PROCEEDINGS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

A Generalization of the Method of Separating Longitudinal and Transverse Waves in Electrodynamics

Sonja Ashauer

Proc. R. Soc. Lond. A 1948 **194**, 206-217
doi: 10.1098/rspa.1948.0074

A generalization of the method of separating longitudinal and transverse waves in electrodynamics

By Sonja Ashauer, *Newnham College, Cambridge*

(*Communicated by P. A. M. Dirac, F.R.S.—Received* 16 *July* 1947)

A generalization of the method of separating longitudinal and transverse waves in electrodynamics is proposed in this paper. It consists in splitting up each Fourier-component of the wave-field 4-vector with respect to two null-vectors $\mathbf{k}$ and $\mathbf{l_k}$, where $\mathbf{k}$ is the Fourier vector of propagation, and $\mathbf{l_k}$ is an arbitrary (real) function of $\mathbf{k}$ satisfying $(\mathbf{l_k}, \mathbf{k}) = 1$ and $\mathbf{l}_{-\mathbf{k}} = -\mathbf{l_k}$. This amounts to referring each Fourier-component of the field to a *different* time-axis in the usual method of splitting up. $\mathbf{l_k}$ may also depend on the co-ordinates of the particles of the system. The longitudinal field variables are eliminated from the Hamiltonian formulation of electrodynamics by a contact transformation; the term which then replaces the longitudinal field depends on the co-ordinates of the particles, and in general also on the transverse field variables, except when $\mathbf{l_k}$ is independent of the co-ordinates of the particles. The work holds both in the classical and in the quantum theory.

For the problem of an electron moving in the field of a nucleus a particular form is chosen for the $\mathbf{l_k}$, which depends on the initial velocities of both the particles. The new longitudinal waves are eliminated, and the classical deflexion formula is derived on the assumption that the interaction with the new transverse waves can be neglected. The new method leads to half the deflexion derived by the usual method of working only with the Coulomb interaction force (and so neglecting radiation damping in a certain way), when the nuclear recoil is neglected and the deflexion angle is small. For a low-velocity electron, the effect of the new transverse waves cannot be small, but for a very high energy electron the new method may be more suitable than the usual one.

Introduction

The main problem of electrodynamics is to treat the interaction of a system of charged particles with an electromagnetic field. In order to enable the direct passage to the quantum theory, it is necessary to put classical electrodynamics into Hamiltonian form. In this form of the theory, each particle of the system is described by a suitable Hamiltonian for the particle and the field. In field theory, the mutual interaction between the charged particles is looked upon as an indirect one: each particle influences the neighbouring field and this influence spreads out to the other particles. The total field consists of a transverse and a longitudinal part, the latter occurring only in the presence of charged particles. It was first shown by Fermi (1930, 1932) that the longitudinal part can be eliminated from the Hamiltonian function and replaced by the static Coulomb interaction of the particles.

A Hamiltonian formulation of electrodynamics which takes into account the effect of radiation reaction has been given by Dirac (1939, 1943) with the help of the λ-limiting process. Each particle is then described by a Hamiltonian equation involving a new field (represented by its 4-vector potential) whose properties differ slightly from those of the Maxwell field and which is called the Wentzel field. This is a wave-field and can be split up into longitudinal and transverse waves according to the usual method in a Lorentz frame having a particular direction as time-axis. This method consists in taking the whole of the time-component of the field to be longitudinal, and separating the spatial part of the field into a rotation-free (longi-

tudinal) part and a divergence-free (transverse) part. It is then shown that the longitudinal waves can be eliminated from the Hamiltonian equations by means of a contact transformation, and replaced by terms giving rise to the Coulomb interaction between the particles. This procedure simplifies the equations of motion considerably, as it reduces the number of field variables in them by half. The relativistic form of the theory, however, is spoilt.

There is great arbitrariness in the method of separating the longitudinal and transverse waves under the requirement that the divergence-free character (in four dimensions) of the transverse waves is preserved. In the usual method, all the Fourier-components of the field are split up with reference to the same selected direction (the time-axis of the particular Lorentz frame). In this paper a general method of separation is proposed, in which each Fourier-component is split up with reference to a different (arbitrary) direction, which may depend on other variables as well as on the Fourier-component. Under certain conditions the new longitudinal waves can be eliminated from the Hamiltonian equations by means of a contact transformation. The work is carried out here in the classical theory and is so arranged that each step can be taken over directly into the quantum theory.

The generalized splitting up of the field may be useful in cases where the usual method is not the most suitable. As an interesting application, the problem of the motion of two interacting charged particles is considered. In this case the usual method of eliminating the longitudinal field and treating the interaction between the particles and the transverse field as small gives satisfactory results for slowly moving particles. The neglect of the terms containing the transverse field variables in the Hamiltonian equations leads to the equations of motion for the particles derived from the Coulomb potential, from which the usual classical formula for the deflexion of an electron by a heavy nucleus is obtained (Darwin 1913). This involves the neglect of radiation damping in a certain way. For particles of large relative velocity, however, some alternative method of treatment of the field may be more suitable. A new way of separating the longitudinal and transverse field based on the proposed generalization is therefore introduced for this case. The neglect of the terms containing the new transverse field in the Hamiltonian equations leads to new equations for the motion of the particles. A new formula for the classical deflexion of a fast electron by a nucleus is thus derived and compared with the usual one.

1. General separation of longitudinal and transverse waves

Relativistic notation will be used, in the metric defined by

$$g_{00} = -g_{11} = -g_{22} = -g_{33} = 1, \quad g_{\mu\nu} = 0 \quad \text{for} \quad \mu \neq \nu,$$

the velocity of light being unity. The scalar product of two 4-vectors $\mathbf{a}$, $\mathbf{b}$ is denoted by

$$(\mathbf{a}, \mathbf{b}) \equiv a_\mu b^\mu \quad (\mu = 0, 1, 2, 3).$$

Consider a number of charged point-particles interacting with an electromagnetic field in the classical theory. Let e_i, $\mathbf{z}_i$, $\mathbf{p}_i$ be the charge, co-ordinates and momenta respectively of the ith particle, and $\mathbf{A}(\mathbf{x})$ the field at $\mathbf{x}$ (4-vector potential). The

equations of motion of the dynamical variables are derived from the many-time Hamiltonian equations written in the usual way in terms of the momentum vector

$$\mathbf{p}_i - e_i \mathbf{A}(\mathbf{z}_i). \tag{1·1}$$

In the theory which takes into account the effect of radiation reaction (Dirac 1939, 1943), the field $\mathbf{A}(\mathbf{x})$ to be used in these Hamiltonian equations is taken to be given by

$$\mathbf{A}(\mathbf{x}) = \tfrac{1}{2}[\mathbf{A}_W(\mathbf{x}+\boldsymbol{\lambda}) + \mathbf{A}_W(\mathbf{x}-\boldsymbol{\lambda})], \tag{1·2}$$

where $\mathbf{A}_W$ is the Wentzel field and $\boldsymbol{\lambda}$ is a small time-like vector which is ultimately made to tend to zero. The $\mathbf{z}_i$ have to satisfy the conditions

$$(\mathbf{z}_i - \mathbf{z}_j \pm \boldsymbol{\lambda})^2 < 0 \quad \text{for} \quad i \neq j. \tag{1·3}$$

