here by the common factor $6lg$ of the ϵ_{nm}'s (cf. (72)), and for the moment regard as the unknown

(74)
$$k^* = -\frac{\epsilon}{6lg},$$

the above equation of the $(l-m)$th degree has the roots

(75) $k^* = \pm(l-m-1), \quad \pm(l-m-3), \quad \pm(l-m-5) \ldots$

where the series stops with ± 1 or 0 (inclusive) according as the *degree* $l-m$ is *even* or *odd*. The proof of this is unfortunately *not* to be found in the appendix, as I have not been successful in obtaining it.

If we form the series (75) for each of the values $m=0, 1, 2 \ldots (l-1)$, then we have in the numbers

(76)
$$\epsilon = -6lgk^*$$

the complete set of *perturbations of the principal quantum number l.*
In order to find the perturbed *proper* values E (term-levels) of the
equation (32), we have only to substitute (76) in

(77)
$$E = -\frac{2\pi^2 me^4}{h^2(l+\epsilon)^2}$$

taking into account the signification of the abbreviations g (see (68))
and a_0 (see (63)).

After reducing this gives

(78)
$$E = -\frac{2\pi^2 me^4}{h^2 l^2} - \frac{3}{8}\frac{h^2 Flk^*}{\pi^2 me}.$$

Comparison with (62) shows that k^* is the *difference* $k_2 - k_1$ of the
parabolic quantum numbers. From (75), bearing in mind the range
of values of m referred to above, we see that k^* may also take the
same values as the difference just mentioned, viz. $0, 1, 2 \ldots (l-1)$.
Also, if we take the trouble to work it out, we will find for the
multiplicity, in which k^* and the difference $k_2 - k_1$ appear, the same
value, viz. $l - |k^*|$.

We have thus obtained the proper value perturbations of the
first order also from the general theory. The next step would be
the solution of the system (21') of linear equations of the general
theory for the κ-quantities. These would then yield, according to
(18) (provisionally putting $\lambda = 0$), the perturbed proper functions
of zero order; this is nothing more than a representation of the
proper functions (64) as linear forms of the proper functions (70).
In our case the solution of (21') would naturally be anything but
unique, on account of the considerable multiplicity of the roots ϵ.
The solution is made much simpler if we notice that the equations
break up into *just as many* groups, viz. $(2l-1)$, or, retaining the
former expression, *half-groups*, with completely separated variables,
as the determinant investigated above contains factors like (73); and
if we further notice that it is allowable, after we have chosen a
definite ϵ-value, to regard only the variables κ of a *single* half-group
as different from zero, of that half-group, in fact, for which the deter-
minant (73) vanishes for the chosen ϵ-value. The definition of this
half-group of variables is then *unique*.

But our object, viz. to illustrate the general method of § 2 by an
example, has been sufficiently attained. Since the continuation of
the calculation is of no special physical interest, I have not troubled
to bring the determinantal quotients, which we immediately obtain
for the coefficients κ, into a clearer form, or to work out the transforma-
tion to principal axes in any other way.

On the whole, we must admit that in the present case the method
of secular perturbations (§ 5) is considerably more troublesome than
the direct application of a system of separation (§ 3). I believe that
this may also be true in other cases. In ordinary mechanics it is,
as we know, usually quite the reverse.

<div style="text-align:center">III.—MATHEMATICAL APPENDIX</div>

Prefatory Note:—It is not intended to supply in uninterrupted detail all the calculations omitted from the text. Without that, the present paper has already become too long. In general, only those methods of calculation will be briefly described which another might utilise with advantage in similar work, if something better does not occur to him—as it may easily do.

