JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 1, NUMBER 1 JANUARY–FEBRUARY, 1960

Gauge Properties of Propagators in Quantum Electrodynamics*

BRUNO ZUMINO†
Lawrence Radiation Laboratory, University of California, Berkeley, California
(Received October 18, 1959)

The effect of a change of gauge on the propagators is studied systematically for quantum electrodynamics. Various gauges are considered, among them the Coulomb, the Landau, the Feynman, and the Yennie gauges. The equivalence of the various formulations of the theory is demonstrated. For the relativistic gauges, the transformation of the wave function renormalization constant is described.

1. INTRODUCTION

THE propagators of quantum electrodynamics are affected by ambiguities because the theory is invariant under gauge transformations. In this note we shall investigate systematically this ambiguity of the unrenormalized propagators and shall give the connection between the various gauges.

We use the Heisenberg equations of motion in the Coulomb gauge. In this gauge the longitudinal part of the magnetic vector potential is a c number. The relativistic covariance of this formulation of the theory (briefly discussed in Appendix B) has been known for a long time. It is also known that the relativistic S-matrix theory of Feynman can be derived directly in the Coulomb gauge. In this note we show how the propagators in other gauges (including the relativistic ones) are connected to those in the Coulomb gauge. It is therefore clear that the Heisenberg equations in the Coulomb gauge provide a complete basis for quantum electrodynamics. The present formulation has the desirable feature that only physical states are considered, no supplementary condition, and no indefinite product in Hilbert space are necessary.[1]

The study of the gauge transformation of the propagators becomes particularly simple and elegant if one employs the method of functional derivatives. This method, which has been largely used by Schwinger,[2] makes use of a generating functional (which we call Z) from which all propagators can be obtained by functional differentiation. The gauge ambiguity of the propagators will be shown here to arise from the gauge ambiguity in the functional Z itself. In previous work this ambiguity has been ignored to a large extent, with a consequent lack of clarity concerning the meaning of the operations to be performed, as for instance the differentiations with respect to the external sources.

The Heisenberg equations in the Coulomb gauge depend upon a c-number gauge function Λ. As shown in Sec. 2, a change in the gauge function induces a gauge transformation in the generating functional Z. We call a quantity gauge-invariant if it is invariant with respect to this c-number gauge transformation. In Sec. 3 we derive the functional differential equations satisfied by Z and extend the definition to more general gauges, characterized by an operator four-vector a_μ. A suitable choice of a_μ gives the Coulomb gauge in any Lorentz frame. Another choice of a_μ gives a relativistic gauge in

* This work was supported in part by the U. S. Atomic Energy Commission and in part by the Office of Naval Research. A preliminary account of it was given in the Bull. Am. Phys. Soc. Ser. II, 4, 280 (1959).

† Permanent address: Physics Department, New York University, New York, New York.

[1] The Coulomb gauge has been used recently by Schwinger and Johnson (Kenneth Johnson, Massachusetts Institute of Technology, private communication, 1959). They have arrived independently at several of the results described in the present note.

[2] J. Schwinger, Proc. Natl. Acad. Sci. U. S. **37**, 452 (1951). For a more detailed discussion see, e.g., K. Symanzik, Z. Naturforsch. **9a**, 809 (1954) and E. S. Fradkin, Doklady Akad. Nauk S.S.S.R. **98**, 47 (1954) and **100**, 897 (1955).

which the zero order photon propagator has the form

$$D_{\mu\nu}=[g_{\mu\nu}-(\partial_\mu\partial_\nu/\partial^2)]D_c.$$

This gauge was widely used by Landau,[3] and we shall call it the Landau gauge. It is very convenient for the study of ultraviolet divergences. In Sec. 4 we proceed to a further generalization, introducing gauges depending in addition upon a function M. If we start from the Landau gauge $D_{\mu\nu}$, we can obtain in this fashion all gauges where the zero-order photon propagator has the form[4]

$$D_{\mu\nu}+\partial_\mu\partial_\nu M.$$

In particular, for $M=\partial^{-2}D_c$, one can cancel the $\partial_\mu\partial_\nu$ term and obtain the ordinary Feynman form of the photon propagator. For $M=3\partial^{-2}D_c$ one obtains a gauge used by Fried and Yennie.[5] They have used it for the study of infrared divergences. Our general gauge is now characterized by the c-number gauge function Λ plus a_μ and M. Quantities invariant with respect to changes in Λ are automatically invariant for changes of a_μ and M. Therefore, these more general gauge transformations do not correspond to new invariance properties of the theory. Rather they allow to establish a connection between different existing formulations of it.

The effect of changes in the function M (within the class of relativistic gauges) on the wave function renormalization constant Z_2 has been studied in particular by Johnson and the present author.[6] They have given an exact transformation formula for Z_2. By using it, one can easily verify that Z_2 (to order e^2) has no ultraviolet divergence in the Landau gauge and no infrared divergence in the Yennie gauge.

2. GENERATING FUNCTIONAL IN THE COULOMB GAUGE

We recall here first the equations for the Heisenberg operators of quantum electrodynamics in the Coulomb gauge.

The Hamiltonian is

$$\mathcal{H}=\int d\mathbf{x}T_{oo} \tag{1}$$

with

$$T_{oo}=\tfrac{1}{2}(\mathbf{E}^2+\mathbf{H}^2)-\mathbf{j}\cdot\mathbf{A}$$
$$+\frac{1}{2}\left[\psi^*,\frac{-i}{2}\boldsymbol{\alpha}\cdot\nabla\psi\right]+\frac{1}{2}\left[\frac{i}{2}\nabla\psi^*,\boldsymbol{\alpha}\psi\right]+\frac{1}{2}[\psi^*,\beta m\psi], \tag{2}$$

and the commutation relations are

$$[E_r^{tr}(\mathbf{x}),A_s(\mathbf{x}')]=i(\delta_{rs}-\nabla_r\nabla_s\nabla^{-2})\delta(\mathbf{x}-\mathbf{x}') \tag{3}$$

and

$$\{\psi_\sigma^*(\mathbf{x}),\psi_\tau(\mathbf{x}')\}=\delta_{\sigma\tau}\delta(\mathbf{x}-\mathbf{x}'), \tag{4}$$

while other commutators vanish. Here we have[7]

$$\mathbf{E}=\mathbf{E}^{tr}-\nabla\phi, \tag{5}$$

$$\phi=-\nabla^{-2}\rho, \tag{6}$$

$$\mathrm{div}\mathbf{E}^{tr}=0, \tag{7}$$

$$\rho=\frac{e}{2}[\psi^*,\psi],\quad \mathbf{j}=\frac{e}{2}[\psi^*,\boldsymbol{\alpha}\psi], \tag{8}$$

$$\mathbf{H}=\mathrm{curl}\mathbf{A}, \tag{9}$$

and

$$\mathbf{A}=\mathbf{A}^{tr}+\nabla\Lambda. \tag{10}$$

Since the longitudinal part of $\mathbf{A}$ commutes with all other operators, $\Lambda(\mathbf{x},t)$ can be taken as an arbitrary c number.

The foregoing Hamiltonian and commutation relations give rise to the correct equations of motion,

$$i(\partial/\partial t)\psi=(m\beta-i\boldsymbol{\alpha}\cdot\nabla-e\boldsymbol{\alpha}\cdot\mathbf{A})\psi+(e/2)\{\phi,\psi\}, \tag{11}$$

$$(\partial/\partial t)\mathbf{E}^{tr}=\mathrm{curl}\mathbf{H}-\mathbf{j}^{tr}, \tag{12}$$

$$(\partial/\partial t)\mathbf{A}^{tr}=-\mathbf{E}^{tr}, \tag{13}$$

so that

$$\mathbf{E}=-(\partial/\partial t)\mathbf{A}^{tr}-\nabla\phi. \tag{14}$$

If we set

$$A^o=\phi-\partial\Lambda/\partial t, \tag{15}$$

we can write

$$\mathbf{E}=-(\partial/\partial t)\mathbf{A}-\nabla A^o. \tag{16}$$

The set of equations presented in the foregoing is invariant under the c-number gauge transformation,

$$\begin{aligned}\Lambda(\mathbf{x},t)&\to\Lambda(\mathbf{x},t)+\tilde{\Lambda}(\mathbf{x},t)\\ \mathbf{A}&\to\mathbf{A}+\nabla\tilde{\Lambda}\\ A^o&\to A^o-(\partial/\partial t)\tilde{\Lambda}\\ \psi&\to\psi\exp ie\tilde{\Lambda}.\end{aligned} \tag{17}$$

The choice $\Lambda=0$ would fix the gauge (transverse Coulomb gauge). However, it is convenient to leave the gauge function arbitrary and to use the gauge transformation to define those operators that are gauge-invariant.

The operator condition (7) is consistent with the commutation relation (3). It is known that the formulation of quantum electrodynamics given above is relativistically covariant, in spite of the apparent asymmetry between space and time variables. The covariance is proven directly in Appendix B. An alternative proof

[3] L. D. Landau, A. A. Abrikosov, and I. M. Khalatnikov, Doklady Akad. Nauk S.S.S.R. **95**, 773 (1954).

[4] The M transformation has been given first by L. D. Landau and I. M. Khalatnikov, J. Exptl. Theoret. Phys. (U.S.S.R.) **29**, 89 (1955); English translation in Soviet Phys. JETP **2**, 69 (1956). Their derivation, however, is based on an operator gauge transformation, the validity of which appears rather questionable.

[5] H. M. Fried and D. R. Yennie, Phys. Rev. **112**, 1391 (1958).

[6] K. Johnson and B. Zumino, Phys. Rev. Letters **3**, 351 (1959).

[7] The transverse part of a vector $\mathbf{D}$ is of course defined by

$$\mathbf{D}^{tr}=\mathbf{D}-\nabla\nabla^{-2}\,\mathrm{div}\mathbf{D}.$$

results from the investigation carried out in Sec. 3 of the covariance of the equations for the propagators.

We consider now the generating functional[8]

$$Z_\Lambda[\eta,\bar\eta,J_\mu]=\left\langle 0\left|T\exp i\int dx(\bar\eta\psi+\bar\psi\eta+J_\mu A^\mu)\right|0\right\rangle \times\exp\frac{1}{2}\int dx J_\sigma\nabla^{-2}J_\sigma. \quad (18)$$

All vacuum expectation values of time-ordered products of Heisenberg operators can be constructed from Z by functional differentiation. Actually the form (18) is somewhat redundant, since A^0 is given through Eqs. (6) and (15) in terms of the spinor field. One could set $J_0=0$ and work with the resulting functional. The form we have chosen, however, is more convenient for the investigation of transformation properties. We wish to emphasize here that it is **not** assumed that the various components of J_μ satisfy a continuity equation like

$$\partial_\mu J^\mu=0. \quad (19)$$

The complete arbitrariness is necessary in order to operate on Z with functional derivatives. Only after all functional differentiations have been performed will one require that J_μ satisfies Eq. (19) (or even that it vanishes) and also, of course, that

$$\eta=\bar\eta=0. \quad (20)$$

The gauge ambiguity of Z and of the propagators is in fact connected with the necessity of extending (in an arbitrary way) the definition of Z to unphysical values of J_μ not satisfying Eq. (19). This fact can be illustrated by exhibiting the dependence of Z upon the gauge function Λ. It follows immediately from Eq. (17) that

$$Z_\Lambda[\eta,\bar\eta,J_\mu]=Z_0[\eta\exp(-ie\Lambda),\bar\eta\exp(ie\Lambda),J_\mu] \times\exp ie\int J_\mu\partial^\mu\Lambda dx. \quad (21)$$

where the right-hand side refers to the value $\Lambda=0$. This relation can be cast into the differential form,

$$i\frac{\delta Z}{\delta\Lambda}=\left(\partial^\mu J_\mu+e\eta\frac{\delta}{\delta\eta}-e\bar\eta\frac{\delta}{\delta\bar\eta}\right)Z. \quad (22)$$

3. FUNCTIONAL EQUATIONS AND THEIR TRANSFORMATION PROPERTIES

By using the definition (18), the equations for the Heisenberg operators and the commutation relations, one can show that the generating functional satisfies the functional differential equations:

$$\left\{\partial^\lambda\left(\partial_\mu\frac{1}{i}\frac{\delta}{\delta J^\lambda}-\partial_\lambda\frac{1}{i}\frac{\delta}{\delta J^\mu}\right)+(\delta_\mu{}^\lambda+a_\mu\partial^\lambda)\times\left(ie\frac{1}{i}\frac{\delta}{\delta\eta}\gamma_\lambda\frac{1}{i}\frac{\delta}{\delta\bar\eta}-J_\lambda\right)\right\}Z=0, \quad (23)$$

$$\left\{\left[\gamma^\mu\left(\partial_\mu-ie\frac{1}{i}\frac{\delta}{\delta J^\mu}\right)+m\right]\frac{1}{i}\frac{\delta}{\delta\bar\eta}-\eta\right\}Z=0, \quad (24)$$

$$\left\{-\frac{1}{i}\frac{\delta}{\delta\eta}\left[-\gamma^\mu\left(\partial_\mu+ie\frac{1}{i}\frac{\delta}{\delta J^\mu}\right)+m\right]-\bar\eta\right\}Z=0, \quad (25)$$

$$\left(a^\mu\frac{1}{i}\frac{\delta}{\delta J^\mu}+\Lambda\right)Z=0. \quad (26)$$

Equation (25) follows from the (Pauli) adjoint of Eq. (11). The source terms J_λ, η, and $\bar\eta$ arise from the time differentiation of the time-ordered products.

The operator vector a_μ, introduced here at first for the sake of concise notation, is defined by

$$a_0=0,\quad a_r=-\nabla_r\nabla^{-2} \quad (27)$$

in the Lorentz frame chosen to define the Coulomb gauge. The form of the equations suggests, however, that we consider more generally operator vectors a_μ (operating on functions of the four-dimensional variable x) satisfying

$$\partial^\mu a_\mu=-1. \quad (28)$$

A suitable choice for a_μ will give, in particular, the Coulomb gauge in any Lorentz frame.[9] Another choice of interest is the limit as ϵ tends to zero of

$$a_\mu=\partial_\mu(-\partial^2-i\epsilon)^{-1}. \quad (29)$$

The different choices of a_μ give rise to different functionals Z, which are not related simply by changes in the gauge function Λ, and yet give rise to physically equivalent formulations of the theory. As shown in Eq. (31) below, a change δa_μ of a_μ which preserves Eq. (28)

$$\partial^\mu(\delta a_\mu)=0, \quad (30)$$

can be considered as a generalized type of gauge transformation, and the various choices of a_μ as various possible gauges. The gauge given by Eq. (29) will be called the Landau gauge. Let us notice that the condition (28) is required for the consistency of the functional Eq. (23), as one can see by operating on this equation with ∂^μ.

From our present more general point of view, the

[8] In order to avoid formal difficulties in connection with the use of the anticommuting spinor sources η and $\bar\eta$, it is best not to interpret the bar as a relation of hermitian conjugation between η and $\bar\eta$. Rather, one should consider $\bar\eta(x)$ and $\eta(x)$ as independent anticommuting symbols and carry out all the formal operations from this point of view. In the final expression one always sets $\eta=\bar\eta=0$, or more correctly, one takes that part of the expression which is independent of η and of $\bar\eta$.

[9] A covariant expression for a_μ corresponding to the Coulomb gauge in the Lorentz frame characterized by the unit time-like vector n_μ is

$$a_\mu=-\frac{\partial_\mu+n_\mu(n\cdot\partial)}{\partial^2+(n\cdot\partial)^2}.$$

Eqs. (23) to (26) are obviously convariant, since we chose to treat a_μ as a four vector. However, one must now investigate how the solution changes with an infinitesimal change δa_μ in a_μ. We show in the following that the corresponding change in Z can be written as

$$\delta Z = i\int (\delta a_\mu)\frac{\delta}{\delta J_\mu}\frac{\delta}{\delta\Lambda}Z. \tag{31}$$

Clearly, if a functional Θ constructed from Z by functional differentiation is gauge invariant in the sense that it does not change when one changes the gauge function Λ, then we have

$$\delta\Theta/\delta\Lambda = 0. \tag{32}$$

From Eq. (31) one then sees that such a functional also does not change in correspondence to a change δa_μ of a_μ.

We proceed now to prove our basic Eq. (31) by showing that the functional differential equations satisfied by Z remain invariant if one performs simultaneous changes of a_μ and of Z. An alternative proof proceeds directly from the explicit expression for Z and is sketched in Appendix A. The invariance of Eqs. (24) and (25) is trivial, since they do not contain J_μ or Λ. The invariance of Eq. (26) is also easily verified. To check the invariance of Eq. (23), one needs a simple identity satisfied by any solution of the functional equations. If one applies $-\delta/i\delta\eta$ to Eq. (24) and $-\delta/i\delta\bar{\eta}$ to Eq. (25) and subtracts, one obtains[10]

$$\{\partial^\lambda[-(\delta/i\delta\eta)\gamma_\lambda(\delta/i\delta\bar{\eta})] - \eta(\delta/i\delta\eta)+\bar{\eta}(\delta/i\delta\bar{\eta})\}Z = 0. \tag{33}$$

Therefore, from Eq. (22), the terms containing a_μ and J_μ in Eq. (23) can be written as

$$-ia_\mu(\delta Z/\delta\Lambda) - J_\mu Z. \tag{34}$$

The invariance of Eq. (23) is now easily verified.

The basic transformation formula (31) allows one, at least in principle, to transform all propagators from one gauge a_μ to another. The transformation is particularly simple in the case $\eta=\bar{\eta}=0$. If one remembers Eq. (22), one has, in this case,

$$\delta L = \int \partial^\rho J_\rho(\delta a_\mu)(\delta/\delta J_\mu)L, \tag{35}$$

where we have defined

$$L[J_\mu] = Z[0,0,J_\mu]. \tag{36}$$

One should observe here that we have

$$[\partial^\rho J_\rho,\ (\delta a_\mu)(\delta/\delta J_\mu)] = 0 \tag{37}$$

as a consequence of Eq. (30). Therefore, the operation to be performed on L is a pure substitution,

$$L'[J_\mu] = L[J_\mu'], \tag{38}$$

with

$$J_\mu' = J_\mu + (\delta a_\mu)\partial^\lambda J_\lambda. \tag{39}$$

We have indicated with a prime the functional in the new gauge. The transformation (38), (39) is correct for finite changes δa_μ also. The functionals in the two gauges coincide when Eq. (19) is satisfied.

The change induced in

$$\mathfrak{A}_\mu = \frac{1}{L}\frac{1}{i}\frac{\delta}{\delta J^\mu}L \tag{40}$$

is a simple gauge transformation. Indicating only the dependence on J_μ, one has

$$\mathfrak{A}_\mu'[J_\rho] = \mathfrak{A}_\mu[J_\rho'] + \partial_\mu(\delta a^\lambda)\mathfrak{A}_\lambda[J_\rho']. \tag{41}$$

Similarly, one obtains for the exact photon propagator

$$\mathcal{G}_{\mu\nu} = \frac{\delta\mathfrak{A}_\nu}{\delta J^\mu} = \frac{\delta\mathfrak{A}_\mu}{\delta J^\nu}, \tag{42}$$

the relation

$$\mathcal{G}_{\mu\nu}'[J] = [\delta_\mu{}^\lambda + \partial_\mu(\delta a^\lambda)][\delta_\nu{}^\rho + \partial_\nu(\delta a^\rho)]\mathcal{G}_{\lambda\rho}[J']. \tag{43}$$

The relations (41) and (43) are of course also valid for finite gauge transformations. An obvious simplification occurs when Eq. (19) is satisfied.

When differentiations with respect to η or $\bar{\eta}$ occur, it does not seem possible to obtain formulas of simplicity comparable to the above. Thus, for the exact electron propagator

$$G(x,y) = \frac{-i}{Z}\frac{\delta^2 Z}{\delta\eta(y)\delta\bar{\eta}(x)}, \tag{44}$$

one obtains from Eqs. (21) or (22), after setting $\eta=\bar{\eta}=0$,

$$L_\Lambda G_\Lambda(x,y) = \exp\left[ie(\Lambda(x)-\Lambda(y)) + i\int J_\mu\partial^\mu\Lambda\right]L_0G_0(x,y), \tag{45}$$

or

$$i\frac{\delta}{\delta\Lambda(z)}[LG(x,y)] = [\partial^\mu J_\mu(z) - e\delta(x-z) + e\delta(y-z)]LG(x,y). \tag{46}$$

Finally, from Eq. (31), we obtain

$$L'G'[x,y;J_\mu(z)] = LG[x,y;J_\mu'(z) - (\delta a_\mu)(e\delta(x-z) - e\delta(y-z))], \tag{47}$$

where J_μ' is given by Eq. (39). This formula, which was given first by Schwinger and Johnson,[1] shows clearly what an intricate connection exists between the electron propagators in two different gauges of the type considered here. In particular, set $J_\mu=0$; one sees that to obtain G' it is necessary to know G for nonvanishing

[10] For $\eta=\bar{\eta}=0$, Eq. (33) gives the conservation of the vacuum currents.

values of the current equal to

$$\mathcal{J}_\mu(z) = -e(\delta a_\mu)[\delta(x-z)-\delta(y-z)]. \quad (48)$$

Notice however that, because of Eq. (30), one has

$$\partial^\mu \mathcal{J}_\mu(z) = 0. \quad (49)$$

4. TRANSITION TO MORE GENERAL GAUGES

In the preceding section, we have considered the generating functional and the propagators in the various gauges specified by different choices of Λ and of a_μ. More general types of gauges can also be considered. Of particular interest are changes of the basic functional given by

$$\delta Z = \frac{i}{2}\int\int \frac{\delta}{\delta\Lambda}(\delta M)\frac{\delta}{\delta\Lambda} Z, \quad (50)$$

where $\delta M(x-y)$ is an arbitrary infinitesimal function even in $x-y$. The generating functional can be considered now as dependent upon a new (even) function $M(x-y)$ in such a way that the infinitesimal change δM induces in Z the change given by Eq. (50). The explicit expression for Z as a function of Λ, a_μ, and M is given in Appendix A.

Clearly a functional $\mathcal{O}$ which is invariant under the original c-number gauge transformations (17), and which therefore satisfies Eq. (32), will also not be affected by the change (50).

We can easily deduce the effect of the change (50) on the first few propagators. First, setting $\eta=\bar{\eta}=0$, we see from Eq. (22) that

$$\delta L = -\frac{i}{2}\int\int \partial^\mu J_\mu(\delta M)\partial^\rho J_\rho L, \quad (51)$$

or, in finite form

$$L'[J_\mu] = \exp\left\{-\frac{i}{2}\int\int \partial^\mu J_\mu(\delta M)\partial^\rho J_\rho\right\} L. \quad (52)$$

From this it follows that

$$\mathcal{C}_\mu'[J] = \mathcal{C}_\mu[J] + \partial_\mu \int (\delta M)\partial^\rho J_\rho. \quad (53)$$

and

$$\mathcal{G}_{\mu\nu}' = \mathcal{G}_{\mu\nu} + \partial_\mu\partial_\nu(\delta M). \quad (54)$$

Similarly, for the electron propagator, we obtain

$$\delta(LG) = \frac{i}{2}\int\int \frac{\delta}{\delta\Lambda}(\delta M)\frac{\delta}{\delta\Lambda}(LG), \quad (55)$$

and, using Eqs. (46) and (52), the result in finite form,

$$G'(x,y;J) = \exp\left\{ie^2[\delta M(x-y)-\delta M(0)] + ie\int[\delta M(x-z)-\delta M(y-z)]\partial^\rho J_\rho(z)dz\right\} \times G(x,y;J). \quad (56)$$

For $J=0$, of course, we have

$$G'(x,y;0) = \exp\{ie^2[\delta M(x-y)-\delta M(0)]\}G(x,y;0). \quad (57)$$

Finally, Eq. (56) can be used to obtain the change in the propagator

$$C_\mu(x,y;z) = [\delta/\delta J^\mu(z)]G(x,y;J), \quad (58)$$

which is closely related to the vertex part. Setting $J_\mu=0$ after the differentiation, one has

$$C_\mu'(x,y;z) = \exp\{ie^2[\delta M(x-y)-\delta M(0)]\} \times\left\{C_\mu(x,y;z) - ieG(x,y)\frac{\partial}{\partial z^\mu}[\delta M(x-z)-\delta M(y-z)]\right\}. \quad (59)$$

As seen in Eq. (54), the gauge transformations considered in this section permit us to operate the transition from what we have called in the "Introduction" the Landau form of the photon propagator to the Feynman and the Yennie forms. To achieve this, one has only to chose

$$\delta M = \gamma(-\partial^2 - i\epsilon)^{-2}\delta(x-y), \quad (60)$$

with a suitable constant γ. Since the choice (60) for δM gives rise to a rather singular function, a regularization procedure is necessary before evaluating consequences of the gauge transformation. This has been done by Johnson and Zumino.[6] From the transformation formula, they have deduced information about the infrared structure of the electron propagator.

If the function δM has a reasonable Fourier transform and vanishes at large distances in coordinate space, one can use Eq. (57) to give the change induced in the wave function renormalization const Z_2 by going from one relativistic gauge to another. It is sufficient to remember that, in a relativistic gauge, Z_2 can be defined from

$$G(x-y) \approx Z_2 G_m(x-y) \quad (61)$$

for large coordinate separation. Here G_m is the zero-order Feynman propagator for a Dirac particle of mass m. On comparing with Eq. (57), we obtain

$$Z_2' = \exp\{-e^2\delta M(0)\}Z_2. \quad (62)$$

ACKNOWLEDGMENT

The author wishes to thank Dr. Kenneth Johnson for many illuminating discussions on the topics treated in this note.

APPENDIX A: EXPLICIT FORM OF THE GENERATING FUNCTIONAL

The physically interesting solution of the functional Eqs. (23) to (26) can be obtained by following methods that were developed first without giving particular consideration to questions of gauge invariance.[2] We shall exhibit here the solution, so as to see its explicit dependence on the gauge.

Consider the electron propagator in an external field B_μ,

$$\tilde{G}[B_\mu; x,y]=\{\gamma^\mu(\partial_\mu-ieB_\mu)+m-i\epsilon\}^{-1}\delta(x-y). \quad \text{(A1)}$$

The vacuum polarization (closed loops) resulting from the external field can be expressed by the functional

$$F[B_\mu]=\exp\{-Tr\log(\tilde{G}[B](\tilde{G}[0])^{-1})\}, \quad \text{(A2)}$$

where we have used an obvious notation of multiplication for integral kernels, and the symbol Tr means the trace taken with respect to space time as well as spinor indices. Notice that

$$\tilde{G}[B_\mu+\partial_\mu\Lambda; x,y]=\exp\{ie[\Lambda(x)-\Lambda(y)]\}\tilde{G}[B_\mu; x,y], \quad \text{(A3)}$$

and therefore

$$F[B_\mu+\partial_\mu\Lambda]=F[B_\mu]. \quad \text{(A4)}$$

It can also be verified, by direct evaluation, that

$$\left(\frac{\delta}{\delta B_\mu}-e\frac{1}{i}\frac{\delta}{\delta\eta}\gamma^\mu\frac{1}{i}\frac{\delta}{\delta\bar{\eta}}\right)\{\exp[i\bar{\eta}\tilde{G}(B_\mu)\eta]F(B_\mu)\}=0. \quad \text{(A5)}$$

The generating functional considered in the text is now given, in its dependence upon Λ, a_μ, and M, by

$$Z[\eta,\bar{\eta},J_\mu]=\exp(i\bar{\eta}\tilde{G}[\delta/i\delta J_\mu]\eta)F[\delta/i\delta J_\mu]\times\exp\left[\frac{i}{2}(J_\mu+a_\mu\partial^\rho J_\rho)D_c(J^\mu+a^\mu\partial^\rho J_\rho)-i\partial^\rho J_\rho\Lambda-\frac{i}{2}\partial^\rho J_\rho M\partial^\lambda J_\lambda\right], \quad \text{(A6)}$$

where D_c is the Feynman function,

$$D_c(x-y)=(-\partial^2-i\epsilon)^{-1}\delta(x-y). \quad \text{(A7)}$$

The expression (A6) satisfies Eqs. (23) to (25). It does not satisfy Eq. (26) unless one sets $M=0$. On the other hand, the dependence on M in Eq. (A6) clearly agrees with Eq. (50).

We indicate now briefly how one can verify that the explicit formula given actually satisfies the functional equations, without however going into the question of the boundary conditions that ensure the uniqueness of the solution. The only equation that is not trivially satisfied is Eq. (23). If we operate with

$$\partial^\lambda\left(\partial_\mu\frac{1}{i}\frac{\delta}{\delta J^\lambda}-\partial_\lambda\frac{1}{i}\frac{\delta}{\delta J^\mu}\right) \quad \text{(A8)}$$

on the last exponential in Eq. (A6) we obtain, after simplifications involving the use of Eq. (28), a factor

$$(J_\mu+a_\mu\partial^\rho J_\rho). \quad \text{(A9)}$$

The terms containing Λ and M give no contribution. Now one must pass the factor to the left of the terms in Eq. (A6) which contain the functional derivatives $\delta/i\delta J_\mu$. If one uses Eq. (A5), one sees that the net effect is to replace J_μ with

$$J_\mu-ie\frac{1}{i}\frac{\delta}{\delta\eta}\gamma_\mu\frac{1}{i}\frac{\delta}{\delta\bar{\eta}}, \quad \text{(A10)}$$

so that Eq. (23) obtains.

By using the explicit form (A6), one can prove again the formulas given in the text for the various changes of gauge. The basic tools are now the relations

$$F\left[\frac{1}{i}\frac{\delta}{\delta J_\mu}\right]\partial^\rho J_\rho=\partial^\rho J_\rho F\left[\frac{1}{i}\frac{\delta}{\delta J_\mu}\right] \quad \text{(A11)}$$

and

$$\tilde{G}\left[\frac{1}{i}\frac{\delta}{\delta J_\mu}; x,y\right]\partial^\rho J_\rho(z)=[\partial^\rho J_\rho(z)-e\delta(x-z)+e\delta(y-z)]\times\tilde{G}\left[\frac{1}{i}\frac{\delta}{\delta J_\mu}; x,y\right], \quad \text{(A12)}$$

which are immediate consequences of (A3) and (A4). The expression for $L[J_\mu]$ is obtained directly from Eq. (A6) by setting $\eta=\bar{\eta}=0$. The expression for the electron propagator (44) is then given by the equation

$$L[J_\mu]G[x,y;J_\mu]=\tilde{G}\left[\frac{1}{i}\frac{\delta}{\delta J_\mu}; x,y\right]L[J_\mu]. \quad \text{(A13)}$$

In this form the evaluation of changes induced by a change in the function M becomes particularly simple.

APPENDIX B: COVARIANCE OF THE OPERATOR FORMALISM

The covariance of the equations for the Heisenberg operators in the Coulomb gauge can be proven by exhibiting the ten fundamental generators of infinitesimal Lorentz transformations P_μ and $M_{\mu\nu}$ and by verifying that they satisfy the correct structure relations. Since the covariance of the operator equations under space-time translations and space rotations is obvious, we shall restrict ourselves to a very brief discussion for the case of actual Lorentz transformations.

The corresponding generators are given by

$$M_{or}=x_oP_r+\int x_rT_{oo}d\mathbf{x}, \quad \text{(B1)}$$

where T_{oo} is the component of the energy-momentum tensor given in Eq. (2). The form (B1) is obtained by analogy from the classical theory. The change induced by an infinitesimal Lorentz transformation in any operator ξ is given by

$$\delta\xi=i[\tfrac{1}{2}\epsilon_{\mu\nu}M^{\mu\nu},\xi], \quad \text{(B2)}$$

where the antisymmetric infinitesimal tensor $\epsilon_{\mu\nu}$ characterizes the Lorentz transformation in question.

The changes induced in the basic operators $\mathbf{A}$ and ψ are easily obtained. It can be shown from the commutation relations (3) and (4) that

$$i[M_{or},A_l]=-(x_o\partial_r-x_r\partial_o)A_l+\delta_{rl}A_o+\partial_l B_r \quad \text{(B3)}$$

and

$$i[M_{or},\psi]=-(x_o\partial_r-x_r\partial_o)\psi-\tfrac{1}{2}\alpha_r\psi+\frac{ie}{2}\{\psi,B_r\}, \quad \text{(B4)}$$

where

$$B_r=\nabla_s\nabla^{-2}(x_rE_s)-x_r\nabla_s\nabla^{-2}E_s. \quad \text{(B5)}$$

Equations (B3) and (B4) show that, in going to a new Lorentz frame, not only do A_μ and ψ transform like a four-vector and a spinor, respectively, but that they also undergo an operator gauge transformation which reestablishes the Coulomb gauge in the new Lorentz frame. It is easy to verify that the operator gauge transformation leaves invariant symmetrized expressions like (8), so that the current and charge densities, for instance, transform like a four vector. It is worth noticing that the symmetrization, necessary for charge conjugation invariance, also appears necessary to ensure the relativistic covariance of the theory. One can give a more convenient form to the gauge operator B, in which the transverse and the longitudinal parts of the electric field are separated. We give only the result

$$B_r=\nabla^{-2}E_r^{tr}-\tfrac{1}{2}[x_r\nabla^{-2}\rho-\nabla^{-2}(x_r\rho)]. \quad \text{(B6)}$$

One can now proceed to verify the structure relations, the expression for the space components M_{rs} and for P_r being well known. This will not be done here. A simpler check on the covariance of the theory is the direct substitution of the transformed quantities obtained from Eqs. (B3) and (B4) into the differential equations. Obviously, only the invariance of the Dirac equation under the gauge part of Eq. (B4) requires detailed examination, since the Lorentz covariance of the unquantized theory is well known. Both procedures result in proving the covariance.[11]

[11] It has been pointed out by Schwinger that the analogous covariance test fails if an anomalous Pauli moment is introduced into the theory. Schwinger has also shown how the covariance of the theory can be saved by the further introduction of a term describing the self-interaction of the magnetic moment density. The author would like to thank Dr. Glashow for an illuminating correspondence on this question of covariance.

Dimensional Changes in Crystals Caused by Dislocations

R. A. TOUPIN, *U. S. Naval Research Laboratory, Washington, D. C.*

AND

R. S. RIVLIN, *Brown University, Providence, Rhode Island*

(Received November 2, 1959)

According to the classical linear elasticity theory, if one or more dislocations are introduced into a body of elastic material, the average value of each of the infinitesimal strain components is zero; in particular, the change in volume is zero. This result seems not to be in accord with experimental data on cold worked metals. In this paper we use nonlinear elasticity theory to show how changes in the average dimensions of elastic bodies, either isotropic or anisotropic, resulting from the introduction of dislocations, can be calculated. In particular, we derive an explicit relation between the resultant change in volume, the stored energy, and the pressure derivatives of the elastic moduli.

1. INTRODUCTION

ACCORDING to classical elasticity theory, if one or more dislocations are introduced into a body of elastic material, the average value of each of the infinitesimal strain components in a rectangular Cartesian coordinate system is zero. It follows that the change of volume is zero; that in a prism of the material the change of cross-sectional area and the average change of length are zero; that in a rectangular block the average changes in the dimensions parallel to the edges are zero. These results do not, of course, prove that the changes are in fact zero physically. They merely imply that the classical elasticity theory provides an inadequate basis for their calculation.

In the present paper we make use of the second-order elasticity theory to show how changes in the average dimensions of elastic bodies, either isotropic or anisotropic, caused by the introduction of dislocations, can be calculated when the displacement gradients produced by the dislocations are sufficiently small.

It is first shown in Sec. 2 that the average value of each of the stress components, in a rectangular Cartesian coordinate system, is zero. This result is then used, within the framework of the second-order theory, to obtain an expression for the average value of each of the infinitesimal strain components as the average value of an expression of second degree in the displacement gradients, which are associated with the dislocations in the body according to the classical theory. From these results, the dimensional changes can in principle be calculated in a number of cases. In Appendices 1 and 2, it is shown how these formulas may be specialized when the material considered has some particular symmetry, by illustrations from the isotropic case and from the case of cubic symmetry of the hextetrahedral, gyroidal, and hexoctahedral classes.

In Sec. 6 we considered the particular problem of the change in volume, resulting from the introduction of dislocations, of a cubic crystal of one of these classes. The formula derived can be easily specialized to the case when the material is isotropic and the result obtained is in agreement with that which Zener[1] obtained by a very different procedure.

Finally, in Sec. 7 we have employed our result for the cubic case to make certain qualitative predictions regarding the effect of dislocations on the volume of crystals of silver, gold, and copper.

2. AVERAGE STRESS

Consider an elastic body which in the undeformed state has the form shown schematically in Fig. 1. Let R_0 denote the region of space occupied by this body and let S_0 denote the complete boundary of R_0. Let the body be deformed in such a way that certain portions of S_0 are brought into contact with each other and bonded together as shown in Fig. 2, the external forces required to bring the surfaces together then being removed. The body is then in a state of deformation without external forces acting on it.

Let R denote the region occupied by the deformed body. Let S denote the external surface of the deformed

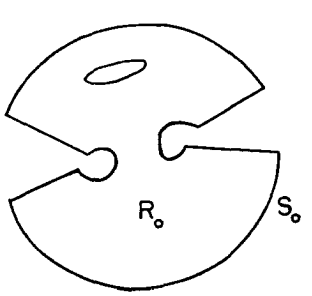

FIG. 1.

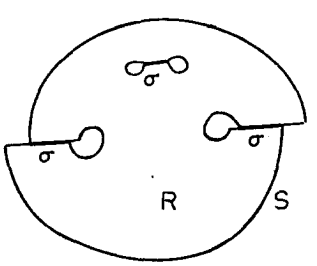

FIG. 2.

[1] C. Zener, Trans. Am. Inst. Mining Met. Engrs. **147**, 361 (1942).

body and let σ denote the bonded surfaces in the deformed body, which we may call the *dislocation surfaces*. Then, the displacement and stress fields have discontinuities on σ, while the stress vector acting on σ is continuous across σ.

Let t_{ij} denote the stress components in a rectangular Cartesian coordinate system x at a point x_i of the deformed body. Since no external forces act on the body, we have*

$$\partial t_{ij}/\partial x_j=0 \quad \text{throughout } R, \tag{2.1}$$

and

$$t_{ij}n_j=0 \quad \text{on } S, \tag{2.2}$$

where n_i denotes the unit normal to S at the point considered.

Let $t_{ij}^{(1)}$ and $t_{ij}^{(2)}$ be the limiting values of the stress t_{ij} at a point of the dislocation surface σ as we approach it from its two sides, which we may call sides 1 and 2. Then, since the stress vector is continuous on σ,

$$(t_{ij}^{(2)}-t_{ij}^{(1)})n_j=0 \quad \text{on } \sigma, \tag{2.3}$$

where n_i is the unit vector to σ drawn in the sense from 2 to 1 (say).

We note that, if dV is an element of volume of the deformed body R and dS and $d\sigma$ are elements of area of the surfaces S and σ, using the divergence theorem and the relations (2.2) and (2.3), we have

$$\int_R \frac{\partial}{\partial x_j}(x_k t_{ij})dV=\int_S x_k t_{ij}n_j dS + \int_\sigma x_k(t_{ij}^{(2)}-t_{ij}^{(1)})n_j d\sigma=0. \tag{2.4}$$

We thus have, with (2.1),

$$\int_R \left(t_{ik}+x_k\frac{\partial t_{ij}}{\partial x_j}\right)dV=\int_R t_{ik}dV=0. \tag{2.5}$$

Equation (2.5) expresses the following theorem: *the average value of each of the stress components in a rectangular Cartesian coordinate system in a body, which is held in equilibrium without external forces being applied, is zero.*

We note that in deriving this result no assumption is made regarding the magnitude of the deformation, nor does the elastic nature of the material of the body enter explicitly.

3. SOME BASIC RESULTS IN SECOND-ORDER ELASTICITY THEORY

We consider a deformation of an elastic material, in which a generic particle initially at X_i in the rectangular Cartesian coordinate system x moves to x_i in the same coordinate system. If u_i are the components of the displacement for the particle, then

$$x_i=X_i+u_i. \tag{3.1}$$

The stored elastic energy W per unit of deformed volume, or strain-energy function, then depends on the displacement gradients $u_{i,j}$, and we shall assume that this dependence is polynomial. We use the notation $_{,j}$ to denote partial differentiation with respect to the coordinate X_j. The components of stress t_{ij} are given by

$$t_{ij}=\frac{1}{|\partial x/\partial X|}x_{j,k}\frac{\partial W}{\partial u_{i,k}}. \tag{3.2}$$

If dV and dV_0 denote corresponding elements of volume in the deformed and undeformed states, we have

$$dV/dV_0=|\partial x/\partial X|. \tag{3.3}$$

From (3.2), (3.3), and (2.5), we see that if a body of elastic material is held in a deformed state, without external forces, by the introduction of dislocations, then†

$$\int |\partial x/\partial X|t_{ij}dV_0=\int x_{j,k}\frac{\partial W}{\partial u_{i,k}}dV_0=0. \tag{3.4}$$

W may be expressed in the form,

$$W=W_1+W_2+W_3+\cdots, \tag{3.5}$$

where $W_1, W_2, W_3, \cdots$ are homogeneous polynomials of the first, second, third, $\cdots$ degrees in the displacement gradients. It can be shown that if the stress components are zero when the displacement gradients are zero, $W_1=0$. Then, if the displacement gradients are sufficiently small, $W=W_2$ provides a first approximation to the strain-energy function (that of classical elasticity theory) and $W=W_2+W_3$ provides a second approximation.

If we take $W=W_2$, we are neglecting in W terms of higher degree than the second in the displacement gradients. From (3.2) it is seen that this implies the neglect, in the expressions for the stress components, of terms of higher degree than the first in the displacement gradients. With this approximation, we obtain from (3.4),

$$\int(\partial W_2/\partial u_{i,j})dV_0=0. \tag{3.6}$$

If we take $W=W_2+W_3$, we are neglecting in W terms of higher degree than the third in the displacement gradients. This implies the neglect, in the expressions for the stress components, of terms of higher degree than the second in the displacement gradients. With this

* Here and throughout this paper, lower case Latin indices take the values 1, 2, 3 and the summation convention is applicable to them.