The field can be represented by its Fourier-components

$$\mathbf{A}(\mathbf{x}) = \int \mathbf{A}_{\mathbf{k}} e^{i(\mathbf{k},\mathbf{x})} d\mathbf{k}, \quad (d\mathbf{k} = dk_0 dk_1 dk_2 dk_3), \tag{1·4}$$

where

$$\mathbf{A}_{\mathbf{k}} = \frac{1}{(2\pi)^4} \int \mathbf{A}(\mathbf{x})\, e^{-i(\mathbf{k},\mathbf{x})} d\mathbf{x}, \tag{1·5}$$

the reality of $\mathbf{A}(\mathbf{x})$ demanding that $\mathbf{A}_{\mathbf{k}} = \bar{\mathbf{A}}_{-\mathbf{k}}$. The Poisson bracket relations satisfied by the field are taken to be

$$[A_{\mathbf{k}\mu}, A_{\mathbf{k}'\nu}] = -\frac{ig_{\mu\nu}}{4\pi^2} \cos(\mathbf{k},\boldsymbol{\lambda})\, \Delta(\mathbf{k})\, \delta(\mathbf{k}+\mathbf{k}'), \tag{1·6}$$

where

$$\Delta(\mathbf{k}) = \frac{1}{|k|}[\delta(|k| - k_0) - \delta(|k| + k_0)], \quad |k| = (k_1^2 + k_2^2 + k_3^2)^{\frac{1}{2}},$$

$$\delta(\mathbf{k}+\mathbf{k}') = \delta(k_0 + k_0')\, \delta(k_1 + k_1')\, \delta(k_2 + k_2')\, \delta(k_3 + k_3').$$

The Wentzel field satisfies the d'Alembertian equation, which gives

$$\mathbf{k}^2 = 0, \tag{1·7}$$

and a divergence equation, which leads to the (supplementary) condition

$$(\mathbf{A}_{\mathbf{k}}, \mathbf{k}) - \frac{1}{4\pi^2} \cos(\mathbf{k},\boldsymbol{\lambda})\, \Delta(\mathbf{k})\, \Sigma_j e_j e^{-i(\mathbf{k},\mathbf{z}_j)} = 0. \tag{1·8}$$

The usual splitting up of $\mathbf{A}(\mathbf{x})$ into longitudinal and transverse waves is carried out in a particular Lorentz frame of reference. Let the (time-like) vector $\mathbf{a}$ specify the direction of the time-axis of this particular frame in a general Lorentz frame. In terms of Fourier-components, the splitting up is given by

$$A_{\mathbf{k}\mu} = L_{\mathbf{k}\mu} + T_{\mathbf{k}\mu} \equiv L^{\mathrm{I}}_{\mathbf{k}} a_\mu + L^{\mathrm{II}}_{\mathbf{k}} k_\mu + T_{\mathbf{k}\mu}, \tag{1·9}$$

where the transverse components $T_{\mathbf{k}\mu}$ satisfy the relations

$$(\mathbf{T}_{\mathbf{k}}, \mathbf{k}) = 0, \quad (\mathbf{T}_{\mathbf{k}}, \mathbf{a}) = 0. \tag{1·10}$$

It can be seen at once that (1·9) and (1·10) give the familiar splitting up equations if one passes to a frame where $a_\mu = (\mathbf{a}^2)^{\frac{1}{2}} \delta_{\mu 0}$, $\delta_{\mu\nu}$ being the Kronecker symbol.

The equations (1·9) and (1·10) give a general form of the splitting up of the field if $\mathbf{a}$ is regarded as a vector that is arbitrary, apart from satisfying the condition that $(\mathbf{a},\mathbf{k})$ must not vanish identically. This condition must be imposed in order that the equations (1·9) and (1·10) can be solved for $L^{\mathrm{I}}_{\mathbf{k}}$, $L^{\mathrm{II}}_{\mathbf{k}}$ and $T_{\mathbf{k}\mu}$. Thus a different direction $\mathbf{a}$ may be selected for each Fourier-component of the field, say $\mathbf{a}_{\mathbf{k}}$, provided that

$$(\mathbf{a}_{\mathbf{k}},\mathbf{k}) \neq 0. \tag{1·11}$$

The transverse waves so defined are gauge-invariant (as should be the case if they are measurable quantities), since the first equation of (1·10) gives in $\mathbf{x}$-space $\partial T_{\mu}(\mathbf{x})/\partial x_{\mu} = 0$. It is seen from (1·9) and (1·10) that the longitudinal and transverse parts of $\mathbf{A}_{\mathbf{k}}$ remain unchanged if $\mathbf{a}_{\mathbf{k}}$ is replaced in (1·9) by $S'_{\mathbf{k}}\mathbf{a}_{\mathbf{k}} + S''_{\mathbf{k}}\mathbf{k}$ with $S'_{\mathbf{k}} \neq 0$. Introducing the null-vector $\mathbf{l}_{\mathbf{k}}$ defined by

$$\mathbf{l}_{\mathbf{k}} = \frac{\mathbf{a}_{\mathbf{k}}}{(\mathbf{a}_{\mathbf{k}},\mathbf{k})} - \frac{1}{2}\frac{\mathbf{a}^2_{\mathbf{k}}\mathbf{k}}{(\mathbf{a}_{\mathbf{k}},\mathbf{k})^2}, \tag{1·12}$$

with the properties

$$(\mathbf{l}_{\mathbf{k}},\mathbf{l}_{\mathbf{k}}) = 0, \quad (\mathbf{l}_{\mathbf{k}},\mathbf{k}) = 1, \tag{1·13}$$

one can put the generalized splitting up into the form

$$A_{\mathbf{k}\mu} = L_{\mathbf{k}\mu} + T_{\mathbf{k}\mu} \equiv L^{\mathrm{I}}_{\mathbf{k}} l_{\mathbf{k}\mu} + L^{\mathrm{II}}_{\mathbf{k}} k_{\mu} + T_{\mathbf{k}\mu}, \tag{1·14}$$

where the transverse components satisfy the relations

$$(\mathbf{T}_{\mathbf{k}},\mathbf{l}_{\mathbf{k}}) = 0, \quad (\mathbf{T}_{\mathbf{k}},\mathbf{k}) = 0. \tag{1·15}$$