§ 1. The Generalised Laguerre Polynomials and Orthogonal Functions

The kth Laguerre polynomial $L_k(x)$ satisfies the differential equation [1]

$$(101) \qquad xy'' + (1-x)y' + ky = 0.$$

If we first replace k by $n+k$, and then differentiate n times, we find that the nth derivative of the $(n+k)$th Laguerre polynomial, which we will always denote by L_{n+k}^n, satisfies the equation

$$(102) \qquad xy'' + (n+1-x)y' + ky = 0.$$

Moreover, by an easy transformation, we find that for $e^{-\frac{x}{2}}L_{n+k}^n(x)$ the following equation holds,

$$(103) \qquad y'' + \frac{n+1}{x}y' + \left(-\frac{1}{4} + \left(k + \frac{n+1}{2}\right)\frac{1}{x}\right)y = 0.$$

This found an application in equation (41') of § 3. The allied generalised Laguerre orthogonal functions are

$$(104) \qquad x^{\frac{n}{2}}e^{-\frac{x}{2}}L_{n+k}^n(x).$$

Their equation, it may be remarked in passing, is

$$(105) \qquad y'' + \frac{1}{x}y' + \left(-\frac{1}{4} + \left(k + \frac{n+1}{2}\right)\frac{1}{x} - \frac{n^2}{4x^2}\right)y = 0.$$

Let us turn to equation (103), and consider there that n is a fixed (real) integer, and k is the proper value parameter. Then, according to what has been said, in the domain $x \geq 0$, at any rate, the equation has the proper functions,

$$(106) \qquad e^{-\frac{x}{2}}L_{n+k}^n(x),$$

belonging to the proper values,

$$(107) \qquad k = 0, 1, 2, 3, \ldots$$

In the text it is maintained that it has no further values, and, above all, that it possesses no continuous spectrum. This seems paradoxical, for the equation

$$(108) \qquad \frac{d^2y}{d\xi^2} + \frac{n+1}{\xi}\frac{dy}{d\xi} + \left(-\frac{1}{(2k+n+1)^2} + \frac{1}{\xi}\right)y = 0$$

[1] Courant-Hilbert, chap. ii. § 11, 5, p. 78, equation (72).

into which (103) is transformed by the substitution

(109)
$$\xi = \left(k + \frac{n+1}{2}\right)x$$

does possess a continuous spectrum, if in it we regard

(110)
$$E = -\frac{1}{(2k+n+1)^2}$$

as proper value parameter, viz. all positive values of E are proper values (cf. Part I., analysis of equation (7)). The reason why *no* proper values k of (103) can correspond to these positive E-values is that by (110) the k-values in question would be complex, and this is impossible, according to general theorems.[1] Each *real* proper value of (103), by (110), gives rise to a *negative* proper value of (108). Moreover, we know (cf. Part I.) that (108) possesses absolutely no negative proper values other than those that arise, as in (110), from the series (107). There thus remains only the one possibility, that in the series (107) certain negative k-values are lacking, which appear on solving (110) for k, on account of the double-valuedness when extracting the root. But this also is impossible, because the k-values in question turn out to be algebraically less than $-\dfrac{n+1}{2}$ and thus, from general theorems,[2] cannot be proper values of equation (103). The series of values (107) is thus complete. Q.E.D.

The above supplements the proof that the functions (70) are the proper functions of (67) (with the perturbing term suppressed), allied to the proper values (69). We have only to write the solutions of (67) as a product of a function of θ, ϕ and a function of η. The equation in η can readily be brought to the form of (105), the only difference being that our present n is there always an odd number, namely, the $(2n+1)$ which is to be found there.

§ 2. Definite Integrals of Products of Two Laguerre Orthogonal Functions

The Laguerre polynomials can all be obtained, in the following manner, as coefficients of the powers of the auxiliary variable t, in the expansion in a series of a so-called " generating function " [3]

(111)
$$\sum_{k=0}^{\infty} L_k(x)\frac{t^k}{k!} = \frac{e^{-\frac{xt}{1-t}}}{1-t}.$$

If we replace k by $n+k$ and then differentiate n times with respect to x, we obtain the generating function of our generalised polynomials,

(112)
$$\sum_{k=0}^{\infty} L_{n+k}^n(x)\frac{t^k}{(n+k)!} = (-1)^n\frac{e^{-\frac{xt}{1-t}}}{(1-t)^{n+1}}.$$

[1] Courant-Hilbert, chap. iii. § 4, 2, p. 115.
[2] Courant-Hilbert, chap. v. § 5, 1, p. 240.
[3] Courant-Hilbert, chap. ii. § 11, 5, p. 78, equation (68).