† Here and subsequently integrals with respect to V_0 are considered to be evaluated over the domain R_0.

approximation, we obtain from (3.4),

$$\int\left[\frac{\partial W_2}{\partial u_{i,j}}+u_{j,k}\frac{\partial W_2}{\partial u_{i,k}}+\frac{\partial W_3}{\partial u_{i,j}}\right]dV_0=0. \tag{3.7}$$

4. THE GENERAL ANISOTROPIC CASE

The strain-energy function W cannot have arbitrary dependence on the displacement gradients, but must depend on them through the components E_{ij} of the finite strain tensor defined by

$$E_{ij}=\tfrac{1}{2}(x_{k,i}x_{k,j}-\delta_{ij})=e_{ij}+\alpha_{ij}, \tag{4.1}$$

where

$$e_{ij}=\tfrac{1}{2}(u_{i,j}+u_{j,i}) \quad \text{and} \quad \alpha_{ij}=\tfrac{1}{2}u_{k,i}u_{k,j}. \tag{4.2}$$

e_{ij} are the components of infinitesimal strain of the classical linear elasticity theory, and it is noted that e_{ij} and α_{ij} are homogeneous of first and second degrees, respectively, in the displacement gradients.

By taking W to be a polynomial in the components E_{ij}, we may write

$$W=a_{ijkl}E_{ij}E_{kl}+b_{ijklmn}E_{ij}E_{kl}E_{mn}+\cdots, \tag{4.3}$$

where a_{ijkl} and b_{ijklmn} are constants. From (3.5), (4.1), and (4.3), we obtain

$$W_2=a_{ijkl}e_{ij}e_{kl}, \tag{4.4}$$

and

$$W_3=b_{ijklmn}e_{ij}e_{kl}e_{mn}+a_{ijkl}(e_{ij}\alpha_{kl}+e_{kl}\alpha_{ij}). \tag{4.5}$$

It is easily seen that in (4.4) we may, without loss of generality, take a_{ijkl} to be unaltered by permutation of i and j, of k and l and of ij and kl. Also, in (4.5), we may take b_{ijklmn} to be unaltered by permutation of i and j, of k and l, of m and n, and of ij, kl and mn. We may then write (4.4) as

$$W_2=a_{ijkl}u_{i,j}u_{k,l}, \tag{4.6}$$

and (4.5) as

$$\begin{aligned}W_3&=b_{ijklmn}e_{ij}e_{kl}e_{mn}+2a_{ijkl}e_{ij}\alpha_{kl}\\&=b_{ijklmn}u_{i,j}u_{k,l}u_{m,n}+a_{ijkl}u_{i,j}u_{m,k}u_{m,l}.\end{aligned} \tag{4.7}$$

By introducing (4.4) into (3.6), we obtain

$$a_{ijkl}\int e_{kl}dV_0=0. \tag{4.8}$$

By bearing in mind the symmetry of e_{kl} and of a_{ijkl}, we see that (4.8) represents six simultaneous equations in the six independent quantities $\int e_{kl}dV_0$. From these we readily obtain

$$\int e_{kl}dV_0=0. \tag{4.9}$$

Equations (4.8) and (4.9) are, of course, valid with the approximation that terms of second and higher degrees in the displacement gradients are neglected; i.e., they are valid within the framework of classical elasticity theory.

We obtain a second approximation to $\int e_{kl}dV_0$ from (4.6), (4.7), and (3.7). This yields

$$\begin{aligned}2a_{ijmn}\int e_{mn}dV_0=&-2\int(a_{ikmn}u_{j,k}+a_{jkmn}u_{i,k})u_{m,n}dV_0\\&-a_{ijmn}\int u_{k,m}u_{k,n}dV_0\\&-3b_{ijklmn}\int u_{k,l}u_{m,n}dV_0.\end{aligned} \tag{4.10}$$

This provides six independent equations for the determination of the six quantities $\int e_{mn}dV_0$. These can be calculated in the following manner. Let us denote the pairs of subscripts 11, 22, 33, 23 or 32, 31 or 13, 12 or 21 by 1, 2, 3, 4, 5, 6, respectively. We define the elastic compliances $s_{\alpha\beta}$ ($\alpha, \beta=1, 2, \cdots, 6$) by‡

$$2s_{\alpha\beta}a_{\beta\gamma}=\delta_{\alpha\gamma}. \tag{4.11}$$

We therefore have

$$\int e_\alpha dV_0=2s_{\alpha\beta}a_{\beta\gamma}\int e_\gamma dV_0, \tag{4.12}$$

where $2a_{\beta\gamma}\int e_\gamma dV_0$ is given by (4.10). This expression can be used to calculate the average values of the changes in the dimensions of the material due to the introduction of dislocations. For example, suppose the body considered has the shape of a prism with its length parallel to the 1 axis. Then, if A_0 is the initial cross-sectional area of the prism, the average change in length is

$$\frac{1}{A_0}\int e_{11}dV_0=\frac{2}{A_0}s_{1\beta}a_{\beta\gamma}\int e_\gamma dV_0, \tag{4.13}$$

$2a_{\beta\gamma}\int e_\gamma dV_0$ being given by (4.10).

It follows immediately from the result given in Appendix 3 that the total change in volume $V-V_0$ undergone by the body as a result of the introduction of dislocations is given by

$$V-V_0=\int e_{kk}dV_0+\frac{1}{2}\int[(u_{k,k})^2-u_{p,q}u_{q,p}]dV_0, \tag{4.14}$$

with the neglect only of terms of higher degree than the second in the displacement gradients. Again, we may substitute from (4.12) and (4.10) for $\int e_{kk}dV_0$ to obtain an expression for the change in volume which is of the second degree in the displacement gradients.

Each of the expressions obtained from (4.12), (4.13), and (4.14) by substituting for $a_{\beta\gamma}\int e_\gamma dV_0$ from (4.10) is of the second degree in the displacement gradients resulting from the introduction of the dislocations according to the second-order elasticity theory for the

‡ Repetition of a Greek subscript indicates summation over the values 1, 2, $\cdots$, 6 for the subscript. This convention will be employed only in the present section.

material. However, to the order of approximation involved in the calculations, we may replace them by those calculated from the first-order theory for the material, and it is in this sense that we shall understand them.

5. EFFECT OF STRAIN ON THE ELASTIC MODULI

In the expression (3.2) for the stress components we take

$$W=W_2+W_3,$$

where W_2 and W_3 are homogeneous polynomials of the second and third degrees respectively in the displacement gradients. Then, neglecting terms of higher degree than the second in the displacement gradients, we obtain

$$t_{ij}=(1-u_{r,r})\frac{\partial W_2}{\partial u_{i,j}}+u_{j,k}\frac{\partial W_2}{\partial u_{i,k}}+\frac{\partial W_3}{\partial u_{i,j}}. \tag{5.1}$$

Now, let us suppose that the displacement field is increased by an infinitesimal displacement $\bar{u}_i$, and let us calculate the change $\dot{t}_{ij}$ in the stress associated with the new displacement field $u_i+\bar{u}_i$ on the assumption that terms of higher degree than the first in the displacement gradients $\bar{u}_{i,j}$ may be neglected. Then,

$$\dot{t}_{ij}=\frac{\partial t_{ij}}{\partial u_{p,q}}\bar{u}_{p,q}. \tag{5.2}$$

By introducing (5.1) into (5.2), we obtain

$$\begin{aligned}\dot{t}_{ij}=&-\frac{\partial W_2}{\partial u_{i,j}}\bar{u}_{p,p}+\frac{\partial W_2}{\partial u_{i,k}}\bar{u}_{j,k}\\&+\left\{(1-u_{r,r})\frac{\partial^2 W_2}{\partial u_{i,j}\partial u_{p,q}}+u_{j,k}\frac{\partial^2 W_2}{\partial u_{i,k}\partial u_{p,q}}\right.\\&\left.+\frac{\partial^2 W_3}{\partial u_{i,j}\partial u_{p,q}}\right\}\bar{u}_{p,q}.\end{aligned} \tag{5.3}$$

By introducing the expressions (4.6) and (4.7) for W_2 and W_3 into (5.3), we obtain

$$\begin{aligned}\dot{t}_{ij}=&-2a_{ijkl}u_{k,l}\bar{u}_{p,p}+2u_{m,n}(a_{ikmn}\bar{u}_{j,k}+a_{jkmn}\bar{u}_{i,k})\\&+2[3b_{pqijmn}u_{m,n}+(a_{pqil}u_{j,l}+a_{pqjl}u_{i,l})\\&+a_{ijql}u_{p,l}+(1-u_{r,r})a_{ijpq}]\bar{u}_{p,q}.\end{aligned} \tag{5.4}$$

This equation may be rewritten as

$$\dot{t}_{ij}=2(a_{ijrs}+c_{ijrs}{}^*)\bar{u}_{r,s}, \tag{5.5}$$

where

$$\begin{aligned}c_{ijrs}{}^*=&-a_{ijkl}u_{k,l}\delta_{rs}+u_{m,n}(a_{ismn}\delta_{jr}+a_{jsmn}\delta_{ir})\\&+(3b_{ijmnpq}u_{m,n}+a_{pqil}u_{j,l}+a_{pqjl}u_{i,l}\\&+a_{ijql}u_{p,l}-u_{r,r}a_{ijpq})\delta_{rp}\delta_{sq}.\end{aligned} \tag{5.6}$$

The relation (5.5) may be rewritten in the form

$$\dot{t}_{ij}=2(a_{ijrs}+c_{ijrs})\partial\bar{u}_r/\partial x_s, \tag{5.7}$$

where

$$c_{ijrs}=c_{ijrs}{}^*+a_{ijrm}u_{s,m}. \tag{5.8}$$

Then, $2(a_{ijrs}+c_{ijrs})$ are the elastic moduli for infinitesimal deformations of the material which is initially subjected to the displacements u_i. From (5.8) and (5.6), employing the notation

$$3\bar{b}_{pqijrs}=\partial c_{ijrs}/\partial u_{p,q}, \tag{5.9}$$

we obtain immediately

$$\begin{aligned}3\bar{b}_{pqijrs}=&-a_{ijpq}\delta_{rs}-a_{ijrs}\delta_{pq}+a_{ispq}\delta_{jr}\\&+a_{jspq}\delta_{ir}+a_{rsiq}\delta_{jp}+a_{rsjq}\delta_{ip}+a_{ijsq}\delta_{pr}\\&+a_{ijrq}\delta_{ps}+3b_{ijpqrs}.\end{aligned} \tag{5.10}$$

By substituting in (4.10) for $3b_{ijklmn}$ from (5.10), we obtain

$$\begin{aligned}2a_{ijmn}\int e_{mn}dV_0=&-2a_{ijkl}\int u_{m,m}u_{k,l}dV_0\\&+a_{ijmk}\int u_{n,k}u_{m,n}dV_0\\&-3\bar{b}_{pqijrs}\int u_{p,q}u_{r,s}dV_0.\end{aligned} \tag{5.11}$$

If the material has some symmetry, then we must express this fact by determining appropriate forms for a_{ijkl} and b_{ijpqrs}. The manner in which this may be done conveniently is illustrated in Appendices 1 and 2 for the cases when the material is isotropic and when it has cubic symmetry of the hextetrahedral, gyroidal, or hexoctahedral classes. In these appendices we also give the special forms taken, in these cases, by the expression (4.10) and the expression (5.4) for the stress corresponding to an infinitesimal strain superposed on an initial deformation.

Meanwhile, in the next section we shall determine the change of volume, caused by the introduction of dislocations, of cubic crystals of the hextetrahedral, gyroidal, or hexoctahedral classes. The result obtained could have been derived from the more general formalism given in Appendix 2. However, the method employed in Sec. 6 takes advantage of certain algebraic simplifications which are possible for this particular problem.

6. CHANGE OF VOLUME FOR CUBIC CRYSTALS (HEXTETRAHEDRAL, GYROIDAL, AND HEXOCTAHEDRAL CLASSES)

In this section we shall determine the change of volume, caused by the introduction of dislocations, in cubic crystals of the hextetrahedral, gyroidal, and hexoctahedral classes. Before doing so, however, we

shall derive certain formulas, which are generally valid, from the results of Secs. 4 and 5.

From (4.10), we obtain, with (4.6),

$$2a_{iimn}\int e_{mn}dV_0=-4\int W_2dV_0-a_{iimn}\int u_{k,m}u_{k,n}dV_0 -3b_{iiklmn}\int u_{k,l}u_{m,n}dV_0. \quad (6.1)$$

By bearing in mind that

$$3\bar{b}_{ijklmn}=\partial c_{klmn}/\partial u_{i,j}, \quad (6.2)$$

we obtain, from (5.10) and (4.6),

$$3b_{iiklmn}u_{k,l}u_{m,n}=3\bar{b}_{iiklmn}u_{k,l}u_{m,n}-W_2+a_{iikl}u_{k,l}u_{m,m} -a_{iikn}u_{k,m}u_{m,n}-a_{iiln}u_{k,l}u_{k,n}. \quad (6.3)$$

By introducing (6.3) into (4.10), we obtain

$$2a_{iimn}\int e_{mn}dV_0=-3\int W_2dV_0 +a_{iikl}\int(u_{k,m}u_{m,l}-u_{k,l}u_{m,m})dV_0 -3\bar{b}_{iiklmn}\int u_{k,l}u_{m,n}dV_0. \quad (6.4)$$

We may obtain a convenient expression for $\bar{b}_{iiklmn}$ as the rate of change of c_{klmn} with volume, when the material is subjected to a uniform dilatation, in the following manner: We suppose that the fractional extensions undergone by the material in the uniform dilatation are β. Then taking $u_{i,j}=\beta\delta_{ij}$ in (5.6) and (5.8), we obtain

$$c_{ijrs}=\beta(-a_{ijpp}\delta_{rs}+a_{ispp}\delta_{jr}+a_{jspp}\delta_{ir} +a_{ijrs}+3b_{ppijrs}). \quad (6.5)$$

By comparing (6.5) and (5.10), we obtain

$$3\bar{b}_{ppijrs}=\partial c_{ijrs}/\partial\beta=3(\partial c_{ijrs}/\partial v), \quad (6.6)$$

where v is the fractional increase of volume of the material in the uniform dilatation.

So far we have made no assumption regarding the symmetry of the material. For cubic crystals of the hextetrahedral, gyroidal, and hexoctahedral classes, it is shown in Appendix 2 that the elastic moduli $2a_{pqrs}$ must be expressible in the form,§

$$2a_{pqrs}=(a_1+2a_2)\delta_{pq}\delta_{rs}-\tfrac{1}{2}a_1(\delta_{sp}\delta_{rq}+\delta_{sq}\delta_{rp}) +2b_1\sum_{\alpha=1}^{3}\delta_{\alpha p}\delta_{\alpha q}\delta_{\alpha r}\delta_{\alpha s}. \quad (6.7)$$

§ The convention will be used that Greek subscripts take the values 1, 2, 3. The summation convention will not be applied to them.

By introducing this result into (6.4), we obtain

$$2(a_1+3a_2+b_1)\int e_{mm}dV_0 =-3\int W_2dV_0 -(a_1+3a_2+b_1)\int[(u_{m,m})^2-u_{m,n}u_{n,m}]dV_0 -3\bar{b}_{iiklmn}\int u_{k,l}u_{m,n}dV_0. \quad (6.8)$$

By using the result obtained in Appendix 3, we see that neglecting terms of higher degree than the second in the displacement gradients, the change in volume $V-V_0$ of the crystal due to the introduction of dislocations is given by

$$2(a_1+3a_2+b_1)(V-V_0) =-3\int W_2dV_0-3\bar{b}_{iiklmn}\int u_{k,l}u_{m,n}dV_0. \quad (6.9)$$

We note that if the cubic crystal undergoes a uniform dilatation, it remains cubic and consequently the elastic moduli $2(a_{pqrs}+c_{pqrs})$ for infinitesimal deformations of this deformed crystal are given by

$$2(a_{pqrs}+c_{pqrs})=(\bar{a}_1+2\bar{a}_2)\delta_{pq}\delta_{rs}-\tfrac{1}{2}\bar{a}_1(\delta_{pr}\delta_{qs}+\delta_{ps}\delta_{qr}) +2\bar{b}_1\sum_{\alpha=1}^{3}\delta_{\alpha p}\delta_{\alpha q}\delta_{\alpha r}\delta_{\alpha s}, \quad (6.10)$$

where $\bar{a}_1$, $\bar{a}_2$, and $\bar{b}_1$ depend on the fractional increase in volume v, and are equal to a_1, a_2, and b_1, when $v=0$. From (6.10) and (6.6), we obtain

$$2\bar{b}_{ppijrs}\int u_{i,j}u_{r,s}dV_0 =\left(\frac{\partial\bar{a}_1}{\partial v}+2\frac{\partial\bar{a}_2}{\partial v}\right)\int(u_{k,k})^2dV_0 -\frac{1}{2}\frac{\partial\bar{a}_1}{\partial v}\int(u_{k,m}u_{k,m}+u_{k,m}u_{m,k})dV_0 +2\frac{\partial\bar{b}_1}{\partial v}\sum_{\alpha=1}^{3}\int(u_{\alpha,\alpha})^2dV_0. \quad (6.11)$$

On defining $\bar{k}$, $\bar{\mu}$, and $\bar{\nu}$ by

$$\bar{k}=\tfrac{2}{3}(\bar{a}_1+3\bar{a}_2+\bar{b}_1),\quad 2\bar{\mu}=2\bar{b}_1-\bar{a}_1 \text{ and } \bar{\nu}=-\bar{b}_1, \quad (6.12)$$

and denoting by k, μ and ν the values of $\bar{k}$, $\bar{\mu}$ and $\bar{\nu}$ when

$v=0$, we can rewrite Eq. (6.11) as

$$\bar{b}_{ppijrs}\int u_{i,j}u_{r,s}dV_0 = \frac{\partial\bar{k}}{\partial v}\frac{1}{k}\int W_D dV_0+\frac{\partial\bar{\mu}}{\partial v}\frac{1}{\mu}\int W_S dV_0 + \frac{\partial\bar{\nu}}{\partial v}\frac{1}{\nu}\int W_S' dV_0, \quad (6.13)$$

where

$$W_D=\tfrac{1}{2}k(u_{m,m})^2,$$
$$W_S=\mu[\tfrac{1}{2}(u_{m,n}u_{m,n}+u_{m,n}u_{n,m})-\tfrac{1}{3}(u_{m,m})^2],$$

and

$$W_S'=\nu[\tfrac{1}{2}(u_{m,n}u_{m,n}+u_{m,n}u_{n,m})-\sum_{\alpha=1}^{3}(u_{\alpha,\alpha})^2]. \quad (6.14)$$

Also, we see from (6.7) and (4.6) that, for the cubic crystals considered,

$$W_2=a_{pqrs}u_{p,q}u_{r,s}=\tfrac{1}{2}(a_1+2a_2)(u_{m,m})^2 -\tfrac{1}{4}a_1(u_{m,n}u_{m,n}+u_{m,n}u_{n,m})+b_1\sum_{\alpha=1}^{3}(u_{\alpha,\alpha})^2. \quad (6.15)$$

It is then easily seen that

$$W_2=W_D+W_S+W_S'. \quad (6.16)$$

By introducing (6.13) and (6.16) into (6.9), we obtain

$$V-V_0=-\frac{1}{k}\left[\left(1+\frac{1}{k}\frac{\partial\bar{k}}{\partial v}\right)\int W_D dV_0 +\left(1+\frac{1}{\mu}\frac{\partial\bar{\mu}}{\partial v}\right)\int W_S dV_0 +\left(1+\frac{1}{\nu}\frac{\partial\bar{\nu}}{\partial v}\right)\int W_S' dV_0\right]. \quad (6.17)$$

We can specialize this result to the case of an isotropic material by taking $b_1=\bar{b}_1=0$. We then obtain Zener's result[1]

$$V-V_0=-\frac{1}{k}\left[\left(1+\frac{1}{k}\frac{\partial\bar{k}}{\partial v}\right)\int W_D dV_0 +\left(1+\frac{1}{\mu}\frac{\partial\bar{\mu}}{\partial v}\right)\int W_S dV_0\right]. \quad (6.18)$$

7. A QUALITATIVE DEDUCTION

In general, in order to obtain an explicit value for $V-V_0$ from Eqs. (6.17) or (6.18), we need to know the displacement field which is associated with the dislocations according to classical elasticity theory. However, it is easily seen that W_S, W_D, and W_S' are essentially positive. Thus, if for a given material the coefficients of all three of the terms $\int W_S dV_0$, $\int W_D dV_0$, and $\int W_S' dV_0$ in (6.17) have the same signs, we can predict whether a body of the material will increase or decrease in volume when dislocations are introduced.

Daniels and Smith[2] have determined the dependence on applied hydrostatic pressure of the speeds of propagation of plane waves in crystals of copper, silver, and gold. These crystals are all of the type considered in Sec. 6. The wave speeds measured in a crystal are, of course, simply related to the values of $\bar{k}$, $\bar{\mu}$, and $\bar{\nu}$ for the crystal.

If p denotes the applied hydrostatic pressure, we have

$$\frac{\partial}{\partial v}=\frac{\partial p}{\partial v}\frac{\partial}{\partial p}=-k\frac{\partial}{\partial p}. \quad (7.1)$$

We can therefore write (6.17) in the alternative form,

$$V-V_0=\frac{1}{k}\left[\left(\frac{\partial\bar{k}}{\partial p}-1\right)\int W_D dV_0 +\left(\frac{k}{\mu}\frac{\partial\bar{\mu}}{\partial p}-1\right)\int W_S dV_0 +\left(\frac{k}{\nu}\frac{\partial\bar{\nu}}{\partial p}-1\right)\int W_S' dV_0\right]. \quad (7.2)$$

From the results of Daniels and Smith, the values given in Table I are obtained for the coefficients $\partial\bar{k}/\partial p-1$, $(k/\mu)\partial\bar{\mu}/\partial p-1$ and $(k/\nu)\partial\bar{\nu}/\partial p-1$.

Since these coefficients are all positive, it follows that for each of the metals, the introduction of dislocations should produce an increase in volume.

We easily can establish upper and lower bounds for the change of volume in terms of the total elastic energy $\int W_2 dV_0$ stored in the deformed body in these cases. For example, for copper and silver, we see from the

TABLE I.

	$\frac{\partial\bar{k}}{\partial p}-1$	$\frac{k}{\mu}\frac{\partial\bar{\mu}}{\partial p}-1$	$\frac{k}{\nu}\frac{\partial\bar{\nu}}{\partial p}-1$
Copper	4.59	2.404	3.685
Silver	5.18	3.332	4.611
Gold	5.43	4.139	7.545

[2] W. B. Daniels and C. S. Smith, ONR. Tech. Rept. No. 1, Contract Nonr-1141(05), Project NR017-309 (1958).

table that

$$\frac{\partial \bar{k}}{\partial p}-1>\frac{k}{\nu}\frac{\partial \bar{\nu}}{\partial p}-1>\frac{k}{\mu}\frac{\partial \bar{\mu}}{\partial p}-1. \tag{7.3}$$

With (7.2) and (6.16), we obtain immediately

$$\frac{1}{k}\left(\frac{\partial \bar{k}}{\partial p}-1\right)\int W_2 dV_0 \geq V-V_0 \geq \frac{1}{k}\left(\frac{k}{\mu}\frac{\partial \bar{\mu}}{\partial p}-1\right)\int W_2 dV_0. \tag{7.4}$$

8. APPENDIX 1. ISOTROPIC MATERIALS

For an isotropic material, W must be expressible as a polynomial in J_1, J_2 and J_3 defined by

$$J_1=E_{ii},\quad J_2=\tfrac{1}{2}[(E_{ii})^2-E_{ij}E_{ji}],$$
$$\text{and} \tag{8.1}$$
$$J_3=\tfrac{1}{6}[2E_{ik}E_{kj}E_{ji}-3E_{kk}E_{ij}E_{ji}+(E_{ii})^3].$$

Thus retaining terms up to the third degree in E_{ij}, the expression (4.3) for W takes, for an isotropic material, the form,‖

$$W=a_1J_2+a_2J_1^2+a_3J_1J_2+a_4J_1^3+a_5J_3. \tag{8.2}$$

From (8.2) and (4.3), we obtain

$$2a_{pqrs}=\frac{1}{4}\left(\frac{\partial}{\partial E_{pq}}+\frac{\partial}{\partial E_{qp}}\right)\left(\frac{\partial}{\partial E_{rs}}+\frac{\partial}{\partial E_{rs}}\right)W\Big|_{E_{kl}=0}$$
$$=(a_1+2a_2)\delta_{pq}\delta_{rs}-\tfrac{1}{2}a_1(\delta_{pr}\delta_{qs}+\delta_{ps}\delta_{qr}), \tag{8.3}$$

and

$$6b_{mnpqrs}=\frac{1}{8}\left(\frac{\partial}{\partial E_{mn}}+\frac{\partial}{\partial E_{nm}}\right)\left(\frac{\partial}{\partial E_{pq}}+\frac{\partial}{\partial E_{qp}}\right)\times\left(\frac{\partial}{\partial E_{rs}}+\frac{\partial}{\partial E_{sr}}\right)W\Big|_{E_{kl}=0}$$
$$=(3a_3+6a_4+a_5)\delta_{mn}\delta_{pq}\delta_{rs}$$
$$-\tfrac{1}{2}(a_3+a_5)[\delta_{mn}(\delta_{sp}\delta_{rq}+\delta_{sq}\delta_{rp})$$
$$+\delta_{pq}(\delta_{ms}\delta_{nr}+\delta_{mr}\delta_{ns})$$
$$+\delta_{rs}(\delta_{mq}\delta_{np}+\delta_{nq}\delta_{mp})]$$
$$+\tfrac{1}{4}a_5[\delta_{rq}(\delta_{sm}\delta_{pn}+\delta_{sn}\delta_{pm})$$
$$+\delta_{sp}(\delta_{qm}\delta_{rn}+\delta_{qn}\delta_{rm})$$
$$+\delta_{rp}(\delta_{sm}\delta_{qn}+\delta_{sn}\delta_{qm})$$
$$+\delta_{sq}(\delta_{pm}\delta_{rn}+\delta_{pn}\delta_{rm})].$$

‖ This is substantially the result given by F. D. Murnaghan, Am. J. Math. **59**, 235 (1937).

By introducing (8.3) into (4.10), we obtain

$$(a_1+2a_2)\delta_{ij}\int e_{mm}dV_0-a_1\int e_{ij}dV_0$$
$$=-\tfrac{1}{2}(2a_1+4a_2-a_3-a_5)\int u_{m,m}(u_{i,j}+u_{j,i})dV_0$$
$$+\tfrac{1}{4}(a_3+a_5)\int u_{m,n}u_{n,m}\delta_{ij}dV_0$$
$$-\tfrac{1}{2}(3a_3+6a_4+a_5)\int (u_{m,m})^2\delta_{ij}dV_0$$
$$+\tfrac{1}{4}(2a_1-a_5)\int (u_{i,m}u_{m,j}+u_{j,m}u_{m,i}+u_{m,i}u_{m,j})dV_0$$
$$+\tfrac{1}{4}(4a_1-a_5)\int u_{i,m}u_{j,m}dV_0$$
$$-\tfrac{1}{4}(2a_1+4a_2-a_3-a_5)\int u_{m,n}u_{m,n}\delta_{ij}dV_0. \tag{8.4}$$

Introducing (8.3) into (5.4), we obtain

$$\begin{aligned}\dot{t}_{ij}=&(a_1+2a_2)\bar{u}_{p,\,p}\delta_{ij}-\tfrac{1}{2}a_1(\bar{u}_{i,\,j}+\bar{u}_{j,\,i})\\&-(2a_1+4a_2-3a_3-6a_4-a_5)\delta_{ij}u_{k,\,k}\bar{u}_{p,\,p}\\&+(\tfrac{3}{2}a_1+2a_2-\tfrac{1}{2}a_3-\tfrac{1}{2}a_5)[(u_{i,\,j}+u_{j,\,i})\bar{u}_{p,\,p}\\&+u_{p,\,p}(\bar{u}_{i,\,j}+\bar{u}_{j,\,i})]\\&+(a_1+2a_2-\tfrac{1}{2}a_3-\tfrac{1}{2}a_5)\delta_{ij}u_{m,\,n}\bar{u}_{m,\,n}\\&-\tfrac{1}{2}(a_3+a_5)\delta_{ij}u_{m,\,n}\bar{u}_{n,\,m}\\&-\tfrac{1}{2}(a_1-\tfrac{1}{2}a_5)[(u_{j,\,k}+u_{k,\,j})(\bar{u}_{i,\,k}+\bar{u}_{k,\,i})\\&+(u_{i,\,k}+u_{k,\,i})(\bar{u}_{j,\,k}+\bar{u}_{k,\,j})]\\&-\tfrac{1}{2}a_1(u_{j,\,k}\bar{u}_{i,\,k}+u_{i,\,k}\bar{u}_{j,\,k}).\end{aligned} \tag{8.5}$$

9. APPENDIX 2. CUBIC CRYSTALS (HEXTETRAHEDRAL, GYROIDAL, AND HEXOCTAHEDRAL CLASSES)

For cubic crystals of the hextetrahedral, gyroidal, and hexoctahedral classes, the axes of which are in the directions of the axes of the coordinate system x, it has been shown[3] that the strain-energy function may be expressed as a polynomial in J_1, J_2, and J_3, defined by (8.1), and I_1, I_2, and I_3, defined by

$$I_1=E_{11}^2+E_{22}^2+E_{33}^2,$$
$$I_2=E_{11}E_{22}E_{33}, \tag{9.1}$$
$$\text{and}$$
$$I_3=E_{11}E_{23}^2+E_{22}E_{31}^2+E_{33}E_{12}^2,$$

together with certain further invariants of higher degree than the third in E_{ij}. Retaining terms up to the third degree in E_{ij}, the strain-energy function W then takes the form

$$W=W'+b_1I_1+b_2J_1I_1+b_3I_2+b_4I_3, \tag{9.2}$$

where W' denotes the expression for W given in (8.2).

[3] G. F. Smith and R. S. Rivlin, Trans. Am. Math. Soc. **88**, 175 (1958).

We obtain immediately

$$\frac{1}{4}\left(\frac{\partial}{\partial E_{pq}}+\frac{\partial}{\partial E_{qp}}\right)\left(\frac{\partial}{\partial E_{rs}}+\frac{\partial}{\partial E_{sr}}\right)\Big|_{E_{kl}=0} W=2a_{pqrs}$$

$$=2a_{pqrs}'+2b_1\sum_{\alpha=1}^{3}\delta_{\alpha p}\delta_{\alpha q}\delta_{\alpha r}\delta_{\alpha s},$$

and

$$\frac{1}{8}\left(\frac{\partial}{\partial E_{mn}}+\frac{\partial}{\partial E_{nm}}\right)\left(\frac{\partial}{\partial E_{pq}}+\frac{\partial}{\partial E_{qp}}\right)$$

$$\times\left(\frac{\partial}{\partial E_{rs}}+\frac{\partial}{\partial E_{sr}}\right)W\Big|_{E_{kl}=0}=6b_{mnpqrs}$$

$$=6b_{mnpqrs}'+2b_2\sum_{\alpha=1}^{3}(\delta_{mn}\delta_{\alpha p}\delta_{\alpha q}\delta_{\alpha r}\delta_{\alpha s} \tag{9.3}$$

$$+\delta_{pq}\delta_{\alpha r}\delta_{\alpha s}\delta_{\alpha m}\delta_{\alpha n}+\delta_{rs}\delta_{\alpha m}\delta_{\alpha n}\delta_{\alpha p}\delta_{\alpha q})$$

$$+b_3\sum_{\alpha,\beta,\gamma=1}^{3}\pi_{\alpha\beta\gamma}\delta_{\alpha m}\delta_{\alpha n}\delta_{\beta p}\delta_{\beta q}\delta_{\gamma r}\delta_{\gamma s}$$

$$+\tfrac{1}{4}b_4\sum_{\alpha,\beta,\gamma=1}^{3}\pi_{\alpha\beta\gamma}[\delta_{\alpha m}\delta_{\alpha n}(\delta_{\beta p}\delta_{\gamma q}+\delta_{\gamma p}\delta_{\beta q})$$

$$\times(\delta_{\beta r}\delta_{\gamma s}+\delta_{\gamma r}\delta_{\beta s})$$

$$+\delta_{\alpha p}\delta_{\alpha q}(\delta_{\beta r}\delta_{\gamma s}+\delta_{\gamma r}\delta_{\beta s})(\delta_{\beta m}\delta_{\gamma n}+\delta_{\gamma m}\delta_{\beta n})$$

$$+\delta_{\alpha r}\delta_{\alpha s}(\delta_{\beta m}\delta_{\gamma n}+\delta_{\gamma m}\delta_{\beta n})(\delta_{\beta p}\delta_{\gamma q}+\delta_{\gamma p}\delta_{\beta q})],$$

where $\pi_{\alpha\beta\gamma}$ is the permutation symbol defined by $\pi_{\alpha\beta\gamma}=1$ if $\alpha\beta\gamma$ is a permutation of 1, 2, 3 and $\pi_{\alpha\beta\gamma}=0$ otherwise¶; a_{pqrs}' and b_{mnpqrs}' are used to denote the values of a_{pqrs} and b_{mnpqrs} given by (8.3).

On introducing (9.3) into (4.10) and using e_{ij}' to denote e_{ij} in Eq. (8.4), we obtain

$$(a_1+2a_2)\delta_{ij}\int e_{mm}dV_0-a_1\int e_{ij}dV_0$$

$$+2b_1\sum_{\alpha=1}^{3}\delta_{\alpha i}\delta_{\alpha j}\int e_{\alpha\alpha}dV_0$$

$$=(a_1+2a_2)\delta_{ij}\int e_{mm}'dV_0-a_1\int e_{ij}'dV_0$$

$$-\sum_{\alpha=1}^{3}\int\{2b_1(\delta_{\alpha i}u_{j,\alpha}u_{\alpha,\alpha}+\delta_{\alpha j}u_{i,\alpha}u_{\alpha,\alpha})$$

$$+b_1\delta_{\alpha i}\delta_{\alpha j}u_{k,\alpha}u_{k,\alpha}+b_2[\delta_{ij}(u_{\alpha,\alpha})^2$$

$$+2\delta_{\alpha i}\delta_{\alpha j}u_{\alpha,\alpha}u_{p,p}]+\tfrac{1}{2}b_3\pi_{\alpha\beta\gamma}\delta_{\alpha i}\delta_{\alpha j}u_{\beta,\beta}u_{\gamma,\gamma}$$

$$+\tfrac{1}{8}b_4\pi_{\alpha\beta\gamma}[\delta_{\alpha i}\delta_{\alpha j}(u_{\beta,\gamma}+u_{\gamma,\beta})^2$$

$$+2u_{\alpha,\alpha}(\delta_{\beta i}\delta_{\gamma j}+\delta_{\gamma i}\delta_{\beta j})(u_{\beta,\gamma}+u_{\gamma,\beta})]\}dV_0. \tag{9.4}$$

¶ We note that Greek subscripts are assumed to take the values 1, 2, 3, and the summation convention is not applied to them.

Introducing (9.3) into (5.4) and using t_{ij}' to denote the expression for t_{ij} given in (8.5), we obtain

$$t_{ij}=t_{ij}'+2b_1\sum_{\alpha=1}^{3}\delta_{\alpha i}\delta_{\alpha j}\bar{u}_{\alpha,\alpha}$$

$$+\sum_{\alpha=1}^{3}[-2(b_1-b_2)\delta_{\alpha i}\delta_{\alpha j}(u_{\alpha,\alpha}\bar{u}_{p,p}+u_{p,p}\bar{u}_{\alpha,\alpha})$$

$$+2b_1(\delta_{\alpha i}u_{j,\alpha}+\delta_{\alpha j}u_{i,\alpha})\bar{u}_{\alpha,\alpha}$$

$$+2b_1u_{\alpha,\alpha}(\delta_{\alpha i}\bar{u}_{j,\alpha}+\delta_{\alpha j}\bar{u}_{i,\alpha})+2b_1\delta_{\alpha i}\delta_{\alpha j}u_{p,\alpha}\bar{u}_{p,\alpha}$$

$$+2b_2\delta_{ij}u_{\alpha,\alpha}\bar{u}_{\alpha,\alpha}]+\sum_{\alpha,\beta,\gamma=1}^{3}\{b_3\pi_{\alpha\beta\gamma}\delta_{\alpha i}\delta_{\alpha j}u_{\beta,\beta}\bar{u}_{\gamma,\gamma}$$

$$+\tfrac{1}{4}b_4\pi_{\alpha\beta\gamma}\delta_{\alpha i}\delta_{\alpha j}(u_{\beta,\gamma}+u_{\gamma,\beta})(\bar{u}_{\beta,\gamma}+\bar{u}_{\gamma,\beta})$$

$$+\tfrac{1}{4}b_4\pi_{\alpha\beta\gamma}(\delta_{\alpha i}\delta_{\beta j}+\delta_{\beta i}\delta_{\alpha j})[(u_{\alpha,\beta}+u_{\beta,\alpha})\bar{u}_{\gamma,\gamma}$$

$$+u_{\gamma,\gamma}(\bar{u}_{\alpha,\beta}+\bar{u}_{\beta,\alpha})]\}. \tag{9.5}$$

10. APPENDIX 3

If $\mathbf{A}$ is any 3×3 matrix, then

$$\det\mathbf{A}=\tfrac{1}{6}[(tr\mathbf{A})^3-3tr\mathbf{A}tr\mathbf{A}^2+2tr\mathbf{A}^3]. \tag{10.1}$$

If in a deformation of a body, a point initially at X_i in a rectangular Cartesian coordinate system x moves to x_i in the same system, the ratio between the volume dV of an element in the deformed state to its volume dV_0 in the undeformed state is given by

$$dV/dV_0=|x_{i,j}|. \tag{10.2}$$

With (10.1), we obtain

$$\frac{dV}{dV_0}=\tfrac{1}{6}[(x_{r,r})^3-3x_{r,r}x_{p,q}x_{q,p}+2x_{p,q}x_{q,r}x_{r,p}]. \tag{10.3}$$

Writing $x_i=X_i+u_i$ and assuming that the displacement gradients $u_{i,j}$ are sufficiently small so that we may neglect terms of higher degree than the second in them, we obtain, from (10.3),

$$\frac{dV}{dV_0}=1+u_{r,r}+\tfrac{1}{2}(u_{r,r})^2-\tfrac{1}{2}u_{p,q}u_{q,p}. \tag{10.4}$$

ACKNOWLEDGMENT

The work described in this paper was started in 1952–53 when both of the authors were members of the staff of the Naval Research Laboratory, Washington, D. C. As far as one of the authors (R.S.R.) is concerned, its completion was sponsored by the Office of Naval Research under Contract Nonr-562(10) with Brown University.

Asymptotic Solutions of Differential Equations with Transition Points or Singularities*

A. ERDÉLYI
Department of Mathematics, California Institute of Technology, Pasadena, California
(Received December 7, 1959)

Asymptotic solutions of $d^2y/dx^2+[\lambda^2p(x)+r(x,\lambda)]y=0$ are found when λ is a large parameter and r is "small" in comparison with λ^2p, except at a single point where either p has a simple zero, or p a pole of the first order and r a pole of the second order. The results are applied to Bessel functions, and to Hermite and Laguerre polynomials. The resulting asymptotic forms are valid uniformly in x.

1. INTRODUCTION

IN many problems of applied mathematics one needs approximations to solutions of the differential equation,

$$d^2y/(dx^2)+[\lambda^2p(x)+r(x,\lambda)]y=0, \qquad (1.1)$$

where λ is a large parameter, and $r(x,\lambda)$ is "small" in comparison with $\lambda^2p(x)$. In this paper, x will be a real variable ranging over a finite or infinite open interval (a,b).

The earliest investigation of such an approximation is due to Liouville (1837) who encountered (1.1) in his work on what is now called the Sturm-Liouville problem. In Liouville's case, $\lambda^2p(x)>0$ throughout (a,b), and the solutions are shown to be *oscillatory*. Liouville's method can also be applied when $\lambda^2p(x)<0$ throughout (a,b) and it turns out that in this case the solutions show a *monotonic* behavior in the sense that no nontrivial solution can vanish more than once. A brief description of Liouville's method will be found in [3, Chap. IV]; here it will be sufficient to note that that method is based on a comparison of (1.1) with a differential equation with constant coefficients, and that the technique of the comparison involves a change of variables followed by a conversion of the differential equation into an integral equation of Volterra type.

The behavior of the solutions of (1.1) is much more involved if $\lambda^2p(x)$ changes its sign in (a,b), say at a point c. This point separates an interval in which the solutions of (1.1) are monotonic from one in which they are oscillatory; and the transition from the one kind of behavior to the other takes place in the immediate vicinity of c. Such a change of sign commonly takes place if c is a zero of $p(x)$, in which case it is called a *transition point* or a *turning point* of the differential equation, or if c is a *singularity* of $p(x)$, and hence of the differential equation.

H. Jeffreys encountered a turning point problem in 1923 when he investigated oscillations of water in an elliptic lake, and he developed approximations for that case [10, Sec. 17.131]. A few years later, turning point problems arose in connection with the one-dimensional Schrödinger equation, and the *WKB* method was invented to solve them. Liouville's method was adapted to differential equations with transition points by R. E. Langer. Among the many investigations carried out in this field those of Langer, T. M. Cherry, and F. W. J. Olver are especially pertinent to our purpose. A brief review of the relevant literature up to 1951 is given in [1], and general descriptions of the comparison method, in [5] and [14].

More recently, singularities of (1.1) at which $\lambda^2p(x)$ changes its sign have also been studied, the principal contributions being due to E. D. Cashwell, N. D. Kazarinoff, and R. McKelvey, Olver, and C. A. Swanson.