With the help of (1·13) and (1·15), the equation (1·14) gives the new $L^{\mathrm{I}}_{\mathbf{k}}$ and $L^{\mathrm{II}}_{\mathbf{k}}$ as

$$L^{\mathrm{I}}_{\mathbf{k}} = (\mathbf{A}_{\mathbf{k}},\mathbf{k}), \quad L^{\mathrm{II}}_{\mathbf{k}} = (\mathbf{A}_{\mathbf{k}},\mathbf{l}_{\mathbf{k}}). \tag{1·16}$$

It is not necessary at this stage to restrict $\mathbf{l}_{\mathbf{k}}$ to be real. However, $\mathbf{l}_{\mathbf{k}}$ must satisfy a suitable condition for $L_{\mu}(\mathbf{x})$ (and therefore also $T_{\mu}(\mathbf{x})$) to be real. This condition is $L_{\mathbf{k}\mu} = \bar{L}_{-\mathbf{k}\mu}$, which gives

$$(\mathbf{A}_{\mathbf{k}},\mathbf{k})\,(l_{\mathbf{k}\mu} + \bar{l}_{-\mathbf{k}\mu}) + (\mathbf{A}_{\mathbf{k}},\mathbf{l}_{\mathbf{k}} + \bar{\mathbf{l}}_{-\mathbf{k}})\,k_{\mu} = 0.$$

On multiplication by $l^{\mu}_{\mathbf{k}}$ and by $\bar{l}^{\mu}_{-\mathbf{k}}$, this leads to

$$(\mathbf{l}_{\mathbf{k}},\bar{\mathbf{l}}_{-\mathbf{k}}) = 0, \quad (\mathbf{A}_{\mathbf{k}},\mathbf{l}_{\mathbf{k}} + \bar{\mathbf{l}}_{-\mathbf{k}}) = 0.$$

Therefore it is necessary to impose on permissible functions $\mathbf{l}_{\mathbf{k}}$ the condition that

$$\mathbf{l}_{\mathbf{k}} = -\bar{\mathbf{l}}_{-\mathbf{k}}. \tag{1·17}$$

The Poisson bracket relations of the variables introduced above will now be considered. If $\mathbf{l}_{\mathbf{k}}$ is a function of $\mathbf{k}$ only, it will have vanishing Poisson brackets with all the dynamical variables. More generally, one could take $\mathbf{l}_{\mathbf{k}}$ to be a function either of the co-ordinates $\mathbf{z}_i$ or of the momenta $\mathbf{p}_i$ of the particles. In either case

$$[l_{\mathbf{k}\mu}, l_{\mathbf{k}'\nu}] = 0, \quad [A_{\mathbf{k}\mu}, l_{\mathbf{k}'\nu}] = 0,$$

$$[L^{\mathrm{I}}_{\mathbf{k}}, l_{\mathbf{k}'\nu}] - 0. \quad [L^{\mathrm{II}}_{\mathbf{k}}, l_{\mathbf{k}'\nu}] = 0 \quad [T_{\mathbf{k}\mu}, l_{\mathbf{k}'\nu}] = 0.$$

From (1·6) and (1·16) one gets, by a straightforward calculation,

$$\left.\begin{aligned}
&[L^{\mathrm{I}}_{\mathbf{k}}, L^{\mathrm{I}}_{\mathbf{k}'}] = 0,\\
&[L^{\mathrm{I}}_{\mathbf{k}}, L^{\mathrm{II}}_{\mathbf{k}'}] = \frac{i}{4\pi^2}\cos(\mathbf{k},\boldsymbol{\lambda})\,\Delta(\mathbf{k})\,\delta(\mathbf{k}+\mathbf{k}'),\\
&[L^{\mathrm{II}}_{\mathbf{k}}, L^{\mathrm{II}}_{\mathbf{k}'}] = -\frac{i}{4\pi^2}(\mathbf{l}_{\mathbf{k}}, \mathbf{l}_{\mathbf{k}'})\cos(\mathbf{k},\boldsymbol{\lambda})\,\Delta(\mathbf{k})\,\delta(\mathbf{k}+\mathbf{k}'),\\
&[L^{\mathrm{I}}_{\mathbf{k}}, T_{\mathbf{k}'\mu}] = 0,\\
&[L^{\mathrm{II}}_{\mathbf{k}}, T_{\mathbf{k}'\mu}] = -\frac{i}{4\pi^2}[l_{\mathbf{k}\mu} + l_{\mathbf{k}'\mu} - (\mathbf{l}_{\mathbf{k}}, \mathbf{l}_{\mathbf{k}'})\,k_\mu]\cos(\mathbf{k},\boldsymbol{\lambda})\,\Delta(\mathbf{k})\,\delta(\mathbf{k}+\mathbf{k}').
\end{aligned}\right\} \tag{1·18}$$

The Poisson brackets of the longitudinal field with the transverse one should vanish (i.e., they should commute in the quantum theory), if they are variables corresponding to different degrees of freedom. The last two equations of (1·18) show that the condition for this is that

$$\mathbf{l}_{\mathbf{k}} = -\mathbf{l}_{-\mathbf{k}}. \tag{1·19}$$

Together with (1·17), (1·19) shows that $\mathbf{l}_{\mathbf{k}}$ must be taken to be a real function of $\mathbf{k}$. This condition and the condition (1·19) will always be assumed to be satisfied in the work that follows. (In terms of $\mathbf{a}_{\mathbf{k}}$ the condition (1·19) is satisfied in particular if

$$a_{\mathbf{k}} = \alpha\mathbf{a}_{-\mathbf{k}}, \tag{1·20}$$

where $\alpha \neq 0$.) With (1·19), one has

$$[L^{\mathrm{II}}_{\mathbf{k}}, L^{\mathrm{II}}_{\mathbf{k}'}] = 0, \quad [L^{\mathrm{II}}_{\mathbf{k}}, T_{\mathbf{k}'\mu}] = 0. \tag{1·21}$$