In order to evaluate with its help integrals such as appeared for the first time in the text in expression (52), or, more generally, such as were necessary in §4 for the calculation of (65), and also in §5, we proceed as follows. We write (112) over again, providing both the fixed index n and the varying index k with a dash, and replacing the undefined t by s. These two equations are then multiplied together, *i.e.* left side by left side, and right side by right. Then we multiply further by

(113) $x^p e^{-x}$

and integrate with respect to x from 0 to ∞. p is to be a positive integer—this being sufficient for our purpose. The integration is practicable by elementary methods on the right-hand side, and we get

$$(114) \quad \sum_{k=0}^{\infty} \sum_{k'=0}^{\infty} \frac{t^k s^{k'}}{(n+k)!\,(n'+k')!} \int_0^{\infty} x^p e^{-x} L_{n+k}^n(x) L_{n'+k'}^{n'}(x)\,dx$$

$$= (-1)^{n+n'} p!\, \frac{(1-t)^{p-n}(1-s)^{p-n'}}{(1-ts)^{p+1}}.$$

We have now, on the left, the desired integrals like pearls on a string, and we merely detach the one we happen to need by searching on the right for the coefficient of $t^k s^{k'}$. This coefficient is always a simple sum, and, in fact, in the cases occurring in the text, always a finite sum with very few terms (up to three). In general, we have

$$(115) \left\{ \begin{array}{l} \displaystyle\int_0^{\infty} x^p e^{-x} L_{n+k}^n(x) L_{n'+k'}^{n'}(x)\,dx = p!\,(n+k)!\,(n'+k')! \\[2mm] \displaystyle\qquad \cdot \sum_{\tau=0}^{\leq k,\,k'} (-1)^{n+n'+k+k'+\tau} \binom{p-n}{k-\tau}\binom{p-n'}{k'-\tau}\binom{-p-1}{\tau}. \end{array} \right.$$

The sum stops after the smaller of the two numbers k, k'. It often, in actual fact, begins at a positive value of τ, as binomial coefficients, whose lower number is greater than the upper, vanish. For example, in the integral in the denominator of (52), we put $p = n = n'$, and $k' = k$. Then τ can take only the *one* value k, and we can establish statement (53) of the text. In the integral of the numerator in (52), only p has another value, namely $p = n+2$. τ now takes the values $k-2$, $k-1$, and k, and after an easy reduction we get formula (54) of the text. In the very same way the integrals appearing in §5 are evaluated by Laguerre polynomials.

We can now, therefore, regard integrals of the type of (115) as known, and we have only to concern ourselves with those occurring in §4 in the calculation of intensities (cf. expression (65) and functions (64) which have to be substituted there). In this type, the two Laguerre orthogonal functions, whose product is to be integrated, *have not the same argument*, but, for example, in our case, have the arguments $\lambda_1/l a_0$ and $\lambda_1/l' a_0$, where l and l' are the principal quantum numbers of the two levels that we have combined. Let us consider, as typical, the integral

$$(116) \qquad J = \int_0^{\infty} x^p e^{-\frac{\alpha+\beta}{2} x} L_{n+k}^n(\alpha x) L_{n'+k'}^{n'}(\beta x)\,dx.$$

Now we can proceed in a superficially different way. At first, the former procedure still goes on smoothly; only on the right-hand side of (114) a somewhat more complicated expression appears. In the denominator occurs the power of a quadrinomial instead of that of a binomial, as before. And this makes the matter somewhat confusing, for the right-hand side of (114) becomes five-fold instead of three-fold, and thus the right side of (115) becomes a three-fold instead of a simple sum. I found that the following substitution made things clearer:

(117)
$$\frac{\alpha+\beta}{2}x=y.$$

Hence

(118)
$$\begin{cases} \alpha x = \left(1+\dfrac{\alpha-\beta}{\alpha+\beta}\right)y \\ \beta x = \left(1-\dfrac{\alpha-\beta}{\alpha+\beta}\right)y. \end{cases}$$

After expanding the two polynomials in their Taylor series, which are finite and have similar polynomials as coefficients, we get, using the abbreviations