In this paper we shall assume that λ is a large positive parameter, (a,b) a finite or infinite open interval, and $p(x)$ is real and changes its sign at c, $a<c<b$. For the sake of definiteness, we mostly take $p(x)<0$ for $a<x<c$, and $p(x)>0$ for $c<x<b$. At c itself, $p(x)$ has either a simple zero or a simple "pole." Comparison with a differential equation with constant coefficients does not lead to adequate approximations in either of these two cases, and it becomes necessary to use comparison equations which themselves have a transition point or a singularity, as the case may be.

The simplest comparison equation with a transition point is

$$d^2w/dt^2+\lambda^2tw=0, \qquad (1.2)$$

which is closely related to Airy's equation. Solutions of (1.2) are

$$Ai(-\lambda^{\frac{2}{3}}\omega^m t), \quad Bi(-\lambda^{\frac{2}{3}}\omega^m t), \qquad (1.3)$$

where Ai and Bi are the Airy functions [15], m is an integer, and $\omega=e^{2\pi i/3}$.

The simplest comparison equation with a singularity would seem to be the differential equation,

$$d^2w/dt^2+\lambda^2t^{-1}w=0,$$

which is closely related to Bessel functions of order unity. However, it turns out that the adoption of this comparison equation necessitates undue restrictions on $r(x,\lambda)$. For this reason, we shall use a somewhat more elaborate comparison equation,

$$\frac{d^2w}{dt^2}+\left(\frac{\lambda^2}{t}+\frac{1-\nu^2}{4t^2}\right)w=0, \qquad (1.4)$$

* This paper is based on several reports by the author and others. These reports were prepared under contract with the Office of Naval Research and are listed with all other references in the Bibliography at the end of this paper under 1, 3, 4, 5, 8, 9, 19, 20. References appear in the text in brackets.

in which ν is a nonnegative constant to be chosen later. (1.4) is also closely related to Bessel's differential equation, and each of the functions,

$$t^{\frac{1}{2}}J_{\pm\nu}(2\lambda t^{\frac{1}{2}}), \quad t^{\frac{1}{2}}Y_{\nu}(2\lambda t^{\frac{1}{2}}), \quad t^{\frac{1}{2}}H_{\nu}^{(1,2)}(2\lambda t^{\frac{1}{2}}) \qquad (1.5)$$

in the usual notation of Bessel functions, is a solution of (1.4).

In this paper λ, x and $p(x)$ are assumed to be real. It may be noted that results for complex values of λ are available in the reports on which this paper is based and results for complex x and $p(x)$ in some of the references given at the end of the paper.

TRANSITION POINTS

2. Outline of the Method and Results

We introduce a new variable x in (1.2) by the substitution,

$$t=\phi(x), \quad w=[\phi'(x)]^{\frac{1}{2}}Y,$$

thus changing (1.2) in

$$(d^2Y/dx^2)+(\lambda^2\phi\phi'^2+\tfrac{1}{2}\{\phi,x\})Y=0, \qquad (2.1)$$

where

$$\{\phi,x\}=(\phi'''/\phi')-\tfrac{3}{2}(\phi''/\phi')^2 \qquad (2.2)$$

is the *Schwarzian derivative* of ϕ with respect to x. The differential equations (1.1) and (2.1) will be nearly identical for large λ if we choose ϕ so that

$$\phi\phi'^2=p(x). \qquad (2.3)$$

With this choice of ϕ, (1.1) may be written in the form,

$$(d^2y/dx^2)+(\lambda^2\phi\phi'^2+\tfrac{1}{2}\{\phi,x\})y=F(x,\lambda)y, \qquad (2.4)$$

where

$$F(x,\lambda)=\tfrac{1}{2}\{\phi,x\}-r(x,\lambda). \qquad (2.5)$$

In this exposition we shall discuss (2.4) rather than (1.1) and will express conditions of validity with reference to ϕ and F rather than p and r. Once ϕ has been determined from (2.3), F is also known and it is usually easy to verify that the conditions of validity are satisfied.

If $F(x,\lambda)$ is small in comparison with $\lambda^2\phi\phi'^2+\frac{1}{2}\{\phi,x\}$, then (2.1) and (2.4) will be nearly identical, and we shall expect the known solutions of (2.1) to be good approximations to solutions of 2.4. In order to estimate the difference $y-Y$, we regard for the moment the right hand side of (2.4) as a given function of x, and apply the method of variation of parameters to (2.4), obtaining

$$y(x)=Y(x)+\int_{x'}^{x}K(x,t)F(t,\lambda)y(t)dt, \qquad (2.6)$$

where x' is some fixed point of the interval, $Y(x)$ is a solution of (2.1), and $K(x,t)$ is the Green's function of (2.1). For a fixed t in (a,b), $K(x,t)$, as a function of x, satisfies (2.1) together with the initial conditions $K(t,t)=0$, $(\partial K/\partial x)(t,t)=1$.

In order to find approximations for a solution of (2.4), which is characterized by initial conditions at an interior point of (a,b), or by its behavior as $x\to a$ or $x\to b$, we choose x' to be the point in question and $Y(x)$ to be that solution of (2.1) which is characterized by the same conditions as $y(x)$. We then establish the existence of a solution of the integral equation (2.6), prove that this solution is differentiable and satisfies (2.4), obtain a rough estimate of $y(x)$, and use this estimate in the integral in (2.6) to obtain an improved estimate of $y(x)-Y(x)$. In order to carry out this program, it will be necessary to introduce some notations.

Clearly, ϕ must be assumed to be three times differentiable on (a,b), and it will also be assumed that ϕ' is either positive throughout (a,b), or else negative throughout (a,b). We set

$$\omega=e^{2\pi i/3}$$

$$\Phi=\Phi(x)=\Phi(x,\lambda)=|\phi'(x)|^{\frac{1}{2}}(1+\lambda^{1/6}|\phi(x)|^{\frac{1}{2}}), \qquad (2.7)$$

$$\theta_m=\theta_m(x)=\theta_m(x,\lambda)=-\lambda^{\frac{1}{3}}\omega^m\phi(x), \quad m=0,\pm1 \qquad (2.8)$$

$$\theta_2=\theta_0=\theta$$

$$\sigma=\tfrac{2}{3}\theta^{\frac{3}{2}}, \quad \sigma_m=\tfrac{2}{3}\theta_m^{\frac{3}{2}}, \quad m=0,\pm1,2. \qquad (2.9)$$

It is seen from (1.3) that

$$\begin{aligned} Y_m(x)&=|\phi'|^{-\frac{1}{2}}Ai(\theta_m) \quad m=0,\pm1\\ Y_2(x)&=|\phi'|^{-\frac{1}{2}}Bi(\theta)\end{aligned} \qquad (2.10)$$

are solutions of (2.1). The approximations to y will be expressed in terms of these solutions. In the error estimates certain complications arise due to the zeros of the Airy functions. In order to cope with these, we set

$$\begin{aligned}\tilde{Y}_m(x)&=\begin{cases} Y_m(x) & \text{except when } \theta_m<0\\ |\phi'|^{-\frac{1}{2}}[|Ai(\theta_m)|^2+|Bi(\theta_m)|^2]^{\frac{1}{2}} & \text{if } \theta_m<0\end{cases}\\ \tilde{Z}_m(x)&=\begin{cases}\dfrac{d}{dx}[|\phi'|^{\frac{1}{2}}Y_m(x)] & \text{except when } \theta_m<0\\ \lambda^{\frac{1}{3}}\phi'[|Ai'(\theta_m)|^2+|Bi'(\theta_m)|^2]^{\frac{1}{2}} & \text{if } \theta_m<0.\end{cases}\end{aligned} \qquad (2.11)$$

These functions will be used in estimates only. $\tilde{Y}_m$ is not a solution of (2.1), and $\tilde{Z}_m(x)$ is not the derivative of $|\phi'|^{\frac{1}{2}}Y_m$, if $\theta_m<0$.

Lastly, we set

$$G(x,\lambda)=\lambda^{\frac{1}{3}}F(x,\lambda)[\Phi(x,\lambda)]^{-2}. \qquad (2.12)$$

x and t range over a,b; and $\lambda>0$. All bounds, including those implied by the O symbols, are independent of x, t, λ. C is a generic bound which may be different at each occurrence.

It is seen from (2.3) that ϕ and p vanish at the same point, and have the same sign where they do not vanish. Thus $\phi'(x)>0$ if $p(x)<0$ for $a<x<c$ and $p(x)>0$ for $c<x<b$, and $\phi'(x)<0$ if $p(x)>0$ to the left, and $p(x)<0$

to the right of c. For the former case, we shall establish the following result:

If (i) ϕ *is three times continuously differentiable, and* $\phi'(x)>0$, *for* $a<x<b$, *and* $\phi(c)=0$; (ii) *for each fixed positive* λ, $F(x,\lambda)$ *is a continuous function of* x *for* $a<x<b$; (iii) $|G(x,\lambda)|\leq\Lambda(\lambda)g(x)$ *where* $(1+\lambda)^{-1}\Lambda(\lambda)$ *is a bounded function of* λ *for* $\lambda>0$, *and* $g(x)$ *is integrable over* (a,b); *and if we set*

$$x_0=a, \quad x_{\pm1}=b, \quad x_2=c, \tag{2.13}$$

then the integral equation,

$$y_m(x)=Y_m(x)+\int_{x_m}^{x}K(x,t)F(t,\lambda)y(t)dt \tag{2.14}$$

has a unique solution for each $m=0,\pm1,2$ *and each* $\lambda>0$, *this solution is twice continuously differentiable and satisfies (2.4) for* $a<x<b$, *and*

$$y_m(x)=Y_m(x)+O\left[\frac{\Lambda(\lambda)}{\lambda}Y_m(x)\int_{x_m}^{x}g(t)dt\right] \tag{2.15}$$

for $\lambda>0$ *and* $a<x<b$.

This result can easily be adapted to the case when $p(x)>0$ for $a<x<c$ and $p(x)<0$ for $c<x<b$.

If $\phi'(x)<0$,

$$x_0=b, \quad x_{\pm1}=a, \quad x_2=c$$

and otherwise the conditions (i) *to* (iii) *hold, then the result enunciated in the foregoing also holds.*

It follows from (2.15) that $Y_m(x)$ is an asymptotic representation of $y_m(x)$ as $\lambda\to\infty$, which holds uniformly in x; and also for fixed λ as $x\to x_m$. This is an especially valuable feature when solutions of (2.1) are characterized in terms of their behavior as $x\to x'$. It can further be proved that y_m' is also represented asymptotically by Y_m' [4].

Y_m may be regarded as a first asymptotic approximation to y_m. Higher approximations may be obtained either by expanding the solution in descending powers of λ (Langer [13], Olver [17,18]) or else by using in place of ϕ a change of variable which itself depends on λ (Cherry [2]).

3. Auxiliary Estimates

In the derivation of (2.15) certain estimates are needed. These are in essence estimates on Airy functions. The reader is referred in this connection to [15] and [16].

$Ai(z)$ and $Bi(z)$ are entire functions. All zeros of $Ai(z)$ and $Ai'(z)$ are situated on the negative real axis and all zeros of $Bi(z)$ or $Bi'(z)$, either on the negative real axis or in one of the two sectors $\pi/3<\pm\arg z<\pi/2$. Moreover, on the negative real axis, the zeros of $Ai(z)$ and $Bi(z)$ interlace, and so do the zeros of $Ai'(z)$ and $Bi'(z)$. The behavior of $Ai(z)$ at infinity follows from the asymptotic formulas,

$$Ai(z)=\tfrac{1}{2}\pi^{-\frac{1}{2}}z^{-\frac{1}{4}}e^{-\zeta}[1+O(\zeta^{-1})] \quad z\to\infty, \quad -\pi<\arg z<\pi \tag{3.1}$$

$$Ai(-z)=\pi^{-\frac{1}{2}}z^{-\frac{1}{4}}\left[\cos\left(\zeta-\frac{\pi}{4}\right)+O(\zeta^{-1})\right] \quad z>0, \quad z\to\infty, \tag{3.2}$$

where

$$\zeta=\tfrac{2}{3}z^{\frac{3}{2}}. \tag{3.3}$$

The behavior of $Bi(z)$ can be deduced by means of the relation,

$$Bi(z)=i[\omega^2Ai(\omega^2z)-\omega Ai(\omega z)]. \tag{3.4}$$

It follows that the functions,

$$(1+|z|^{\frac{1}{2}})Ai(z)e^{\zeta}, \quad (1+|z|^{\frac{1}{2}})^{-1}Ai'(z)e^{\zeta} \quad |\arg z|\leq\pi$$
$$(1+|z|^{\frac{1}{2}})Bi(z)e^{-\zeta}, \quad (1+|z|^{\frac{1}{2}})^{-1}Bi'(z)e^{-\zeta} \quad |\arg z|\leq\pi/3$$
$$(1+|z|^{\frac{1}{2}})Bi(z)e^{\zeta}, \quad (1+|z|^{\frac{1}{2}})^{-1}Bi'(z)e^{\zeta} \quad \pi/3\leq|\arg z|\leq\pi,$$

and the reciprocals of the functions,

$$(1+|z|^{\frac{1}{2}})[|Ai(z)|^2+|Bi(z)|^2]^{\frac{1}{2}}e^{\zeta}$$
$$(1+|z|^{\frac{1}{2}})^{-1}[|Ai'(z)|^2+|Bi'(z)|^2]^{\frac{1}{2}}e^{\zeta} \quad \arg z=\pm\pi$$

are bounded.

As a consequence we have the following:

Lemma 1

The following functions are bounded:

$$\Phi(x)Y_m(x)\exp[\sigma_m(x)], \quad \{[\Phi(x)\tilde{Y}_m(x)\exp[\sigma_m(x)]\}^{\pm1}$$

for $m=0,\pm1$ *and* $|\arg\theta_m|\leq\pi$ *or* $m=2$ *and* $\theta<0$

$$[\Phi(x)Y_2(x)e^{-\sigma(x)}]^{\pm1} \quad \textit{for} \quad |\arg\theta|\leq\pi/3.$$

There are also corresponding results involving the $\tilde{Z}_m$ and the derivatives of the Y_m, but these will not be noted in detail (see [4], lemma 2).

Since Y_m and $Y_{m'}$ are linearly independent solutions of (2.1) when m, $m'=0,\pm1,2$ and $m\neq m'$, the Green's function can be expressed as a linear combination of any pair of these, indeed,

$$K(x,t)=\frac{1}{\Delta}[Y_m(x)Y_{m'}(t)-Y_m(t)Y_{m'}(x)], \tag{3.5}$$

where

$$\Delta=Y_m'(t)Y_{m'}(t)-Y_m(t)Y_{m'}'(t)$$

is the Wronskian of Y_m and $Y_{m'}$, which is independent of t and can be computed from the known values of $Ai(0)$, $Ai'(0)$, $Bi(0)$, and $Bi'(0)$.

For the Green's function, we need the following result:

Lemma 2

If $m=0$ and $a<t\leq x<b$, or $m=\pm 1$ and $a<x\leq t<b$, or else $m=2$ and either $a<x\leq t\leq c$ or $c\leq t\leq x<b$, then

$$|K(x,t)\tilde{Y}_m(t)|\leq C\lambda^{-\frac{1}{3}}|\Phi(t)|^{-2}|\tilde{Y}_m(x)| \qquad (3.6)$$

and

$$\left|\frac{\partial}{\partial x}[|\phi'(x)|^{\frac{1}{3}}K(x,t)]\tilde{Y}_m(t)\right| \leq C\lambda^{-\frac{1}{3}}|\Phi(t)|^{-2}|\tilde{Z}_m(x)|. \qquad (3.7)$$

We shall prove (3.6) for $m=0$. The proof of (3.7) and the discussion of the cases $m=\pm 1$, 2 are similar.

We use (3.5) with $m=0$, $m'=1$, say, obtaining

$$K(x,t)=2\pi\omega^{\frac{1}{3}}\lambda^{-\frac{1}{3}}[Y_0(x)Y_1(t)-Y_0(t)Y_1(x)]$$

and consequently,

$$\frac{1}{2\pi}\lambda^{\frac{1}{3}}\Phi^2(t)\left|K(x,t)\frac{\tilde{Y}_0(t)}{\tilde{Y}_0(x)}\right|\leq\left|\frac{Y_0(x)}{\tilde{Y}_0(x)}\right||\Phi^2(t)\tilde{Y}_0(t)Y_1(t)| + \left|\frac{Y_1(x)}{\tilde{Y}_0(x)}\right||\Phi^2(t)Y_0(t)\tilde{Y}_0(t)|. \qquad (3.8)$$

Here, we take $\arg\theta=0$ and $\arg\theta_1=2\pi/3$ if $x<c$; and $\arg\theta=-\pi$ and $\arg\theta_1=-\pi/3$ if $x>c$; noting that in either case $\sigma_1=-\sigma$, and $\mathrm{Re}\,\sigma(x)$ is a decreasing function of x. Clearly, $|Y_0(x)/\tilde{Y}_0(x)|\leq 1$, and $\Phi^2(t)\tilde{Y}_0(t)Y_1(t)$ is bounded by lemma 1. Thus, the first term on the right hand side of (3.8) is bounded. Furthermore,

$$\frac{Y_1(x)}{\tilde{Y}_0(x)}e^{-2\sigma(x)} \quad\text{and}\quad \Phi^2(t)Y_0(t)\tilde{Y}_0(t)e^{2\sigma(t)}$$

are bounded (lemma 1), and $|e^{2\sigma(x)-2\sigma(t)}|\leq 1$ since $t\leq x$ and $\mathrm{Re}\,\sigma(x)$ is a decreasing function of x. Thus, the second term on the right hand side of (3.8) is also bounded, and (3.6) is proved in this case.

4. The Integral Equation and Its Solution

We shall now solve the integral equation (2.14) by the method of successive approximations. Under assumptions (i) to (iii) of Sec. 2, we fix m and λ, and set

$$u_0(x)=Y_m(x) \qquad (4.1)$$

$$u_{n+1}(x)=\int_{x_m}^{x}K(x,t)F(t,\lambda)u_n(t)dt \quad n=0,1,2,\cdots. \qquad (4.2)$$

In order to prove the convergence of $\sum u_n(x)$, we first remark that according to assumption (iii) and lemma 2,

$$|K(x,t)F(t,\lambda)\tilde{Y}_m(t)|\leq(C/\lambda)|\tilde{Y}_m(x)||G(t,\lambda)| \leq(C\Lambda/\lambda)|\tilde{Y}_m(x)|g(t) \qquad (4.3)$$

Next, we prove by induction that

$$|u_n(x)|\leq 1/(n!)\left((C\Lambda/\lambda)\left|\int_{x_m}^{x}g(t)dt\right|\right)^n|\tilde{Y}_m(x)| \qquad n=0,1,2,\cdots. \qquad (4.4)$$

Indeed, (4.4) is true for $n=0$. If it is true for any n, then it follows from (4.2), (4.3), (4.4)

$$|u_{n+1}(x)|\leq\frac{1}{n!}\left(\frac{C\Lambda}{\lambda}\right)^{n+1} \times|\tilde{Y}_m(x)|\left|\int_{x_m}^{x}g(t)\left|\int_{x_m}^{t}g(s)ds\right|^n dt\right|.$$

Since

$$\int_{x_m}^{x}g(t)\left[\int_{x_m}^{t}g(s)ds\right]^n dt=(n+1)^{-1}\left[\int_{x_m}^{x}g(s)ds\right]^{n+1},$$

this proves (4.4) for $n+1$, and hence for all n.

After proving (4.4), we note that, by assumption (iii), $\Lambda(\lambda)/\lambda$ is a bounded function of λ for $\lambda\geq\lambda_0>0$, and

$$\int_{x_m}^{x}g(t)dt$$

is a bounded function of x, so that

$$|u_n(x)|\leq\frac{1}{n!}A^n|\tilde{Y}_m(x)| \quad n=0,1,2,\cdots, \qquad (4.5)$$

where A is independent of x and λ. It follows that

$$\sum u_n(x)/\tilde{Y}_m(x)$$

converges uniformly and absolutely for $a<x<b$, and

$$y_m(x)=\sum_{n=0}^{\infty}u_n(x) \qquad (4.6)$$

defines a function which is clearly continuous for $a<x<b$, and satisfies (2.14). Also,

$$|y_m(x)|\leq|\tilde{Y}_m(x)|\exp\left(\frac{C\Lambda}{\lambda}\left|\int_{x_m}^{x}g(t)dt\right|\right) \leq|\tilde{Y}_m(x)|e^{A}. \qquad (4.7)$$

The uniqueness of the solution can be proved in the usual manner.

To prove the differentiability of $y_m(x)$, we first form

$$d/(dx)[u_n(x)\phi'^{\frac{1}{3}}(x)] =\int_{x_m}^{x}\partial/(\partial x)[K(x,t)\phi'^{\frac{1}{3}}(x)]F(t,\lambda)u_{n-1}(t)dt.$$

Here, we use assumption (iii), lemma 2, and (4.4) to

show that

$$\left|\frac{d}{dx}[u_n(x)\phi'^{\frac{1}{2}}(x)]\right| \leq \frac{1}{n!}\left(\frac{C\Lambda}{\lambda}\left|\int_{x_m}^{x} g(t)dt\right|\right)^n |\bar{Z}_m(x)|,$$

and conclude that the series

$$\sum (u_n\phi'^{\frac{1}{2}})'/\bar{Z}_m(x)$$

converges uniformly, and hence that $\sum u_n\phi'^{\frac{1}{2}}$ may be differentiated term by term to yield $(y_m\phi'^{\frac{1}{2}})$. Finally,

$$y_m'(x)=\phi'^{-\frac{1}{2}}[y_m(x)\phi'^{\frac{1}{2}}(x)]'-\tfrac{1}{2}\frac{\phi''}{\phi'}y_m(x)$$

is continuous for $a<x<b$.

Next, we remark that it is seen from (4.2) that

$$u_n''(x)=F(x,\lambda)u_{n-1}(x)+\int_{x_m}^{x}\frac{\partial^2}{\partial x^2}K(x,t)F(t,\lambda)u_{n-1}(t)dt$$

$$=F(x,\lambda)u_{n-1}(x)-[\lambda^2\phi\phi'^2+\tfrac{1}{2}\{\phi,x\}]u_n(x)$$

since $K(x,t)$ satisfies (2.4). As before, it follows that (4.6) may be differentiated a second time, $y_m''(x)$ is continuous, and $y_m(x)$ satisfies (2.4) for $a<x<b$.

To estimate y_m-Y_m, we use (2.14), (4.7), and (4.3).

$$|y_m(x)-Y_m(x)| \leq e^A\left|\int_{x_m}^{x}|K(x,t)F(t,\lambda)\bar{Y}_m(t)|\,dt\right|$$

$$\leq \frac{C\Lambda}{\lambda}e^A|\bar{Y}_m(x)|\left|\int_{x_m}^{x} g(t)dt\right|.$$

This completes the proof.

5. Application to Bessel Functions

We shall now apply our results to Bessel functions of large order. The asymptotic forms of these functions were determined from the differential equation they satisfy by Langer, Cherry, Olver, and also in [3], Sec. 4.8. Here, we shall show that the leading term of the asymptotic expansion follows from our result without much effort.

The functions,

$$x^{\frac{1}{2}}J_\lambda(\lambda x),\quad x^{\frac{1}{2}}H_\lambda^{(1)}(\lambda x),\quad x^{\frac{1}{2}}H_\lambda^{(2)}(\lambda x), \tag{5.1}$$

satisfy the differential equation,

$$\frac{d^2y}{dx^2}+\left[\lambda^2\left(1-\frac{1}{x^2}\right)+\frac{1}{4x^2}\right]y=0. \tag{5.2}$$

This equation is of the form (1.1) and it will turn out that all our conditions are satisfied on the interval $0<x<\infty$. According to (2.3) we set

$$\phi\phi'^2=1-(1/x^2). \tag{5.3}$$

and obtain by integration

$$\tfrac{2}{3}[-\phi(x)]^{\frac{3}{2}}=\alpha(x)=\int_x^1\left(\frac{1}{t^2}-1\right)^{\frac{1}{2}}dt$$

$$=-(1-x^2)^{\frac{1}{2}}+\log[1+(1-x^2)^{\frac{1}{2}}]-\log x$$

$$0<x\leq 1 \tag{5.4}$$

$$\tfrac{2}{3}[\phi(x)]^{\frac{3}{2}}=\beta(x)=\int_1^x\left(1-\frac{1}{t^2}\right)^{\frac{1}{2}}dt$$

$$=(x^2-1)^{\frac{1}{2}}-\cos^{-1}\frac{1}{x}\quad 1\leq x<\infty. \tag{5.5}$$

We note that

$$\alpha(x)=-\log x+\log 2-1+O(x^2)\quad \text{as}\quad x\to 0$$

$$\alpha'(x)=-\frac{1}{x}+O(x)\quad \text{as}\quad x\to 0 \tag{5.6}$$

$$\beta(x)=x-\frac{\pi}{2}+O(1/x)\quad \text{as}\quad x\to\infty,$$

$$\beta'(x)=1+O(1/x^2)\quad \text{as}\quad x\to\infty. \tag{5.7}$$

The differential equation (5.2) may now be written in the form (2.4) with

$$F(x,\lambda)=F(x)=\tfrac{1}{2}\{\phi,x\}-1/(4x^2). \tag{5.8}$$

We proceed to verify that the conditions (i) to (iii) of Sec. 2 are satisfied. Clearly, $\phi(x)$ is three times continuously differentiable (indeed it is analytic). Also $\phi'(x)>0$, and it follows that $F(x)$ is continuous (in fact, analytic). We also have

$$|G(t,\lambda)|=\left|\frac{\lambda^{\frac{1}{2}}F(t)}{\Phi^2(t)}\right|\leq\frac{|F(t)|}{|\phi(t)|^{\frac{1}{2}}\phi'(t)}=\frac{t|F(t)|}{|1-t^2|^{\frac{1}{2}}},$$

and in order to verify assumption (iii), we shall investigate the behavior of $F(t)$ as $t\to 0$ and as $t\to\infty$. To do this, we use the chain rule

$$\{\phi,x\}=\{\phi,\chi\}\chi'^2+\{\chi,x\}$$

for the Schwarzian derivative, setting $\chi=\alpha$ when $0<x<1$, and $\chi=\beta$ when $1<x<\infty$. Thus,

$$F(x)=\frac{5\alpha'^2}{18\alpha^2}+\tfrac{1}{2}\{\alpha,x\}-\frac{1}{4x^2}=\frac{5\alpha'^2}{18\alpha^2}-\frac{4+x^2}{4(1-x^2)^2}$$

$$0<x<1 \tag{5.9}$$

$$F(x)=\frac{5\beta'^2}{18\beta^2}-\frac{4+x^2}{4(1-x^2)^2}\quad 1<x<\infty. \tag{5.10}$$

From these formulas, in conjunction with (5.6) and (5.7), it is seen that

$$F(x)=O[(x\log x)^{-2}]\quad \text{as}\quad x\to 0$$
$$F(x)=O(x^{-2})\quad \text{as}\quad x\to\infty. \tag{5.11}$$

It follows that

$$\left[x\log\left(2+\frac{1}{x}\right)\right]^2 F(x) \tag{5.12}$$

is bounded when $0<x<\infty$, and we may set

$$g(t)=\left[t|1-t^2|^{\frac{1}{2}}\log^2\left(2+\frac{1}{t}\right)\right]^{-1} \tag{5.13}$$

and Λ some fixed number (independent of t and λ) in assumption (iii). Clearly

$$\int_0^\infty g(t)dt<\infty,$$

and we note in passing that

$$\begin{aligned}\int_0^x g(t)dt&=O(1/\log x) \quad \text{as} \quad x\to 0\\ \int^\infty g(t)dt&=O(1/x) \quad \text{as} \quad x\to\infty.\end{aligned} \tag{5.14}$$

On verifying all assumptions, we know that there exist solutions $y_m(x)$ of (5.2) for which (2.15) holds. It remains to identify these solutions in terms of the functions (5.1).

Let us start with $y_0(x)$. By fixing $\lambda>0$, and making $x\to 0$, we have from (2.15) and (5.14),

$$y_0(x)=\phi'^{-\frac{1}{2}}Ai(-\lambda^{\frac{2}{3}}\phi)\left[1+O\left(\frac{1}{\log x}\right)\right] \quad \text{as} \quad x\to 0.$$

Since $-\phi\to\infty$ as $x\to 0$, we may use (3.1) and (5.6) to write

$$y_0(x)=\frac{\exp[-\lambda\alpha(x)]}{2(\pi\phi')^{\frac{1}{2}}(-\phi)^{\frac{1}{2}}}\left[1+O\left(\frac{1}{\log x}\right)\right] \quad \text{as} \quad x\to 0.$$

Here, we use (5.6) and (5.3) to obtain the asymptotic behavior of $y_0(x)$ as $x\to 0$, λ being fixed.

$$y_0(x)=\tfrac{1}{2}\pi^{-\frac{1}{2}}2^{-\lambda}\lambda^{-1/6}e^{\lambda}x^{\lambda+\frac{1}{2}}\times\left[1+O\left(\frac{1}{\log x}\right)\right] \quad \text{as} \quad x\to 0 \tag{5.15}$$

In order to avoid inessential complications, let us assume that λ is not an integer, so that J_λ and $J_{-\lambda}$ are linearly independent. Then, $y_0(x)$, being a solution of (5.2), must be of the form,

$$y_0(x)=x^{\frac{1}{2}}[c_1(\lambda)J_\lambda(\lambda x)+c_2(\lambda)J_{-\lambda}(\lambda x)]. \tag{5.16}$$

Now,

$$J_{\pm\lambda}(\lambda x)=\frac{(\lambda x/2)^{\pm\lambda}}{\Gamma(1\pm\lambda)}[1+O(x)] \quad \text{as} \quad x\to 0,$$

and substituting in (5.16) the asymptotic form of each term as $x\to 0$, we see that

$$c_1(\lambda)=\tfrac{1}{2}\pi^{-\frac{1}{2}}\lambda^{-\lambda-1/6}e^{\lambda}\Gamma(\lambda+1), \quad c_2(\lambda)=0,$$

so that

$$x^{\frac{1}{2}}J_\lambda(\lambda x)=\frac{2\pi^{\frac{1}{2}}}{\Gamma(\lambda+1)}e^{-\lambda}\lambda^{\lambda+1/6}y_0(x). \tag{5.17}$$

We then have

$$J_\lambda(\lambda x)=\frac{2\pi^{\frac{1}{2}}}{\Gamma(\lambda+1)}e^{-\lambda}\lambda^{\lambda+1/6}(x\phi')^{-\frac{1}{2}}Ai(-\lambda^{\frac{2}{3}}\phi)\times\left[1+O\left(\frac{1}{\lambda}\right)\right] \tag{5.18}$$

for $\lambda>0$ and $0<x<\infty$, except that in case $x>1$, the error term must be modified in order to take account of the zeros of Ai. When $x<1$, $O(\lambda^{-1})$ in (5.18) may be replaced by $O[(\lambda\log x)^{-1}]$. The asymptotic form of the derivative may be obtained similarly.

A somewhat simpler asymptotic form results if Stirling's formula is used for $\Gamma(\lambda+1)$.

$$J_\lambda(\lambda x)=(\tfrac{1}{2}\lambda^{\frac{2}{3}}x\phi')^{-\frac{1}{2}}Ai(-\lambda^{\frac{2}{3}}\phi)[1+O(\lambda^{-1})] \tag{5.19}$$

for $\lambda>0$ and $0<x<\infty$, with the same modification of the error term near zeros of $Ai(-\lambda^{\frac{2}{3}}\phi)$ as before. In (5.19) the O symbol cannot be strengthened for $0<x<1$.

We now proceed to $y_1(x)$, fix $\lambda>0$, and make $x\to\infty$. From (2.15) and (5.14),

$$y_1(x)=\frac{1}{\phi'^{-\frac{1}{2}}}Ai(-\lambda^{\frac{2}{3}}\omega\phi)\left[1+O\left(\frac{1}{x}\right)\right] \quad \text{as} \quad x\to\infty.$$

Since $\phi\to\infty$ as $x\to\infty$, we may use (3.1) and (5.7)

$$\begin{aligned}y_1(x)&=\frac{\exp[i\lambda\beta(x)+i\pi/12]}{2(\pi\phi')^{\frac{1}{2}}\phi^{\frac{1}{4}}}\left[1+O\left(\frac{1}{x}\right)\right] \quad \text{as} \quad x\to\infty\\ &=\tfrac{1}{2}\pi^{-\frac{1}{2}}\lambda^{-1/6}\exp\left[i\left(\lambda x-\frac{\lambda\pi}{2}+\frac{\pi}{12}\right)\right]\left[1+O\left(\frac{1}{x}\right)\right]\end{aligned}$$

as $x\to\infty$.

$y_1(x)$, being a solution of (5.2), must be of the form,

$$y_1(x)=x^{\frac{1}{2}}[c_3(\lambda)H_\lambda^{(1)}(\lambda x)+c_4(\lambda)H_\lambda^{(2)}(\lambda x)].$$

We know that

$$x^{\frac{1}{2}}H_\lambda^{(1,2)}(\lambda x)=\left(\frac{2}{\pi\lambda}\right)^{\frac{1}{2}}\exp\left[\pm i\left(\lambda x-\frac{\lambda\pi}{2}-\frac{\pi}{4}\right)\right]\times\left[1+O\left(\frac{1}{x}\right)\right] \quad \text{as} \quad x\to\infty,$$

where $+$ is for $H_\lambda^{(1)}$ and $-$ for $H_\lambda^{(2)}$, and conclude, similarly to the previous case, that

$$x^{\frac{1}{2}}H_\lambda^{(1)}(\lambda x)=2^{\frac{1}{2}}\lambda^{-\frac{1}{2}}e^{-\pi i/3}y_1(x), \tag{5.20}$$

and hence

$$H_\lambda^{(1)}(\lambda x)=2^{\frac{1}{3}}\lambda^{-\frac{1}{3}}e^{-\pi i/3}(x\phi')^{-\frac{1}{4}}Ai(-\lambda^{\frac{2}{3}}\omega\phi)\times\left[1+O\left(\frac{1}{\lambda}\right)\right] \quad (5.21)$$

for $\lambda>0$ and $0<x<\infty$. If $x>1$, $O(\lambda^{-1})$ may be replaced by $O[(\lambda x)^{-1}]$. The corresponding asymptotic form for the second Hankel function for $\lambda>0$ and $0<x<\infty$ is

$$H_\lambda^{(2)}(\lambda x)=2^{\frac{1}{3}}\lambda^{-\frac{1}{3}}e^{i\pi/3}(x\phi')^{-\frac{1}{4}}Ai(-\lambda^{\frac{2}{3}}\omega^{-1}\phi)\times\left[1+O\left(\frac{1}{\lambda}\right)\right] \quad (5.22)$$

with the same remarks about the error term as before.

6. Application to Hermite Polynomials

Next we shall apply our results to Hermite polynomials of large order. The asymptotic forms of these have been the subject of several investigations. For references and a summary of results see [21]. Since Hermite polynomials are connected with parabolic cylinder functions, results regarding the latter functions are also relevant [8], [11]. Here, we shall follow the discussion given by H. Skovgaard [19].

n is a nonnegative integer, and $N=2n+1$ throughout this section. It is known [7, Vol. II, p. 193 Eq. (13)] that

$$\exp(-\tfrac{1}{2}Nx^2)H_n(N^{\frac{1}{2}}x) \quad (6.1)$$

satisfies the differential equation,

$$\frac{d^2y}{dx^2}+N^2(1-x^2)y=0, \quad (6.2)$$

whose general solution may also be written in the form,

$$c_1D_n[(2N)^{\frac{1}{2}}x]+c_2D_{-n-1}[(2N)^{\frac{1}{2}}x]. \quad (6.3)$$

(6.2) is of the form (1.1) on (a,∞), where $-1<a<0$, and N takes the place of λ. Since $p(x)$ changes here from positive to negative values as x increases, we shall have $\phi'(x)<0$ and will use the second form of the results summarized in Sec. 2. Accordingly, we set

$$\phi\phi'^2=1-x^2$$

and obtain

$$\tfrac{2}{3}\phi^{\frac{3}{2}}=\alpha(x)=\int_x^1(1-t^2)^{\frac{1}{2}}dt$$

$$=-\tfrac{1}{2}x(1-x^2)^{\frac{1}{2}}+\tfrac{1}{2}\cos^{-1}x \quad a<x\leq 1, \quad (6.4)$$

$$\tfrac{2}{3}(-\phi)^{\frac{3}{2}}=\beta(x)=\int_1^x(t^2-1)^{\frac{1}{2}}dt$$

$$=\tfrac{1}{2}x(x^2-1)^{\frac{1}{2}}-\tfrac{1}{2}\cosh^{-1}x \quad 1\leq x<\infty, \quad (6.5)$$

where $\cos^{-1}$ and $\cosh^{-1}$ denote the principal branches of the inverse functions. We then have (2.4) with

$$F(x,\lambda)=F(x)=\tfrac{1}{2}\{\phi,x\}$$

$$=\tfrac{5}{16}\frac{1-x^2}{\phi^3}-\frac{3x^2+2}{4(1-x^2)^2} \quad x\neq 1, \quad (6.6)$$

and, if we define $F(1)=\lim F(x)$ as $x\to 1$, $F(x)$ is continuous for $a\leq x<\infty$. Moreover, it can be shown that $(1+x)^2F(x)$ is bounded on (a,∞). Thus all the conditions of Sec. 2 will be satisfied if we take $\Lambda(\lambda)$ to be some suitable number independent of x and λ, and set

$$g(x)=(1+x)^{-2}|1-x^2|^{-\frac{1}{2}}. \quad (6.7)$$

Since $g(x)=O(x^{-3})$ as $x\to\infty$, we also have

$$\int_x^\infty g(t)dt=O\left(\frac{1}{1+x^2}\right) \quad a<x<\infty. \quad (6.8)$$

We now apply the results of Sec. 2 with $m=0$, N fixed, and $x\to\infty$. Since $-\phi(x)>0$ in this case, $\tilde{Y}_0=Y_0$ and we have from (6.8)

$$y_0(x)=|\phi'|^{-\frac{1}{2}}Ai(-N^{\frac{2}{3}}\phi)\left[1+O\left(\frac{1}{1+x^2}\right)\right].$$

Since $-\phi\to\infty$, we use (3.1) to obtain

$$y_0(x)=\frac{\exp[-N\beta(x)]}{2\pi^{\frac{1}{2}}N^{1/6}(-\phi\phi'^2)^{\frac{1}{4}}}[1+O(x^{-2})] \quad \text{as} \quad x\to\infty.$$

From (6.5)

$$\beta(x)=\tfrac{1}{2}x^2-\tfrac{1}{2}\log(2x)-\tfrac{1}{4}+O(x^{-2}) \quad x\to\infty$$

and hence,

$$y_0(x)=\frac{(2x)^n}{(2\pi)^{\frac{1}{2}}N^{1/6}}\exp[-\tfrac{1}{2}N(x^2-\tfrac{1}{2})]\times[1+O(x^{-2})] \quad x\to\infty \quad (6.9)$$

In particular, it is seen that $y_0(x)\to 0$ as $x\to\infty$. $y_0(x)$, being a solution of (6.2) is of the form (6.3). Now, in (6.3), the first term vanishes, and the second term is unbounded, as $x\to\infty$ [7, vol. II p. 122, Eq. (1)]. It follows that $c_2=0$ and $y_0(x)$ is a constant multiple of (6.1). From the explicit formula for Hermite polynomials [7, vol II, p. 193, Eq. (9)],

$$\exp(-\tfrac{1}{2}Nx^2)H_n(N^{\frac{1}{2}}x)=(2N^{\frac{1}{2}}x)^n\exp(-\tfrac{1}{2}Nx^2)[1+O(x^{-2})] \quad (6.10)$$

as $x\to\infty$ and a comparison of (6.9), and (6.10) shows that

$$H_n(N^{\frac{1}{2}}x)=(2\pi)^{\frac{1}{2}}N^{n/2+1/6}\exp[\tfrac{1}{2}N(x^2-\tfrac{1}{2})]y_0(x).$$

If we combine this result with (2.15), we obtain the

asymptotic formula,

$$H_n(N^{\frac{1}{2}}x)=(2\pi)^{\frac{1}{2}}N^{n/2+1/6}\exp[\tfrac{1}{2}N(x^2-\tfrac{1}{2})]$$
$$\times|\phi'|^{-\frac{1}{2}}Ai(-N^{\frac{2}{3}}\phi)\left[1+O\left(\frac{1}{n(1+x^2)}\right)\right], \quad (6.11)$$

which is valid uniformly in $-1<a<x<\infty$ as $n\to\infty$, except that the error term must be modified to take care of the zeros of the Airy function when $a<x<1$. This result can be extended to $-\infty<x<0$ by using $H_n(-x)=(-1)^nH_n(x)$.

Simpler asymptotic formulas holding for restricted ranges of x may be deduced from (6.11). In [19] all the known formulas are so deduced, in part with improved conditions of validity. Here, we shall restrict ourselves to a brief indication of two of these results.