It will now be assumed that the $\mathbf{l}_{\mathbf{k}}$ depend on $\mathbf{k}$ and on the co-ordinates $\mathbf{z}_i$ $(i = 1, 2, \ldots)$. Then $L^{\mathrm{I}}_{\mathbf{k}}$, $L^{\mathrm{II}}_{\mathbf{k}}$ and $\mathbf{T}_{\mathbf{k}}$ depend on the $\mathbf{z}_i$. The following Poisson brackets hold

$$[l_{\mathbf{k}\mu}, z_{i\nu}] = 0, \quad [l_{\mathbf{k}\mu}, p_{i\nu}] = -\frac{\partial l_{\mathbf{k}\mu}}{\partial z^\nu_i}. \tag{1·22}$$

2. Elimination of the longitudinal waves

If all the dynamical variables of the system are transformed to new ones by means of a contact transformation, the new Hamiltonian equations describing the system will be just the old ones expressed in terms of the new variables. It is possible to eliminate the longitudinal waves from the Hamiltonian equations by such a transformation, the supplementary condition (1·8) being at the same time automatically satisfied and so dropping out from the system of field equations. The method of elimination follows that given by Dirac (1943) for ordinary longitudinal waves, adapted to the resolution of the field into Fourier-components.

The required contact transformation can be built up from infinitesimal contact transformations, in which each dynamical variable ξ is transformed according to

$$\frac{d\xi^{(\tau)}}{d\tau} = [\xi^{(\tau)}, G^{(\tau)}] \quad (0 \leqslant \tau \leqslant 1), \tag{2·1}$$

from its initial value $\xi^{(0)} = \xi$ to its final value $\xi^{(1)} = \xi'$ say, with

$$G^{(\tau)} = i\Sigma_j e_j \int (\mathbf{A}_{\mathbf{k}}^{(\tau)}, \mathbf{l}_{\mathbf{k}}^{(\tau)}) \exp[i(\mathbf{k}, \mathbf{z}_j^{(\tau)})]\, d\mathbf{k}. \tag{2·2}$$

From the Poisson brackets of the last section one finds that $\mathbf{l}_{\mathbf{k}}^{(\tau)} = \mathbf{l}_{\mathbf{k}}$ and $\mathbf{z}_i^{(\tau)} = \mathbf{z}_i$. Since $G^{(\tau)}$ is expressible in terms of $L_{\mathbf{k}}^{\mathrm{II}(\tau)}$ (by (1·16)), which has vanishing Poisson brackets with the transverse waves, $\mathbf{T}_{\mathbf{k}}^{(\tau)}$ is also independent of τ. Taking $\xi^{(\tau)} = A_{\mathbf{k}\mu}^{(\tau)}$ in (2·1), one has

$$\begin{aligned} \frac{dA_{\mathbf{k}\mu}^{(\tau)}}{d\tau} &= i\Sigma_j e_j \int [A_{\mathbf{k}\mu}, A_{\mathbf{k}'\rho}]^{(\tau)} l_{\mathbf{k}'}^{\rho} e^{i(\mathbf{k}', \mathbf{z}_j)}\, d\mathbf{k}' \\ &= -\frac{1}{8\pi^2} \Delta(\mathbf{k})\, l_{\mathbf{k}\mu} \Sigma_j e_j \{e^{-i(\mathbf{k}, \mathbf{z}_j + \boldsymbol{\lambda})} + e^{-i(\mathbf{k}, \mathbf{z}_j - \boldsymbol{\lambda})}\}, \end{aligned} \tag{2·3}$$

where (1·6) and (1·19) have been used. By integration with respect to τ this gives

$$A_{\mathbf{k}\mu}^{(\tau)} = A_{\mathbf{k}\mu} - \frac{\tau}{8\pi^2} \Delta(\mathbf{k})\, l_{\mathbf{k}\mu} \Sigma_j e_j \{e^{-i(\mathbf{k}, \mathbf{z}_j + \boldsymbol{\lambda})} + e^{-i(\mathbf{k}, \mathbf{z}_j - \boldsymbol{\lambda})}\}. \tag{2·4}$$

The transformation of the $\mathbf{p}_i$ is given by

$$\frac{dp_{i\mu}^{(\tau)}}{d\tau} = -e_i \int (\mathbf{A}_{\mathbf{k}}^{(\tau)}, \mathbf{l}_{\mathbf{k}})\, k_\mu e^{i(\mathbf{k}, \mathbf{z}_i)}\, d\mathbf{k} + i\Sigma_j e_j \int \left(\mathbf{A}_{\mathbf{k}}^{(\tau)}, \frac{\partial \mathbf{l}_{\mathbf{k}}}{\partial z_i^\mu}\right) e^{i(\mathbf{k}, \mathbf{z}_j)}\, d\mathbf{k}. \tag{2·5}$$

The properties (1·13) of $\mathbf{l}_{\mathbf{k}}$ show that $\left(\mathbf{l}_{\mathbf{k}}, \dfrac{\partial \mathbf{l}_{\mathbf{k}}}{\partial z_i^\mu}\right) = \left(\mathbf{k}, \dfrac{\partial \mathbf{l}_{\mathbf{k}}}{\partial z_i^\mu}\right) = 0$. Thus, according to (2·4), $\mathbf{A}_{\mathbf{k}}^{(\tau)}$ can be replaced by $\mathbf{A}_{\mathbf{k}}$ in the two integrals of (2·5). Also, from the splitting up equations (1·14) and (1·16), one has

$$(\mathbf{A}_{\mathbf{k}}, \mathbf{l}_{\mathbf{k}})\, k_\mu = A_{\mathbf{k}\mu} - (\mathbf{A}_{\mathbf{k}}, \mathbf{k})\, l_{\mathbf{k}\mu} - T_{\mathbf{k}\mu}$$

and

$$\left(\mathbf{A}_{\mathbf{k}}, \frac{\partial \mathbf{l}_{\mathbf{k}}}{\partial z_i^\mu}\right) = \left(T_{\mathbf{k}}, \frac{\partial \mathbf{l}_{\mathbf{k}}}{\partial z_i^\mu}\right).$$