(119)
$$\sigma=\frac{2}{\alpha+\beta},\quad \gamma=\frac{\alpha-\beta}{\alpha+\beta},$$

the following,

(120)
$$J=\sigma^{p+1}\sum_{\lambda=0}^{k}\sum_{\mu=0}^{k'}(-1)^{\mu}\frac{\gamma^{\lambda+\mu}}{\lambda!\,\mu!}\int_{0}^{\infty}y^{p+\lambda+\mu}L_{n+k}^{n+\lambda}(y)L_{n'+k'}^{n'+\mu}(y)dy.$$

Thus the calculation of J is reduced to the simpler type of integral (115). In the case of the Balmer lines, the double sum in (120) is comparatively tractable, for one of the two k-values, namely, the one referring to the two-quantum level, never exceeds unity, and thus λ may have two values at most, and, as it turns out, μ four values at most. The circumstance that out of the polynomials referring to the two-quantum level, none but

$$L_0=1,\quad L_1=-x+1,\quad L_1^1=-1,$$

appear, permits further simplifications. Nevertheless we must calculate out a number of tables, and it is much to be regretted that the figures given in the tables of the text for the intensities do not allow their general construction to be seen. By good fortune the additive relations between the ||- and the ⊥ components hold good, so that we may, with some *probability*, feel ourselves safe from arithmetical blunders at least.

§ 3. Integrals with Legendre Functions

There are three simple integral relations between associated Legendre functions, which are necessary for the calculations in § 5. For the convenience of others, I will state them here, because I was

not able to discover them in any of the places I searched.   We use the customary definition,

(121) $P_n^m (\cos \theta) = \sin^m \theta \, \dfrac{d^m P_n (\cos \theta)}{(d \cos \theta)^m}.$

Then the following holds,

(122) $\displaystyle\int_0^\pi [P_n^m (\cos \theta)]^2 \sin \theta \, d\theta = \dfrac{2}{2n+1} \dfrac{(n+m)!}{(n-m)!}$

(the normalising relation).

Moreover,

(123) $\left\{ \begin{array}{l} \displaystyle\int_0^\pi P_n^m (\cos \theta) P_{n'}^m (\cos \theta) \cos \theta \sin \theta \, d\theta = 0 \\[2mm] \qquad\qquad \text{for } |n - n'| \neq 1. \end{array} \right.$

On the other hand,

(124) $\displaystyle\int_0^\pi P_n^m (\cos \theta) P_{n-1}^m (\cos \theta) \cos \theta \sin \theta \, d\theta$

$$= \frac{n+m}{2n+1} \int_0^\pi [P_{n-1}^m (\cos \theta)]^2 \sin \theta \, d\theta = \frac{2(n+m)!}{(4n^2-1)(n-m-1)!}.$$

The last two relations decide the " selection " of the determinantal terms on page 95 of the text.   They are, moreover, of fundamental importance for the theory of spectra, for it is obvious that the selection principle for the azimuthal quantum number depends on them (and on two others which have $\sin^2 \theta$ in place of $\cos \theta \sin \theta$).

Addition at Proof Correction

Hr. W. Pauli, jun., informs me that he has arrived at the following closed formulae for the total intensity of the lines in the Lyman and Balmer series, through a modification of the method given in section 2 of the Appendix.   For the Lyman series these are

$$\nu_{l,1} = R\Big(\frac{1}{1^2} - \frac{1}{l^2}\Big); \quad J_{l,1} = \frac{2^7 \cdot (l-1)^{2l-1}}{l \cdot (l+1)^{2l+1}};$$

and for the Balmer series

$$\nu_{l,2} = R\Big(\frac{1}{2^2} - \frac{1}{l^2}\Big); \quad J_{l,2} = \frac{4^3 \cdot (l-2)^{2l-3}}{l \cdot (l+2)^{2l+3}} (3l^2 - 4)(5l^2 - 4).$$

The total emission intensities (square of amplitudes into fourth power of the frequency) are proportional to these expressions, within the series in question.   The numbers obtained from the formula for the Balmer series are in complete agreement with those given on pp. 91, 92.

Zürich, Physical Institute of the University.