If we take $0\leq x<1$ and set $x=\cos\theta$, $0<\theta\leq\pi/2$, we have

$$\alpha=\tfrac{1}{4}(2\theta-\sin 2\theta)$$

from (6.4). By using (3.2) in (6.11) we then obtain

$$H_n(N^{\frac{1}{2}}\cos\theta)$$
$$=2^{\frac{1}{2}}N^{n/2}\sin^{-\frac{1}{2}}\theta\exp(\tfrac{1}{4}N\cos 2\theta)$$
$$\times\{\cos[\tfrac{1}{4}N(2\theta-\sin 2\theta)-\tfrac{1}{4}\pi]+O(n^{-1}\theta^{-3})\}. \quad (6.12)$$

For $x>1$, we set $x=\cosh\theta$, $\theta>0$, and obtain similarly

$$H_n(N^{\frac{1}{2}}\cosh\theta)=(2\sinh\theta)^{-\frac{1}{2}}N^{n/2}\exp[\tfrac{1}{4}N(2\theta+e^{-2\theta})]$$
$$\times\left[1+O\left(n^{-1}\sinh^{-3}\frac{2\theta}{3}\right)\right]. \quad (6.13)$$

SINGULARITIES

7. Outline of the Method and Results

Here, we shall give only a brief outline of the method following Swanson and state the result. The proofs are similar to those given in Sec. 3 and 4 for the transition point case and can be found in [20]. By using (1.4) as our comparison equation, we introduce a new variable x by the substitution,

$$t=\psi(x), \quad w=|\psi'(x)|^{\frac{1}{2}}U,$$

thus changing (1.4) into

$$\frac{d^2U}{dx^2}+\left[\left(\frac{\lambda^2}{\psi}+\frac{1-\nu^2}{4\psi^2}\right)\psi'^2+\tfrac{1}{2}\{\psi,x\}\right]U=0, \quad (7.1)$$

an equation whose solutions are known from (1.5). $\{\psi,x\}$ is defined by (2.2). The differential equations (1.1) and (7.1) will be nearly equal for large λ if we choose ψ so that

$$\psi^{-1}\psi'^2=p(x). \quad (7.2)$$

With this choice of ψ, (1.1) may be written in the form,

$$\frac{d^2y}{dx^2}+\left[\left(\frac{\lambda^2}{\psi}+\frac{1-\nu^2}{4\psi^2}\right)\psi'^2+\tfrac{1}{2}\{\psi,x\}\right]y=F(x,\lambda)y, \quad (7.3)$$

where

$$F(x,\lambda)=\frac{1-\nu^2}{4\psi^2}\psi'^2+\tfrac{1}{2}\{\psi,x\}-r(x,\lambda). \quad (7.4)$$

As in the case of a transition point, we shall assume that the differential equation has been transformed to this form and, moreover, that ν has been so chosen that $|p(x)|^{-\frac{1}{2}}F(x,\lambda)$ is an integrable function of x for each $\lambda>0$.

As a typical case, let us assume that $(x-c)p(x)$ is a twice continuously differentiable function of x and $(x-c)^2r(x,\lambda)$, for each $\lambda>0$, a continuously differentiable function of x, for $a<x<b$ so that

$$p(x)=q_0(x-c)^{-1}[1+O(x-c)]$$
$$r(x,\lambda)=r_0(x-c)^{-2}[1+O(x-c)] \quad \text{as} \quad x\to c$$

and assume $q_0>0$, $r_0\leq\frac{1}{4}$, r_0 independent of λ. Then $\psi(x)=q_0(x-c)[1+O(x-c)]$ from (7.2), $\psi'(x)/\psi(x)=(x-c)^{-1}[1+O(x-c)]$, $\{\psi,x\}=O[(x-c)^{-1}]$ and

$$F(x,\lambda)=\frac{1-\nu^2-4r_0}{4(x-c)^2}+O[(x-c)^{-1}] \quad x\to c,$$

and the integrability condition is satisfied if and only if $\nu=(1-4r_0)^{\frac{1}{2}}$.

If $F(x,\lambda)=0$, (7.3) is identical with (7.1). In general, we obtain from (7.3) by the method of variation of parameters, the integral equation of Volterra type,

$$y(x)=U(x)+\int_{x'}^{x}K(x,t)F(t,\lambda)y(t)dt \quad (7.5)$$

satisfied by solutions of (7.3). Here, U is a solution of (7.1), x' is a fixed point of the interval, and $K(x,t)$ is the Green's function of (7.1). For a fixed t in (a,b), $K(x,t)$, considered as a function of x, satisfies (7.1) for $x\neq t$, and also satisfies the initial conditions $K(t,t)=0$, $(\partial K/\partial x)(t,t)=1$.

The significant solutions of (7.1) are obtained from (1.5) as

$$U_0(x)=(\psi/\psi')^{\frac{1}{2}}J_\nu(2\lambda\psi^{\frac{1}{2}}),$$
$$U_1(x)=(\psi/\psi')^{\frac{1}{2}}H_\nu^{(1)}(2\lambda\psi^{\frac{1}{2}}), \quad (7.6)$$
$$U_{-1}(x)=(\psi/\psi')^{\frac{1}{2}}H_\nu^{(2)}(2\lambda\psi^{\frac{1}{2}}).$$

For the error estimates, it is desirable to modify U_0 in the region where it has zeros. Let $\Delta>0$ be so chosen that $J_\nu(z)>0$ for $0<z\leq\Delta$, and set

$$\tilde{U}_0(x)=U_0(x) \quad \text{if} \quad 4\lambda^2\psi\leq\Delta^2,$$
$$=(\psi/\psi')^{\frac{1}{2}}[|J_\nu(2\lambda\psi^{\frac{1}{2}})|^2+|Y_\nu(2\lambda\psi^{\frac{1}{2}})|^2]^{\frac{1}{2}} \quad \text{if} \quad 4\lambda^2\psi>\Delta^2, \quad (7.7)$$
$$\tilde{U}_{\pm1}(x)=U_{\pm1}(x).$$

It is known from the theory of Bessel functions that the functions

$$z^{-\nu}(1+|z|^{\frac{1}{2}+\nu})e^{-|\mathrm{Im}z|}J_\nu(z)$$
$$z^{\nu}(1+|z|^{\frac{1}{2}-\nu})e^{\mp iz}H_\nu^{(1,2)}(z)$$

are bounded when $|\arg z|\leq\pi$, except that a slight modification of the second function is necessary when $\nu=0$ and $H_0^{(1,2)}(z)$ has a logarithmic singularity. Also, in the second function the upper sign is taken for $H_\nu^{(1)}$ and the lower sign for $H_\nu^{(2)}$. Moreover, the second function has a bounded reciprocal.

This information can be used to develop results which correspond to lemma 1 and give bounds for the U_m, U_m', $\tilde{U}_m$, $1/\tilde{U}_m$. These results are in turn used to prove under certain conditions [see (iv) below],

$$|K(x,t)\tilde{U}_m(t)|\leq C\lambda^{-1}|p(x)|^{-\frac{1}{4}}|\tilde{U}_m(x)|,$$

together with a corresponding estimate for $\partial K(x,t)/\partial x$; in fact, to establish an analogue of lemma 2 in this case.

From here on, the investigation resembles that described in Sec. 4. The integral equation is solved by successive approximations, and the uniqueness and differentiability of the solution is proved. The estimates arising out of the successive approximations are used under the integral sign to provide a better estimate of $y(x)-U(x)$. There is one new feature, though. $x=c$ is a singular point of the integral equation and it is impossible to continue a general solution of (7.3) across this point. Accordingly, solutions corresponding to $x'=a$ or $x'=b$ exist only on the intervals (a,c) or (c,b), respectively. Again, for $x'=c$, the integral equation has a solution only if $U(x)$ in (7.5) is a multiple of $U_0(x)$. This solution can be continued across $x=c$, but the continuation is not unique: indeed, in $U_0(x)$ itself we may take $\arg\psi^{\frac{1}{4}}=\pi/2$, or $\arg\psi^{\frac{1}{4}}=-\pi/2$ when $x<c$.

Apart from this, the discussion of the integral equation resembles the discussion given in Sec. 4 and leads to the following result:

If (i) $\psi(x)$ *is three times continuously differentiable and* $\psi'(x)>0$ *for* $a<x<b$, *and* $\psi(c)=0$; (ii) *for each fixed* $\lambda>0$, $F(x,\lambda)$ *is a continuous function of* x *for* $a<x<c$ *and* $c<x<b$; (iii) $|p(x)|^{-\frac{1}{4}}|F(x,\lambda)|\leq\Lambda(\lambda)g(x)$ *for* $x\neq c$ *where* $(1+\lambda)^{-1}\Lambda(\lambda)$ *is a bounded function of* λ *and* $g(x)$ *is integrable over* (a,b); *and* (iv) *either* $m=0$, $x_0=c$, *and* $a<x<b$ *or* $m=\pm1$, $x_m=a$, $\arg\phi=m\pi$, *and* $a<x<c$ *or* $m=\pm1$, $x_m=b$, $\arg\phi=0$, *and* $c<x<b$; *then the integral equation,*

$$u_m(x)=U_m(x)+\int_{x_m}^{x}K(x,t)F(t,\lambda)u_m(t)dt \qquad (7.8)$$

has a unique solution for each $\lambda>0$; *this solution is twice continuously differentiable and satisfies* (7.3), *except possibly for* $x=c$ *in case* $m=0$, *and on the intervals specified in* (iv),

$$u_m(x)=U_m(x)+O\left[\frac{\Lambda(\lambda)}{\lambda}\tilde{U}_m(x)\int_{x_m}^{x}g(t)dt\right]. \qquad (7.9)$$

There is a corresponding formula for $y_m'(x)$.

8. Application to Laguerre Polynomials

Laguerre polynomials of large order have been investigated by several writers. References and a summary of results will be found in [21]. Of more recent investigations especially those by Tricomi [22] are noteworthy. Since these polynomials are connected with confluent hypergeometric functions, results regarding the latter functions are also pertinent; see [9] and the references given there.

Laguerre polynomials present an especially interesting feature in that the differential equation which they satisfy has both a singularity and a transition point. Accordingly, we shall obtain two asymptotic representations, one in terms of Bessel functions and the other in terms of Airy functions. The regions of validity of the two representations overlap, and in the overlapping part the two representations are asymptotically equivalent.

We shall consider Laguerre polynomials $L_n^\alpha(z)$ with fixed $\alpha\geq0$, large n, and unrestricted real z. Throughout this section, $N=n+(\alpha+1)/2$. It is known [7, Sec. 10.12] that

$$x^{(\alpha+1)/2}e^{-2Nx}L_n^\alpha(4Nx) \qquad (8.1)$$

satisfies the differential equation,

$$\frac{d^2y}{dx^2}+\left[4N^2\left(\frac{1}{x}-1\right)+\frac{1-\alpha^2}{4x^2}\right]y=0. \qquad (8.2)$$

Here, N is the large parameter, and the differential equation has a singularity at $x=0$ and a transition point at $x=1$. We shall first study the asymptotic representation on an interval including the singularity.

From (7.2),

$$\psi^{-1}\psi'^2=4[(1/x)-1],$$

and we obtain

$$\begin{aligned}\psi(x)&=-[(x^2-x)^{\frac{1}{2}}+\sinh^{-1}(-x)^{\frac{1}{2}}]^2 && x\leq0,\\ &=[(x-x^2)^{\frac{1}{2}}+\sin^{-1}x^{\frac{1}{2}}]^2 && 0\leq x<1.\end{aligned} \qquad (8.3)$$

Here, all square roots are nonnegative, and $\sin^{-1}$, $\sinh^{-1}$ denote the principal branches of the respective functions. Also, $r(x,\lambda)=(1-\alpha^2)/(4x^2)$, $r_0=(1-\alpha^2)/4\leq\frac{1}{4}$ here, and we obtain $\nu=(1-4r_0)^{\frac{1}{2}}=\alpha$. (8.2) can then be written in the form (7.3) with

$$F(x,\lambda)=F(x)=\frac{1-\alpha^2}{4\psi^2}\psi'^2+\tfrac{1}{2}\{\psi,x\}-\frac{1-\alpha^2}{4x^2} \qquad x\neq0. \qquad (8.4)$$

From the behavior of $\psi(x)$ near $x=0$ and $x=-\infty$, it is easy to show that $|x|^{\frac{1}{4}}F(x)$ is bounded in some neighborhood of $x=0$ and $x^2F(x)$, for say, $x\leq-1$. We can now verify conditions (i), (ii), (iii) of the preceding section on the interval $-\infty<x<b<1$, the last condition with $\Lambda(\lambda)$ a constant and $g(x)=[|x|(1-x)]^{-\frac{1}{4}}$ so that

$$\int_0^x g(t)dt=\pm\left(\frac{|x|}{1-x}\right)^{\frac{1}{4}}. \qquad (8.5)$$

We now apply the results of the preceding section with $m=0$. By fixing N and making $x\to0$, we have

$\psi(x)=4x[1+O(x)]$, $\psi'(x)=4+O(x)$ and from (7.9) in combination with the known behavior of $J_\alpha(z)$ near the origin,

$$u_0(x)=\left(\frac{\psi}{\psi'}\right)^{\frac{1}{2}}J_\alpha(2N\psi^{\frac{1}{2}})[1+O(|x|^{\frac{1}{2}})]$$

$$=\frac{(2N)^\alpha x^{(\alpha+1)/2}}{\Gamma(\alpha+1)}[1+O(|x|^{\frac{1}{2}})] \quad x\to 0. \quad (8.6)$$

From the behavior of the general solution of (8.2) at $x=0$, it follows that (8.1) and $u_0(x)$ are numerical multiples of each other. Moreover, from the explicit representation of $L_n^\alpha(z)$ [7, Vol. II, p. 188 Eq. (7)],

$$x^{(\alpha+1)/2}e^{-2Nx}L_n^\alpha(4Nx)$$

$$=\frac{\Gamma(n+\alpha+1)}{n!\Gamma(\alpha+1)}x^{(\alpha+1)/2}[1+O(x)] \quad x\to 0 \quad (8.7)$$

and from a comparison of (8.6) and (8.7),

$$L_n^\alpha(4Nx)=\frac{\Gamma(n+\alpha+1)}{n!(2N)^\alpha}x^{-(\alpha+1)/2}e^{2Nx}u_0(x). \quad (8.8)$$

From this formula in combination with (7.9), we obtain the asymptotic representation,

$$L_n^\alpha(4Nx)=\frac{\Gamma(n+\alpha+1)}{n!(2N)^\alpha}x^{-(\alpha+1)/2}e^{2Nx}\left(\frac{\psi}{\psi'}\right)^{\frac{1}{2}}$$

$$\times J_\alpha(2N\psi^{\frac{1}{2}})\left[1+O\left(\frac{1}{N}\left|\frac{x}{1-x}\right|^{\frac{1}{2}}\right)\right] \quad (8.9)$$

valid, with $N=n+(\alpha+1)/2$ and ψ given by (8.3), for $n=0, 1, 2, \cdots$ and $-\infty<x<b<1$, except that the error term must be modified in the vicinity of zeros of J_α, which occur when $x>0$. For $x<0$, $\arg x=\arg\psi$ must be taken in (8.9): with this convention, the right hand side of (8.9) is defined unambiguously.

We now turn to the transition point. According to (2.3), we set

$$\phi\phi'^2=4[(1/x)-1]$$

and obtain

$$\alpha(x)=\tfrac{2}{3}\phi^{\frac{3}{2}}=\cos^{-1}x^{\frac{1}{2}}-(x-x^2)^{\frac{1}{2}} \qquad 0<x\le 1$$

$$\beta(x)=\tfrac{2}{3}(-\phi)^{\frac{3}{2}}=(x^2-x)^{\frac{1}{2}}-\cosh^{-1}x^{\frac{1}{2}} \quad 1\le x<\infty \quad (8.10)$$

with the same conventions about many valued functions as in (8.3). In this case

$$F(x,\lambda)=F(x)=\tfrac{1}{2}\{\phi,x\}-(1-\alpha^2)/(4x^2)$$

is continuous for $0<x<\infty$ and is $O(x^{-2})$ as $x\to\infty$, and on any interval $0<a<x<\infty$, the conditions of Sec. 2, with $\phi'(x)<0$, can easily be verified with $\Lambda(\lambda)$ a suitable constant, and $g(x)=x^{-\frac{1}{2}}|1-x|^{-\frac{1}{2}}$. It follows also that

$$\int_x^\infty g(t)dt=O\left(\frac{1}{x}\right) \quad a<x<\infty. \quad (8.11)$$

We now apply the results of Sec. 2 with $m=0$, $x_0=b$. By fixing n and making $x\to\infty$, we see that

$$y_0(x)/Y_0(x)\to 1.$$

Moreover, in the expression (2.10) for Y_0, we may use (3.2), and in the resulting asymptotic form,

$$\beta(x)=x-\log(2x^{\frac{1}{2}})-\tfrac{1}{2}+O(x^{-\frac{1}{2}}) \quad x\to\infty,$$

and we finally find

$$y_0(x)=\frac{(4x)^Ne^{-2Nx+N}}{2(2\pi)^{\frac{1}{2}}N^{1/6}}[1+O(x^{-\frac{1}{2}})] \quad x\to\infty \quad (8.12)$$

In particular, $y_0(x)\to 0$ as $x\to\infty$. It being known that numerical multiples of (8.1) are the only solutions of (8.2) vanishing as $x\to\infty$, we conclude that (8.1) is a numerical multiple of $y_0(x)$. Moreover, from the explicit formula for Laguerre polynomials [7, vol. II, p. 188, Eq. (7)].

$$x^{(\alpha+1)/2}e^{-2Nx}L_n^\alpha(4Nx)$$

$$=\frac{(-4N)^n}{n!}x^Ne^{-2Nx}[1+O(x^{-1})] \quad x\to\infty \quad (8.13)$$

and from a comparison of (8.12) and (8.13), $L_n^\alpha(4Nx)$ can be expressed in terms of $y_0(x)$. Lastly, we use (2.14) and (8.11) to obtain

$$L_n^\alpha(4Nx)=\frac{(-1)^n}{n!}\left(\frac{2\pi}{-\phi'}\right)^{\frac{1}{2}}2^{-\alpha}N^{n+1/6}x^{-(\alpha+1)/2}$$

$$\times e^{2Nx-N}Ai(-N^{\frac{2}{3}}\phi)\left[1+O\left(\frac{1}{Nx}\right)\right] \quad (8.14)$$

valid, with $N=n+(\alpha+1)/2$ and ϕ given by (8.10), for $n=0, 1, 2, \cdots$ and $0<a<x<\infty$, except that the error term must be modified in the vicinity of zeros of Ai which occur when $0<x<1$.

Since we may choose $0<a<b<1$, the regions of validity of the two asymptotic representations (8.9) and (8.14) between them cover the entire real axis. Moreover, these regions overlap. In the common part, $a<x<b$, both forms are valid, the known asymptotic forms may be used for the Bessel function in (8.9) and the Airy function in (8.14), and the asymptotic equivalence of (8.9) and (8.14) for $a<x<b$ follows from the relation $\alpha+\psi^{\frac{1}{2}}=\pi/2$ which is an immediate consequence of (8.3) and (8.10).

Simpler asymptotic forms may be derived from (8.9) and (8.14). If one is satisfied with an error term $O(N^{-1})$, the gamma function and the factorials may be replaced by their approximations obtained from Stirling's formula. In another direction, restriction of x to narrower intervals will make it possible to use approximations for the Bessel functions in (8.9) or the Airy function in (8.14), and will result in asymptotic representations

involving elementary functions only. Some of the resulting formulas are noted in [6]. Compared with the host of simpler representations, the results developed here have the by no means inconsiderable advantage that two formulas suffice to cover all cases, and that they describe the behaviour of $L_n^{\alpha}(x)$ both for unrestricted n as $x \to 0$ or $x \to \infty$, and for unrestricted x as $n \to \infty$.

BIBLIOGRAPHY

1. H. F. Bohnenblust *et al.*, "Asymptotic solutions of differential equations with turning points, review of the literature," Tech. Rept. 1, Contract Nonr-220(11), Department of Mathematics, California Institute of Technology, 1953.
2. T. M. Cherry, "Uniform asymptotic formulas for functions with transition points," Trans. Am. Math. Soc. **68**, 224–257 (1950).
3. A. Erdélyi, "Asymptotic expansions," Tech. Rept. **3**, Contract Nonr-220(11), Department of Mathematics, California Institute of Technology, 108 (1955), Reprint, Dover (1956).
4. A. Erdélyi, "Differential equations with transition points. I. The first approximation," Tech. Rept. **6**, Contract Nonr-220(11), Department of Mathematics, California Institute of Technology, 22 (1955).
5. A. Erdélyi, "Asymptotic solutions of differential equations with transition points," Proc. Intern. Congr. Mathematicians (1954), **3**, 92–101 (1956).
6. A. Erdélyi, "Asymptotic forms for Laguerre polynomials," Golden Jubilee Volume of the Indian Mathematical Society.
7. A. Erdélyi *et al.*, *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., 1953–55), 3 Vols.
8. A. Erdélyi, M. Kennedy, and J. L. McGregor, "Parabolic cylinder functions of large order," J. Rat. Mech. and Analysis **3**, 459–485 (1954).
9. A. Erdélyi and C. A. Swanson, "Asymptotic forms of Whittaker's confluent hypergeometric functions," Memoirs of the Am. Math. Soc. No. **25** (1957).
10. H. Jeffreys and B. Swirles Jeffreys, *Methods of mathematical physics* (Cambridge University Press, New York, 1956), 3d ed.
11. N. D. Kazarinoff, "Asymptotic theory of second order differential equations with two simple turning points," Arch. Rat. Mech. Anal. **2**, 129–150 (1958).
12. R. E. Langer, "The asymptotic solutions of ordinary linear differential equations of the second order with special reference to the Stokes' phenomenon," Bull. Am. Math. Soc. **40**, 545–582 (1934).
13. R. E. Langer, "The asymptotic solutions of ordinary linear differential equations of the second order, with special reference to turning points," Trans. Am. Math. Soc. **67**, 461–490 (1949).
14. R. E. Langer, "Asymptotic theories for linear ordinary differential equations depending upon a parameter," J. Soc. Indust. Appl. Math. **7**, 298–305 (1959).
15. J. C. P. Miller, "The Airy integral" in B.A.A.S. Math. Tables, Part-Volume B, (Cambridge University Press, New York, 1946).
16. F. W. J. Olver, "The asymptotic solutions of linear differential equations of the second order for large values of a parameter and the asymptotic expansion of Bessel functions of large order," Phil. Trans. Roy. Soc. (London) **A247**, 307–368 (1954).
17. F. W. J. Olver, "The asymptotic solution of linear differential equations of the second order in a domain containing one transition point," Phil. Trans. Roy. Soc. (London) **A249**, 65–97 (1956).
18. F. W. J. Olver, "Uniform asymptotic expansions of solutions of linear second-order differential equations for large values of a parameter," Phil. Trans. Roy. Soc. (London) **A250**, 479–517 (1958).
19. H. Skovgaard, "Asymptotic forms of Hermite polynomials," Technical Report **18**, Contract Nonr-220(11), Department of Mathematics, California Institute of Technology, 20 (1959).
20. C. A. Swanson, "Differential equations with singular points," Tech. Rept. **16**, Contract Nonr-220(11), Department of Mathematics, California Institute of Technology, 24 (1956).
21. G. Szegö, *Orthogonal Polynomials* (American Mathematical Soc. Colloquium Publications, 1959), Vol. 23, 2d ed.
22. F. G. Tricomi, "Sul comportamento asintotico dei polinomi di Laguerre," Ann. Mat. Pura Appl. (4) **28**, 263–289 (1949).

Minimum Property in the Hulthén-Type Variational Methods

TAKASHI OHMURA*
Division of Pure Physics, National Research Council, Ottawa, Canada
(Received October 19, 1959)

The nature of the extremum of the phase shift in the variation principles, which are all based on the Schrödinger differential equation directly, is investigated in detail. If the system has no bound state, the scattering length determined by the original Hulthén method (5, 29) is proved to have the minimum character. The minimum nature is maintained even for the system which has several bound states, if the trial function is taken to be orthogonal to the bound state wave functions. The Kohn method and the second Hulthén method (15) give neither an upper bound nor a lower bound for the phase shift in general. The sufficient condition is, however, obtained for each case, under which the stationary expression for the scattering length has the minimum character. The condition does not hold true for the electron-hydrogen atom scattering for the Kohn method and for the second Hulthén method. It is pointed out that a comparison of approximate values obtained by different methods using the same trial function does not afford any information about the "proper source of errors" of the results.

I. INTRODUCTION

THE Rayleigh-Ritz variation principle provides a very powerful method of calculating the energy value of the ground state in the quantum mechanical systems. A remarkable feature of this method is the minimum nature of the stationary expression for the energy eigenvalues. The calculated values are always larger than the true eigenvalue, so far as the lowest eigenvalue is concerned. A better trial function would give a lower value for the eigenvalue. On the other hand, no such a simple rule is known in the variational methods[1-4] for the scattering problems. The stationary expressions for phase shifts have been believed to give neither minimum nor maximum. One exception of this statement is found in the Schwinger method. Kato[5] has proved that the Schwinger method gives an upper (lower) bound for $k\cot\eta$ (η=phase shift), if the central short-range potential is everywhere attractive (repulsive), and if the absolute value of phase shift is less than π. The Schwinger method, which is based on an inhomogeneous integral equation, can, in principle, be extended to more complex cases, such as the scattering of electrons by hydrogen atoms. However, actual application would be extremely difficult because of the unclosed form of the kernel of the integral equation. On the contrary, the Hulthén type variational methods,[1-3] which are all based on the differential equations, have been (and will be) widely used for many complicated problems. The purpose of the present paper is to investigate the nature of the extremum of the phase shift in the Hulthén type variation principles.

* National Research Council Postdoctorate Fellow, on leave of absence from Department of Physics, University of Tokyo, Japan.

[1] L. Hulthén, K. Fysiogr. Sällsk. Lund Föohandl. **14**, 21 (1944).

[2] L. Hulthén, Arkiv. Mat. Astron. Fysik **35A**, 25 (1948); T. Kato, Phys. Rev. **80**, 475 (1950).

[3] W. Kohn, Phys. Rev. **74**, 1763 (1948).

[4] J. Schwinger, Phys. Rev. **72**, 742 (1947); **78**, 135 (1950); J. M. Blatt and J. D. Jackson, Phys. Rev. **76**, 18 (1949).

[5] T. Kato, Progr. Theoret. Phys. **6**, 394 (1951).

II. ORIGINAL HULTHÉN METHOD. ONE-DIMENSIONAL PROBLEM

The simplest example in scattering problems, s wave scattering by a central potential, is considered in this section. Let $u(r)$ be the (regular, i.e., $u(0)=0$) solution of the equation,

$$Lu(r)\equiv(d^2/dr^2+W(r)+k^2)u(r)=0, \quad r\geqq 0, \tag{1}$$

where $W(r)$ is a short range potential function. The potential is attrative (repulsive) when $W(r)>0$ ($W(r)<0$). The normalization of the wave function can be fixed arbitrarily. For convenience in later discussions, the normalization is specified by the following asymptotic form,

$$u(r)\rightarrow \cos kr+\lambda\sin kr, \quad \lambda=\cot\eta, \quad (r\rightarrow\infty). \tag{2}$$

We seek an accurate value for λ by the variational method. The trial function $u_t(r)$ has an approximate asymptotic form, such as,

$$u_t(r)\rightarrow \cos kr+(\lambda+\Delta\lambda_t)\sin kr, \quad \lambda+\Delta\lambda_t=\cot\eta_t, \quad (r\rightarrow\infty).$$

Then we put $u(r)$ and $u_t(r)$ as follows,

$$u(r)=y(r)+\cos kr+\lambda\sin kr, \quad y(0)=-1, \tag{3}$$

$$u_t(r)=u(r)+\Delta y(r)+\Delta\lambda_t\sin kr, \quad \Delta y(0)=0, \tag{4}$$

where $y(r)$ and $\Delta y(r)$ will tend to zero rapidly as r increases. The original Hulthén method imposes the following condition on the trial function:

$$\int_0^\infty u_t L u_t dr=0. \tag{5}$$

Equation (5) guarantees the extremum property of the approximate phase shift η_t as first shown by Hulthén, that is, $\Delta\lambda=\cot\eta_t-\cot\eta$ is a quantity of $O[(\Delta y)^2]$. In order to know the nature of this extremum

property more clearly, (4) is substituted into Eq. (5).

$$\int_0^\infty u_t L u_t dr = (\Delta\lambda_t)^2 \int W \sin^2 kr dr + \Delta\lambda_t \int W \sin kr(u+2\Delta y)dr + \int \Delta y L \Delta y dr = 0. \quad (6)$$

The error $\Delta\lambda$ of the Hulthén method is calculated by the formula (6). By using the relation,

$$\int W(r) \sin kr u(r) dr = k, \quad (7)$$

we derive the following result:

$$\Delta(k\cot\eta) = -\left\{\int \Delta y L \Delta y dr + (k\Delta\lambda)^2 \times \int k^{-2} W \sin^2 kr dr\right\}(1+2\epsilon)^{-1}, \quad (8)$$

where $\epsilon k \equiv \int W(r) \sin kr \Delta y(r) dr$ and $\Delta(k\cot\eta) \equiv k\Delta\lambda$. $|\epsilon|$ should be a small quantity in the practical calculation. The proof of (7) is easy. The solution of (1) with the normalization specified by (2) satisfies the integral equation,

$$u(r) = \lambda \sin kr + k^{-1} \int_0^\infty \sin kr_< \cos kr_> W(r') u(r') dr'. \quad (9)$$

The symbol $<$ represents the smaller one of r and r', and $>$ the larger one of r and r'. By comparing (2) and (9) for large r, we get the relation (7). It is clear from (8) that the extremal nature at the vicinity of $u_t = u$ is determined by the sign of the main term: $\int_0^\infty \Delta y L \Delta y dr$, since the factor $(k\Delta\lambda)^2$ is a quantity of order $(\Delta y)^4$. By defining $\bar{H}$ as an "energy" expectation value of the function $\Delta y(r)$,

$$\bar{H} \equiv -\int_0^\infty \Delta y(r)\left(\frac{d^2\Delta y(r)}{dr^2} + W(r)\Delta y(r)\right)dr \times \left\{\int_0^\infty (\Delta y(r))^2 dr\right\}^{-1}, \quad (10)$$

we obtain an expression for $\Delta(k\cot\eta) = k\cot\eta_t - k\cot\eta$ correct up to the second order of Δy,

$$\Delta(k\cot\eta) = -\int_0^\infty \Delta y L \Delta y dr = (\bar{H} - k^2)\int_0^\infty (\Delta y)^2 dr. \quad (11)$$

Since $\Delta y(0) = 0$, and $\Delta y(r)$ tends to zero rapidly at infinity, therefore Δy certainly belongs to a Hilbert space. If the system has no negative eigenvalue for the hermitian operator $-(d^2/dr^2 + W(r))$, $\bar{H}$ is necessarily positive. Thus the right-hand side of (11) is always positive at zero incident energy ($k^2 = 0$), if there is no bound state. $k\cot\eta$ at zero energy is related to the scattering length a by the relation,

$$(k\cot\eta)_{k=0} = -a^{-1}, \quad (k\cot\eta_t)_{k=0} = -a_t^{-1}.$$

Then (11) tells us the maximum nature of $(1/a_t)$,

$$\Delta\left(\frac{1}{a}\right) \equiv \frac{1}{a_t} - \frac{1}{a} = \frac{a - a_t}{a a_t} = -\bar{H}\int_0^\infty (\Delta y)^2 dr \leqq 0. \quad (12)$$

Accordingly, the original Hulthén method gives an upper bound for the scattering length except for the critical case where aa_t is negative. When there are discrete eigenstates, which correspond to bound states, the minimum property of the scattering length does not hold any more, because the "energy" expectation value $\bar{H}$ can be negative as well. In order to avoid this undesirable situation, the trial function must be taken to be orthogonal to the bound state wave functions. The proof will be given in Appendix I. We do not discuss this further, since this additional condition on the trial function may cause a troublesome calculation in actual application.

Numerical Example

To illustrate the results obtained in the preceding arguments, we will give numerical examples, assuming the square well for $W(r)$,

$$W(r) = \begin{cases} V, & 0 \leqq r \leqq 1, \\ 0, & 1 < r. \end{cases} \quad (13)$$

The trial function is of the simplest form,

$$u_t = \begin{cases} (2 - a_t^{-1})r - r^2, & 0 \leqq r \leqq 1, \\ 1 - a_t^{-1} r, & 1 < r. \end{cases} \quad (14)$$

The original Hulthén method determines a_t by the condition,

$$\int_0^\infty u_t L u_t dr = (30a_t^2)^{-1} \times \{10V - (25V - 30)a_t + (16V - 40)a_t^2\} = 0.$$

The present example, can, of course, be solved exactly. The exact scattering length is given by

$$a = \begin{cases} 1 - \tan(V)^{\frac{1}{2}}/[(V)^{\frac{1}{2}}], & \text{if} \quad V > 0, \\ 1 - \tan(-V)^{\frac{1}{2}}/[(-V)^{\frac{1}{2}}], & \text{if} \quad V < 0. \end{cases}$$

There is no bound state in the system only if $V < \pi^2/4 = 2.4674$. The values of the scattering length determined by the original Hulthén variational method are tabulated in Table I together with the exact values. It is seen from Table I that the approximate values are always larger than the exact values when the system has no bound state, that is, $V < \pi^2/4$. The same is true for

$V=2.5\sim4$. The latter fact is due to the positive nature of $\bar{H}$ in the present example even for $V=4$, although the positive nature of $\bar{H}$ is not guaranteed by the theorem which is applicable for all trial functions. The approximate values give lower bounds between $V=\pi^2/4$ and 2.5. This is not due to the existence of the bound state but the different signs between exact and approximate values. Thus the lower bound nature between $V=\pi^2/4$ and 2.5 essentially originates from the singular nature of the scattering length at $V=\pi^2/4$. At the vicinity of this singularity, the statement of the theorem has to be replaced by, "the inverse of the scattering length calculated by an approximate function gives a lower bound for the true value," as the inequality (12) means properly.

TABLE I. Scattering length calculated by the original Hulthén variational method is compared with the exact ones. The Hulthén method gives an upper bound for the scattering length if the system has no bound state, that is, $V<\pi^2/4$.

V	Original Hulthén	Exact value
-1	0.24091	0.23841
ϵ[a]	$-(\epsilon/3)-0.1296\epsilon^2$	$-(\epsilon/3)-0.1333\epsilon^2$
1	-0.5497	-0.5574
2.25	-7.330	-8.401
↓	↓	↓
$\pi^2/4-0$	-61.516	$-\infty$
$\pi^2/4+0$	-61.516	$+\infty$
↓	↓	↓
$2.5-0$	$-\infty$	62.161
$2.5+0$	$+\infty$	62.161
↓	↓	↓
4	2.1321	2.0925

[a] $|\epsilon|\ll1$.

III. SECOND HULTHÉN AND KOHN METHODS. ONE-DIMENSIONAL PROBLEM

The ambiguity and inconvenience appearing in the actual calculation connected with the use of the quadratic equation $\int u_t L u_t dr=0$ in the original Hulthén method were removed independently by Hulthén[2] and Kohn.[3] Although the principal ideas of these authors for the improvement of the method are almost the same, the difference in the normalization of the trial wave function affects the results in approximate calculations. We refer to these two methods as the second Hulthén and the Kohn method. The extremum nature in those methods will be examined in the present section.

In the second Hulthén method, the way of normalizing the wave function is given by (3) and (4). The stationary expression of $k\cot\eta$ is given by

$$k\cot\eta_H=k\lambda_H=k\lambda_t-\int_0^\infty u_t(r)Lu_t(r)dr. \qquad (15)$$

The error of the stationary value $(k\lambda_H)$ for $k\lambda=k\cot\eta$, and the error of $k\lambda_t=k\cot\eta_t$ in the trial wave function are defined by $k\Delta\lambda_H=k\lambda_H-k\lambda$, and $k\Delta\lambda_t=k\lambda_t-k\lambda$, respectively. It should be noted in this case that $k\Delta\lambda_t$ is a quantity of $O(\Delta y)$ rather than $O((\Delta y)^2)$. $k\Delta\lambda_H$ can readily be rewritten in the following way:

$$\Delta(k\cot\eta_H)\equiv k\Delta\lambda_H=k\Delta\lambda_t-\int_0^\infty u_tLu_tdr$$
$$=-\int_0^\infty \Delta yL\Delta ydr-2(k\Delta\lambda_t)\int_0^\infty k^{-1}W\sin kr\Delta ydr$$
$$-(k\Delta\lambda_t)^2\int_0^\infty k^{-2}W\sin^2krdr. \qquad (16)$$

Use has been made of (15) and (6) in the course of derivation of (16). The sign of the first term on the right hand side of the Eq. (16) is positive at low energies if the system has no bound state. The sign of the second term, however, can be negative as well as positive, if the sign of $k\Delta\lambda_t$ is independent of the sign of the integral $\int k^{-1}W\sin kr\Delta ydr$. The sign of the third term depends on the sign of the potential $W(r)$. Thus $\Delta(k\cot\eta_H)$ can be positive or negative, and on that account the stationary behavior of the phase shift reached for $u_t=u$ in the second Hulthén method has no definite maximum and minimum, when we just insert some prepared function u_t into the expression (15). However, according to the revised procedure by Hulthén, the sign of $\Delta\lambda_t$ is not completely independent of the sign of $\int k^{-1}W\sin kr\Delta ydr$. Let Δy be a given function. On following Hulthén,[6] we determine $k\Delta\lambda_t$ by

$$\frac{\partial(k\Delta\lambda_H)}{\partial(k\Delta\lambda_t)}=-2\left(\int k^{-1}W\sin kr\Delta ydr+(k\Delta\lambda_t)\int k^{-2}W\sin^2krdr\right)=0. \qquad (17)$$

We notice that the Eq. (17) is equivalent to the condition imposed on u_t,

$$\int k^{-1}u_tW\sin krdr=1, \qquad (18)$$

which means that

$$\int k^{-1}\Delta uW\sin krdr=0.$$

On substituting the value of $k\Delta\lambda_t$ determined by (17), we get an expression for $k\Delta\lambda_H$,

$$k\Delta\lambda_H=-\int\Delta yL\Delta ydr+\left(\int k^{-1}W\sin kr\Delta ydr\right)^2\Big/\left(\int k^{-2}W\sin^2krdr\right). \qquad (19)$$

[6] In ordinary notations, Hulthén proposes the equation,

$$\partial\int u_tLu_tdr/\partial(k\lambda_t)=1,$$

which is equivalent to (17).

We can see from (19) that $k\lambda_H$ has minimum character at zero energy provided that there is no bound state in the system and that $\int r^2 W dr > 0$. In the case of $\int k^{-2} W \sin^2 kr dr < 0$, the extremum nature of the second Hulthén method is that of the saddle point.

The situation is somewhat similar in the Kohn method. The stationary quantity in the Kohn method is $\tan\eta$ rather than $\cot\eta$. The normalization of the wave function is different from the one in the second Hulthén method, and is specified by the unit amplitude of the $\sin kr$ term,

$$u_k \longrightarrow \lambda_t^{-1}(\cos kr + \lambda_t \sin kr), \quad \lambda_t = \cot\eta_t, \quad (r \longrightarrow \infty).$$

Then the Kohn trial function u_k may be set, correct to the first order of Δ,

$$u_k = \lambda^{-1}(u + \Delta y + \Delta\lambda_t \sin kr)\{1 - (\Delta\lambda_t/\lambda)\},$$

where $u = y + \cos kr + \lambda \sin kr$, and $\Delta\lambda_t \equiv \lambda_t - \lambda$. The stationary expression for $\tan\eta$ in the Kohn method is given by

$$\lambda_k^{-1} = \lambda_t^{-1} + k^{-1}\int_0^\infty u_k(r) L u_k(r) dr. \tag{20}$$

The error of λ_k is calculated in a similar way as in the Hulthén method,

$$(k\Delta\lambda_k) = (k\Delta\lambda_t)^2/(k\lambda) - \left\{\int_0^\infty \Delta y L \Delta y dr + 2(k\Delta\lambda_t)\int_0^\infty k^{-1} W \sin kr \Delta y dr + (k\Delta\lambda_t)^2 \times \int_0^\infty k^{-2} W \sin^2 kr dr\right\} + O(\Delta^3). \tag{21}$$

The sign of the first term depend on the sign of $k\lambda$. The following three terms have the same forms and same signs as in Eq. (16) for the second Hulthén method. Hence the Kohn expression, again, represents the saddle point character of extremum, when we just insert a prepared function u_t into the expression (20).

However, if we determine λ_t according to the method proposed by Kohn, which is written in terms of $\Delta\lambda_k$ and $\Delta\lambda_t$ as

$$\partial(k\Delta\lambda_k)/\partial(k\Delta\lambda_t) = 0, \tag{22}$$

the resulting expression for $\Delta(k\cot\eta_k)$ is again similar to (19), namely

$$k\Delta\lambda_k = -\int \Delta y L \Delta y dr + \left(\int k^{-1} W \sin kr \Delta y dr\right)^2 \Big/ \left(\int k^{-2} W \sin^2 kr dr - (k\lambda)^{-1}\right) + O[(\Delta y)^3]. \tag{23}$$

From (23) we recognize that $k\lambda_k$ has minimum character at zero energy, if there is no bound state and if $\int r^2 W dr + a > 0$. (a; scattering length).

We shall summarize the conditions under which the second Hulthén and the Kohn procedures give upper bounds for the scattering length;

no bound state, and

$$\int_0^\infty r^2 W dr > 0, \qquad \text{(Hulthén)} \tag{24}$$

no bound state, and

$$\int_0^\infty r^2 W dr + a > 0, \qquad \text{(Kohn)} \tag{25}$$

although we must pay special attention to the location of the singular point of the scattering length. For energies near the singular point, where the absolute value of the scattering length may tend to infinity, the theory guarantee the lower bound of the inverse of the scattering length under the condition (24), or (25).

If we take the unperturbed function as a trial function, the Kohn method gives the (first) Born approximation result. Our discussion, however, does not apply to this case, because we always assume that the trial function is close to the actual function. With the Born trial function, $\Delta\lambda_t$ is no longer small. It is $\pm\infty$ indeed! Nevertheless, it has been proved[7] that the scattering length calculated by the Born approximation a_B gives an upper bound, if there is no bound state.

$$a_B \geqq a \geqq c a_B,$$

where c is a constant which relates to the convergence radius of the Born expansion.

It may be worthwhile to remark that the error coming from the term $-\int \Delta y L \Delta y dr$ is commonly contained in three different methods treated in this paper. Consequently, a comparison of the three results would not provide any idea about the magnitude of $\int \Delta y L \Delta y dr$. On the other hand, the errors due to terms which have the factors $\Delta\lambda_t$ and $(\Delta\lambda_t)^2$ in (16) and (21) can be estimated roughly by varying the value of λ_t a little bit and by noting that $\lambda_t - \lambda_k \fallingdotseq \lambda_t - \lambda_H \fallingdotseq \Delta\lambda_t$. The term $\int \Delta y L \Delta y dr$ should be considered as the proper source of errors.