With the help of these equations, (2·5) becomes

$$\frac{dp_{i\mu}^{(\tau)}}{d\tau} = -e_i A_\mu(\mathbf{z}_i) + e_i T_\mu(\mathbf{z}_i) + e_i \int (\mathbf{A}_{\mathbf{k}}, \mathbf{k})\, l_{\mathbf{k}\mu} e^{i(\mathbf{k}, \mathbf{z}_i)}\, d\mathbf{k} + i\Sigma_j e_j \int \left(\mathbf{T}_{\mathbf{k}}, \frac{\partial \mathbf{l}_{\mathbf{k}}}{\partial z_i^\mu}\right) e^{i(\mathbf{k}, \mathbf{z}_j)}\, d\mathbf{k}. \tag{2·6}$$

After integrating (2·6) over τ from 0 to 1, and expressing $(\mathbf{A}_{\mathbf{k}}, \mathbf{k})$ with the help of the supplementary condition (1·8), one gets

$$\begin{aligned} p_{i\mu} - e_i A_\mu(\mathbf{z}_i) = p'_{i\mu} - e_i T_\mu(\mathbf{z}_i) - i\Sigma_j e_j \int \left(\mathbf{T}_{\mathbf{k}}, \frac{\partial \mathbf{l}_{\mathbf{k}}}{\partial z_i^\mu}\right) e^{i(\mathbf{k}, \mathbf{z}_j)}\, d\mathbf{k} \\ - \frac{e_i}{8\pi^2} \Sigma_j e_j \int \Delta(\mathbf{k})\, l_{\mathbf{k}\mu} \{e^{i(\mathbf{k}, \mathbf{z}_i - \mathbf{z}_j + \boldsymbol{\lambda})} + e^{i(\mathbf{k}, \mathbf{z}_i - \mathbf{z}_j - \boldsymbol{\lambda})}\}\, d\mathbf{k}. \end{aligned} \tag{2·7}$$

The left-hand side of (2·7) is the expression (1·1), in function of which the Hamiltonian equation of the ith particle is written (both in the classical and in the quantum theory). The right-hand side of (2·7) is expressed in terms of the new dynamical variables introduced by the contact transformation; the longitudinal field variables have been eliminated and replaced by the two terms containing integrals. The first integral contains in general transverse field variables referring to *all* the particles

(not only to the ith one), but it vanishes if $\partial l_{\mathbf{k}\nu}/\partial z_i^\mu = 0$. The second integral depends on the co-ordinates of the particles and the choice of $\mathbf{l}_\mathbf{k}$, but not on any field variables. It is interesting to examine when the term corresponding to $j = i$ ('self-energy') in the summation over j in this integral vanishes. This term is of the form

$$\int \Delta(\mathbf{k})\, \mathbf{l}_\mathbf{k} \cos(\mathbf{k}, \boldsymbol{\lambda})\, d\mathbf{k} = \iiint \frac{1}{|k|} \mathbf{l}_\mathbf{k}^+ \cos b\, |k|\, dk_1 dk_2 dk_3 - \iiint \frac{1}{|k|} \mathbf{l}_\mathbf{k}^- \cos c\, |k|\, dk_1 dk_2 dk_3, \tag{2·8}$$

where $\mathbf{l}_\mathbf{k}^+$, $\mathbf{l}_\mathbf{k}^-$ is $\mathbf{l}_\mathbf{k}$ with k_0 replaced by $\pm |k|$ respectively,

$$b = \lambda_0 - |\lambda| \cos\beta, \quad c = \lambda_0 + |\lambda| \cos\beta$$

and β is the angle between the 3-vectors λ_s and k_s. Since $\boldsymbol{\lambda}$ is a time-like vector, b and c are different from zero. Now transform the right-hand side of (2·8) to polar co-ordinates $r = |k|$, θ, ϕ, and suppose $\mathbf{l}_\mathbf{k}^+$ and $\mathbf{l}_\mathbf{k}^-$ expansible in power series of r with powers $\geqslant -1$ in the form $\mathbf{l}_\mathbf{k}^+ = \sum_0^\infty \mathbf{b}_n r^{2n-1}$, $\mathbf{l}_\mathbf{k}^- = \sum_0^\infty \mathbf{c}_n r^{2n-1}$ with $\mathbf{b}_n$ and $\mathbf{c}_n$ independent of r. These series can only contain odd powers of r due to the condition (1·19). The integrals over r in (2·8) become

$$\int_0^\infty \sum_0^\infty \mathbf{b}_n r^{2n} \cos br\, dr, \quad \int_0^\infty \sum_0^\infty \mathbf{c}_n r^{2n} \cos cr\, dr. \tag{2·9}$$

Each of the two integrals of (2·9) may be evaluated by term by term integration, whenever this procedure, applied to the integral in question and to the corresponding integral obtained by replacing cos by sin, leads to a series that is summable (in the Cesàro sense) (Hardy 1908). Since

$$\int_0^\infty r^{2n} \cos br\, dr = 0, \quad \int_0^\infty r^{2n} \sin br\, dr = (-1)^n (2n)!/b^{2n+1},$$

the 'self-energy' vanishes if $\sum_0^\infty (-1)^n (2n)!\, \mathbf{b}_n/b^{2n+1}$ and $\sum_0^\infty (-1)^n (2n)!\, \mathbf{c}_n/c^{2n+1}$ are

summable. In particular, the 'self-energy' vanishes if $\mathbf{a}_\mathbf{k}$ depends only on the direction of $\mathbf{k}$, so that $\mathbf{a}_{m\mathbf{k}} = \alpha \mathbf{a}_\mathbf{k}$ for any m, where m and α are non-vanishing scalars, since the power series above reduced to their respective first terms in this case.

If the $\mathbf{l}_\mathbf{k}$ are such that the corresponding $\mathbf{a}_\mathbf{k}$ may be taken independent of $\mathbf{k}$, the last term of (2·7) gives, after passing to the limit $\boldsymbol{\lambda} \to 0$,

$$-\frac{e_i}{2} \sum_{j \neq i} e_j \frac{[2(\mathbf{a}, \mathbf{z}_i - \mathbf{z}_j)^2 - \mathbf{a}^2(\mathbf{z}_i - \mathbf{z}_j)^2]\, a_\mu - (\mathbf{a}, \mathbf{z}_i - \mathbf{z}_j)\, \mathbf{a}^2 (z_{i\mu} - z_{j\mu})}{[(\mathbf{a}, \mathbf{z}_i - \mathbf{z}_j)^2 - \mathbf{a}^2(\mathbf{z}_i - \mathbf{z}_j)^2]^{\frac{3}{2}}}. \tag{2·10}$$

If the $\mathbf{a}$ are also independent of all the $\mathbf{z}_i$, the first integral on the right-hand side of (2·7) vanishes, and (2·10) agrees with the result given by Dirac (1943) in a reference frame where $a_\mu = (\mathbf{a}^2)^{\frac{1}{2}} \delta_{\mu 0}$. In the single-time Hamiltonian theory, (2·10) then leads to the electrostatic Coulomb potential of the particles.

One could also consider the case where the $\mathbf{l}_\mathbf{k}$ are functions of $\mathbf{k}$ and of the $\mathbf{p}_i$. The contact transformation (2·1) with (2·2) then makes the new $\mathbf{z}_i$ depend on the $\mathbf{A}_\mathbf{k}$, thus leading to difficulties. This case is therefore abandoned.

Up to this point, each step taken in the classical theory above can be carried over into the quantum theory. The right-hand side of the supplementary condition (1·8) must be applied onto the wave-function. One arrives at the same result (2·7) with both sides of that equation applied onto the wave-function.

3. Classical deflexion of an electron by a nucleus

As an application of the general theory above, consider the problem of the interaction of two charged particles, say an electron and a nucleus, in the classical theory.

The usual method, where the field is split up with reference to a selected time-axis, taken to be the direction of the initial velocity 4-vector of the nucleus, leads to the Coulomb interaction. If the interaction with the transverse field and the nuclear recoil are neglected, the following (relativistic) formula for the angle of deflexion ϕ of an electron, which is supposed coming from infinity, is obtained (Darwin 1913; Schonland 1922)

$$tg[\tfrac{1}{2}(\pi+\phi)\sqrt{(1-(b_0/b)^2)}] = \frac{b}{b_0}\frac{v_1}{\sqrt{(1+v_1^2)}}\sqrt{(1-(b_0/b)^2)}. \tag{3·1}$$

Here ϕ is taken to be positive when the deflexion is towards the nucleus, $b_0 = Ee/mv_1$, E = charge of the nucleus, $-e$ = charge of the electron, m = rest-mass of the electron, v_1 = spatial part of the initial 4-velocity of the electron, and b is the impact parameter. For small angles of deflexion one can put $\sqrt{(1-(b_0/b)^2)} \approx 1$, and (3·1) gives

$$\phi \approx \frac{2Ee}{mb}\frac{\sqrt{(1+v_1^2)}}{v_1^2} \approx \frac{2Ee}{mbv_1} \quad \text{if} \quad v_1 \gg 1. \tag{3·2}$$

A different method of splitting up the field will now be introduced. The initial velocity 4-vectors of the nucleus and of the electron determine a set of parallel two-dimensional planes in space-time, which is taken as reference for defining the $\mathbf{a_k}$. These may be defined by the following geometrical method.

Only such functions $\mathbf{a_k}$ will be considered whose direction depends only on the direction of $\mathbf{k}$, so that $\mathbf{a}_{m\mathbf{k}} = \alpha\mathbf{a_k}$ for any m, where m and α are non-vanishing scalars. The direction of $\mathbf{a_k}$ (which determines the splitting up) is specified by its point at infinity, $A(K)$ say, where K is the point at infinity of the direction of $\mathbf{k}$. All the points at infinity of space-time form a 3-dimensional projective space S_∞^3. The $g_{\mu\nu}$-metric in space-time defines a null-quadric in S_∞^3. The selected set of parallel planes has a line r in S_∞^3, that intersects the null-quadric. The problem of defining the direction of $\mathbf{a_k}$ is equivalent to the problem of defining in S_∞^3 a point-function $A(K)$, where K is a variable point on the null-quadric, by means of projective operations, the line r being used as reference. Since the $\mathbf{a_k}$ have to satisfy (1·11), the point $A(K)$ must not lie on the polar plane of K. The condition (1·20) is automatically satisfied.

The function $A(K)$ may be defined as follows. Consider the plane $[rK]$ through the line r and the point K (figure 1). The polar line of r, $\bar{r}$ say, does not lie on this plane and it therefore intersects this plane in a point, O say. Take $A(K) = O$ for all the points K of the null-quadric which lie on the plane $[rK]$.

To express $A(K)$ analytically, introduce homogeneous co-ordinates X_μ, X_4 ($\mu = 0, 1, 2, 3$) into space-time by putting $x_\mu = X_\mu/X_4$. The points of S_∞^3 have $X_4 = 0$,

and the null-quadric is given by $X_\mu X^\mu \equiv (\mathbf{X}, \mathbf{X}) = 0$. Let P, Q be a pair of polar points on r, such that P lies outside and Q inside the null-quadric. The co-ordinates of P and Q satisfy the equation $(\mathbf{P}, \mathbf{Q}) = 0$. The equation of $\bar{r}$ is obtained by looking upon $\bar{r}$ as the intersection of the polar planes of P and Q, namely $(\mathbf{P}, \mathbf{X}) = 0$ and $(\mathbf{Q}, \mathbf{X}) = 0$. Any point of the plane $[rK] \equiv [PQK]$ may be represented parametrically by $X_\mu = \gamma_1 P_\mu + \gamma_2 Q_\mu + \gamma_3 K_\mu$. The point O in which $\bar{r}$ intersects the plane $[PQK]$ can therefore be obtained from the solution of the equations

$$O_\mu = \gamma_1 P_\mu + \gamma_2 Q_\mu + \gamma_3 K_\mu, \quad (\mathbf{P}, \mathbf{O}) = 0, \quad (\mathbf{Q}, \mathbf{O}) = 0.$$

Hence

$$O_\mu = (\mathbf{P}, \mathbf{K})(\mathbf{Q}, \mathbf{Q}) P_\mu + (\mathbf{Q}, \mathbf{K})(\mathbf{P}, \mathbf{P}) Q_\mu - (\mathbf{P}, \mathbf{P})(\mathbf{Q}, \mathbf{Q}) K_\mu.$$