In the following we shall bring out the merit of the original Hulthén method for the calculation of the scattering length. Actual calculational effort involved in the original method is by no means enormous when compared with the revised procedures, so far as the linear trial function is used. The situation may be explained below. Let the trial function be set in the form,

$$u_t(r) = \sum_{i=1}^{n} a_i f_i(r) + \cos kr + \lambda_t \sin kr. \tag{26}$$

[7] T. Ohmura (formerly Kikuta), Progr. Theoret. Phys. 12, 225 (1954); See especially p. 230.

This leads to the simultaneous linear equations for a_i and λ_t with the calculated constants c_{ij}, b_j, and d_j.

$$\sum_{i=1}^{n} c_{ij}a_i=b_j+\lambda_t d_j. \quad (j=1, \cdots n).$$

Then the unknown variable a_i is expressed by the form,

$$a_i=B_i+D_i\lambda_t.$$

On inserting a_t in Eq. (26), we get, with some known simple constants A_{ij},

$$\int u_t L u_t dr=(A_{11}-\sum_{i=1}^{n} b_iB_i)$$
$$+[A_{12}-\sum_{i=1}^{n}(d_iB_i+b_iD_i)]\lambda_t$$
$$+(A_{22}-\sum_{i=1}^{n} d_iD_i)\lambda_t^2=0. \quad (27)$$

All the coeff c_{ij} do not appear any more in the formula (27). The variational calculation of the original Hulthén method will be easily completed by finding the value of λ_t from (27). On the other hand, the second Hulthén and the Kohn method lead to a similar type of equation,

$$\sum_{i=1}^{n+1} c_{ij}a_i=c_j, \quad (j=1, \cdots n, n+1).$$

Thus the main trouble in the original method over the revised ones is to solve two inhomogeneous linear equation of order n instead of one equation of order $n+1$. The electronic computers will eliminate this trouble. The calculating effort of the matrix element is just the same for the three methods. On that account, if there is no bound state, the original Hulthén method is most superior for the calculation of the scattering length over other methods due to the following two reasons. If the second Hulthén (or Kohn) method is not guaranteed to give an upper bound for a, namely, when the condition (24) (or (25)) is not satisfied, only the original Hulthén method gives an upper bound for a. Even when the second Hulthén (or Kohn) method is guaranteed to give an upper bound for a, the original Hulthén method will produce the best result or at least the same order of accuracy as the other methods. This is understandable if we compare Eq. (11) with Eq. (19) and Eq. (23), and assume that the Δy contained in (11), (19), and (23) are all the same,

$$k\Delta\lambda_{\text{original}}=k\Delta\lambda_H-(\Delta^2/\int Wr^2dr)$$
$$=k\Delta\lambda_k-\Delta^2/(\int Wr^2dr+a)>0,$$

where $\Delta\equiv\int k^{-1}W \sin kr\Delta y dr$.

Numerical Example

We shall take the same example as given in II in order to understand the general situation of the subject we already discussed for the one dimensional problem. The potential (13) and the trial function (14) are the same. The second Hulthén and the Kohn scattering lengths are calculated by (15) and (20). We shall write the results only for completeness.

$$a_H=(V/80+7/6-3/V)^{-1}, \quad a_B=-V/3,$$
$$a_k=-(V/3)+[125V^2/192(2V-5)].$$

The calculated scattering lengths for various depths of the potential in various methods are tabulated in Table II, where the values of $\int Wr^2dr$ and $\int Wr^2dr+a$ are added. The latter quantities appear in the conditions (24) and (25). We can see from Table II that the second Hulthén results are not superior to the original Hulthén when the former is assured to give an upper bound for a. The Born approximation is not good, though it gives an upper bound. All values of $\int Wr^2dr+a$ have the negative sign, so we cannot say the Kohn method would give an upper bound. We also understand from Table II that $-\int \Delta y L\Delta y dr$ may be the main source of error, because the approximate values in three methods are much the same, but the exact values are definitely different from the approximate values.

IV. ELECTRON-HYDROGEN S WAVE SCATTERING

The arguments developed in II and III can be extended without difficulty to more complicated cases. In this section, we shall take up a quantum mechanical three-body problem, the S wave scattering of electrons by hydrogen atoms. This problem was systematically

TABLE II. The scattering lengths calculated by various variational methods are compared to each other and with the exact ones.

V	Exact value	Original Hulthén	Second Hulthén	Kohn	First Born	$\int Wr^2dr$	$a+\int Wr^2dr$
-1	0.23841	0.24091	0.24072[a]	0.24033[a]	0.3333	-0.3333	-0.0949
ϵ	$\frac{\epsilon}{3}-0.1333\epsilon^2$	$\frac{\epsilon}{3}-0.1296\epsilon^2$	$\frac{\epsilon}{3}-0.1296\epsilon^2$	$\frac{\epsilon}{3}-0.1302\epsilon^2$ [a]	$\frac{\epsilon}{3}$	$\frac{\epsilon}{3}$	$\frac{2\epsilon^2}{15}$
1	-0.5574	-0.5497	-0.5492	-0.5503[a]	-0.3333	0.3333	-0.2241
2.25	-8.401	-7.330	-7.218	-7.342[a]	-0.7500	0.7500	-7.651

[a] Indicates that these values are not theoretically assured to give an upper bound or a lower bound for the scattering length.

studied by Massey and Moiseiwitch[8] among others,[9] using the Hulthén and the Kohn methods. The so-called "effective range approximation" has been shown[10] to produce the Massey-Moiseiwitch results reasonably well even for high energies. In that approximation, the scattering length and the effective range are the only quantities which describe the scattering. In this consideration, we shall confine our attention to the case of zero incident electron energy. The effective range can be calculated if the wave function at zero energy is known.

The wave function u of the system having two electrons and one infinitely heavy proton can be written in the following form,

$$u_\pm(r_1,r_2,r_{12}) = y_\pm(r_1,r_2,r_{12})+e^{-r_1}/2\pi r_2(\cos kr_2+\lambda_\pm \sin kr_2) \pm e^{-r_2}/2\pi r_1(\cos kr_1+\lambda_\pm \sin kr_1), \quad (28)$$

where the atomic unit is used, and y tends to zero rapidly as r_1 (or r_2) increases. λ_+ and λ_- are cotangent of the s phase shifts for the singlet and the triplet states. The wave function u satisfies the fundamental equation of the system,

$$Lu\equiv\left(\tfrac{1}{2}\Delta_1+\tfrac{1}{2}\Delta_2+\frac{1}{r_1}+\frac{1}{r_2}-\frac{1}{r_{12}}+E\right)u=0.$$

The trial function u_t differs from u by Δy and $\Delta\lambda_t$ term as follows,

$$u_t=u+\Delta y+\left(\frac{e^{-r_1}}{2\pi}\frac{\sin kr_2}{r_2}\pm\frac{e^{-r_2}}{2\pi}\frac{\sin kr_1}{r_1}\right)\Delta\lambda_t.$$

The original Hulthén method imposes the condition,

$$\int u_t L u_t d\tau_1 d\tau_2=0, \quad (29)$$

on the trial function u_t. Let us first estimate the error of $k\lambda_t$ in the original Hulthén method. With arguments similar to those in II, we readily obtain the following formula,

$$\Delta(k\cot\eta)=k\Delta\lambda_t = -\int \Delta y L\Delta y d\tau_1 d\tau_2 \Big/ \left(k^{-1}\int u L v d\tau_1 d\tau_2\right), \quad (30)$$

where v is defined by

$$v=\frac{e^{-r_1}}{2\pi}\frac{\sin kr_2}{r_2}\pm\frac{e^{-r_2}}{2\pi}\frac{\sin kr_1}{r_1}. \quad (30a)$$

The denominator of the right hand side of (30) can be shown to be unity. The proof will be given in Appendix II. Equation (30) may be brought into a form very similar to Eq. (11),

$$\Delta(k\cot\eta)=\left(\bar{W}-\frac{k^2}{2}\right)\int(\Delta y)^2 d\tau_1 d\tau_2,$$

if we define that $\bar{W}-\frac{1}{2}$ is the expectation value of the energy for a state Δy, $-\frac{1}{2}$ being the ground state energy of hydrogen atoms. The singlet system of the electron-hydrogen atom has one bound state with very small "binding energy" ($\sim$0.028). Therefore, the value of $\bar{W}$ can be negative, only if the unnormalized "wave function" Δy is so close to the bound state that the usual Ritz variational expression gives a positive binding energy. $\bar{W}$ is always positive for the triplet state. A very similar consideration as in II leads to the following conclusion: the scattering length computed by the original Hulthén method gives an upper bound in the triplet state of the electron-hydrogen atom system. In the singlet state, it will be only presumed that the same method would also give an upper bound for the scattering length because of smallness of the electron affinity.

Consideration about the errors in the second Hulthén and the Kohn methods can be done with closely parallel arguments to III. We shall just write down the corresponding equations to (19) and (23),

$$\Delta(k\cot\eta_H)=-\int\Delta y L\Delta y d\tau_1 d\tau_2 + \frac{\left(\int k^{-1}\Delta y L v d\tau_1 d\tau_2\right)^2}{\int k^{-2} v L v d\tau_1 d\tau_2}, \quad (31)$$

$$\Delta(k\cot\eta_k)=-\int\Delta y L\Delta y d\tau_1 d\tau_2 + \frac{\left(\int k^{-1}\Delta y L v d\tau_1 d\tau_2\right)^2}{\int k^{-2} v L v d\tau_1 d\tau_2-(k\cot\eta)^{-1}}. \quad (32)$$

The denominator of the second term of (31) has been calculated at $k^2=0$. This is -5 for the triplet state and 7 for the singlet state. The scattering length, which equals to $-(k\cot\eta)^{-1}$ at $k=0$, is roughly estimated[9,10] to be 2.33 for the triplet state, and 6$\sim$7 for the singlet state. Hence, the second term for each Eq. (31) or (32) is negative in the triplet state, though the first term is positive. While the second term for each equation is positive in the singlet state, the first term has no definite sign because of the existence of one bound

[8] H. S. W. Massey and B. Moiseiwitch, Proc. Roy. Soc. (London) **A205**, 483 (1951).

[9] M. J. Seaton, Proc. Roy. Soc. (London) **A241**, 522 (1957); S. Borowitz and H. Greenberg, Phys. Rev. **108**, 716 (1957); Bransden, Dalgarno, John, and Seaton, Proc. Phys. Soc. (London) **71**, 877 (1958).

[10] Ohmura, Hara, and Yamanouchi, Progr. Theoret. Phys. **20**, 82 (1958); **22**, 152 (1959).

state. Thus the second Hulthén and the Kohn method give neither an upper bound nor a lower bound for the scattering lengths of electron-hydrogen scattering.

V. SUMMARY OF THE RESULTS

The extremum nature of the ordinarily used variational methods in scattering problem is investigated. In general, these methods give neither an upper bound nor a lower bound for the phase shifts. In the scattering states corresponding to the lowest energy state, namely, at zero incident energy of particles if the system has no "bound state," however, we can find some cases where the maximum or minimum nature of the phase shift is guaranteed.

It has been proved in II that the original Hulthén method guarantees the minimum nature of the scattering length, if there is no bound state. In other words, the Hulthén (original) method gives an upper bound for the scattering length when we take an approximate trial function which is close to the exact function and will converge to the exact value from above as the trial function is improved. By applying this theorem to various problems, we understand that the original Hulthén method gives an upper bound for the scattering length in proton-proton collision, electron scattering by hydrogen atoms in the triplet state, electron scattering by helium atoms, nucleon-deuteron scattering in the quartet state, nucleon-helium nucleus scattering, etc. In order to preserve the minimum nature of the scattering length for the system which has bound states, such as the singlet state of the negative hydrogen ion, the trial function must be taken to be orthogonal to the bound state wave functions. This additional condition may cause complicated calculation.

The sufficient condition (24), (25) has been derived in III, under which the stationary expression for the scattering length in the second Hulthén method and in the Kohn method is minimum. It is shown in IV, however, that this condition is not satisfied in the case of electron-hydrogen scattering.

It is customary to compare the results obtained by different methods but using the same trial function, in order to get a rough idea of the accuracy of the results. Concerning this matter, it is pointed out in III that a comparison of three methods considered in the present paper does not provide any information about the term: $-\int_0^\infty \Delta y L \Delta y dr$.

The original Hulthén method, which is not very complicated in actual application when compared with the Kohn method or the revised Hulthén method, is considered to be the best one for the scattering length calculation, because it has the same nature of extremum as the usual Ritz method does.

Note added in Proof. If the phase shift approaches zero at the zero energy limit, $\beta_{\pi/2}$ defined in footnote 5 becomes infinity at zero energy. Consequently, the formula (23) in footnote 5 tells us that the Kohn method, as well as the original Hulthén method, give an upper bound for the scattering length in the one-dimensional problem, if the phase shift vanishes at zero energy. A special consideration is needed concerning an associated eigenvalue problem, if this (Kato's) method is applied to more complicated cases.

ACKNOWLEDGMENT

The author wishes to express his sincere thanks to Dr. Ta-You Wu for hospitality at the National Research Council, and for giving many useful advices on preparation of the manuscript.

APPENDIX I. BOUND STATE CONTRIBUTION TO $\int \Delta y L \Delta y d\tau$

It will be shown here that the minimum nature of the scattering length in the original Hulthén method is maintained even for the system which has several bound states, if the trial function is set to be orthogonal to the bound state wave functions.

We shall consider the one dimensional problem as treated in II and III. Let φ_i be the normalized wave function of ith bound state with the energy E_i. The assumed conditions imposed on the trial function u_t are

$$\int_0^\infty u_t \varphi_i dr = 0. \tag{A.1}$$

The correct wave function $u(r)$ is, of course, orthogonal to φ_i.

$$\int_0^\infty u \varphi_i dr = 0. \tag{A.2}$$

By using Eq. (4), we get from (A.1) and (A.2) the following relation,

$$\int_0^\infty \Delta y(r) \varphi_i(r) dr = -\Delta\lambda_t \int_0^\infty \sin kr \varphi_i(r) dr. \tag{A.3}$$

It should be noticed that the magnitude of the right hand side of (A.3) is of order $(\Delta y)^2$. The error in the original Hulthén method is given by (11). Let us rewrite (11) in order to understand the role of the bound states more clearly. $\Delta y(r)$ may be expanded in φ_i and $\varphi(E)$, the latter representing the correct wave function normalized to energy E,

$$\Delta y = \sum f_j \varphi_j + \int_0^\infty f(E) \varphi(E) dE. \tag{A.4}$$

By substituting (A.4) to (11), we get,

$$-\int_0^\infty \Delta y L \Delta y dr \underset{(k=0)}{=} \sum E_i f_i^2 + \int_0^\infty E f^2(E) dE, \tag{A.5}$$

where all E_i are negative. If Δy in (A.3) is replaced by (A.4), all terms disappear except one term by virtue

of the orthogonality,

$$f_i = -\Delta\lambda_i \int_0^\infty \sin kr\,\varphi_i(r)dr. \tag{A.6}$$

Thus the coefficients f_i turn out to be quantities of second order in Δy. This means that the sum appearing in the right hand side of (A.5) is of the fourth order in Δy. Therefore, the sign of (A.5) is determined by the second term which is a positive quantity, when Δy tends to zero. The scattering length is thus a minimum.

APPENDIX II. PROOF OF $\int uLvd\tau = k$

We shall prove that the denominator of Eq. (30) in its right-hand side term is unity.

By following Mott and Massey,[11] the (unsymmetrized) wave function of one electron and one hydrogen atom may be expanded in terms of the normalized wave functions of the hydrogen atom ψ_n with coeff F_n (or G_n).

$$\begin{aligned}\varphi(1,2) &= \left(\Sigma + \int\right) F_n(1)\psi_n(2), \\ &= \left(\Sigma + \int\right) G_n(2)\psi_n(1),\end{aligned} \tag{A.7}$$

where $\mathbf{r}_1$ represents the coordinate of the incident electron with incident energy $k^2/2$. The function $F_0(1)$, which is the coefficient of the ground state function $\psi_0(2)$, is expressed by $\varphi(1,2)$, as

$$F_0(\mathbf{r}_1') = e^{ikz_1'} + \int \frac{e^{ik|\mathbf{r}_1-\mathbf{r}_1'|}}{4\pi|\mathbf{r}_1-\mathbf{r}_1'|}\left(\frac{2}{r_1} - \frac{2}{r_{12}}\right) \times \psi_0(2)\varphi(1,2)d\tau_1 d\tau_2. \tag{A.8}$$

The corresponding expression for G_0 has no incoming wave,

$$G_0(\mathbf{r}_1') = \int \frac{e^{ik|\mathbf{r}_1-\mathbf{r}_1'|}}{4\pi|\mathbf{r}_1-\mathbf{r}_1'|}\left(\frac{2}{r_1} - \frac{2}{r_{12}}\right) \times \psi_0(2)\varphi(2,1)d\tau_1 d\tau_2. \tag{A.9}$$

For the S wave part in (A.8) and (A.9), and for large r_1', F_0 and G_0 have the following asymptotic behavior,

$$F_0(1') \to \frac{\sin kr_1'}{kr_1'} + \frac{(f_+ + f_-)}{2}\frac{e^{ikr_1'}}{kr_1'}, \tag{A.10}$$

$$G_0(1') \to \frac{(f_+ - f_-)}{2}\frac{e^{ikr_1'}}{kr_1'}, \tag{A.11}$$

where

$$f_\pm \equiv \int \frac{\sin kr_1}{2\pi r_1}\left(\frac{1}{r_1} - \frac{1}{r_{12}}\right)\psi_0(2) \times \{\varphi(1,2) \pm \varphi(2,1)\}d\tau_1 d\tau_2. \tag{A.12}$$

Hence, the asymptotic form of the symmetrized wave function $\varphi(1,2) \pm \varphi(2,1)$ is, when only elastic scattering is possible,

$$\begin{aligned}\varphi(1,2) \pm \varphi(2,1) &\to \psi_0(2)\{(\sin kr_1/kr_1) \\ &\quad + f_\pm e^{ikr_1}/kr_1\}, \quad (r_1 \to \infty), \\ &\to \pm\psi_0(1)\{(\sin kr_2/kr_2) \\ &\quad + f_\pm e^{ikr_2}/kr_2\}, \quad (r_2 \to \infty).\end{aligned} \tag{A.13}$$

It is clear from (A.13) that

$$f_\pm = \sin\eta_\pm e^{i\eta_\pm}. \tag{A.14}$$

On comparing (28) and (A.13, 14), we can set

$$\varphi(1,2) \pm \varphi(2,1) = \pm(2\pi \sin\eta_\pm e^{i\eta_\pm}/(\pi)^{\frac{1}{2}}k)u_\pm(r_1,r_2,r_{12}). \tag{A.15}$$

By substituting (A.15) into (A.12), and making use of (A.14), we obtain the result,

$$\int \frac{\sin kr_1}{2\pi r_1}\left(\frac{1}{r_1} - \frac{1}{r_{12}}\right)e^{-r_2}u_\pm(r_1,r_2,r_{12})d\tau_1 d\tau_2 = \pm\frac{k}{2}. \tag{A.16}$$

Finally, we calculate Lv by the definition of v, (30a),

$$Lv = \left(\frac{1}{r_2} - \frac{1}{r_{12}}\right)\frac{e^{-r_1}\sin kr_2}{2\pi r_2} \pm \left(\frac{1}{r_1} - \frac{1}{r_{12}}\right)\frac{e^{-r_2}\sin kr_1}{2\pi r_1}, \tag{A.17}$$

and by comparing (A.16) and (A.17), get

$$\int uLvd\tau_1 d\tau_2 = k.$$

[11] N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions* (Clarendon Press, Oxford, England, 1949), 2nd edition, p. 136.

Stationary Expression for Effective Range

TAKASHI OHMURA*
Pure Physics Division, National Research Council, Ottawa, Canada
(Received October 23, 1959)

A variation principle is presented for the effective range of s wave scattering. This principle applies both to the bound state and to the scattering state. It is proved that some of the stationary expressions for the effective range has the minimum value for the exact solution, thus the calculated value by using an approximate wave function gives an upper bound for the effective range.

A modified method is also proposed which may simplify the calculation in an actual application. The stationary expression gives, however, neither an upper bound nor a lower bound.

I. INTRODUCTION

THE so-called effective range theory was put forward originally by Schwinger[1] with reference to nucleon-nucleon scattering problems, and afterwards derived by Bethe[2] in an illuminating manner. According to the theory, the scattering in a low energy region is described by two parameters, the scattering length a, and the effective range, r_0 or ρ. Hence, all potentials are equivalent[3] as far as the low energy nucleon-nucleon scattering is concerned, provided they correspond to the a and r_0. The course of analyzing the problems is reversed in the case of atomic scatterings. Here, we already know the interactions between electrons and nuclei. We can calculate the cross section in a variety of approximate methods. For this case, the low energy scattering is again described by a and r_0, if these quantities[1] are calculated by using the zero energy wave function or by "bound state" wave function. The purpose of the present paper is to present a variational method for the effective range. We shall take up the simplest problem, the s scattering from a central field. The direct numerical method is applicable for this one dimensional solution, hence nothing is saved by an application of variational method for the effective range in this simplest problem. But the variational method can be extended to more complicated cases, where the direct numerical solution is almost impossible.

It has been found that the s scattering of electrons by hydrogen atoms may be described by the effective range approximation[4] up to quite high energies. The scattering lengths of electron-hydrogen scattering are now being calculated in our group by the Hulthén variational method using many parameters. In the present work, the construction of a variational method for the effective range is attempted.

II. ZERO ENERGY STATE

Let $u(r)$ be the radial wave function multiplied by r, for a s state. Then, $u(r)$ satisfies the Schrödinger equation, with a short range potential $[-W(r)]$.

$$d^2u/dr^2+[W(r)+k^2]u\equiv Lu=0. \tag{1}$$

Let $u(r)$ be normalized by,

$$u(r)\rightarrow \cos kr+\cot\eta\,\sin kr, \quad u(0)=0. \tag{2}$$

We expand $u(r)$ in powers of k^2 at $k^2=0$,

$$u(r)=u_0(r)+k^2v(r)+\cdots, \tag{3}$$

then $u_0(r)$ is the solution of (1) at $k^2=0$, and $v(r)$ satisfies the following equation:

$$d^2v/dr^2+W(r)v+u_0\equiv L_0v+u_0=0. \tag{4}$$

The asymptotic forms and the initial conditions for u_0 and v are derived from (2) by expanding in power series of k^2,

$$k\cot\eta=-1/a+(r_0/2)k^2+\cdots,$$

$$u_0\rightarrow 1-\frac{r}{a}, \quad (r\rightarrow\infty), \quad u_0(0)=0, \tag{5}$$

$$v\rightarrow \frac{r_0}{2}r-\frac{r^2}{2}+\frac{r^3}{6a}, \quad (r\rightarrow\infty), \quad v(0)=L_0v(0)=0, \tag{6}$$

where a and r_0 are the scattering length and the effective range, respectively. We now consider the expression,

$$\int_0^\infty wL_0^2w\,dr, \quad (w=2v),$$

where $w(r)$ satisfies the following equation and the conditions,

$$L_0^2w=0, \tag{7a}$$

$$w(0)=L_0w(0)=0, \quad w(r)\rightarrow r_0r-r^2+(r^3/3a). \tag{7b}$$

The function $w(r)$ is completely determined by (7a) and (7b). L_0^2 is a fourth order differential operator, so there are four independent solutions for w, if (7b) is not imposed on w. The last condition of (7b) means the lack of a constant term at large distances. The three

* National Research Council Postdoctorate Fellow, on leave of absence from Department of Physics, University of Tokyo.

[1] J. Schwinger, Phys. Rev. **72**, 742A (1947); J. M. Blatt and J. D. Jackson, Phys. Rev. **76**, 18 (1949).

[2] H. A. Bethe, Phys. Rev. **76**, 38 (1949); G. F. Chew and M. L. Goldberger, Phys. Rev. **75**, 1637 (1949).

[3] Breit, Condon, and Present, Phys. Rev. **50**, 825 (1936).

[4] Ohmura, Hara, and Yamanouchi, Progr. Theoret. Phys. **22**, 152 (1959).

conditions of (7b) select one solution among the possible four. The normalization of w is fixed so as to give -1 as the coefficient of r^2 at large distances. By taking variations from the exact solution w and making use of the following boundary conditions,

$$\delta w(0)=0, \quad \delta w \to r\delta r_0+(r^3/3)\delta(1/a),$$
$$d(\delta w)/dr \to \delta r_0+r^2\delta(1/a), \quad L_0 w \to -2+(2/a)r,$$
$$L_0\delta w(0)=0, \quad L_0\delta w \to 2r\delta(1/a)$$
$$d(L_0\delta w)/dr \to 2\delta(1/a),$$

which are all derivable from the conditions (7b), we get a simple result,

$$\delta\int_0^\infty wL_0^2w\,dr = w\frac{d(L_0\delta w)}{dr}-\frac{dw}{dr}L_0(\delta w) + L_0 w\frac{d(\delta w)}{dr}-\frac{d(L_0 w)}{dr}\delta w\Big|_0^\infty = -2\delta r_0.$$

Thus the following quantity is a stationary expression for the effective range,

$$r_0{}^S = r_0+\tfrac{1}{2}\int_0^\infty wL_0^2w\,dr. \tag{8}$$

Our variational principle for r_0 may be used as follows. Let w_t be a trial function with n adjustable parameters c_i, and satisfy the conditions (7b), where r_0 and a should be considered two additional adjustable parameters, r_{0t} and a_t. By differentiating $r_0{}^S$ by c_i, a_t and r_{0t}, we get $n+2$ equations,

$$\frac{\partial}{\partial c_i}\int_0^\infty w_tL_0^2w_t\,dr=0, \quad (i=1,\cdots n), \tag{9a}$$

$$\frac{\partial}{\partial a_t}\int_0^\infty w_tL_0^2w_t\,dr=0, \tag{9b}$$

$$\frac{\partial}{\partial r_{0t}}\int_0^\infty w_tL_0^2w_t\,dr=-2. \tag{9c}$$

Equations (9) will determine the $n+2$ unknown parameters c_i, a_t and r_{0t}. By inserting the determined values of c_i, a_t and r_{0t} into w_t, we get the final approximate value $r_0{}^T$ for the effective range.

$$r_0{}^T = r_{0t}+\tfrac{1}{2}\int_0^\infty w_tL_0^2w_t\,dr. \tag{10}$$

(8) can also be used in a slightly different way. If we impose the condition,

$$\int_0^\infty w_tL_0^2w_t\,dr=0, \tag{11}$$

instead of (9c), we have the result, $r_0{}^T=r_{0t}$. This corresponds to the original Hulthén method[5] for the phase shift. If we know an accurate value for a beforehand, a_t is no longer to be adjusted, and (9b) can be removed from the set of Eqs. (9). (9a) and (11) may be conveniently used to get an upper bound for the effective range, as explained in V.

In order to obtain a clear picture of the general behavior of $w(r)$, we write down the exact solution where the potential is an attractive square well.

$$W(r)=\begin{cases} V>0, & 0\leqq r\leqq 1, \\ 0, & r>1. \end{cases}$$

The exact values of a and r_0, together with the solutions u_0 and w, are given by (12),

$$a=1-(V^{-1})^{\frac{1}{2}}\tan(V)^{\frac{1}{2}}, \tag{12a}$$

$$r_0=1+\frac{1}{(V)^{\frac{1}{2}}[\tan(V)^{\frac{1}{2}}-(V)^{\frac{1}{2}}]}-\frac{V}{3[\tan(V)^{\frac{1}{2}}-(V)^{\frac{1}{2}}]^2}, \tag{12b}$$

$$u_0=\frac{\sin[(V)^{\frac{1}{2}}r]}{\sin(V)^{\frac{1}{2}}-(V)^{\frac{1}{2}}\cos(V)^{\frac{1}{2}}}=-\tfrac{1}{2}L_0w; \quad 0\leqq r\leqq 1, \tag{12c}$$

$$w=\frac{1}{(V)^{\frac{1}{2}}[\tan(V)^{\frac{1}{2}}-(V)]^{\frac{1}{2}}\cos(V)^{\frac{1}{2}}}\left\{r\cos[(V)^{\frac{1}{2}}r]-\frac{V[\sin[(V)^{\frac{1}{2}}r]}{3[\tan(V)^{\frac{1}{2}}-(V)^{\frac{1}{2}}]}\right\}; \quad 0\leqq r\leqq 1. \tag{12d}$$

The solutions u_0 and w are given by the asymptotic

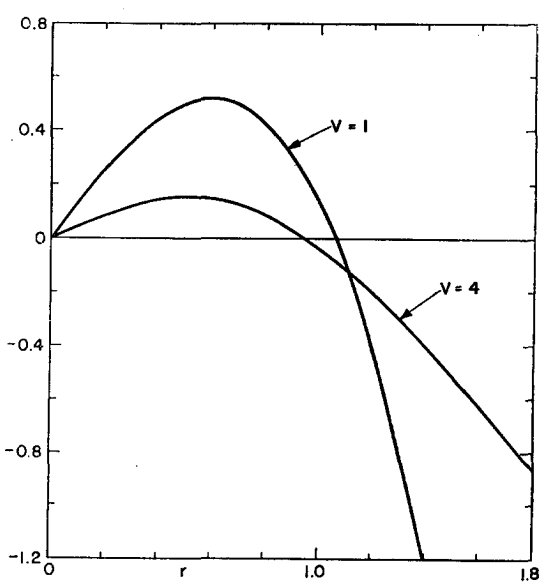

FIG. 1. $w(r)$ for the square well potential with the range; unity, and $V=1$ and 4.

[5] L. Hulthén, Kungl. Fysiograf. Sällskap. Lund Förh. 14, 21 (1944).

forms of (5) and (7b) for $r>1$. $w(r)$ is shown in Fig. 1 for $V=1$ and $V=4$.

It is worth noting that the third term of the $k\cot\eta$ expansion can be calculated from u_0 and w,

$$T=-\frac{1}{2}\int_0^\infty(\psi_0\varphi-u_0w)dr, \tag{13}$$

where $\psi_0=1-(r/a)$, $\varphi=r_0r-r^2+(r^3/3a)$ and T is defined by

$$k\cot\eta=-\frac{1}{a}+\frac{r_0}{2}k^2+Tk^4+\cdots.$$

Only when the value of $|T|$ is small, can the effective range approximation be considered good for the energy, the value of which is by no means small.

III. MODIFIED METHOD

We have constructed a variational method in II for the effective range of zero energy state. We must first evaluate the integral, $\int_0^\infty w_tL_0^2w_tdr$. The operator L_0^2, however, may be a very complicated one in an actual problem. The variational method will be modified in this section so as to be easier for the actual application.

Consider the following integral:

$$\int_0^\infty(wL_0u_0+2u_0^2+u_0L_0w)dr,$$

where w and u_0 satisfy Eq. (7a) and $L_0u_0=0$, together with (5) and (7b), respectively. By taking infinitesimal variations of w and u_0 independently, we get

$$\begin{aligned}\delta\int_0^\infty(wL_0u_0+2u_0^2+u_0L_0w)dr&=\int(wL_0\delta u_0+2u_0\delta u_0+u_0L_0\delta w)dr\\&=w\frac{d(\delta u_0)}{dr}-\frac{dw}{dr}\delta u_0+u_0\frac{d(\delta w)}{dr}-\frac{du_0}{dr}\delta w\Big|_0^\infty\\&=\delta r_0.\end{aligned}$$

Use has been made of the following boundary conditions in the course of derivation,

$$w(0)=\delta u_0(0)=u_0(0)=\delta w(0)=0,$$

$$\delta u_0\to-r\delta(1/a),\quad \delta w\to r\delta r_0+(r^3/3)\delta(1/a),\quad (r\to\infty).$$

Then we have a stationary expression for the effective range,

$${}^Sr_0=r_0-\int_0^\infty(wL_0u_0+2u_0^2+u_0L_0w)dr, \tag{14}$$

because the integrand of (14) vanishes, at exact solution, by the equation (4) and $L_0u_0=0$. The expression (14) is identical with (8), if we replace u_0 by $-L_0w/2$. But it is essential that w and u_0 in (14) are independent, so the nature of (8) and (14), as a functional, is quite different from each other. (14) does not contain the operator L_0^2. This fact may reduce the effort of calculation, even though the number of terms in the stationary expression is increased. As to the application of (14), the differentiation of ${}^Sr_0$ with respect to the parameter, say b_i, contained in w_t does not give any equation for b_i. Therefore, an explicit explanation of the method will be needed.

In view of computing labor involved, a linear trial function may be most convenient. We shall assume the following trial function:

$$u_t(r)=\sum_{i=1}^m a_if_i(r)+1-\alpha_tr,\quad u_t(0)=0,\quad \alpha_t=1/a_t, \tag{15a}$$

$$w_t(r)=\sum_{i=1}^n b_ig_i(r)+r_{0t}r-r^2+(\alpha_t/3)r^3,\qquad w_t(0)=L_0w_t(0)=0, \tag{15b}$$

where $f_i(r)$ and $g_i(r)$ are well damped functions at large distances. By substituting (15) into (14), we have, in terms of a number of known coeficients, $c_0,\cdots s_i$,

$$\begin{aligned}{}^Sr_0=r_{0t}+c_0+\sum_{i=1}^m c_ia_i+\sum_{i,j=1}^m c_{ij}a_ia_j+\sum_{i=1}^n d_ib_i\\+\sum_{i,j=1}^{\substack{j=m\\i=n}}d_{ij}b_ia_j+h_1\alpha_t+h_2\alpha_t^2+\sum_{i=1}^m l_ia_i\alpha_t+\sum_{i=1}^n o_ib_i\alpha_t\\+pr_{0t}\alpha_t+qr_{0t}+\sum_{i=1}^m s_ia_ir_{0t}.\end{aligned}$$

Since ${}^Sr_0$ should be stationary for variations of all adjustable parameters, we get a set of simultaneous linear equations,

$$\frac{\partial{}^Sr_0}{\partial b_i}=d_i+\sum_{j=1}d_{ij}a_j+o_i\alpha_t=0,\quad (i=1,\cdots n), \tag{16a}$$

$$\frac{\partial{}^Sr_0}{\partial r_{0t}}=1+p\alpha_t+q+\sum_{i=1}^m s_ia_i=0, \tag{16b}$$

$$\frac{\partial{}^Sr_0}{\partial a_i}=c_i+2\sum_{j=1}^m c_{ij}a_j+\sum_{j=1}^n d_{ji}b_j+l_i\alpha_t+s_ir_{0t}=0,\quad (i=1,\cdots m) \tag{16c}$$

$$\frac{\partial{}^Sr_0}{\partial\alpha_t}=h_1+2h_2\alpha_t+\sum_{i=1}^m l_ia_i+\sum_{i=1}^n o_ib_i+pr_{0t}=0. \tag{16d}$$

The variables b_i are not contained in (16a) and (16b), hence the number of parameters (b_i,r_{0t}) should be

restricted by the number of Eqs. (16c) and (16d),

$$m+1 \geqq n+1.$$

If $m=n$, Eqs. (16a, b) determine a_i $(i=1, \cdots m)$ and α_t, and Eqs. (16c, d) determine b_i $(i=1, \cdots n)$ and r_{0t}. If $m>n$, Eqs. (16) should be solved as a whole.

Since the function $u_0(r)$ is the solution of the Schrödinger equation at zero energy, it is possible that $u_t(r)$ has already been solved by the ordinary variational method to evaluate the scattering length. If this is so, $n+1$ Eqs., (16a) and (16b), are unnecessary. We can obtain the values of b_i $(i=1, \cdots m)$ and r_{0t} from Eqs. (16c) and (16d).

The final approximate value for the effective range is simply calculated by

$$^{T}r_0=c_0+(r_{0t}+qr_{0t}+h_1\alpha_t+\sum_{i=1}^{m} c_ia_i+\sum_{i=1}^{n} d_ib_i)/2,$$

where use has been made of Eqs. (16).

The difference between $^{T}r_0$ and the exact value of r_0 is of the second order of errors in u_{0t} and w_t. An explicit calculation shows that

$$\begin{aligned}\Delta^{T}r_0=\Delta r_0-\int_0^\infty &[(w+\Delta w)L_0(u_0+\Delta u)\\ &+2(u_0+\Delta u)^2+(u_0+\Delta u)L_0(w+\Delta w)]dr\\ =-\int_0^\infty &\{[2(\Delta u)^2+\Delta uL_0\Delta w]+\Delta wL_0\Delta u\}dr,\end{aligned}$$

where $\Delta^{T}r_0={}^{T}r_0-r_0$ (exact value), $\Delta r_0=r_{0t}-r_0$ (exact value), $\Delta u=u_t-u_0$, $\Delta w=w_t-w$. If the accuracy of w_t is inferior to the one of u_t, the error of $^{T}r_0$ is of the order $\Delta w\times\Delta u$.

IV. BOUND STATE

The effective range can be also defined for bound states. If the binding energy (γ^2) is small compared with the characteristic depth of the potential, the effective range ρ thus defined is useful to describe the low energy scattering. The effective range defined by Schwinger in the triplet neutron-proton scattering was really based on the ground state wave function of the deuteron.

For negative energies, k in (1) is purely imaginary, $k=-i\epsilon$. Equation (1) is now

$$(d^2u/dr^2)+[W(r)-\epsilon^2]u\equiv L_\epsilon u=0, \quad u(0)=0. \tag{17}$$

The solution of (17) has the asymptotic form,

$$u\rightarrow \cosh\epsilon r+\beta\sinh\epsilon r. \tag{18}$$

We can easily see from (2), by extending to negative energy states, that

$$k\cot\eta_{k^2=-\epsilon^2}=\epsilon\beta.$$

The condition of ϵ being a bound state is $e^{2i\eta}=0$, namely $\cot\eta=-i$, or $\beta=-1$. Let the binding energy be $\gamma^2(\gamma>0)$. The condition of bound state leads to,

$$k\cot\eta_{k^2=-\gamma^2}=-\gamma.$$

On assuming an expansion in powers of (γ^2+k^2) at $k=-i\gamma$,

$$k\cot\eta=-\gamma+\frac{\rho}{2}(\gamma^2+k^2)+O[(\gamma^2+k^2)^2], \quad (\gamma\text{; small}).$$

The asymptotic form (18) and the solution of (17) are also expanded at $k=-i\gamma$, namely, $\epsilon=\gamma$.

$$\begin{aligned}u\rightarrow &\cosh\epsilon r+(k\cot\eta/\epsilon)\sinh\epsilon r\\ &=e^{-\gamma r}+(\gamma^2+k^2)\{1/2\gamma[\rho-(1/\gamma)]\sinh\gamma r\\ &\qquad +re^{-\gamma r}/2\gamma\}+O[(\gamma^2+k^2)^2],\\ u(r)=&u_\gamma(r)+(\gamma^2+k^2)v(r)+O((\gamma^2+k^2)^2).\end{aligned}$$

$v(r)$ satisfies an equation similar to (4),

$$L_\gamma v+u_\gamma=0. \tag{19}$$

We see from (19) that $u_\gamma(0)=0$ leads to $L_\gamma v(0)=0$, and $L_\gamma u_\gamma=0$ leads to $L_\gamma^2 v=0$. The analogy of the variational expression at zero energy state suggests that the consideration of the integral, $\int_0^\infty wL_\gamma^2wdr$, may provide a variational method at bound states. w is the solution of the following equation with the specified conditions,

$$L_\gamma^2w(r)=0, \tag{20a}$$

$$\begin{aligned}w(0)=L_\gamma w(0)=0, \quad &w\rightarrow 1/\gamma(\rho-1/\gamma)\sinh\gamma r\\ &+re^{-\gamma r}/\gamma, \quad (r\rightarrow\infty).\end{aligned} \tag{20b}$$

Let us take variations of w from the exact solution, and we get

$$\begin{aligned}\delta\int_0^\infty wL_\gamma^2wdr=&w\frac{d(L_\gamma\delta w)}{dr}-\frac{dw}{dr}L_\gamma\delta w\\ &+L_\gamma w\frac{d(\delta w)}{dr}-\frac{d(L_\gamma w)}{dr}\delta w\Big|_0^\infty\\ =&-2e^{-\gamma r}\cosh\gamma r\delta\rho-2e^{-\gamma r}\sinh\gamma r\delta\rho\\ =&-2\delta\rho.\end{aligned}$$

The last term of (20b) is important in the derivation, because we have used the relations,

$$\begin{aligned}&\delta w\rightarrow(\delta\rho/\gamma)\sinh\gamma r, \quad d(\delta w)/dr\rightarrow\delta\rho\cosh\gamma r,\\ &L_\gamma w\rightarrow-2e^{-\gamma r}, \quad L_\gamma\delta w\rightarrow 0,\\ &d(L_\gamma w)/dr\rightarrow 2\gamma e^{-\gamma r}, \quad d(L_\gamma\delta w)/dr\rightarrow 0.\end{aligned}$$

The stationary expression for the effective range ρ is now given by

$$\rho^S=\rho+\tfrac{1}{2}\int_0^\infty wL_\gamma^2wdr. \tag{21}$$

The binding energy γ^2 has been assumed known beforehand. Usually, a sufficiently accurate value of the energy is available, so this assumption may not cause a restriction from a practical point of view. It will be shown in the Appendix that the inaccuracy of the binding energy will only give rise to a very

small effect which is of the second order in the inaccuracy of the energy.