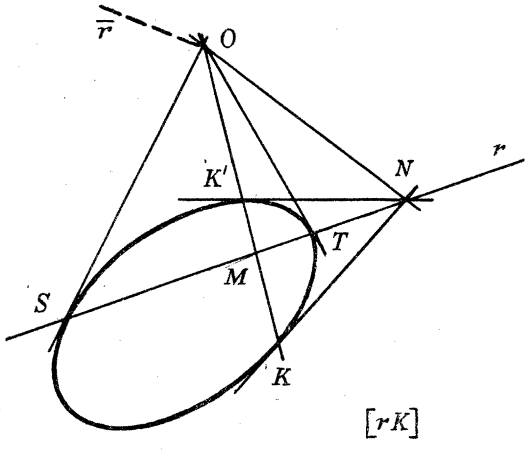

FIGURE 1

On returning to non-homogeneous co-ordinates in space-time, and dropping the part parallel to $\mathbf{k}$, which makes no change in the splitting up of the field, this gives

$$\mathbf{a}_\mathbf{k} = (\mathbf{p}, \mathbf{k})\mathbf{p} - (\mathbf{q}, \mathbf{k})\mathbf{q}, \tag{3·3}$$

where

$$\mathbf{p}^2 = -1, \quad \mathbf{q}^2 = 1, \quad (\mathbf{p}, \mathbf{q}) = 0. \tag{3·4}$$

The choice of $\mathbf{a}_\mathbf{k}$ above is the only simple one. For instance, the intersection of OK with r, $[OK, r] = M$ say, taken as $A(K)$ leads to the same result (3·3), and so does the choice as $A(K)$ of any point on the line OK. Further, if S, T are the two points where r intersects the null-quadric, one might think of taking the harmonic conjugate N of M with respect to S, T as $A(K)$, but this is not permissible since N can be shown to lie on the polar plane of K. Other choices for $\mathbf{a}_\mathbf{k}$ are more complicated and artificial, and so are not of interest.

Take a Lorentz frame of reference in which the nucleus is initially at rest at the origin, with the z_1-axis parallel to the spatial part of the initial 4-velocity of the electron, and the z_3-axis orthogonal in three-space to the plane of motion. The points at infinity of the z_0-axis and the z_1-axis are two polar points on r, which can be taken as the points Q and P introduced above, so that $q_\mu = \delta_{\mu 0}$ and $p_\mu = \delta_{\mu 1}$ in (3·3). Then (1·12) gives

$$l_{\mathbf{k}\mu} = (k_0 \delta_{\mu 0} + k_1 \delta_{\mu 1} - \tfrac{1}{2} k_\mu)/(k_0^2 - k_1^2). \tag{3·5}$$

This defines the new way of splitting up the field. The new longitudinal field is replaced in the Hamiltonian equations by the last term on the right-hand side of (2·7), as the $\mathbf{l}_{\mathbf{k}}$ of (3·5) are independent of the $\mathbf{z}_i$. The 'self-energy' term $j = i$ vanishes as discussed earlier. One may now put $\boldsymbol{\lambda} = 0$ in (2·7), and pass to the single-time Hamiltonian theory by putting the time variables of the two particles equal, $z_{i0} = z_0$, and adding the Hamiltonian equations of the two particles in the usual way to form the total Hamiltonian function. The three spatial components of the last term in (2·7) vanish. Thus the last term of (2·7) contributes the following 'interaction potential' $\mathscr{V}$ to the total Hamiltonian

$$\begin{aligned}\mathscr{V} &= \frac{1}{8\pi^2}\sum_i e_i \sum_{j\neq i} e_j \int \Delta(\mathbf{k})\frac{k_0}{k_0^2-k_1^2} e^{-i\Sigma_t k_t(z_{it}-z_{jt})}\,d\mathbf{k} \quad (t=1,2)\\ &= -\frac{Ee}{\pi}\delta(z_1-Z_1)\int_0^{2\pi} d\phi \int_0^{\infty}\frac{1}{r}e^{-i|z_2-Z_2|r\cos\phi}\,dr, \end{aligned} \tag{3·6}$$

where the co-ordinates of the nucleus are denoted by capital letters, and those of the electron by small letters. The integral over r in (3·6) diverges near the point $r = 0$ independently of $z_2 - Z_2$, since

$$\int_0^{\infty}\frac{1}{r}e^{-izr\cos\phi}\,dr = \int_0^{\rho}\frac{1}{r}dr + \int_{\rho}^{\infty}\frac{1}{r}e^{-izr\cos\phi}\,dr,$$

where $\rho > 0$ is small. Now

$$\frac{\partial}{\partial z}\int_{\rho}^{\infty}\frac{1}{r}e^{-\epsilon r - izr\cos\phi}\,dr = -i\int_{\rho}^{\infty}\cos\phi\, e^{-\epsilon r - izr\cos\phi}\,dr \to -\frac{1}{z},$$

where $e^{-\epsilon r}$ is a Cesàro convergence factor. Thus the constant infinity can be separated out in (3·6), and one gets

$$\mathscr{V} = 2Ee\,\delta(z_1-Z_1)\,[\log|z_2-Z_2| + \text{const.}] \tag{3·7}$$

as the expression replacing the Coulomb potential. It will be found that the infinite constant has no physical effect on the motion of the particles.

On the assumption that the interaction with the new transverse field may be neglected, the equations of motion of the two particles are as follows, where S and s are the proper-times of the nucleus and of the electron respectively,

$$M\frac{d^2Z_t}{dS^2} = -\frac{dz_0}{dS}\frac{\partial\mathscr{V}}{\partial Z_t}, \quad m\frac{d^2z_t}{ds^2} = -\frac{dz_0}{ds}\frac{\partial\mathscr{V}}{\partial z_t} \quad (t=1,2), \tag{3·8}$$

$$M\frac{d^2z_0}{dS^2} = -\Sigma_t\frac{dZ_t}{dS}\frac{\partial\mathscr{V}}{\partial Z_t}, \quad m\frac{d^2z_0}{ds^2} = -\Sigma_t\frac{dz_t}{ds}\frac{\partial\mathscr{V}}{\partial z_t}, \tag{3·9}$$

and the trivial equations

$$\left(\frac{dz_0}{dS}\right)^2 - \left(\frac{dZ_1}{dS}\right)^2 - \left(\frac{dZ_2}{dS}\right)^2 = 1, \quad \left(\frac{dz_0}{ds}\right)^2 - \left(\frac{dz_1}{ds}\right)^2 - \left(\frac{dz_2}{ds}\right)^2 = 1. \tag{3·10}$$

The system (3·8), (3·9) and (3·10) may be regarded as giving Z_t and z_t $(t = 1, 2)$ as functions of z_0. Let z_0^* be the value of z_0 for which $z_1 - Z_1 = 0$. Since the potential $\mathscr{V}$ contains a δ-function, which vanishes everywhere except at z_0^*, the motion of the

two particles is straightline and uniform before and after z_0^*. In a relativistic theory, where the velocity of the particles is always finite, they will change their direction of motion by a finite angle at the instant z_0^* due to the δ-potential. It may be pointed out that with such a δ-potential it would be physically wrong to neglect the nuclear recoil. The two sets of equations (3·8) give the following momentum equations

$$M\frac{dZ_1}{dS}+m\frac{dz_1}{ds}=mv_1, \tag{3·11}$$

where v_1 is the spatial part of the initial 4-velocity of the electron, and

$$M\frac{dZ_2}{dS}+m\frac{dz_2}{ds}=0. \tag{3·12}$$

The two equations (3·9) give the conservation of energy equation

$$M\frac{dz_0}{dS}+m\frac{dz_0}{ds}=\mathscr{V}+M+m\sqrt{(1+v_1^2)}. \tag{3·13}$$

The second equation of (3·8) for $t=2$ is

$$m\frac{d^2z_2}{ds^2}=-\frac{2Ee}{z_2-Z_2}\delta(z_1-Z_1)\frac{dz_0}{ds}. \tag{3·14}$$

Try a solution of the form

$$\frac{dz_2}{ds}=v_2'\epsilon(z_1-Z_1), \tag{3·15}$$

where v_2' is a constant and

$$\epsilon(z_1-Z_1)=\begin{cases}0 & \text{for} \quad -\infty<z_1-Z_1<0,\\ 1 & \text{for} \quad 0<z_1-Z_1<\infty.\end{cases}$$

The differentiation of (3·15) gives

$$\begin{aligned}\frac{d^2z_2}{ds^2}&=v_2'\delta(z_1-Z_1)\left\{\frac{dz_1}{ds}-\frac{dZ_1}{dS}\frac{dz_0/ds}{dz_0/dS}\right\}\\&=v_2'\delta(z_1-Z_1)\left\{\frac{dz_1}{ds}\left(1+\frac{m}{M}\frac{dz_0/ds}{dz_0/dS}\right)-\frac{m}{M}\frac{dz_0/ds}{dz_0/dS}v_1\right\},\end{aligned} \tag{3·16}$$

where (3·11) has been used. In a small neighbourhood of z_0^*, the equations (3·10), (3·11) and (3·12) give

$$\frac{dz_0/ds}{dz_0/dS}=\left\{\frac{1+(dz_1/ds)^2+(dz_2/ds)^2}{1+\dfrac{m^2}{M^2}(v_1-dz_1/ds)^2+\dfrac{m^2}{M^2}(dz_2/ds)^2}\right\}^{\frac{1}{2}}\sim\frac{dz_1/ds}{\dfrac{m}{M}\,|\,v_1-dz_1/ds\,|}\sim\frac{M}{m}.$$

Omitting in (3·16) the term which remains finite in a small neighbourhood of z_0^*, one has

$$\frac{d^2z_2}{ds^2}\sim 2v_2'\delta(z_1-Z_1)\frac{dz_0}{ds}, \tag{3·17}$$

so that, from (3·14),

$$v_2'=-Ee/mb, \tag{3·18}$$

where b is the impact parameter. The z_1-component v_1' of the final 4-velocity of the electron is obtained from the conservation of energy

$$M\sqrt{(1+V_1'^2+V_2'^2)}+m\sqrt{(1+v_1'^2+v_2'^2)}=M+m\sqrt{(1+v_1^2)},$$

where V_1' and V_2' are the z_1- and z_2-components of the final 4-velocity of the nucleus, namely,

$$V_1' = \frac{m}{M}(v_1 - v_1'), \quad V_2' = -\frac{m}{M}v_2'.$$

Thus v_1' satisfies the equation

$$[m^2v_1^2 - (M + m\sqrt{(1+v_1^2)})^2]\,v_1'^2 + 2m(m + M\sqrt{(1+v_1^2)})\,v_1v_1' + (m + M\sqrt{(1+v_1^2)})^2 - (1+v_2'^2)\,(M + m\sqrt{(1+v_1^2)})^2 = 0. \quad (3{\cdot}19)$$

The angle of deflexion ϕ of the electron, taken to be positive when the deflexion is towards the nucleus, is given by

$$\sin\phi = \frac{Ee}{mb}\frac{1}{\sqrt{(v_1'^2 + v_2'^2)}}, \quad (3{\cdot}20)$$

where v_1' satisfies (3·19) and v_2' is given by (3·18). The deflexion (3·20) can never exceed a right angle, as can be seen directly from the nature of the potential $\mathscr{V}$.

If $M \gg m$, the equation (3·19) gives as first approximation $v_1' = \sqrt{(v_1^2 - v_2'^2)}$, and (3·20) becomes

$$\sin\phi = Ee/mbv_1 \approx \phi$$

if ϕ is small. This is half the deflexion (3·2) derived from the Coulomb potential for ϕ small and $v_1 \gg 1$. Thus, for small velocities v_1, where the deflexion (3·1) agrees with experiment, the method proposed above of splitting up the field and neglecting the new transverse waves does not lead to agreement with experiment, which means that the effect of the new transverse waves is not negligible for small velocities of the electron. However, for a very high energy electron, the usual method of working only with the Coulomb interaction force involves neglecting radiation damping in a certain way, and in this case the new method proposed above may be a better one.

In conclusion, I wish to express my deepest gratitude to Professor P. A. M. Dirac, F.R.S. for his supervision and guidance during the course of the above work. I am also indebted to the British Council for the award of a scholarship.

References

Darwin, C. G. 1913 *Phil. Mag.* **25**, 201.
Dirac, P. A. M. 1939 *Ann. Inst. Poincaré*, **9**, 13.
Dirac, P. A. M. 1943 *Comm. Dubl. Inst. Adv. Studies*, A, no. 1.
Fermi, E. 1930 *R.C. Accad. Lincei*, **12**, 431.
Fermi, E. 1932 *Rev. Mod. Phys.* **4**, 87.
Hardy, G. H. 1908 *Trans. Camb. Phil. Soc.* **21**, 1.
Schonland, B. F. J. 1922 *Proc. Roy. Soc.* A, **101**, 299.