The variational calculations based on (21) is very similar to those described in II. On assuming $n+1$ adjustable parameters a_i and ρ_t in the trial function $w_t(r)$, which satisfies the condition (20b), we calculate

$$\frac{\partial}{\partial a_i}\int_0^\infty w_t L_\gamma^2 w_t dr=0, \quad (i=1, \cdots n), \tag{22a}$$

$$\frac{\partial}{\partial \rho}\int_0^\infty w_t L_\gamma^2 w_t dr=-2. \tag{22b}$$

(22a) and (22b) determine a_i and ρ_t, and we finally get the approximate value of ρ,

$$\rho^T=\rho_t+\tfrac{1}{2}\int_0^\infty w_t L_\gamma^2 w_t dr. \tag{23}$$

Alternatively, we can set instead of (22), as

$$\frac{\partial}{\partial a_i}\int_0^\infty w_t L_\gamma^2 w_t dr=0, \quad (i=1, \cdots n), \tag{24a}$$

$$\int_0^\infty w_t L_\gamma^2 w_t dr=0. \tag{24b}$$

The stationary value of ρ is nothing but ρ_t on account of (24b). The variational method (24) is very useful in that an upper bound for ρ is given as proved in the next section.

A modified method corresponding to (14) can also be developed. By replacing u_0, L_0, r_0 in (14) by u_γ, L_γ, ρ, and imposing $u_\gamma \to e^{-\gamma r}$, $(r \to \infty)$, we easily have the corresponding variational method for the case of bound states.

V. MINIMUM NATURE OF EFFECTIVE RANGE

The minimum nature in the usual Ritz principle is well-known. The Ritz method always gives an upper bound for the ground state energy. The original method of Hulthén has a similar property[6] for the scattering length. In this section, it will be proved that the method (24) guarantees the minimum nature for the effective range. Consideration will also be made for the method (11).

The exact solution of (20), w, and the trial function w_t can be set as

$$w(r)=y(r)+1/\gamma(\rho-1/\gamma)\sinh\gamma r+re^{-\gamma r}/\gamma, \qquad w(0)=L_\gamma w(0)=0,$$

$$\Delta w(r)=w_t(r)-w(r)=\Delta y(r)+\frac{\Delta\rho}{\gamma}\sinh\gamma r, \qquad \Delta w(0)=L_\gamma \Delta w(0)=0,$$

where $y(r)$ and $\Delta y(r)$ must tend to zero faster than $re^{-\gamma r}$. Let us evaluate the integral, $\int_0^\infty w_t L_\gamma^2 w_t dr$, using the fundamental equation for w and the boundary conditions,

$$\begin{aligned}\int_0^\infty w_t L_\gamma^2 w_t dr &= \int_0^\infty (wL_\gamma^2\Delta w+\Delta w L_\gamma^2 \Delta w)dr \\ &= -2\Delta\rho+\int_0^\infty \Delta w L_\gamma^2 \Delta w dr \\ &= -2\Delta\rho\left\{1-\frac{1}{2\gamma}\int_0^\infty W\sinh\gamma r L_\gamma \Delta y dr\right\} \\ &\quad +\int_0^\infty \Delta y L_\gamma^2 \Delta y dr+\frac{(\Delta\rho)^2}{\gamma^2}\int_0^\infty W^2\sinh^2\gamma r dr.\end{aligned}$$

By applying the condition (24b), we have the expression for $\Delta\rho$,

$$2\Delta\rho=\frac{\int_0^\infty \Delta y L_\gamma^2\Delta y dr+(\Delta\rho/\gamma)^2\int_0^\infty W^2\sinh^2\gamma r dr}{1-\frac{1}{2\gamma}\int_0^\infty W\sinh\gamma r L_\gamma \Delta y dr}. \tag{25}$$

The term, $\int_0^\infty \Delta y L_\gamma^2 \Delta y dr$, is positive, because $\Delta y(0)=0$, $\Delta y e^{\gamma r}/r \to 0$ $(r\to\infty)$, consequently $\Delta y(r)$ belongs to a Hilbert space. The right hand side of (25) is, therefore, positive so far as the denominator remains positive. The main term is of course $\int_0^\infty \Delta y L_\gamma^2 \Delta y dr$. The minimum nature for the effective range at bound states is thus proved.

$$\begin{aligned}\rho^T-\rho&=\rho_t-\rho=\Delta\rho \\ &=\tfrac{1}{2}\int_0^\infty \Delta y L_\gamma^2\Delta y dr\geqq 0, \quad (\Delta y\to 0).\end{aligned} \tag{26}$$

A similar argument is possible for the method (11). On assuming first that the scattering length a is a precisely known quantity, we put

$$w=y+r_0 r-r^2+Ar^3, \quad A=1/(3a),$$
$$w_t-w=\Delta w=\Delta y+r\Delta r_0+r^3\Delta A,$$

where Δr_0 is a quantity of $O[(\Delta y)^2]$ and ΔA is also assumed to be $O[(\Delta y)^2]$. We calculate $\int_0^\infty w_t L_0^2 w_t dr$.

$$\int_0^\infty w_t L_0^2 w_t dr=\Delta r_0\int_0^\infty wL_0^2 r dr+\Delta A\int_0^\infty wL_0^2 r^3 dr + \int_0^\infty \Delta y L_0^2\Delta y dr+O[(\Delta y)^3].$$

The second term vanishes by the fundamental equation

[6] T. Ohmura (to be published).

and the boundary conditions,

$$\int_0^\infty wL_0^2r^3dr=w\frac{d(L_0r^3)}{dr}-\frac{dw}{dr}L_0r^3 +L_0w\frac{dr^3}{dr}-\frac{d(L_0w)}{dr}r^3\Big|_0^\infty=0.$$

We also have

$$\int_0^\infty wL_0^2rdr=-2.$$

The condition $\int_0^\infty w_tL_0^2w_tdr=0$ leads to

$$r_0{}^T-r_0=\Delta r_0=\tfrac{1}{2}\int_0^\infty \Delta yL_0^2\Delta ydr\geqq 0. \quad (\Delta y\to 0). \quad (27)$$

The minimum nature of r_0 will be valid only for a restricted class of potentials, when the scattering length is determined by (9a, b) and (11). We just write down the expression for the error Δr_0.

$$\Delta r_0=\tfrac{1}{2}\int_0^\infty \Delta yL_0^2\Delta ydr-\frac{1}{2c}\left(\int_0^\infty r^3L^2\Delta ydr\right)^2, \quad (28)$$

where

$$c=\int_0^\infty r^3L_0^2r^3dr=\int_0^\infty (12r^4W+r^6W^2)dr.$$

From (28) we can see that $r_0{}^T$ has the minimum nature only for $c<0$.

VI. SUMMARY

Low energy scatterings can be described by two parameters, the scattering length and the so-called effective range. The scattering length a is defined by the asymptotic form of the wave function of zero energy. The effective range r_0 (or ρ) is also evaluated by the zero energy solution u_0 (or by the bound state wave function u_γ),

$$r_0=2\int_0^\infty\left[\left(1-\frac{r}{a}\right)^2-u_0^2\right]dr, \quad (29a)$$

$$\rho=2\int_0^\infty (e^{-2\gamma}-u_\gamma{}^2)dr, \quad u_\gamma\to e^{-\gamma}, \quad (r\to\infty). \quad (29b)$$

The expression (29a) or (29b) is, however, not stationary with respect to small deviations from u_0 (or u_γ). The stationary expressions for r_0 and ρ are given in (8) and (21). It has been proved that the condition (24) guarantees the minimum nature for ρ. The same is true for the conditions (9a, 11) provided that the scattering length is known beforehand. In view of practical applications, these methods, however, seem to be not always adequate because of a complicated operator L^2 in the integral. A stationary expression (14) for r_0 (or ρ) is devised, which has the operator L in it instead of L^2. (14) may be used conveniently for approximate evaluation of r_0 at the expense of the minimum character.

ACKNOWLEDGMENT

The author would like to thank Dr. Ta-You Wu for his hospitality and for many advices in the preparation of the manuscript.

APPENDIX. EFFECT OF THE INACCURACY OF THE BINDING ENERGY

We have assumed in IV that the binding energy is a known quantity. It will be shown here that a deviation Δ from the exact energy value causes a corresponding change of the effective range of the order Δ^2.

Let the exact value of the binding energy be γ^2, and let an approximate value, which we take as the binding energy, be ϵ^2. The asymptotic form of w at $k^2=-\epsilon^2$ can be written as

$$w\to 1/\epsilon(\rho'-1/\epsilon)\sinh\epsilon r+(re^{-\epsilon r}/\epsilon)+(\Delta/2\epsilon^2)(\rho'-1/\epsilon) \times\{(1/\epsilon)\sinh\epsilon r-r\cosh\epsilon r\}+O(\Delta^2), \quad (A.1)$$

where $\Delta=\gamma^2-\epsilon^2$. ρ' is the effective range at $k^2=-\epsilon^2$, namely

$$\frac{d(k\cot\eta)}{dk^2{}_{(k^2=-\epsilon^2)}}=\frac{\rho'}{2}.$$

ρ' is not sensitive to the energy, so it is nearly the same as ρ. According to the variational method in IV, we take a trial function w_t of the form,

$$w_t\to 1/\epsilon(\rho_t-1/\epsilon)\sinh\epsilon r+re^{-\epsilon r}/\epsilon. \quad (A.2)$$

Thus we understand that the trial function w_t has an error of $O(\Delta)$ in the asymptotic form. By allowing this kind of variation together with the variation $\delta\rho$ in the asymptotic form, we calculate the variation of $\int wL_\epsilon{}^2wdr$.

$$\delta\int_0^\infty wL_\epsilon{}^2wdr=w\frac{d(L_\epsilon\delta w)}{dr}-\frac{dw}{dr}L_\epsilon\delta w +L_\epsilon w\frac{d(\delta w)}{dr}-\frac{d(L_\epsilon w)}{dr}\delta w\Big|_0^\infty \quad (A.3)$$

$$=-2\delta\rho'.$$

There is no term proportional to Δ in (A.3). This means that the effect of Δ to the stationary quantity,

$$\rho'^T=\rho_t'+\tfrac{1}{2}\int_0^\infty w_tL_\epsilon{}^2w_tdr,$$

is that of the second order in Δ.

Mandelstam Representation for Potential Scattering

ABRAHAM KLEIN*†
University of Pennsylvania, Philadelphia, Pennsylvania
(Received October 19, 1959)

A proof of the Mandelstam representation for the nonrelativistic scattering amplitude is given when the potential is of the Yukawa form or (by obvious extension) a suitable linear combination of such forms. The analytic properties of the scattering amplitude as a function of momentum transfer are established by using only a finite sequence of equivalent definitions of the scattering amplitude. By studying the Born series for individual partial waves, it is shown in addition that there cannot be an essential singularity at infinity. Together, these results imply both dispersion relations for individual partial waves and the Mandelstam representation.

I. INTRODUCTION

IN view of both the intuitive appeal and practical significance of Mandelstam's conjecture concerning the analytic properties of elementary particle scattering amplitudes,[1] any inroads into the problem of proof must be accounted of some interest. In the present note, we show that the Mandelstam representation is valid for potential scattering with a Yukawa force law.[2] This work may then be considered an extension of previous studies,[3] which established dispersion relations in the energy variable for fixed (and limited) momentum transfer, albeit for a wider class of potentials. Though the latter work in some respect played little more than the role of camp follower to the field theoretical hosts, the present attempts may aspire to providing some heuristic insight into the more difficult and essentially untouched field theoretical problem. Of the methods exposed in this paper, we feel that at least the method of Sec. II, which establishes the analytic properties of the scattering amplitude in the momentum transfer variable for all finite values of the variable may be at least faintly suggestive of an approach to the field theoretical problem.

In detail we shall deal exclusively with a potential of the form,

$$\lambda V(r)=\lambda \exp(-\mu r)/r, \tag{1}$$

though the work can obviously be extended to a suitable linear combination of potentials of the type (1). Let t be the negative of the square of the momentum transfer variable, s the energy. In I, we have shown that for real s the exact scattering amplitude $f(s,t)$, exclusive of its first Born approximation, is an analytic function of t for $|t|<4\mu^2$. In Sec. II, this proof is repeated in a convenient form and extended to the following theorem: For real s the (exact) difference between $f(s,t)$ and its nth Born approximation, $f_{(n)}(s,t)$ is an analytic function of t for $|t|<(n+1)\mu^2$. It is moreover, straight forward to show[4] that $f_{(n)}(s,t)$ itself is analytic for all t, except for branch points at $t=4\mu^2$, $9\mu^2$, $\cdots n\mu^2$ and a pole due to the potential at $t=\mu^2$. Together, these facts constitute a proof that $f(s,t)$ is, exclusive of the pole and the cut from $4\mu^2$ to ∞ analytic in t for all finite t. It follows also that the dispersion relation in the energy variable is valid for all t.

To establish the Mandelstam or double dispersion representation it is still necessary to show that $f(s,t)$ does not have an essential singularity as $|t|\to\infty$. We have so far been unable to carry through this part of the program directly for the three dimensional representation of the scattering amplitude. On the other hand, it is easy to see that such a property would imply a similar property for the partial wave amplitude $A_l(s)$,

$$A_l(s)=\exp(i\delta_l(s))\,\sin\delta_l(s)/s^{\frac{1}{2}}, \tag{2}$$

and it is shown that this is not the case. We first remark that the properties of $f(s,t)$ imply that $A_l(s)$ is analytic for all finite s except for cuts from $s=0$ to $s=\infty$ and from $s=-(\mu^2/4)$ to $s=-\infty$. It is then demonstrated in Sec. III (in detail only for the S wave) that $A_l(s)$ approaches zero as $|s|\to\infty$ like its first Born approximation. Indeed, it is shown that the Born series for $A_l(s)$ converges uniformly in s for sufficiently large s exclusive of the cuts.

* Supported in part by the U. S. Atomic Energy Commission.

† Part of this work was done while the author was a summer visitor at RIAS, inc., Baltimore, Maryland.

[1] S. Mandelstam, Phys. Rev. **112**, 1344 (1958).

[2] At the time of composition, the author had in hand the following manuscripts bearing on this subject: J. Bowcock and A. Martin, *A Mandelstam Representation in Potential Scattering* (CERN preprint); Blankenbecler, Goldberger, Khuri, and Treiman, *Mandelstam Representation for Potential Scattering* (Princeton University preprint); T. Regge, *Introduction to Complex Orbital Momenta*, (University of Munich preprint). Of these, the first establishes only the analytic properties of the individual terms of the Born series, the second contains essentially the results of Sec. II of this paper (obtained by studying the Fredholm solution of the scattering problem), but fails to establish the required behaviour for infinite momentum transfer, and the third appears to contain the elements of a full proof by methods radically different from those of the present paper. Also of relevance to this problem is a paper by A. Martin, *On the Analytic Properties of Partial Wave Scattering Amplitudes* . . . (CERN preprint). During the conception and execution of this work, the author had in hand only the first and last-named works. He is indebted to the various authors for their preprints.

[3] The most complete of these is the paper, hereafter referred to as I, by A. Klein and C. Zemach, Ann. Phys. **7**, 440 (1959); see also the previous work of N. N. Khuri, Phys. Rev. **107**, 1148 (1957); S. Gasiorowicz and H. P. Noyes, Nuovo Cimento **10**, 78 (1958), and L. D. Faddeev, JETP **35**(8), 229 (1959). Results similar to those of I have been obtained in unpublished work of R. Jost and W. Hunziker.

[4] J. Bowcock and A. Martin, reference 2.

From our study of the partial wave amplitudes we may infer that $f(s,t)$ does not have an essential singularity in t as $|t|\to\infty$, and that a Mandelstam representation follows for it (Sec. IV). We cannot, however, conclude the exact behavior for large $|t|$ and in this sense our work is incomplete. The concluding remarks of Sec. V are concerned with the determination of the so-called spectral function in the new representation and with a possible application, especially of the partial wave relations.

II. ANALYTIC PROPERTIES OF THE SCATTERING AMPLITUDE IN THE MOMENTUM TRANSFER VARIABLE

With units in which $s=k^2$, where k is the wave number, we may write for the scattering amplitude the integral representation (see Sec. I)

$$F(\mathbf{k}',\mathbf{k})=f(s,t)=f^{(0)}(t)$$
$$-\frac{\lambda^2}{4\pi}\int \exp(-i\mathbf{k}'\cdot\mathbf{r}')V(r')G(\mathbf{r}',\mathbf{r};s)\times V(r)\exp(i\mathbf{k}\cdot\mathbf{r})d^3r'd^3r$$
$$=f^{(0)}(t)-4\pi\lambda^2\int\frac{1}{\mu^2+(\mathbf{k}'-\mathbf{q}')^2}\times G(\mathbf{q}',\mathbf{q};s)\frac{1}{\mu^2+(\mathbf{k}-\mathbf{q})^2}d^3q'd^3q, \quad (3)$$

where

$$f^{(0)}(t)=-\lambda[\mu^2-t]^{-1} \quad (4)$$

is essentially the Fourier transform of the potential[1] and

$$G(\mathbf{r}',\mathbf{r};s)=\langle\mathbf{r}'|[s+i\epsilon-(-\nabla^2+\lambda V)]^{-1}|\mathbf{r}\rangle$$
$$=\int e^{i\mathbf{q}'\cdot\mathbf{r}'}G(\mathbf{q}',\mathbf{q};s)\exp(-i\mathbf{q}\cdot\mathbf{r})d^3q'd^3q \quad (5)$$

is the full Green's function satisfying the (symbolic) integral equations,

$$G=G^{(0)}+\lambda G^{(0)}VG=G^{(0)}+\lambda GVG^{(0)}. \quad (6)$$

We recall that

$$G^{(0)}(\mathbf{r}',\mathbf{r};k^2)=-\exp[ik|\mathbf{r}'-\mathbf{r}|]/4\pi|\mathbf{r}'-\mathbf{r}|. \quad (7)$$

We first use the form (3) to show that for real s $f(s,t)-f^{(0)}(t)$ is analytic for $|t|<4\mu^2$ and that the first singularity actually occurs at $t=4\mu^2$. Toward this end, we study the possible zeros of a factor such as $\mu^2+(\mathbf{k}-\mathbf{q})^2$. We choose a coordinate system in which

$$\mathbf{k}=k(\cos\tfrac{1}{2}\theta,-\sin\tfrac{1}{2}\theta,0),$$
$$\mathbf{q}=(q_1,q_2,q_3)=q(\sin\gamma\sin\varphi,\cos\gamma\sin\varphi,\cos\varphi), \quad (8)$$

where θ is the scattering angle. We are looking for the zeros of

$$\mu^2+(\mathbf{k}-\mathbf{q})^2=\mu^2+k^2+q^2-2kq\sin\varphi\sin(\gamma-\tfrac{1}{2}\theta). \quad (9)$$

Since no solution exists for real θ, we write $\theta=\theta_1+i\theta_2$ and find directly as necessary conditions that

$$\cos(\gamma-\tfrac{1}{2}\theta_1)=0,$$
$$\cosh\tfrac{1}{2}\theta_2=(\mu^2+k^2+q^2)/2kq\sin\varphi, \quad (10)$$

or

$$2k\sin\tfrac{1}{2}\theta=(q^2\sin^2\varphi)^{-1}\{-q_2(\mu^2+k^2+q^2)+iq_1[(\mu^2+k^2+q^2)^2-4k^2q^2\sin^2\varphi]^{\frac{1}{2}}\}. \quad (11)$$

The expression (11) must be varied with respect to q_1, q_2, and $\sin\varphi$ to find that solution with the minimum value of $|t|$, where the latter is given, according to (11), by the equation,

$$|t|=|2k\sin\tfrac{1}{2}\theta|^2=(q^2\sin^2\varphi)^{-1}\times\{(\mu^2+k^2+q^2)^2-4k^2q_1^2\}. \quad (12)$$

The solution is obviously $\sin\varphi=1$ and, by straightforward differentiation, $q_2=0$, $q_1^2=k^2+\mu^2$. This yields from (12), (11)

$$|t|\,(\text{minimum})=4\mu^2, \quad t^{\frac{1}{2}}=\pm 2\mu. \quad (13)$$

Thus, as long $|t|<4\mu^2$, the integral (3) as well as its derivatives with respect to t are guaranteed to exist.

We now show how the argument may be extended to yield more complete information on the analytic properties of $f(s,t)$ in t. We write (3) symbolically,

$$f=f^{(0)}-\lambda^2VGV \quad (14)$$

and by repeated use of the alternative forms of (6) find

$$f=\sum_{i=0}^{2n}f^{(i)}+R_{(2n+1)}$$
$$=f_{(2n+1)}+R_{(2n+1)}, \quad (15)$$

where

$$f^{(i)}=-(\lambda)^i(VG^{(0)})^iV, \quad (16)$$

and

$$R_{(2n+1)}(s,t)=-(\lambda)^{2n+2}(VG^{(0)})^nVGV(G^{(0)}V)^n$$
$$=-(4\pi)^{2n+1}\lambda^{2n+2}\int\frac{1}{\mu^2+(\mathbf{k}'-\mathbf{q}_1')^2}\times\frac{(2\pi)^{-3}}{s+i\epsilon-(q_1')^2}\frac{1}{\mu^2+(\mathbf{q}_1'-\mathbf{q}_2')^2}\times\frac{(2\pi)^{-3}}{s+i\epsilon-(q_2')^2}\cdots\frac{1}{\mu^2+(\mathbf{q}_n'-\mathbf{q}')^2}\times G(\mathbf{q}',\mathbf{q};s)\frac{1}{\mu^2+(\mathbf{q}-\mathbf{q}_n)^2}\times\frac{(2\pi)^{-3}}{s+i\epsilon-q_n^2}\cdots\frac{1}{\mu^2+(\mathbf{q}_1-\mathbf{k})^2}\times d^3q_1'\cdots d^3q_n'd^3q_n\cdots d^3q_1d^3q'd^3q. \quad (17)$$

We shall prove that $R_{(2n+1)}(s,t)$ is, for real s, analytic in t for $|t|<(2n+2)^2\mu^2$, whereas $f^{(i)}(s,t)$, $i\neq 0$ is analytic for all t except for a cut from $t=(i+1)^2\mu^2$ to ∞.

Toward this end, we perform first the angular integrations,

$$\int \frac{1}{\mu^2+(\mathbf{k}-\mathbf{q}_1)^2}d\Omega_1\frac{1}{\mu^2+(\mathbf{q}_1-\mathbf{q}_2)^2} \times d\Omega_2\cdots d\Omega_n\frac{1}{\mu^2+(\mathbf{q}_n-\mathbf{q})^2} \quad (18)$$

by successive application of the formula[5]

$$\int d\Omega'\frac{1}{\alpha_1-\mathbf{n}_1\cdot\mathbf{n}'}\frac{1}{\alpha_2-\mathbf{n}_2\cdot\mathbf{n}'} =-4\pi\int_{-1}^{1}\frac{dy}{(1-y^2)(\alpha_{12}-\mathbf{n}_1\cdot\mathbf{n}_2)}, \quad (19)$$

where

$$\alpha_{12}=(1-y^2)^{-1}\{\tfrac{1}{2}[\alpha_1(1+y)+\alpha_2(1-y)]^2-(1+y^2)\}, \quad (20)$$

since (19) clearly lends itself to iteration. Now let us suppose (19) to represent the integration over the direction of $\mathbf{q}_n$, the last of the integrations in (18). In that event,

$$\alpha_2=(\mu^2+q^2+q_n^2)/2qq_n,$$
$$\alpha_1=\alpha_1(q_1,q_2,\cdots q_{n-1},y_1,\cdots y_{n-1};\, q_n,k). \quad (21)$$

If we look for the solution of $\alpha_{12}-\mathbf{n}_1\cdot\mathbf{n}_2=0$, we are first led to repeat the reasoning involved in the transition from Eqs., (8) to (12) and (with $\sin\varphi=1$) find that possible solutions are given by

$$\sin\tfrac{1}{2}\theta=-\cos\gamma\cosh\tfrac{1}{2}\theta_2+i\sin\gamma(\cosh^2\tfrac{1}{2}\theta_2-1)^{\frac{1}{2}}, \quad (22)$$

where γ is defined as before (Eq. (8)), but now

$$\cosh\tfrac{1}{2}\theta_2=[q(1-y^2)]^{-1} \times\{\tfrac{1}{2}[\alpha_1(1+y)+\alpha_2(1-y)]^2-(1+y^2)\} =\alpha_{12}/q. \quad (23)$$

Equivalently (22) implies that

$$|t|=4k^2[\cosh^2\tfrac{1}{2}\theta_2-\sin^2\gamma]. \quad (24)$$

We must now minimize (24) with respect to its host of variables. This first of all requires the minimization of α_1 with respect to all variables, except q_n and k, i.e., with respect to the ones on which it alone depends. We shall assume (and prove imminently) that this minimum is in fact given by

$$\alpha_1=[(n\mu)^2+q_n^2+k^2]/2kq_n. \quad (25)$$

Next we minimize α_{12}, Eq. (20) with respect to q_n and y, in that order, for fixed q. This yields straightforwardly

$$\alpha_{12}=[(n+1)^2\mu^2+q^2+k^2]/2kq \quad (26)$$

and incidentally proves (25) if one gives a little thought to the structure of α_1. This sequence of minimizations has now reduced (24) to the form

$$|t|=q^{-2}\{[(n+1)^2\mu^2+k^2+q^2]^2-4k^2q^2\sin^2\gamma\}. \quad (26)$$

But the latter is exactly of the same structure as (12) with μ replaced by $(n+1)\mu$. We have thus proved that the minimum value of t for which $R_{(2n+1)}$ may possibly have a singularity is $(2n+2)^2\mu^2$.

It is now simple to see that our reasoning has also established the analytic properties in t of $f^{(n)}$, the $(n+1)st$ term in the Born series. For its dependence on t is given essentially by an integral of the form (18) with q replaced by k'. In the integral (19), then, if this is again the result of the last angular integration, $\mathbf{n}_1\cdot\mathbf{n}_2=\cos\theta$. Since α_{12} is now a real positive quantity, we deduce that $f^{(n)}$ is an analytic function of $\cos\theta$ except for a cut along the positive real axis starting at the minimum value of α_{12}. According to (26), with $q=|k'|=k$, the cut starts at

$$\cos\theta=1+[(n+1)^2\mu^2/2k^2]=1+(t/2k^2), \quad (27)$$

or $t=(n+1)^2\mu^2$, as asserted previously. On combining this result with the previous one for $R_{(2n+1)}$, we conclude that $f(s,t)$ is for real s analytic in t except along the indicated cuts.

We now apply these results to the dispersion relation,

$$f(s,t)=f^{(0)}(t)+\frac{1}{\pi}\int_0^\infty ds'Imf(s',t)/(s'-s-i\epsilon) +\sum_{l,i}p_{li}(t)/(s_{li}-s), \quad (28)$$

where the last term represents the contribution from the bound states of angular momentum l and energy s_{li}, $p_{li}(t)$ being polynomials in t.[6] Equation (28) has previously been established only for $-t<4\mu^2$. But both sides of (28) are analytic in t for all t exclusive of the poles and cut; Equation (28) therefore represents for real s a relation valid for all finite t. Moreover, since the right hand side of (28) can be continued in s from real s, it constitutes a proper analytic continuation of $f(s,t)$ for complex s, t. We have thus arrived at a complete knowledge of the analytic properties of $f(s,t)$ in its two variables, except that we still have no information on the limit of $f(s,t)$ as $|t|\to\infty$.

In the next section, we shall ascertain that there cannot be an essential singularity at ∞. At the moment, however, let us not presume this knowledge. By using the contour of Fig. 1, we can conclude only that

$$Imf(s,t)=\frac{1}{\pi}\int_{4\mu^2}^{T}\rho(s,t')/(t'-t) +\frac{1}{2\pi i}\oint Imf(s,t')/(t'-t), \quad (29)$$

[5] In part, the argument here follows that of Bowcock and Martin, reference 2.

[6] The definition of $p_{li}(t)$ is given in reference 3. It is a polynomial in t of degree l.

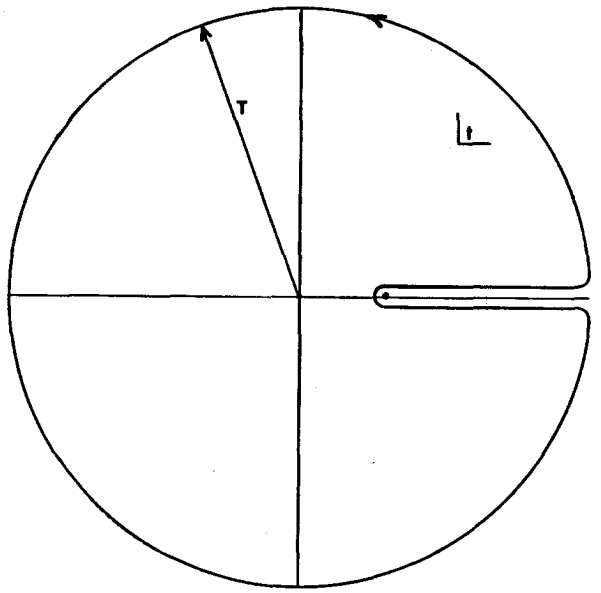

FIG. 1. Contour for the application of Cauchy's theorem to express the analytic properties of the scattering amplitude in the momentum transfer variable.

where T is the radius of a large circle in the t plane which represents the contour in the second integral.

Of course, the spectral function $\rho(s,t)$ gives the discontinuity across the cut of the function $Imf(s,t)$. In actuality, the lower limit of the first integral also depends on s and approaches $4\mu^2$ as $s\to\infty$. The only point about writing (29) and inserting it into (28) at this stage is to have an explicit representation from which can be deduced information on the partial waves. We thus find

$$f(s,t)=f^{(0)}(t)+\frac{1}{\pi^2}\int_0^\infty ds'\int_{4\mu^2}^T dt'\frac{\rho(s',t')}{(s'-s)(t'-t)}+\frac{1}{2i\pi^2}\int_0^\infty ds'\oint dt'\frac{Imf(s',t')}{(s'-s)(t'-t)}+\sum_{l,i}\frac{p_{li}(t)}{s_{li}-s}. \quad (30)$$

It need hardly be emphasized here that in general $Imf(s,t)$ is not the imaginary part of $f(s,t)$, though it is convenient to continue the labeling of the function. It should also be remarked that in writing (30), there is no a priori assumption that ultimately we shall establish a double representation with no subtractions other than those already exhibited explicitly in (30). This matter will in fact be discussed in Sec. IV.

III. ANALYTIC PROPERTIES OF PARTIAL WAVE AMPLITUDES

We now turn to the examination of the analytic properties of the partial wave amplitudes, defined as usual by the equation,

$$f(s,t)=\sum_{l=0}^\infty(2l+1)A_l(s)P_l(\cos\theta),$$
$$A_l(s)=\frac{1}{2}\int_{-1}^1 d(\cos\theta)P_l(\cos\theta)f(s,t) \quad (31)$$
$$=\exp[i\delta_l(s)]\sin\delta_l(s)/s^{\frac{1}{2}}.$$

By remembering that $t=-2s(1-\cos\theta)$ and recalling the definition (4) of $f^{(0)}(t)$, a representation for $A_l(s)$ is readily obtained by applying (31) to (30). We find that

$$A_l(s)=-\frac{\lambda}{2s}Q_l\left(\frac{\mu^2+2s}{2s}\right)+\frac{1}{\pi^2}\int_0^\infty ds'\int_{4\mu^2}^T dt'\frac{\rho(s',t')}{2s(s'-s)}Q_l\left(\frac{t'+2s}{2s}\right)+\frac{1}{2i\pi^2}\int_0^\infty ds'\oint dt'\frac{Imf(s',t')}{2s(s'-s)}Q_l\left(\frac{t'+2s}{2s}\right)+\sum_{l,i}\frac{\frac{1}{2}\int_{-1}^1 d(\cos\theta)P_l(\cos\theta)p_{li}(t)}{s_{li}-s}. \quad (32)$$

Here, $Q_l(v)$ are the Legendre functions of the second kind, defined by the equations,

$$Q_0(v)=\tfrac{1}{2}\ln(v+1)/(v-1),$$
$$Q_1(v)=-\tfrac{1}{2}+\tfrac{1}{2}v\ln(v+1)/(v-1), \quad (33)$$
$$(l+1)Q_{l+1}(v)=(2l+1)vQ_l(v)-lQ_{l-1}(v),$$

and the form of the bound state contributions follows from the remarks in footnote 6.

From (32), we now conclude that $A_l(s)$ is analytic for all finite s except for branch points at the origin (cut from $s=0$ to $s=\infty$) and at the points $s=-(\frac{1}{4})\mu^2$, $-\mu^2$, $-(9/4)\mu^2$, $-(16/4)\mu^2$, $\cdots$ (cut from $s=-\frac{1}{4}\mu^2$ to $s=-\infty$). The first branch point arises from the original behavior of $f(s,t)$ in s, the others from the behavior in t. Because of the additional integral in (32), the possibility is present that $A_l(s)$ has a ring of singularities on or beyond the circle $s=(\frac{1}{4})T$. The remainder of this section is concerned with establishing the actual behavior of $A_l(s)$ as $|s|\to\infty$.

We shall illustrate the method by a study of the s wave only. A separate paper devoted entirely to the study of the analytic properties of the partial waves will contain details of the general case. On dropping the subscript zero, we write for this case[7]

$$A(s)=\frac{f(\lambda;k)-f(\lambda;-k)}{2ikf(\lambda;-k)}, \quad (34)$$

[7] Equations (34)–(38) may be verified, for example, from the paper of W. Kohn, Revs. Modern Phys. 26, 292 (1954).

where

$$f(\lambda;k)=f(\lambda;k,r=0) \tag{35}$$

and $f(\lambda;k,r)$ is that solution of the wave-equation,

$$[(d^2/dr^2)+k^2-\lambda V(r)]f(r)=0, \tag{36}$$

which satisfies the boundary condition,

$$\lim_{r\to\infty} e^{ikr}f(\lambda;k,r)=1. \tag{37}$$

It is given equivalently by the solution of the integral equation,

$$f(\lambda;k,r)=e^{-ikr} +\frac{\lambda}{k}\int_r^\infty dr' \sin k(r'-r)V(r')f(\lambda;k,r'). \tag{38}$$

Our method will entail the study of the Born series for $f(\lambda;k,r)$. By iterating (38), we readily find that

$$f(\lambda;k,r)=\sum_{n=0}^{\infty} f^{(n)}(\lambda;k,r), \tag{39}$$

where

$$f^{(0)}(\lambda;k,r)=e^{-ikr},$$

$$f^{(n)}(\lambda;k,r)=\lambda^n e^{-ikr}\int_\mu^\infty \frac{d\alpha_1 \exp(-\alpha_1 r)}{\alpha_1(\alpha_1+2ik)} \times\int_\mu^\infty \frac{d\alpha_2 \exp(-\alpha_2 r)}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+2ik)} \times\cdots\int_\mu^\infty \frac{d\alpha_n \exp(-\alpha_n r)}{(\alpha_1+\alpha_2+\cdots+\alpha_n)(\alpha_1+\cdots+\alpha_n+2ik)}. \tag{40}$$

The series (39), (40) makes evident the branch points of $A(s)$ at $s=-(\frac{1}{4})\mu^2$, $-\mu^2$, $\cdots$. It is now easy to show (appendix A) that for sufficiently large $|k|$,

$$f^{(n)}(\lambda;k)=(\lambda \ln k/2ik)^n[1+O(1/\ln k)], \tag{41}$$

or

$$\left|\frac{f^{(n)}(\lambda;k)}{f^{(n-1)}(\lambda;k)}\right|=\frac{|\lambda|\ln|k|}{2|k|}\left[1+O\left(\frac{1}{\ln|k|}\right)\right]. \tag{42}$$

It is thus clear that for sufficiently large $|k|$, the series defined by (39), (40) converges uniformly in k. The same is true of the series with the sign of k reversed.[8] We can conclude from (34) that $A(s)$ is analytic in s for large s, except for the cuts excluded at the beginning of this section. Moreover, for sufficiently large s, it is dominated by its first Born approximation. A similar result can be proved for arbitrary l. We can thus derive the dispersion relation

$$A_l(s)=A_l^{(0)}(s)+\sum_i \frac{c_{li}^2}{s-s_{li}}+\frac{1}{\pi}\int_0^\infty \frac{ImA_l(s')ds'}{s'-s} +\frac{1}{\pi}\int_{-\infty}^{-\mu^2} \frac{Im[A_l(s')-A_l^{(0)}(s')]ds'}{s'-s}, \tag{43}$$

$$A_l^{(0)}(s)=-(\lambda/2s)Q_l[(\mu^2+2s)/2s].$$

The derivation, in particular of the bound state contributions, is given in Appendix B. We emphasize again that the convergence of the integral is guaranteed by our results on the Born Series.

IV. DEDUCTION OF THE MANDELSTAM REPRESENTATION

From our results on the behavior of $Al(s)$ for large s, we may be assured that $f(s,t)$ does not have an essential singularity in t for large t. We cannot in general conclude that $\mathrm{Im}f(s,t)\to 0$ as $|t|\to\infty$. If this were so (and in the absence of bound states), we could write

$$f(s,t)=f^{(0)}(t)+\frac{1}{\pi^2}\int_0^\infty ds\int_{4\mu^2}^\infty dt'\frac{\rho(s',t')}{(s'-s)(t'-t)}. \tag{44}$$

In the absence of bound states, this result is at least consistent with unitarity. If bound states occur, however, Blankenbecler *et al*[2] have shown that the corresponding form of (44) including the bound state terms, but excluding subtractions in t, is inconsistent with the unitary requirement.[9] The exact number of subtractions required in any given instance remains to be established, though it is intuitively clear that there must be an intimate connection between the answer to this question and the number and kinds of bound states.

V. CONCLUDING REMARKS

Several aspects of our study are worthy of additional comment. First, concerning the construction of the spectral function $\rho(s,t)$, we have nothing to add to previous remarks which make use of what is essentially perturbation theory or the combination of the latter with unitarity.[9] In any event it does not appear that the Mandelstam representation *per se* will find extensive application in the immediate future. Rather, it seems that the dispersion relations for the partial waves, Eq. (43) and its analogues in field theory are most relevant in this connection. Chew and Mandelstam[10] have formulated an approximate method of solving Eq. (13). It might be of considerable interest to test this approxima-

[8] For real k, these results are stated without proof by W. Kohn reference 7.

[9] See Blankenbecler *et al.*, footnote 2. The author is indebted to N. N. Khuri for emphasizing in a private communication that several assertions in the original version of this paper about the question of subtractions were incompatible with the results of this reference.

[10] G. F. Chew and S. Mandelstam, *Theory of the Low-Energy Pion-Pion Interaction* (Lawrence Radiation Laboratory preprint).

tion for the problem considered in the present paper, where exact numerical solutions of the scattering problem are easily obtained.

Finally, perhaps the most significant aspect of the present paper is the suggestion that some analogue of the method of Sec. II may yield an approach to the field theoretical problem. We hope to be able to amplify this remark on a future occasion.

ACKNOWLEDGMENTS

It is a pleasure to thank Mr. W. Bender, director, as well as the various members of the physics group, of RIAS, for their hospitality during the month of August, 1959, when this work was conceived and partially executed.

APPENDIX A

We seek the asymptotic value for large k of the integral, Eq. (40), with $r=0$. We set $\alpha_i=k\beta_i$ and find

$$f^{(n)}(\lambda;k)=\left(\frac{\lambda}{k}\right)^n\int_{(\mu/k)}^{\infty}\frac{d\beta_1}{\beta_1(\beta_1+2i)}\cdots \times\int_{(\mu/k)}^{\infty}\frac{d\beta_n}{(\beta_1+\cdots+\beta_n)(\beta_1+\cdots+\beta_n+2i)}. \quad \text{(A1)}$$

We perform the β_1 integral first. Since for large k the main contribution to (A1) comes from the lower limit, to obtain the leading term we may set $\beta_1=0$ throughout the integral, except for the factor β_1^{-1}, of course. In this approximation, the integral must be cut off at its upper limit. This gives the result after performing the β_1 integration but ignoring the contribution from the upper limit, that

$$f^{(n)}(\lambda;k)\sim\left(\frac{\lambda}{k}\right)^n\frac{1}{2i}\ln\left(\frac{k}{\mu}\right)\int_{\mu(/k)}^{\infty}\frac{d\beta_2}{\beta_2(\beta_2+2i)}\cdots \times\int_{(\mu/k)}^{\infty}\frac{d\beta_n}{(\beta_2+\cdots+\beta_n)(\beta_2+\cdots+\beta_n+2i)}. \quad \text{(A2)}$$

The same argument may now be repeated $n-1$ times to give

$$f^{(n)}(\lambda;k)\sim(\lambda\ln k/2ik)^n\{1+O[(\ln k)^{-1}]\}, \quad \text{(A3)}$$

as asserted in the text.

APPENDIX B

To derive the dispersion relation (43) for $A_l(s)$ we evaluate the integral,

$$A_l(s)-A_l^{(0)}(s)=\frac{1}{2\pi i}\oint\frac{[A_l(s')-A_l^{(0)}(s')]ds'}{s'-s} \quad \text{(B1)}$$

taken around the contour of Fig. 2 which excludes both poles and branch lines. In the limit as the large circle

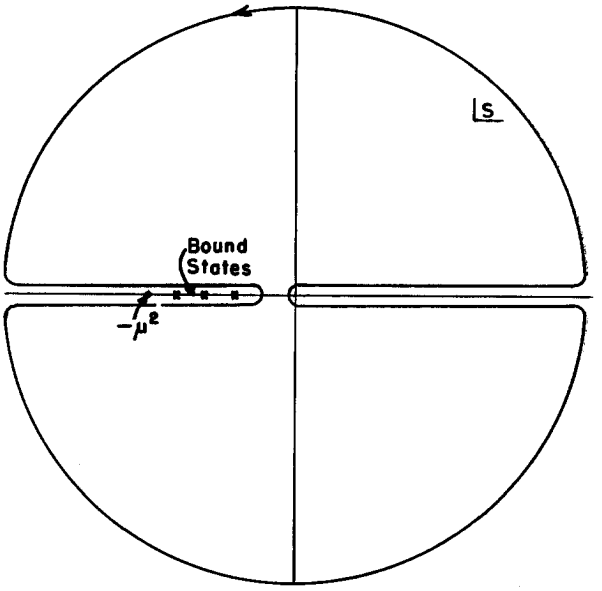

FIG. 2. Contour in the energy plane used in the derivation of a dispersion relation for the lth partial wave scattering amplitude.

recedes to infinity, it gives, of course, no contributions. As the remainder of the contour is deformed into an integral enclosing the cuts and encircling the poles due to bound states, we thus obtain in a standard way

$$A_l(s)=A_l^{(0)}(s)+\frac{1}{\pi}\int_0^{\infty}\frac{ImA_l(s')ds'}{s'-s} +\frac{1}{\pi}\int_{-\infty}^{-\mu^2}\frac{Im[A_l(s')-A_l^{(0)}(s')]ds'}{s'-s} -\frac{1}{2\pi i}\sum_i\int_j\frac{A_l(s')ds'}{s'-s}, \quad \text{(B2)}$$

where the last sum involves the typical integral $\int_j$ taken counterclockwise about an infinitesimal circle enclosing the jth pole.

To find the residue at this pole, we consider the formula for $A_l(s)$ which is the one-dimensional analog of Eq. (3), namely,

$$A_l(s)=A_l^{(0)}(s)-\frac{\lambda^2}{s}\int_0^{\infty}dr\int_0^{\infty}dr'U_l(kr)V(r) \times K_l(r,r';s)V(r')U_l(kr'), \quad \text{(B3)}$$

where $U_l(kr)$, proportional to the spherical Bessel function of order l,

$$U_l(kr)=krj_l(kr), \quad \text{(B4)}$$

is the solution of

$$[(d^2/dr^2)-(l(l+1)/r^2)+k^2]U_l(kr)=0, \quad \text{(B5)}$$

and $K_l(r,r';s)$ is the Green's function defined by the bilinear formula, the sum over bound and continuum states

$$K_l(r,r';s)=\sum_i \frac{\varphi_{li}(r)\varphi_{li}^*(r')}{s-s_{li}} + \int ds' \frac{\varphi_{ls'}(r)\varphi_{ls'}^*(r')}{s'-s}, \quad \text{(B6)}$$

where, for example $\varphi_{li}(r)$ satisfies

$$[(d^2/dr^2)-(l(l+1)/r^2)+s_{li}]\varphi_{li}(r)=\lambda V(r)\varphi_{li}(r). \quad \text{(B7)}$$

From (B2) and (B7), we thus find that the contribution of the bound states to (B2) is

$$\sum_i \frac{1}{s_{li}(s_{li}-s)}\left|\lambda\int U_{li}[(s_{li})^{\frac{1}{2}}r]V(r)\varphi_{li}(r)dr\right|^2. \quad \text{(B8)}$$

To evaluate the integral in (B8) we use the following formula obtainable from (B5) and (B7):

$$I_l=\lambda\int_0^\infty drU_l[(s_{li})^{\frac{1}{2}}r]V(r)\varphi_{li}(r) = \lim_{r\to\infty}[U_l(d/dr)\varphi_{li}-\varphi_{li}(d/dr)U_l]. \quad \text{(B9)}$$

Since we have

$$U_l[(s_{li})^{\frac{1}{2}}r]\sim(2i)^{-1}\{\exp[i(s_{li})^{\frac{1}{2}}r-i\tfrac{1}{2}l\pi] -\exp[-i(s_{li})^{\frac{1}{2}}r+i\tfrac{1}{2}l\pi]\}, \quad \text{(B10)}$$

and defining c_{li} by the equation,

$$\varphi_{li}(r)\sim c_{li}\exp[i(s_{li})^{\frac{1}{2}}r-i\tfrac{1}{2}l\pi] \quad \text{(B11)}$$

(remember that s_{li} is negative) we readily compute that

$$I_l=c_{li}s_{li}^{\frac{1}{2}}. \quad \text{(B12)}$$

We thus find that (B8) is equal to

$$\sum_i|c_{li}|^2/(s-s_{li}), \quad \text{(B13)}$$

as required in the text.

Integration in Functional Spaces and its Applications in Quantum Physics*†‡

I. M. GEL'FAND AND A. M. YAGLOM

(Received December 7, 1959)

This translation of the survey article by I. M. Gel'fand and A. M. Yaglom on the theory and applications of integration in functional spaces in problems of quantum physics was prepared because it was felt that such a review would be of interest and of use to mathematical physicists working in several different fields.

The article begins with a discussion of Wiener measure, after which the extension is made to the complex measure introduced by Feynman in his formulation of quantum mechanics, and examples are given of the use of these methods in quantum mechanics, quantum field theory, and quantum statistical physics. A comprehensive bibliography of works devoted to the theory and applications of functional integration methods is included.

AT the present time, the methods of measure theory and integration in functional spaces are widely applied only in the theory of random processes. It is already apparent today, however, that these methods can have great importance for a number of other branches of mathematics and theoretical physics (primarily for the theory of the differential equations of quantum physics and hydrodynamics). In the present paper, we attempt briefly to throw light on the question of the utilization of these methods in quantum physics[1]; at the same time, certain mathematical questions connected with this application will also be examined.

The beginning of the penetration of the methods of integration in functional spaces into quantum physics, apparently must be considered as having occurred in 1942 when the dissertation of R. Feynman on the principle of least action in quantum physics, containing a new derivation of Schrödinger's equation, was defended in Princeton. This dissertation was never published in full, but the part of greatest interest to us is contained in Feynman's article [1a§], which appeared in 1948 and recently was translated into the Russian language. In 1949, an interesting mathematical work of Kac appeared [2a] (cf. also [2b]), written, as the author indicates, under the strong influence of Feynman's dissertation. It is devoted to the calculation of the mean values of certain functionals over the trajectories of a Brownian particle with the help of the reduction of this problem to the solution of differential equations related to Schrödinger's equation. A special measure in the space of continuous functions which gives the distribution of probabilities of distinct trajectories of a Brownian particle (with neglect of its inertia) plays an extremely essential role in Kac's work. This measure had already been thoroughly studied in the twenties by N. Wiener, and is generally called "Wiener measure." Following Kac's article a series of works by scholars in different countries appeared developing these same ideas further (cf., for example, [4]–[8]). We note further that independently of these works, beginning in 1944, a long series of works was published by Cameron, Martin, and their collaborators devoted to the investigation of separate questions connected with Wiener measure in functional space (cf., for example, [9]–[11], as well as a series of other works by these same authors); in recent years, these investigations also partially coincided with works based on the ideas of Feynman and Kac (cf., references [10b] and [2], [4]).

We enumerated above a series of mathematical works developing the ideas of Feynman's work, [1a]. Among theoretical physicists, however, this work was not sufficiently appreciated in the beginning, apparently due to the novelty and unusualness of the idea of the possibility of integration in functional space. In any event, in the first several years after the appearance of reference [1a] containing in it a new mathematical apparatus, its utilization in physical investigations was almost only by its author. However, in 1954, suddenly two works appeared almost simultaneously which in fact were devoted to the transference of the methods which were utilized in [1a] for the case of nonrelativistic quantum mechanics to the quantum theory of fields (these are the works of Edwards and Peierls [13] and I. M. Gel'fand and R. A. Minlos [14]). It is curious to note that there are no references to [1a] in both of these works; in both cases, the authors were led by a completely different approach[2] to the necessity of introducing the apparatus of integration in functional spaces. Following [13] and [14], there immediately appeared a series of analogous works devoted to the exposition of the fundamental facts of quantum field theory with the utilization of functional integrals (references [15]–[23a] and host of others); notwith-

* A review article, presented at the Third All-Union Conference on the Theory of Probability and its Applications, in Leningrad, May 30 to June 4, 1955, under the title "Methods of the Theory of Random Processes in Quantum Physics."

† Translated from Uspekhi Matematicheskikh Nauk, Vol. XI (Jan.–Feb., 1956), p. 77, by A. A. Maradudin, U. of Maryland.

‡ This translation was partially supported by the U. S. Air Force Office of Scientific Research under Contract AF18(600)1315.

[1] Certain ways of utilizing analogous methods in statistical hydrodynamics (in the theory of turbulence) were mentioned in reference [12].

§ Numbers in brackets refer to references in the Bibliography at the end of the paper.

[2] In essence, the apparatus of functional integration appeared in the these works in solving differential equations in the infinite dimensional space of quantum field theory (the so-called "Schwinger equations" in variational derivatives).

standing the fact that no particularly valuable results have been obtained as yet with the aid of these methods, at the present time it is becoming ever more apparent that the future development of quantum field theory (connected, apparently, with the overcoming of a number of serious difficulties of a physical character) will find its mathematical language in just this apparatus.

We will emphasize again that after the appearance of Feynman's work [1a] investigations related to it were in fact carried out independently by mathematicians and physicists. In the present article, we will examine, in parallel as far as possible, works which are relevant here.

1. INTEGRATION OVER WIENER MEASURE IN FUNCTIONAL SPACE

Wiener Measure

We will begin with a definition of Wiener measure in the space of continuous functions and with an examination of certain purely mathematical questions connected with integration over this measure; such an order of exposition will be most convenient for all that follows. We examine a particle undergoing Brownian motion along the x axis under the influence of random impulses (for example, molecular) in the absence of any kind of systematic forces and being found at time $t=0$ at the origin of coordinates. Then, if we neglect the particle's inertia the distribution of probabilities for its position at time $t\neq 0$ will be described by the fundamental solution of the simplest diffusion equation,

$$\partial\psi/\partial t = D\partial^2\psi/\partial x^2 \tag{1.1}$$

(D, the diffusion coefficient, is connected in the case of Brownian motion under the influence of molecular forces to the mass and dimensions of the particle, the temperature and viscosity of the medium by the well-known Einstein relation), i.e., it will have a density equal to

$$p_t(x)=\frac{1}{(4\pi Dt)^{\frac{1}{2}}}\exp\left(-\frac{x^2}{4Dt}\right). \tag{1.2}$$

For simplicity, we will everywhere in the following assume that $D=\frac{1}{4}$; clearly, this can always be achieved with the help of a choice of a suitable system of units. Then, for the probability that the coordinate $x(t_1)$ of the particle at time t_1 will be found within the limits $a_1<x(t_1)<b_1$, at time t_2 within the limits $a_2<x(t_2)<b_2$, $\cdots$, at time t_n within the limits $a_n<x(t_n)<b_n$, where $0<t_1<t_2\cdots<t_n$, we obtain the formula,

$$\frac{1}{[\pi^n t_1(t_2-t_1)\cdots(t_n-t_{n-1})]^{\frac{1}{2}}}\int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_n}^{b_n}\exp\left[-\frac{x_1^2}{t_1}-\frac{(x_2-x_1)^2}{t_2-t_1}-\cdots-\frac{(x_n-x_{n-1})^2}{t_n-t_{n-1}}\right]dx_1dx_2\cdots dx_n. \tag{1.3}$$

The measure (1.3) of the cylindrical set ("quasi-intervals" in the terminology of Wiener) in the space of all functions, dictated by the conditions

$$a_1<x(t_1)<b_1,\ \cdots,\ a_n<x(t_n)<b_n,$$

can be extended (by a well-known theorem of A. N. Kolmogorov [24]) to the class of all functions $x(\tau)$ of a continuous parameter τ; this is Wiener measure. It was demonstrated by N. Wiener [3] that the corresponding measure is completely concentrated on the class of continuous (but not differentiable) functions $x(\tau)$, equal to zero at $\tau=0$, and satisfying the Lipshitz condition with an exponent of $\frac{1}{2}-\epsilon$ for any $\epsilon>0$. It was also demonstrated by him that for a wide class of functionals $F[x(\tau)]$ (in particular for all bounded and continuous functionals) in a space C of continuous functions in the interval $[0,t]$, such that $x(0)=0$, there exists an integral over Wiener measure, which can be calculated in the following manner: the curve $x(\tau)$ is replaced by the broken one $x_n(\tau)$, which coincides with $x(\tau)$ at the points,

$$x(0)=0,\quad x(t_1)=x_1,\quad x(t_2)=x_2,\quad \cdots,\quad x(t_n)=x_n$$

(where $t_n=t$ and the points t_1, t_2, $\cdots$, t_{n-1} divide the interval $[0,t]$ into n equal parts of length $\Delta t=t/n$), after which the integral of $F[x(\tau)]$ over Wiener measure is defined as

$$\int_C F[x(\tau)]d_wx = \lim_{n\to\infty}\frac{1}{(\pi\Delta t)^{n/2}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}F(x_1,\cdots,x_n)\times\exp\left[-\frac{x_1^2}{\Delta t}-\sum_{j=1}^{n-1}\frac{(x_{j+1}-x_j)^2}{\Delta t}\right]dx_1\cdots dx_n, \tag{1.4}$$

where $\int_C\cdots d_wx$ denotes an integral over the Wiener measure d_wx, extended over the entire space C, while the function $F(x_1\cdots x_n)$ of n variables in the right hand side of the formula is the value of our functional on substituting in place of $x(\tau)$ the broken curve $x_n(\tau)$,

$$F(x_1,\cdots,x_n)=F[x_n(\tau)]. \tag{1.5}$$

Symbolically, the expression (1.4) of the integral over Wiener measure can be written in the form,

$$\int_C F[x(t)]d_wx = \frac{1}{N}\int_{-\infty}^{\infty}\cdots\int F[x(t)]\times\exp\left\{-\int_0^t\left[\frac{dx(\tau)}{d\tau}\right]^2d\tau\right\}\prod_0^t dx(t), \tag{1.4'}$$

where N is a constant normalization factor which is determined by the condition that

$$\int_C d_w x = 1.$$

Indeed, if

$$\prod_0^t dx(\tau)$$

is replaced by the product of differentials of the "coordinates" $x(\tau)$ at the finite number of points $\tau = t_1, t_2, \cdots, t_n$, and the integral

$$\int_0^t [dx(\tau)/d\tau]^2 d\tau$$

in the exponent by the corresponding sum of difference ratios

$$\sum_{j=0}^{n} [x(t_{j+1}) - x(t_j)]^2/(t_{j+1} - t_j),$$

then (1.4′) goes over into (1.4).

Examples

1. In the simplest case, when $F[x(\tau)]$ depends only on the value of the function $x(\tau)$ at a finite number of points, the integral (1.4) by definition becomes an ordinary finite dimensional integral; its evaluation in this case does not involve any difficulties in principle. In the particular case when $F[x(\tau)] = x(t_1)x(t_2)\cdots x(t_k)$, we obtain the well known expressions for the moments of the Brownian motion process which we have described,

$$\int_C x(t_1) d_w x \equiv 0$$

and in general

$$\int_C x(t_1)x(t_2)\cdots x(t_{2k+1}) d_w x \equiv 0, \tag{1.6}$$

$$\int_C x(t_1)x(t_2) d_w x \equiv b(t_1,t_2) = \tfrac{1}{2}\min(t_1,t_2), \tag{1.7}$$

$$\int_C x(t_1)x(t_2)\cdots x(t_{2k}) d_w x \equiv b(t_1,t_2,\cdots,t_{2k}) = \sum b(t_{i_1},t_{i_2})\cdot b(t_{i_3},t_{i_4})\cdots b(t_{i_{2k-1}},t_{i_{2k}}), \tag{1.8}$$

where the summation in the last formula extends over all possible partitions of the $2k$ indices $1, 2, \cdots, 2k$ into k pairs; (i_1, i_2), (i_3, i_4), $\cdots$, (i_{2k-1}, i_{2k}).

2. We examine now one instructive example in which $F[x(\tau)]$ depends on all the values of the function $x(\tau)$. Let

$$F[x(\tau)] = \exp\left\{\lambda \int_0^t p(\tau)x^2(\tau) d\tau\right\},$$

where λ is a real number and $p(\tau) \geqslant 0$. By substituting into (1.4),

$$F(x_1,x_2,\cdots,x_n) = \exp\{\lambda \sum_{j=1}^{n} p(j\Delta t)x^2(j\Delta t)\Delta t\} \equiv \exp\{\lambda \Delta t \sum_{j=1}^{n} p_j x_j^2\}, \tag{1.9}$$

where

$$\Delta t = t/n, \quad p_j = p(j\Delta t), \quad x_j = x(j\Delta t) \tag{1.10}$$

and making use of the well-known formula,

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \exp(-\sum_{j,k=1}^{n} a_{jk}x_j x_k) dx_1\cdots dx_n = \pi^{n/2}[\det(a_{jk})]^{-\frac{1}{2}}, \tag{1.11}$$

(which is valid on the assumption that (a_{jk}) is a positive definite matrix) we easily find that

$$\int_C \exp\left\{\lambda \int_0^t p^2(\tau)x^2(\tau) d\tau\right\} d_w x = \lim_{n\to\infty} [D_1^{(n)}]^{-\frac{1}{2}}, \tag{1.12}$$

where $D_1^{(n)}$ is the nth order determinant,

$$D_1^{(n)} = \begin{vmatrix} 2-\lambda(\Delta t)^2 p_1 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2-\lambda(\Delta t)^2 p_2 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 2-\lambda(\Delta t)^2 p_3 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 2-\lambda(\Delta t)^2 p_{n-1} & -1 \\ 0 & 0 & 0 & \cdots & -1 & 1-\lambda(\Delta t)^2 p_n \end{vmatrix}. \tag{1.13}$$

To find the limit on the right hand side of (1.12), one can make use of the following argument. We denote by $D_k^{(n)}$ the principal minor of $(n-k+1)$st order of the determinant (1.13) situated in the lower right hand corner; then, expanding the determinant $D_k^{(n)}$ in terms of the elements of the first row, we readily obtain the recurrence relation,

$$D_k^{(n)} = [2-\lambda(\Delta t)^2 p_k] D_{k+1}^{(n)} - D_{k+2}^{(n)}, \quad (1 \leqslant k \leqslant n-2). \tag{1.14}$$

If instead of $D_k^{(n)}$ we write $D^{(n)}(k\Delta t)$, then Eq. (1.14) can be rewritten in the form,

$$\frac{D^{(n)}(k\Delta t)-2D^{(n)}((k+1)\Delta t)+D^{(n)}((k+2)\Delta t)}{(\Delta t)^2} = -\lambda p(k\Delta t)D^{(n)}((k+1)\Delta t). \quad (1.14')$$

We note further that $D_n^{(n)}=1-\lambda(\Delta t)^2 p_n$ and $D_{n-1}^{(n)}=1-\lambda(\Delta t)^2(2p_n+p_{n-1})+\lambda^2(\Delta t)^4 p_n p_{n-1}$, so that

$$D^{(n)}(n\Delta t)=1-\lambda(\Delta t)^2 p_n$$
$$\frac{D^{(n)}(n\Delta t)-D^{(n)}((n-1)\Delta t)}{\Delta t}=\lambda(\Delta t)[p_{n-1}+p_n-\lambda(\Delta t)^2 p_n p_{n-1}]. \quad (1.15)$$

In this way, $D^{(n)}(k\Delta t)$ will be the solution of the difference Eq. (1.14') satisfying the conditions (1.15). From this, it may be derived that as $n\to\infty$ the value of $D_k^{(n)}$ approaches the value at the point $k\Delta t$ of the continuous function $D(\tau)$ given by the solution to the differential equation,

$$[d^2D(\tau)/d\tau^2]+\lambda p(\tau)D(\tau)=0 \quad (1.16)$$

subject to the conditions,

$$D(t)=1, \quad D'(t)=0. \quad (1.17)$$

Corresponding to this, in view of Eq. (1.12),

$$\int_C \exp\left\{\lambda\int_0^t p(\tau)x^2(\tau)d\tau\right\}d_w x=[D(0)]^{-\frac{1}{2}}. \quad (1.18)$$

It is natural that all of these arguments will be legitimate only when, in representing (1.4) in the form (1.11), we indeed have in the exponent a positive definite quadratic form, i.e., when $\lambda<\min(\lambda_1,\lambda_2,\cdots,\lambda_n)$ where $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the zeros of the determinant (1.13); it can be shown that in the limit as $n\to\infty$, these zeros become the characteristic values of Eq. (1.16) subject to the boundary conditions $D'(t)=0$, $D(0)=0$, so that λ must be smaller than the smallest of these characteristic values. In the simplest particular case $p(\tau)\equiv 1$, we will have that $D(\tau)=\cos(\lambda)^{\frac{1}{2}}(t-\tau)$ from which we obtain

$$\int_C \exp\left\{\lambda\int_0^t x^2(\tau)d\tau\right\}d_w x=[\sec t(\lambda)^{\frac{1}{2}}]^{\frac{1}{2}} \quad \text{for } \lambda^{\frac{1}{2}}<\pi/2t. \quad (1.19)$$

The result (1.18) was first obtained by Cameron and Martin [9c] with the aid of a method of which we will say more below. The derivation presented here was pointed out in the work of Montroll [25], which also contains certain other examples of the direct evaluation of integrals over Wiener measure.

"Conditional" Wiener Measure

If in Eq. (1.3) we put $t_n=t$, fix the value of $x_n=x(t)=X$ and do not integrate over this coordinate, then the corresponding integral defines a certain measure of the cylindrical set,

$$a_1<x(t_1)<b_1, \cdots, a_{n-1}<x(t_{n-1})<b_{n-1}$$

of a functional space, which can then be extended to the class of all functions $x(\tau)$, $0<\tau<t$. It is not difficult to verify that the measure of all space here equals $(1/(\pi t)^{\frac{1}{2}})\exp(-X^2/t)$; the class of functions corresponding to the measure in this case will be the class of continuous functions $x(\tau)$, which are equal to zero for $\tau=0$, equal to X at $\tau=t$, and satisfy the very same conditions of smoothness as in the case of ordinary Wiener measure. We will denote the integral of a functional $F[x(\tau)]$ over this new measure extending over the class $C_{t,X}$ of all continuous functions satisfying the conditions $x(0)=0$, $x(t)=X$, by

$$\int_{C_{t,X}} F[x(\tau)]d_{w(t,X)}x; \quad (1.20)$$

this integral can be evaluated with the help of a formula differing from (1.4) only in the absence of an integration over dx_n on the right-hand side. It is clear that

$$\int F[x(\tau)]d_w x=\int_{-\infty}^{\infty}\left\{\int_{C_{t,X}} F[x(\tau)]d_{w(t,X)}x\right\}dX; \quad (1.21)$$

so that knowing the integral over $d_{w(t,X)}x$, we can determine the integral of the same functional over the measure $d_w x$ with the help of a single quadrature. Together with the measure $d_{w(t,X)}x$, we will also investigate the normalized measure

$$d_{w(t,X)}{}^*x=(\pi t)^{\frac{1}{2}}\exp(X^2/t)d_{w(t,X)}x; \quad (1.22)$$

the integral of a functional over the measure $d_{w(t,X)}{}^*x$ will have the meaning of a conditional mathematical expectation of the corresponding functional over the trajectory of a Brownian particle under the condition that this trajectory passes through the given point X at the time t.

Together with the measures $d_{w(t,X)}x$ and $d_{w(t,X)}{}^*x$ assigned on the class $C_{t,X}$ of continuous functions satisfying the conditions $x(0)=0$ and $x(t)=X$, it sometimes is necessary to investigate likewise the analogous measures $d_{w(t_0,X_0;\,t,X)}x$ and $d_{w(t_0,X_0;\,t,X)}{}^*x$, assigned on the classes $C_{t_0,X_0;\,t,X}$ of continuous functions satisfying the conditions $x(t_0)=X_0$ and $x(t)=X$. Integration over these new measures, however, easily reduces to the integration studied by us over the set of curves starting

at the point $x(0)=0$, namely[3]

$$\int_{C_{t_0,X_0;t,X}} F[x(\tau)]d_{w(t_0,X_0;\ t,X)}x = \int_{C_{t-t_0;X-X_0}} F[X_0+x(\tau)]d_{w(t-t_0;\ X-X_0)}x \quad (1.23)$$

and analogously for the normalized measures "with an asterisk" (these equations can be taken, if necessary, as the definition of the measures $d_{w(t_0,X_0;\ t,X)}x$ and $d_{w(t_0,X_0;\ t,X)}x$).

Examples

1. Just as in the case of ordinary Wiener measure the integral over $d_{w(t,X)}x$ or $d_{w(t,X)}{}^*x$ of a functional $F[x(\tau)]$, which depends only on the value of the function $x(\tau)$ at a finite number of points, reduces to a simple finite dimensional integral. In the case when $F[x(\tau)]=x(t_1)x(t_2)\cdots x(t_k)$, the integrals over the measure $d_{w(t,X)}{}^*x$ are easily handled and give conditional moments of the trajectory of a Brownian particle (under the condition that at the time t this trajectory passes through the point X):

$$\int_{C_{t,X}} x(t_1)d_{w(t,X)}{}^*x = \frac{t_1}{t}X, \quad (1.24)$$

and in general,

$$\int_{C_{t,X}} \left[x(t_1)-\frac{t_1}{t}X\right]\left[x(t_2)-\frac{t_2}{t}X\right]\cdots \times\left[x(t_{2k+1})-\frac{t_{2k+1}}{t}X\right]d_{w(t,X)}{}^*x \equiv 0, \quad (1.24')$$

$$\int_{C_{t,X}} \left[x(t_1)-\frac{t_1}{t}X\right]\left[x(t_2)-\frac{t_2}{t}X\right]d_{w(t,X)}{}^*x \equiv b_t(t_1,t_2)=\frac{t_1(t_1-t_2)}{2t}, \quad t_1\leqslant t_2, \quad (1.25)$$

$$\int_{C_{t,X}} \left[x(t_1)-\frac{t_1}{t}X\right]\left[x(t_2)-\frac{t_2}{t}X\right]\cdots \times\left[x(t_{2k})-\frac{t_{2k}}{t}X\right]d_{w(t,X)}{}^*x \equiv b_t(t_1,t_2,\cdots,t_{2k}) = \sum b_t(t_{i_1},t_{i_2})b_t(t_{i_3},t_{i_4})\cdots b_t(t_{i_{2k-1}},t_{i_{2k}}), \quad (1.26)$$

where the summation in the last formula extends over all possible partitions of the $2k$ indices $1, 2, \cdots, 2k$ into k pairs: (i_1,i_2), (i_3,i_4), $\cdots$, (i_{2k-1},i_{2k}).

[3] If the functional $F[x(\tau)]=F[x(\tau),\tau]$ depends explicitly on τ, then $F[X_0+x(\tau), t_0+\tau]$ must appear on the right side of this formula under the integral sign.

2. We also examine the integral

$$\int_{C_{t,X}} \exp\left[\lambda\int_0^t p(\tau)x^2(\tau)d\tau\right]d_{w(t,X)}x = I, \quad (1.27)$$

where

$$A=(a_{jk}), \quad c^2=\sum_{j,k} a_{jk}b_jb_k, \quad (a_{jk})^{-1}=A^{-1},$$

similar to (1.18). By making use again of formulas (1.9) and (1.4) (in formula (1.4) it must now be considered that $x_n=X$, and that it is not necessary to carry out an integration over x_n), we are led to a finite dimensional integral differing from (1.11) only in the presence of the term $2(\Delta t)^{-1}Xx_{n-1}$, linearly dependent on x_{n-1}, in the exponential under the integral sign. We now make use of the general formula,

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} F\Big(\sum_{j=1}^{n-1} b_jx_j\Big)\exp\Big(-\sum_{j,k=1}^{n-1} a_{jk}x_jx_k\Big)dx_1\cdots dx_{n-1} = \frac{\pi^{(n-2)/2}}{[c^2\det A]^{\frac{1}{2}}}\int_{-\infty}^{\infty} F(u)\exp(-u^2/c^2)du \quad (1.28)$$

(this formula is most easily obtained from probability theoretic considerations[4]). With its aid we introduce a finite dimensional integral of interest to us in the form,

$$I_n=\frac{\exp(-X^2/\Delta t)}{\pi[c^2D_1{}^{(n-1)}]}\int_{-\infty}^{\infty}\exp[u-(u^2/c^2)]du = \frac{\exp[-(X^2/\Delta t)+(c^2/4)]}{[\pi D_1{}^{(n-1)}]^{\frac{1}{2}}}, \quad c^2=\frac{4X^2}{\Delta t}\frac{D_1{}^{(n-2)}}{D_1{}^{(n-1)}},$$

where $D_1{}^{(n-1)}$ is the minor of $(n-1)$st order lying in the upper left hand corner of the determinant (1.13), the lower row of which is moreover multiplied by Δt, while $D_1{}^{(n-2)}$ is the analogously transformed upper left hand principal minor of $(n-2)$nd order of the same determinant (1.13). The subsequent limiting process is carried out analogously to the limiting process in formula (1.12); by taking into account that the exponent in the formula for I_n after the substitution of the value of c^2 transforms into

$$-\frac{X^2}{D_1{}^{(n-1)}}\frac{D_1{}^{(n-1)}-D_1{}^{(n-2)}}{\Delta t},$$

[4] After the appropriate normalization the integral (1.28) expresses the mean value of the function $F(u)$ where $u=\sum b_jx_j$ is a linear combination of random variables distributed according to a multidimensional normal law with a zero vector of mean values and a matrix of second moments equal to $\frac{1}{2}(a_{jk})^{-1}$; from this, it develops that u is also normally distributed with a mean value 0 and a dispersion equal to $c^2/2$.

we easily obtain the final result,

$$I=[\pi D_1(0)]^{-\frac{1}{2}}\exp\left[-\frac{D(0)}{D_1(0)}X^2\right], \qquad (1.29)$$

where $D(0)$ has the same meaning as in formula (1.18), while $D_1(0)$ is the value at zero of the solution to the differential equation (1.16) satisfying the conditions,

$$D_1(t)=0, \quad D_1'(t)=-1. \qquad (1.30)$$

By integrating the left side of formula (1.29) over X from $-\infty$ to $+\infty$, in view of (1.21), we obtain (1.18) again.

In the special case $p(\tau)=1$, we have

$$D_1(\tau)=\frac{1}{(\lambda)^{\frac{1}{2}}}\sin(\lambda)^{\frac{1}{2}}(t-\tau)$$

from which

$$\int_{C_{t,X}}\exp\left[\lambda\int_0^t x^2(\tau)d\tau\right]d_{w(t,X)}x$$

$$=\left[\frac{(\lambda)^{\frac{1}{2}}}{\pi}\operatorname{cosec}t(\lambda)^{\frac{1}{2}}\right]^{\frac{1}{2}}\exp\{-(\lambda)^{\frac{1}{2}}\operatorname{ctg}t(\lambda)^{\frac{1}{2}}X^2\} \qquad (1.31)$$

(it is obvious that this formula is correct only for $\lambda^{\frac{1}{2}}<\pi/t$).

Integral over Wiener Measure of the Functional $\exp\{-\int_0^t V[x(\tau)]d\tau\}$

The direct evaluation of the integral over Wiener measure of the functional $F[x(\tau)]$ with the aid of formula (1.4) can be accomplished only in extremely special isolated cases. As a rule, an explicit formula for the finite dimensional integral on the right hand side of (1.4) will be absent, and it will not be possible to carry out the limiting process in this formula in an effective way.[5] For this reason, those results are very interesting which relate the evaluation of such integrals to more traditional mathematical problems. The most important of such results is the result of Kac [2], which is concerned with the connection of integrals over Wiener measure of functionals of the type

$$\exp\left\{-\int_0^t V[x(\tau)]d\tau\right\},$$

where $V(x)$ is a sufficiently "good" function (for example, continuous and bounded from below[6]) with the solution of a certain special parabolic equation.

We will examine firstly the integral of our functional over the "conditional" Wiener measure $d_{w(t,X)}x$:

$$\psi(X,t)=\int_{C_{t,X}}\exp\left\{-\int_0^t V[x(\tau)]d\tau\right\}d_{w(t,X)}x. \qquad (1.32)$$

Then, according to reference [2] the function $\psi(X,t)$ for $t>0$ will be the solution which approaches zero as $X\to\pm\infty$ of the following parabolic differential equation,

$$\frac{\partial\psi(X,t)}{\partial t}=\frac{1}{4}\frac{\partial^2\psi(X,t)}{\partial X^2}-V(X)\psi(X,t), \qquad (1.33)$$

satisfying the condition $\psi(X,0)=\delta(X)$, where $\delta(X)$ is Dirac's δ function. In view of (1.21), we also obtain from this a simple formula for the integral of

$$\exp\left\{-\int_0^t V[x(\tau)]d\tau\right\}$$

over ordinary Wiener measure,

$$\int_C\exp\left\{-\int_0^t V[x(\tau)]d\tau\right\}d_wx$$

$$\equiv\frac{1}{N}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left(-\int_0^t\{V[x(\tau)]+\dot{x}^2(\tau)\}d\tau\right)$$

$$\times\prod_0^t dx(\tau)=\int_{-\infty}^{\infty}\psi(X,t)dX. \qquad (1.34)$$

We note further that the solution $\psi(X,t;X_0)$ of Eq. (1.33) satisfying the condition $\psi(X,0)=\delta(X-X_0)$ may also be introduced,

$$\psi(X,t;X_0)$$

$$=\int_{C_{t;X-X_0}}\exp\left\{-\int^t V[X_0+x(\tau)]d\tau\right\}d_{w(t;\,x-x_0)x}, \qquad (1.35)$$

and, in general, the fundamental solution $\psi(X,t;X_0,t_0)$ of the differential Eq. (1.33) satisfying the condition $\psi(X,t_0)=\delta(X-X_0)$ for $t>t_0$ will be equal to

$$\psi(X,t;X_0,t_0)$$

$$=\int_{C_{t_0,X_0;t,X}}\exp\left\{-\int_{t_0}^t V[x(\tau)]d\tau\right\}d_{w(t_0,X_0;\,t,X)x} \qquad (1.36)$$

[cf. (1.23)]. A general solution to Eq. (1.33) for arbitrary initial values $\psi(X,t_0)=\psi_0(X)$ is also easily constructed from the function (1.36); such a solution is

$$\psi(X,t)=\int_{-\infty}^{\infty}\psi_0(X_0)\psi(X,t;X_0,t_0)dX_0. \qquad (1.37)$$

[5] It is possible, of course, to raise the question of the numerical integration of integrals over Wiener measure; the first results in this direction are contained in work of Cameron [10a]. However, we will not concern ourselves with this question here.

[6] We note that it is entirely possible to consider that the function $V(x)$ also depends explicitly on τ; in what follows we will write $V[x(\tau)]$, and not $V[x(\tau),\tau]$ only for simplification of the notation.

The simplest method of deriving the results indicated here apparently is the method proposed by Feynman in applying it to Schrödinger's equation; a rigorous proof of this method (in application to the case examined by us here) was shown by Blanc-Lapierre and Fortet [6] and independently by E. B. Dynkin [7a]. Moreover, underlying this method is the circumstance that in view of the independence of the increments of the process $x(\tau)$, the Brownian motion we have described, for non-intersecting time intervals, the function (1.36), which gives the mean value of the functional,

$$\exp\left\{-\int_{t_0}^{t} V[x(\tau)]d\tau\right\}$$

along all trajectories of this process starting at time t_0 at point X_0 and ending at time t at point X, will for $t_0<t_1<t$ satisfy the relation,

$$\psi(X,t;\,X_0,t_0) = \int_{-\infty}^{\infty} \psi(X_1,t_1;\,X_0,t_0)\psi(X,t;\,X_1,t_1)dX_1, \quad (1.38)$$

analogous to the well-known Smoluchowski-Kolmogorov relation for the transition probabilities of a random Markov process. From this with certain simple conditions of regularity placed on $V(x)$, it is possible to obtain Eq. (1.33) in a completely analogous fashion to Kolmogorov's derivation of the basic differential equations for the transition probabilities of Markov random processes (cf., for example, [26], §52); the term $V(X)\psi(X,t)$, which distinguishes (1.33) from the ordinary diffusion equation (the Fokker-Planck-Kolmogorov equation for the process of Brownian motion) arises here from the fact that the function $\psi(X_1, t+\Delta t;\, X,t)$ does not satisfy the normalization condition,

$$\int_{-\infty}^{\infty} \psi(t+\Delta t,\, X_1;\, t,\, X)dX_1 \neq 1$$

(but on the other hand in view of (1.36),

$$\lim_{\Delta t\to 0}\left[1-\int_{-\infty}^{\infty}\psi(X_1,\, t+\Delta t;\, X,\, t)dX_1\right]=V(x)\Big).$$

Another very general way of obtaining this and a series of results related to it is found in the work of E. B. Dynkin [7b].

It may be thought that the possibility of representing the solution to the differential Eq. (1.33) in the form of a functional integral, following from formulas (1.35) and (1.37), has a very general character. If we first replace the differential equation by a difference equation and solve it step by step starting from the initial values, then the value of the solution at the point (X,t) appears in the form of a sum extending over all preceding points of the net; this sum may be considered as the sum of the contributions to the solution from the individual broken curves connecting the region of initial values with the point (X,t) through a sequence of neighboring points of the net.

The final expression of the solution in the form of a functional integral can be considered as a similar sum, obtained after a limiting process, of contributions to the solution from all possible continuous paths connecting the region of initial values with the final point. It is natural to suppose that such an expression for the solution will be possible for a wide class of problems and may present a definite interest for the theory of differential equations.[7] It would be interesting to examine the question of a similar expression for the solution of the Cauchy problem for a number of differential equations differing from (1.33), for which this problem is well-posed (for example, for the solution of the equation $\partial u/\partial t = -\partial^4 u/\partial x^4$). In the following, in the example of Schrödinger's equation, we will see that investigations of even simple differential equations by this means can meet with very serious difficulties.

The derivation of Eq. (1.33), which has been indicated here, will not in fact be changed if instead of the integral of

$$\exp\left\{-\int_{0}^{t} V[x(\tau)]d\tau\right\}$$

over Wiener measure, i.e., of the mean value of this functional over the trajectories of the process $x(\tau)$, of the previously described Brownian motion, we examine the mean value of the same functional over the trajectories of an arbitrary continuous Markov process $x(\tau)$. In this case, and in the general case, we arrive at a parabolic equation of the second order differing from the differential equation for the transition probabilities of the corresponding process only in the term $-V(X)\psi(X,t)$ on the right side; it is significant that the equation obtained (as is apparent from its derivation) depends only on the first and second moments of the increments of the process for small time intervals, but does not depend on all the other statistical characteristics of $x(\tau)$. This circumstance makes it possible to utilize results relating to integrals of functionals over Wiener measure for the proof of a wide class of meaningful limit theorems for the sums of independent terms (see the remarks on the "arcsine law" at the end of example 1 below, and also the interesting work of U. V. Prohorov [27], which is devoted to a general analysis of such an approach to the proof of limit theorems). On the other hand, from here it is possible to obtain one more method of proof of theorems about the connection of the function (1.32) with the solution to the differential

[7] It is possible, moreover, that in a number of cases the measure itself will not exist, and there will be only a certain method of integration for a wide class of "sufficiently good" functionals, so that in this case instead of speaking of measure we must speak of "distributions" in functional spaces in the sense of Schwartz, i.e., of "generalized measure".

equation (1.33): by replacing the process $x(\tau)$ by a sum of discrete independent random values, the increments of this process for nonintersecting small time intervals, and arguing that for the limit relations only the first two moments of the corresponding quantities are important but not their distribution functions, one can assume that all of these quantities assume only two values $+\Delta$ and $-\Delta$ (each with probability one-half), after which the equation for $\psi(X,t;X_0,t_0)$ is easily obtained from a probability theoretic argument in terms of finite differences, going over in the limit into the sought for differential equation. This method of deriving Eq. (1.33) was actually made use of in the work of Kac [2a].

In the case that the function $V(X)$ does not depend explicitly on τ, the calculation of the solutions $\psi(X,t;X_0)=\psi(X,t;X_0,0)$ and

$$\psi(X,t;X_0,t_0)=\psi(X,t-t_0;X_0,0)=\psi(X,t-t_0;X_0)$$

of the parabolic equation (1.33) can be related to the solution of the ordinary differential equation,

$$\frac{1}{4}\frac{d^2\varphi(X)}{dX^2}-[V(X)+E]\varphi(X)=0. \tag{1.39}$$

Actually, solving Eq. (1.33) with the aid of Laplace transforms, we find that the function,

$$\varphi_E(X)=\int_0^\infty \psi(X,t;X_0)e^{-Et}dt \tag{1.40}$$

will satisfy the differential equation (1.39); the conditions that $\psi(X,t;X_0)$ approach zero as $X\to\pm\infty$ and that it reduce to $\delta(X-X_0)$ as $t\to 0$ go over into the following conditions on $\varphi_E(X)$:

$$\varphi_E(X)\to 0 \text{ as } X\to\pm\infty;$$
$$\left.\frac{d\varphi_E}{dX}\right|_{X=X_0+0}-\left.\frac{d\varphi_E}{dX}\right|_{X=X_0-0}=-4. \tag{1.41}$$

(The function $\varphi(X)$ itself will be continuous at the point X_0.[8]) The conditions (1.41) uniquely specify the solution of Eq. (1.39) we need; after it is found, the function $\psi(X,t;X_0)$ is determined from the formulas for Laplace integral transforms.

[8] Making use of formula (1.40) and the conditions which $\psi(X,t;X_0)$ satisfies, it is easy to prove that

$$|\varphi(X_0+\epsilon)-\varphi(X_0-\epsilon)|\to 0$$

as $\epsilon\to 0$; after that from the equality,

$$\int_{X_0-\epsilon}^{X_0+\epsilon}\int_0^\infty\left[\frac{\partial\psi(X,t;X_0)}{\partial t}-\frac{1}{4}\frac{\partial^2\psi}{\partial X^2}+V(X)\psi\right]e^{-Et}dtdX=0$$

and the condition $\psi(X,t;X_0)\to\delta(X-X_0)$ as $t\to 0$, the relation

$$\lim_{\epsilon\to 0}\left[\frac{\partial\varphi_E(X_0+\epsilon)}{\partial X}-\frac{\partial\varphi_E(X_0-\epsilon)}{\partial X}\right]=-4$$

is derived.

A different method of reducing the problem of finding the function $\psi(X,t;X_0)$ to the solution of the differential Eq. (1.39) is connected with the use of the expansion of this function in a series of characteristic functions. Let us suppose for simplicity that Eq. (1.39) has a purely discrete spectrum of characteristic values $E_1, E_2, \cdots, E_n, \cdots$, to which correspond normalized characteristic functions $\varphi_1(X), \varphi_2(X), \cdots, \varphi_n(X), \cdots$. Then, the system of functions $\{\varphi_n(X)\}$ constitutes an orthonormal basis in L^2, while the system of functions $\{\varphi_n(X)\cdot\varphi_m(X_0)\}$ is an orthonormal basis in the analogous space of functions of the two variables X and X_0. By expanding $\psi(X,t;X_0)$ in a series on this basis and taking into account that $L_X\psi=\partial\psi/\partial t, L_{X_0}\psi=-\partial\psi/\partial t$, where $L_X=\frac{1}{4}d^2/dX^2-V(X)$, so that $L_X\varphi_n(X)=-E_n\varphi_n(X)$, $L_{X_0}\varphi_m(X_0)=-E_m\varphi_m(X_0)$, we easily obtain that the coefficient $a_{nm}(t)$ of $\varphi_n(X)\varphi_m(X_0)$ in this expansion will have the form $a_{nm}(t)=\delta_{nm}a_ne^{-E_nt}$.[9]
Moreover, inasmuch as it must be that

$$\psi(X,0,X_0)=\delta(X-X_0)=\sum_n\varphi_n(X)\varphi_n(X_0),$$

then $a_n\equiv 1$, i.e.,

$$\psi(X,t;X_0)=\sum_{n=1}^{\infty}e^{-E_nt}\varphi_n(X)\varphi_n(X_0). \tag{1.42}$$

This formula permits us to determine $\psi(X,t;X_0)$ in terms of the known characteristic functions and characteristic values of the differential equation (1.39).

Examples (Kac [2a])

1. Let

$$V(X)=\frac{1+\text{sign}X}{2}=\begin{cases}1 & \text{for } X>0,\\ 0 & \text{for } X<0.\end{cases} \tag{1.43}$$

Equation (1.39) here takes the form,

$$\left.\begin{aligned}\frac{1}{4}\frac{d^2\varphi}{dX^2}-(1+E)\varphi=0, \quad X>0\\ \frac{1}{4}\frac{d^2\varphi}{dX^2}-E\varphi=0, \quad X<0.\end{aligned}\right\} \tag{1.44}$$

[9] With the aid of integrations by parts we easily obtain

$$\begin{aligned}\frac{\partial a_{nm}(t)}{\partial t}&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\partial\psi(X,t;X_0)}{\partial t}\varphi_n(X)\varphi_m(X_0)dXdX_0\\ &=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}L_X\psi(X,t;X_0)\varphi_n(X)\varphi_m(X_0)dXdX_0\\ &=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi(X,t;X_0)L_X\varphi_n(X)\varphi_m(X_0)dXdX_0\\ &=-E_na_{nm}(t),\end{aligned}$$

and analogously (using the equality $-\partial\psi/\partial t=L_{X_0}\psi$)

$$\frac{\partial a_{nm}(t)}{\partial t}=-E_ma_{nm}(t).$$

From this follows the relation written down.

The solution to this equation, which goes to zero as $X \to \pm\infty$, will be:

$$\varphi(X)=\begin{cases} A\exp(-2(1+E)^{\frac{1}{2}}X) & \text{for } X>0 \\ B\exp(2(E)^{\frac{1}{2}}X) & \text{for } X<0. \end{cases} \tag{1.45}$$

The condition of continuity of $\varphi(X)$ at $X=0$ and the condition that $\varphi'(+0)-\varphi'(-0)=-4$ leads to the following values for the constants A and B:

$$A=B=2/[(E)^{\frac{1}{2}}+(1+E)^{\frac{1}{2}}]. \tag{1.46}$$

In virtue of (1.40), (1.32), and (1.34) we will have

$$\int_0^\infty\left\{\int_C \exp\left\{-\int_{\substack{x(\tau)>0\\ \tau\leqslant t}} d\tau\right\}d_w x\right\}e^{-Et}dt$$

$$=\frac{2}{(E)^{\frac{1}{2}}+(1+E)^{\frac{1}{2}}}\left\{\int_0^\infty \exp[-2(1+E)^{\frac{1}{2}}X]dX + \int_{-\infty}^0 \exp[2(E)^{\frac{1}{2}}X]dX\right\}$$

$$=\frac{1}{[E(1+E)]^{\frac{1}{2}}}=\sum_{k=0}^\infty\binom{-\frac{1}{2}}{k}\frac{1}{E^{k+1}} \quad (E>1). \tag{1.47}$$

From this, we immediately obtain

$$\int \exp\left\{-\int_{\substack{x(\tau)>0\\ \tau\leqslant t}} d\tau\right\}d_w x=\sum_{k=0}^\infty \frac{1}{k!}\binom{-\frac{1}{2}}{k}t^k$$

$$=\frac{2}{\pi}\int_0^{\pi/2}\exp(-t\cos^2\theta)d\theta. \tag{1.48}$$

But, inasmuch as

$$\int_C \exp\left\{-\int_{\substack{x(\tau)>0\\ \tau\leqslant t}} d\tau\right\}d_w x=\int_0^t e^{-t_1}dF(t_1), \tag{1.49}$$

where $F(t_1)$ is the probability that the point $x(\tau)$, undergoing Brownian motion along the x axis and found at the origin at time $t=0$, after a time t will be found on the right (positive) semi-axis during a time not less than t_1, we see that $F(t_1)$ must coincide with the distribution function of the magnitude of $t\cos^2\theta$, where θ has a uniform distribution in the interval $0\leqslant\theta\leqslant\pi/2$, i.e.,

$$F(t_1)=(2/\pi)^{\frac{1}{2}}\arcsin(t_1/t)^{\frac{1}{2}}. \tag{1.50}$$

In accordance with that stated on p. 54, this will also be the limiting (as $t\to\infty$) distribution law for the value of positive partial sums of any sequence of t mutually independent uniformly distributed random values with a zero mean value and finite dispersion ("the arcsine law"); see, for example [28], ch XII, §5).

2. Let $V(X)=CX^2$. In this case, Eq. (1.39) becomes

$$\tfrac{1}{4}\varphi''(X)-CX^2\varphi(X)=E\varphi(X); \tag{1.51}$$

its characteristic values and characteristic functions, as is well known (see the solution of Schrödinger's equation for an oscillator in any course on quantum mechanics), are equal to

$$E_n=(n+\tfrac{1}{2})(C)^{\frac{1}{2}} \quad (n=0, 1, 2\cdots) \tag{1.52}$$

and correspondingly,

$$\varphi_n(X)=\left(\frac{4C}{\pi^2}\right)^{\frac{1}{8}}\frac{1}{(2^n n!)^{\frac{1}{2}}}H_n[(4C)^{\frac{1}{4}}X]\exp[-(C)^{\frac{1}{2}}X^2], \tag{1.53}$$

where $H_n(z)$ are Hermite polynomials. In virtue of (1.42), we obtain from this

$$\psi(X,t;X_0)$$

$$=\left(\frac{4C}{\pi^2}\right)^{\frac{1}{4}}\sum_{n=0}^\infty\frac{\exp[-(n+\frac{1}{2})(C)^{\frac{1}{2}}t]}{2^n n!}H_n[(4C)^{\frac{1}{4}}X] \times H_n[(4C)^{\frac{1}{4}}X_0]\exp[-(C)^{\frac{1}{2}}(X_0^2+X^2)]. \tag{1.54}$$

By making use now of a well-known formula from the theory of Hermite polynomials (see, for example, [29], p. 104),

$$\sum_{n=0}^\infty\frac{\exp\{\frac{1}{2}(z^2+z_0^2)\}}{2^n n!}a^n H_n(z)H_n(z_0) =\frac{1}{(1-a^2)^{\frac{1}{2}}}\exp\left\{\frac{z^2-z_0^2}{2}-\frac{(z-az_0)^2}{1-a^2}\right\} \tag{1.55}$$

and assuming here $z=(4C)^{\frac{1}{4}}X$, $z_0=(4C)^{\frac{1}{4}}X_0$, $a=e^{-(C)^{\frac{1}{2}}t}$ we will have

$$\psi(X,t;X_0) =(C/\pi^2)^{\frac{1}{4}}[\operatorname{cosech}(C)^{\frac{1}{2}}t]^{\frac{1}{2}}\exp\{-(C)^{\frac{1}{2}}\operatorname{cth}(C)^{\frac{1}{2}}tX^2 +2(C)^{\frac{1}{2}}\operatorname{cosech}(C)^{\frac{1}{2}}tXX_0-(C)^{\frac{1}{2}}\operatorname{cth}(C)^{\frac{1}{2}}tX_0^2\}. \tag{1.56}$$

In the particular case $X_0=0$, the last formula gives

$$\psi(X,t)=(C/\pi^2)^{\frac{1}{4}}[\operatorname{cosech}(C)^{\frac{1}{2}}t]^{\frac{1}{2}} \times\exp\{-(C)^{\frac{1}{2}}\operatorname{cth}(C)^{\frac{1}{2}}tX^2\}; \tag{1.57}$$

for $C=-\lambda$ it goes over into formula (1.31) as it should. From (1.31), by integration over X, it is also possible to obtain formula (1.18).

Many Dimensional Generalization

If we now understand $x(\tau)$ to be a many dimensional random process $\{x_1(\tau), x_2(\tau), \cdots, x_N(\tau)\}$ describing the Brownian motion of a particle without inertia in an N-dimensional space, then all of our arguments and all of the formulas (1.32)–(1.42) will remain valid with the only difference being that the integration over Wiener measure now must be understood everywhere as an integration over an N-dimensional measure

$d_w x = d_w x_1 d_w x_2 \cdots d_w x_N$, the quantities X, X_0, etc. as N-dimensional vectors, while Eqs. (1.33) and (1.39) and the last of their conditions (1.41) change correspondingly into

$$\frac{\partial \psi(X,t)}{\partial t} = \tfrac{1}{4}\Delta\psi(X,t) - V(X)\psi(X,t), \qquad (1.33')$$

$$\tfrac{1}{4}\Delta\varphi(X) - [V(X)+E]\varphi(X) = 0, \qquad (1.39')$$

and

$$\lim_{\epsilon \to 0} \int_{|X-X_0|=\epsilon} \frac{\partial \varphi(x)}{\partial n} ds = -4, \qquad (1.41')$$

where Δ is the N-dimensional Laplace operator, $\partial\varphi/\partial n$ the derivative of the function φ along the normal to the sphere $|X-X_0| = \epsilon$, while ds is a surface element of this sphere (see reference [4], which is specially devoted to the N-dimensional case). If we wish to obtain the solution to Eq. (1.33') for a finite region Ω (for zero boundary conditions on the boundary of Ω), then we only have to stipulate that $V = \infty$ outside Ω, i.e., in (1.33) and (1.34)–(1.36) integration is carried out only over those trajectories $x(\tau)$, which do not intersect the boundary of Ω (see [5]); choosing the set of trajectories along which the integration is carried out in a different way it is possible to obtain other boundary conditions. Later on we will again return to the question of how it is possible to utilize these results for obtaining certain mathematical derivations concerning the behavior of the differential equation (1.33'); now, however, we will not delay ourselves on this point, but will proceed immediately to an examination of analogous methods in quantum mechanics.

2. INTEGRATION IN A FUNCTIONAL SPACE IN PROBLEMS OF QUANTUM MECHANICS

Case of Quantum Mechanics of a System with a Finite Number of Degrees of Freedom

In an investigation of problems of quantum mechanics, it is natural to begin with the simplest case of the nonrelativistic mechanics of a system with a finite number of degrees of freedom; we will proceed in just this way. As is well known, the state of the system is completely given by a wave function ψ, a complex function of the generalized coordinates of the system; the integral of the square of the modulus of this function over some region gives the probability of finding the system in this region. The change of the wave function with time is determined by Schrödinger's linear differential equation, which from the value of ψ at time $t = t_0$ uniquely allows the wave function to be determined at all subsequent times; for the case of the motion of a particle of mass m in a force field with potential $V(X)$ the equation has the form,

$$-\hbar i \partial\psi(X,t)/(\partial t) = \hbar^2/(2m)\Delta\psi - V(X)\psi, \qquad (2.1)$$

i.e., in fact it differs from Eq. (1.33') only in the factor $-i$ before the time derivative (the constants $\hbar$ and $\hbar^2/2m$ clearly can be made equal to 1 and $\frac{1}{4}$ by choosing appropriate units for measuring mass, length, and time). The value of the wave function $\psi(X,t)$ at time $t > t_0$ can be determined by the value $\psi(X,t_0)$ with the aid of the formula,

$$\psi(X,t) = \int_{-\infty}^{\infty} \psi(X_0,t_0)\psi(X,t;\, X_0,t_0)\, dX_0, \qquad (2.2)$$

analogous to (1.37) (by $\psi(X,t;\, X_0,t_0)$ we denote here the fundamental solution of Schrödinger's equation (2.1), i.e., the value of the wave function under the condition that at time t_0 the system was found at the point X_0.[10] The basic problem of quantum mechanics consists of finding the function $\psi(X,t)$ or $\psi(X,t;\, X_0,t_0)$.

By following from the fact that $\psi(X,t)$ is a solution of the differential equation (2.1), it can be supposed that a representation of this function in the form of some "integral over a functional space," which gives the contribution of each separate trajectory in the final solution, is also possible here. This supposition was adopted by Feynman [1a] in the nature of a postulate. After this, it is already easy to understand how one must choose the numerical contribution from the separate trajectories in order to obtain a result equivalent to that following from Schrödinger's equation: since Eq. (1.33') changes into (2.1) by a change of the time t into $it/\hbar$ and the coordinate X into $(m)^{\frac{1}{2}}X/(2)^{\frac{1}{2}}\hbar$, then we will obtain the correct formula for the wave function $\psi(X,t;\, X_0,t_0)$ by carrying out the same substitutions in formula (1.36),

$$\psi(X,t;\, X_0,t_0) = \frac{1}{N}\int\cdots\int \exp\left(\frac{i}{\hbar}\int_{t_0}^{t}\left\{\frac{m\dot{x}^2(\tau)}{2} - V[x(\tau)]\right\}d\tau\right) \times \prod_{t_0}^{t} dx(\tau). \qquad (2.3)$$

The integral over $dx(\tau)$ extends over all continuous paths connecting the points (X_0,t_0) and (X,t); N is a certain normalization constant, determined, for ex-

[10] We note that according to (2.2) the function $\psi(X,t)$ transforms analogously to the probability distribution density for a Markov process in which $\psi(X,t;\, X_0,t_0)$ plays the role of a "transition probability." By examining the state of the physical system at successive moments of time $t_0, t_1, t_2, \cdots$, we will have a chain analogous to a Markov chain with the only difference being that the role of the probability density in this case is played by a complex function $\psi(X)$ determined at time t_n from the known states of the system at preceding instants of time $t_0, t_1, \cdots, t_{n-1}$ by only one value of ψ, at the last of these instants t_{n-1}; the square of the modulus of this complex function will be equal to the probability density for locating the system. Discussions of the physical meaning of the differences indicated here between "quantum chains" and Markov chains basic for problems of classical statistical physics can be found in references [1a], [25].

ample, by the "unitarity condition"

$$\int_{-\infty}^{\infty} |\psi(X,t)|^2 dX = 1.$$

The meaning of the "symbolic expression" (2.3) is the same as of (1.4'): it is necessary to break up the interval $[t_0,t]$ into n equal parts by the points $t_0, t_1, \cdots, t_n=t$, to consider the n fold integral as analogous to that which figures in the right side of (1.4) (with the addition in appropriate places of the factors i, $\hbar$, and m), and to pass then to the limit as $n \to \infty$. We note immediately that in an exact accomplishment of this program, one will come up against definite difficulties: the integrals of rapidly oscillating functions which arise, generally speaking will not be convergent, and in order to assign to them a definite meaning it will be necessary to resort to certain devices, for example, to consider $\hbar$ as having a small purely imaginary addition, $-i\delta$, in the lower half-plane, and only after carrying out all integrations to pass to the limit as $\delta \to 0$ (compare the footnote on p. 184 of the Russian translation of reference [1a]).

The convention indicated here for the evaluation of oscillating integrals (2.3) can be illustrated in the following manner. On replacing $\hbar$ by a complex quantity $\hbar - i\delta$, we arrive at the following "complex measure" in functional space: to the quasi-interval given by the conditions $a_i < x(t_i) < b_i$ $(i=1, 2, \cdots, n)$ is related a quantity differing from (1.3) only by the presence in the exponents of the exponential under the integral sign of a constant complex multiplier $2m^{-1}(\delta + i\hbar)$ in the right half-plane (and an appropriate change in the normalization constant in front of the integral). It is natural that such a complex measure for arbitrary $m>0$, $\delta>0$, and R will be just as "good" as Wiener measure, i.e., it will have just as precise a meaning as measure in the space C of continuous functions, and it will allow integration over it of a wide class of functionals, including all continuous and bounded functionals.[11] Moreover, it will be possible to prove that a functional integral over this new measure, defined by the symbolic expression (2.3) (with complex $\hbar$), will represent a fundamental solution of an equation of the type of (1.33), namely Eq. (2.1) with a complex $\hbar$. If now we let the imaginary part of $\hbar$ tend to zero, then the corresponding fundamental solutions will tend to the fundamental solution of the original Eq. (2.1). In this way, although from a parabolic equation with a purely imaginary "diffusion coefficient" (as, for example, Schrödinger's equation) it is hardly possible to construct any arbitrary reasonable measure in a functional space, however, the solution of the corresponding equation can be obtained as the limit of a sequence of functional integrals corresponding to complex diffusion coefficients, as the imaginary part of this coefficient tends to zero.[12] This limit here will play the role of a representation of the solution in the form of a sum of "contributions" from separate trajectories, of which we have spoken earlier, and will "symbolically" be expressed by formula (2.3).

We turn now from the question of a precise mathematical meaning for the expression (2.3) and will argue further purely formally. We note first of all that the expression under the integral sign in the exponent in (2.3) completely coincides with Lagrange's function,

$$L(\dot{x},x) = (m\dot{x}^2/2) - V(x)$$
$$= \text{kinetic energy} - \text{potential energy} \quad (2.4)$$

for a particle of mass m in a force field with a potential $V(x)$; therefore, the integral from t_0 to t_1 is equal to the classical action $S(t_0,t_1)$ (to the integral of the Lagrangeian) along the trajectory $x(\tau)$. From this we see that formula (2.3) can be rewritten in a still more concise form,

$$\psi(X,t; X_0,t_0) \sim \int_{-\infty}^{\infty} \cdots \int \exp\left(\frac{i}{\hbar} S[x(\tau)]\right) \prod_{t_0}^{t} dx(\tau). \quad (2.5)$$

In other words, it should be considered that all trajectories give the same contribution to $\psi(X,t; X_0,t_0)$ in absolute magnitude and differ only in the argument ("phase") which is equal to the classical action along the corresponding trajectory taken with a factor $i/\hbar$. After this, just as in the case of integrals over Wiener measure, it may be proved that $\psi(X,t,X_0,t_0)$ is in fact a fundamental solution of Schrödinger's equation, so that the determination of this function indicated here is equivalent to the usual construction of quantum mechanics (see reference [1a]).

Until now, we have spoken only of Schrödinger's equation (2.1) for the simplest case of the motion of a particle in a potential field. However, the formulated method of writing the wave function in the form of an "integral over trajectories" requires only a knowledge of the classical action function for the problem, and can easily be extended to the case of motion of any system with a finite number of degrees of freedom, for which the Lagrangeian is known. Thus, for example, in the case of the motion of a charged particle (with a charge e) in a constant electromagnetic field in a three dimensional space, which is characterized by a scalar potential $\varphi(X) = 1/(e)V(X)$, $X = (X_1,X_2,X_3)$, and a vector potential $A(X) = (c/e)a(X)$ (where c is the speed of light), Schrödinger's equation also contains terms with a first

[11] A strict proof of this fact does not differ from the corresponding proof for the case of Wiener measure.

[12] It is probable, in general, that for a wide class of functionals the integrals over a complex measure related to a complex diffusion coefficient $D = D_1 + iD_2$, $D_1 > 0$ will tend to definite limits as $D_1 \to 0$. In this way, even for $D_1 = 0$ we will have a certain method for the integration of functionals but the corresponding "integrals" will no longer be related to a completely additive measure.

derivative with respect to the coordinates,

$$-i\hbar\frac{\partial\psi}{\partial t}=\frac{\hbar^2}{2m}\Delta\psi-\frac{i\hbar}{m}\left(a_1\frac{\partial\psi}{\partial X_1}+a_2\frac{\partial\psi}{\partial X_2}+a_3\frac{\partial\psi}{\partial X_3}\right)-\left(\frac{i\hbar}{2m}\operatorname{div}a+\frac{1}{2m}a^2+V\right)\psi. \quad (2.6)$$

The expression for the solution of this equation in the form of a functional integral obtained with the aid of the expression, known from electrodynamics, for the corresponding action function, has the form

$$\psi(X,t;X_0,t_0)=\frac{1}{N}\int_{-\infty}^{\infty}\cdots\int\exp\left[\frac{i}{\hbar}\int_{t_0}^{t}\left\{\frac{m\dot{x}^2}{2}-V(x)+a\dot{x}\right\}d\tau\right]\times\prod_{t_0}^{t}dx(\tau), \quad (2.7)$$

where $dx(\tau)=dx_1(\tau)dx_2(\tau)dx_3(\tau)$. It is curious to note that passing here anew from Schrödinger's equation to an ordinary parabolic equation with the aid of a change of variables, the reverse of the one which was used above for the passage from (1.33′) to (2.1), we find that the solution of the equation,

$$\frac{\partial\psi}{\partial t}=\tfrac{1}{4}\Delta\psi-a_1\frac{\partial\psi}{\partial X_1}-a_2\frac{\partial\psi}{\partial X_2}-a_3\frac{\partial\psi}{\partial X_3}-V\psi, \quad (2.8)$$

satisfying the condition $\psi(X,0)=\delta(X)$, is given by the functional integral

$$\psi(X,t)=\int_{C_{t,x}}\exp\left(-\int_0^t\{V([x(\tau)]-\tfrac{1}{2}\operatorname{div}a[x(\tau)]+\tfrac{1}{2}a^2[x(\tau)])d\tau+adx(\tau)\}\right)d_{w(t,x)}x, \quad (2.9)$$

where $d_{w(t,x)}x$ is an integration over "conditional" Wiener measure in the space of continuous vector functions $x(\tau)$ in three dimensional space. The last result apparently has not been encountered in the literature; it may be of interest in investigations of a series of mathematical questions concerning the solution of Eq. (2.8).[13]

[13] An analogous result for a one dimensional parabolic equation containing a term with a first derivative with respect to the coordinate can be obtained from (1.32) with the aid of the simple substitution $\psi(X,t)=e^{-a(X)}\varphi(X,t)$. This fact is closely related with that circumstance that in electrodynamics the vector potential having a single nonvanishing component A_1, depending only on the coordinate x_1, does not have a physical meaning and can always be eliminated with the aid of a "gradient transformation."

Case of the Quantum Field Theory[14]

We pass now to an investigation of the quantum mechanics of systems with an infinite number of degrees of freedom, quantum field theory. The most important example of such a theory is quantum electrodynamics, the theory of the interaction one with another of an electromagnetic and an electron field; in this case, however, the original classical equations (the systems of Maxwell and Dirac equations) are comparatively cumbersome and cannot be considered as common knowledge among mathematicians. Inasmuch as all investigations for arbitrary fields (in any event if the case of half-integer spin is excluded) are carried out completely analogously, then in the beginning we restrict ourselves to an investigation of the simplest model of quantum field theory, which retains all of the most characteristic features of such theories, in particular the presence of an infinite number of degrees of freedom and nonlinearity. A quantum theory of a field $u(x)$ in three-dimensional space $x=(x_1,x_2,x_3)$ whose potential energy is given by the formula,

$$V[u(x)]=\int_{-\infty}^{\infty}\int\int[(\operatorname{grad}u)^2+\kappa^2u^2+\lambda u^4]dx \quad (2.10)$$

is obtained with such a model. In the case of a one-dimensional space x the analogous field $u(x)$ with potential energy

$$V[u(x)]=\int_{-\infty}^{\infty}\left[\left(\frac{\partial u}{\partial x}\right)^2+\kappa^2u^2+\lambda u^4\right]dx \quad (2.10')$$

admits of a simple mechanical interpretation: $u(x)$ can be regarded as the displacement from equilibrium of an elastic string in tension, at every point of which a force is applied, which depends nonlinearly on displacement (with a density of force equal to $\kappa^2u+2\lambda u^3$). In the three dimensional case for descriptiveness, we may analogously speak of the oscillations of a three dimensional elastic volume with an applied nonlinear force. The corresponding mechanical "equation of vibration" has the form

$$\frac{\partial^2u}{\partial t^2}=\Delta u+\kappa^2u+2\lambda u^3; \quad (2.11)$$

this will be the "classical" (not quantum) equation of the problem to be investigated. The model of a field having a potential energy given by (2.10) is very often

[14] The construction of a quantum field theory by Feynman's method [1a] presented here was described in a report of I. M. Gel'fand and R. A. Minlos at an All Union Congress on Quantum Electrodynamics and the Theory of Elementary Particles in Moscow, in March of 1955. In very recent times an approach to quantum field theory similar to this one is reported also in references [21] and [22]. Another approach to the introduction of functional integrals into the quantum theory of fields is contained in the works of N. N. Bogolyubov [15], E. C. Fradkin [16] and Matthews and Salem [17].

made use of in contemporary physical literature in the nature of a very simple model very well suited for an initial investigation of new methods and for the clarification of questions of principle in the general theory (it is commonly called "Thirring's model" after the name of the physicist who first studied this model).

We turn now to the formulation of the quantum theory of this field. Instead of a wave function $\psi(u_1,u_2,\cdots,u_n)=\psi(u)$ (for the case of a system with n degrees of freedom), we will have here a "wave functional" $\psi[u(x)]$, depending on the function $u(x)$.[15] Further relations are likewise completely analogous to the relations of quantum mechanics with a finite number of degrees of freedom, and we emphasize this by writing down the relations for ordinary quantum mechanics and for the quantum mechanics of our system in parallel (for simplicity we assume here a system of units in which $\hbar=1$ and we consider the "masses" corresponding to all degrees of freedom to be equal to 2):

Laplacian Operator

$$\Delta\psi=\sum_{k=1}^{n}\frac{\partial^2\psi(u)}{\partial u_k^2}, \tag{2.12}$$

$$\Delta\psi=\int_{-\infty}^{\infty}\frac{\delta^2\psi[u(x)]}{\delta[u(x)]^2}dx, \tag{2.12'}$$

where in (2.12') the expression under the integral sign represents the second variational derivative of the functional $\psi[u(x)]$.

Wave Equation (Schrödinger's Equation)

$$i\frac{\partial\psi}{\partial t}=-\tfrac{1}{4}\Delta\psi+V(u)\psi, \tag{2.13}$$

$$i\frac{\partial\psi}{\partial t}=-\tfrac{1}{4}\Delta\psi+V[u(x)]\psi, \tag{2.13'}$$

where in (2.13') $\Delta\psi$ is given by (2.12') while $V[u(x)]$ is given by (2.10).

Representation of ψ in the Form of a Functional Integral

$$\psi(u)=\int\exp\left(i\int_{t_0}^{t}\left\{\sum_{k=1}^{n}\dot{u}_k^2(\tau)-V[u(\tau)]\right\}d\tau\right)\times\prod_{(k,\tau)=(1,t_0)}^{(n,t)}du_k(\tau), \tag{2.14}$$

$$\psi[u(x)]=\int\exp\left(i\int_{t_0}^{t}\int\int_{-\infty}^{\infty}\int\left\{\left(\frac{\partial u}{\partial t}\right)^2-(\mathrm{grad}u)^2-\kappa^2u^2+\lambda u^4\right\}d\tau dx\right)\times\prod_{(x,\tau)=(-\infty,t_0)}^{(+\infty,t)}du(x,\tau). \tag{2.14'}$$

We note that the expression in the exponent under the integral sign in (2.14'),

$$S[u]=\int\int\int\int(\partial u/\partial t)^2dxd\tau-\int V[u(x,\tau)]d\tau, \tag{2.15}$$

is none other than the "classical action" for our field which can be written in the form

$$S[u]=\int\int\int\int Ldxd\tau, \tag{2.16}$$

where

$$L=(\partial u/\partial t)^2-(\mathrm{grad}u)^2-\kappa^2u^2-\lambda u^4 \tag{2.17}$$

is the so called Lagrangeian function for the field which is equal (making use of a mechanical interpretation) to the density of kinetic energy of the field minus the density of potential energy. The "classical" equations of the field (2.11) arise from the variational principle $\delta S[u]=0$ in precisely the same way as in the case of a system with a finite number of degrees of freedom.

The value of the wave functional $\psi[u(x)]$ completely determines all physical quantities connected with the quantum field $u(x)$. In particular, the so called matrix elements of the S matrix which are of the greatest interest to physicists and which give the transition probabilities from one state into another appear as variational derivatives of the functional $\psi[u(x)]$ (see, for example, [14]).

An analogous expression for the wave functional is obtained for an arbitrary field with integer spin. In particular, if we have a certain system of coupled scalar, vector, or tensor fields $u(x)$, $v(x)$, $\cdots$, $w(x)$ in three dimensional space with a given classical Lagrangeian $L[u,v,\cdots,w]$, then the corresponding wave functional will be given by the formula,

$$\psi[u,v,\cdots,w]=\int\exp\left\{i\int\int\int\int L[u,v,\cdots,w]dxd\tau\right\}\times\prod_{x,\tau}du(x,\tau)dv(x,\tau)\cdots dw(x,\tau). \tag{2.18}$$

In the case of the so called half-integer spin (i.e., in the presence of spinor fields), the matter is considerably more complicated. The additional difficulties encountered here were overcome in the works of N. N. Bogolyubov [15], E. C. Fradkin [16], Matthews and

[15] It would be more precise to write $\psi(u,t)$ and $\psi[u(x),t]$ inasmuch as both the wave function and the wave functional depend explicitly on t.

Salam [17], and in a work especially devoted to this question by I. M. Khalatnikov [23a].

3. CHANGE OF VARIABLES IN FUNCTIONAL INTEGRALS

Certain Mathematical Applications

Change of Variables in Functional Integrals

In the preceding part of this report, it was shown that just as the solution to the ordinary parabolic equation (1.33′) can be represented in the form of a functional integral, so also can the solution to Schrödinger's equation (2.1); it was also pointed out that this circumstance to all appearance has a rather general character. Nevertheless, the usefulness of this notation for the theory of equations was by no means asserted: because we have written the solution in the form of a special symbol, the meaning of which is not very simply presented, it appears that there is little more which can be extracted.[16] In what follows, we will show in several examples that in fact the situation is not quite like this; it appears to us, however, that all of the usefulness which is gained by the introduction of functional integrals can really be appraised only later on when the algorithmic rules of operating with functional integrals will be sufficiently worked out and a sufficient experience in utilizing these rules will be accumulated.[17] At the present time, the technique of operating with functional integrals is only beginning to be developed; nevertheless, a series of interesting results in this direction has already been obtained in certain cases by mathematicians (by Cameron, Martin, and their collaborators) and independently by physicists (by the latter usually without strong proofs). It is on these results that we now pause.

It is natural to begin with the question of the rules for transforming functional integrals by a change of the "variable of integration" $x(\tau)$; this question is the only one that we will examine here. For definiteness, we will examine only integrals over Wiener measure $d_w x$; at the same time, it will be convenient for us to utilize a "symbolic" notation for this measure, which is indicated in (1.4′). We begin with the simplest question of the transformation of integrals over Wiener measure by means of a *parallel translation* in the functional space C, i.e., under the following transformation of functions in C

$$x(\tau) \to y(\tau) = x(\tau) + x_0(\tau) \tag{3.1}$$

[$x_0(\tau)$ is a fixed function in C]. Formally carrying out the substitution (3.1) into the right side of the formula (1.4′), we obtain

$$\frac{1}{N}\int\cdots\int F[x(\tau)]\exp\left\{-\int_0^t \dot{x}^2(\tau)d\tau\right\}\prod_0^t dx(\tau)$$
$$=\frac{1}{N}\int\cdots\int F[x+x_0]\exp\left\{-\int_0^t [\dot{x}^2(\tau)$$
$$+2\dot{x}(\tau)\dot{x}_0(\tau)+\dot{x}_0^2(\tau)]\right\}\prod_0^t dx(\tau). \tag{3.2}$$

On removing

$$\exp\left\{-\int_0^t \dot{x}_0^2(\tau)d\tau\right\}$$

outside the integral sign on the right-hand side of this formula, replacing $\dot{x}(\tau)d\tau$ by $dx(\tau)$, and transforming back into the usual notation for an integral over Wiener measure, we have

$$\int_S F[x(\tau)]d_w x$$
$$=\exp\left[-\int_0^t \dot{x}_0^2(\tau)d\tau\right]\int_{TS} F[x+x_0]$$
$$\times\exp\left[-2\int_0^t x_0(\tau)dx(\tau)\right]d_w x. \tag{3.3}$$

[For generality, we assumed that the integral on the left-hand side is extended only over a subset S of C; then on the right-hand side we will have an integral over the set TS, which transforms into S under the transformation (3.1).] Formula (3.3) defines the transformation of Wiener integrals under a parallel translation; it is correct when the right-hand side of (3.3) has a meaning, i.e., when $x_0(\tau)$ is a differentiable function on C, having a derivative $\dot{x}_0(\tau)$ with a bounded variation. (If these conditions are not satisfied, then the transformation (3.1), generally speaking, will transform sets which are measurable over Wiener measure into nonmeasurable sets; see [10c].) The proof of formula (3.3) presented here at first glance appears weak; it is not difficult to realize, however, that a strong proof is obtained only by a detailed deciphering of the "symbolic transformations" contained in (3.2) and (3.3) (see reference [9a]).

For the next example, we examine the transformation of Wiener integrals under a general *linear transformation* in the space C,

$$x(\tau) \to y(\tau) = x(\tau) + \int_0^t K(\tau,\sigma)x(\sigma)d\sigma \equiv x + Ax. \tag{3.4}$$

[16] It is rather clear that on the other hand the connection of functional integrals with differential equations can be useful for the study of functional integrals themselves (see p. 53 and the examples on p. 55).

[17] In certain respects, the passage from the task of solving partial differential equations with the aid of difference schemes to the task of obtaining them in the form of functional integrals can be compared to the passage from finite sums to integrals in finding areas, the solutions to an ordinary differential equation $dy/dx = f(x)$. It is natural, however, that in order that this passage be utilized properly it is necessary first of all to learn to handle the resulting expressions.

By reasoning just as in the derivation of (3.3), it is necessary in this case to keep in mind that under a linear transformation of the "coordinates" $x(\tau)$ according to formula (3.4), the "volume element" $\prod_0^t dx(\tau)$ is multiplied by the "Jacobian" for the transformation; the role of such a Jacobian, which is obtained by a limiting process from the Jacobian of a finite dimensional linear transformation of coordinates $x(t_1), x(t_2), \cdots, x(t_n)$, will be played, as is well known, by the Fredholm determinant D of the kernel $K(t,s)$,

$$D=1+\sum_{n=1}^{\infty}\frac{1}{n!}\int_0^t\cdots\int_0^t\begin{vmatrix}K(s_1,s_1), & \cdots, & K(s_1,s_n)\\ \cdots & \cdots & \cdots\\ K(s_n,s_1), & \cdots, & K(s_n,s_n)\end{vmatrix}ds_1\cdots ds_n. \quad (3.5)$$

(In fact, it is in this very connection that the Fredholm determinant appears in the theory of integral equations.) By taking this circumstance into account, we have

$$\int F[x(\tau)]\exp\left[-\int_0^t \dot{x}^2(\tau)d\tau\right]\prod_0^t dx(\tau)$$

$$=|D|\int F[x+Ax]\exp\left\{-\int_0^t[\dot{x}(\tau)+A\dot{x}(\tau)]^2d\tau\right\}\prod_0^t dx(\tau)$$

$$=|D|\int F\left[x(\tau)+\int_0^t K(\tau,\sigma)x(\sigma)d\sigma\right]\exp\left\{-\int_0^t\left[\frac{d}{d\tau}\int_0^t K(\tau,\sigma)x(\sigma)d\sigma\right]^2 d\tau\right.$$

$$\left.-2\int_0^t\left[\frac{d}{d\tau}\int_0^t K(\tau,\sigma)x(\sigma)d\sigma\right]dx(\tau)-\int_0^t\dot{x}^2(\tau)d\tau\right\}\prod_0^t dx(\tau), \quad (3.6)$$

i.e.,

$$\int_S F[x(\tau)]d_w\tau=|D|\int_{TS}F[x+Ax]\exp\left\{-\int_0^t[A\dot{x}]^2dx-2\int_0^t A\dot{x}(\tau)dx(\tau)\right\}d_wx. \quad (3.7)$$

[The symbol TS has a meaning here analogous to its meaning in (3.3).] A strong proof of formula (3.7) [under reasonable restrictions on the kernel $K(t,s)$] is contained in reference [9b]; it is rather long, but not complicated in principle [in fact it consists of proving the validity of the transformation (3.6) in formulas which occur prior to passing to the limit $n\to\infty$]. In applications, it is essential that the result which is obtained also remains valid in the case of a kernel $K(t,s)$, which undergoes a discontinuity on the diagonal $t=s$ (for example in the case of a Volterra kernel); for this, it is only necessary during the construction of Fredholm's determinant to equate the value of the kernel on the diagonal to half the sum of its right and left hand limiting values.

We point out further that formula (3.7) is found to be correct also for a wide class of *nonlinear transformations* of the form,

$$x(\tau)\to y(\tau)=x(\tau)+A[x(\tau);\tau], \quad (3.8)$$

where, together with certain other restrictions of the type of "conditions of regularity," we also need to require that the transformation $A[x;\tau]$ has a nondegenerate "linear part," i.e., that $A[x;\tau]$ has a "functional derivative" (variation),

$$\delta A=\frac{\partial}{\partial h}A[x+hy;\tau]_{h=0}, \quad (3.9)$$

which can be written in the form,

$$\delta A=\int_0^t K[x;\tau,\sigma]y(\sigma)d\sigma. \quad (3.10)$$

The role of the "Jacobian" in this case is played by the Jacobian of the "linear part" of the transformation, i.e., the Fredholm determinant of the kernel $K[x;\tau,\sigma]$, which is clearly a functional of $x=x(\tau)$. Therefore, in the corresponding formula (3.7) it is no longer possible to remove $D=D[x]$ outside the integral sign. In all other respects, the outward form of formula (3.7) even for nonlinear A, as a rule will remain unchanged; in connection with this see references [9d] and [11], which are devoted to the derivation of this formula for different restrictions on $A[x;\tau]$ (and also the applications at the end of reference [20a]).

In a number of problems, the utilization of only the simplest rules of transformation of functional integrals which have been presented here (which, besides, are the only known rules at the present time) allows a great simplification and formalization of all derivations.

Example. We examine the problem of evaluating the functional integral

$$J=\int_C\exp\left[\lambda\int_0^t p(\tau)x^2(\tau)d\tau\right]d_wx \quad (p(\tau)\geqslant 0). \quad (3.11)$$

We have already studied this problem previously and saw that a direct evaluation of this integral with the aid of formula (1.4) leads to cumbersome investigations, particularly if it is desired to make such an evaluation sufficiently rigorous. We will show now that making use of formula (3.7) enables us to obtain the result (1.18) very quickly with the aid of purely formal transformations. Actually, we examine the following linear transformation of continuous functions on the interval $0 \leqslant \tau \leqslant t$,

$$x(\tau) \to y(\tau) - \int_0^\tau \frac{D'(\sigma)}{D(\sigma)} x(\sigma) d\sigma \quad (0 \leqslant \tau \leqslant t), \quad (3.12)$$

where $D(\sigma)$ is the same as in formulas (1.16) and (1.17). [We assume that λ is smaller than the smallest of the characteristic values of (1.16) subject to the boundary conditions $D(0)=0$, $D'(t)=0$; in such a case, $D(\sigma)$ will not be identically zero on the interval $0 \leqslant \tau \leqslant t$.] Fredholm's determinant here will be equal to

$$D = \exp\left[\tfrac{1}{2}\int_0^t - \frac{D'(\sigma)}{D(\sigma)} d\sigma\right]$$

$$= \exp\{-\tfrac{1}{2}[\log D(t) - \log D(0)]\} = [D(0)]^{\frac{1}{2}} \quad (3.13)$$

(because for Volterra's kernel, Fredholm's determinant is equal to

$$\exp\left\{\tfrac{1}{2}\int_0^t K(\sigma,\sigma) d\sigma\right\}\Big),$$

and

$$-\int_0^t \left[\frac{d}{d\tau}\int_0^\tau - \frac{D'(\sigma)}{D(\sigma)} x(\sigma) d\sigma\right]^2 d\tau$$

$$-2\int_0^t \left[\frac{d}{d\tau}\int_0^\tau - \frac{D'(\sigma)}{D(\sigma)} x(\sigma) d\sigma\right] dx(\tau)$$

$$= -\int_0^t \left[\frac{D'(\tau)}{D(\tau)} x(\tau)\right]^2 d\tau + \int_0^t \frac{D'(\tau)}{D(\tau)} d[x^2(\tau)]$$

$$= -\int^t \frac{D'^2(\tau)}{D^2(\tau)} x^2(\tau) d\tau - \int_0^t \left[\frac{D'(\tau)}{D(\tau)}\right]' x^2(\tau) d\tau$$

$$= -\int_0^t \frac{D''(\tau)}{D(\tau)} x^2(\tau) d\tau = \lambda \int_0^t p(\tau) x^2(\tau) d\tau \quad (3.14)$$

[in view of (1.16)]. Inasmuch as it is very easy to demonstrate that the transformation (3.12) reciprocally and single valuedly reflects the space C of continuous functions on itself (the reverse transformation here has the form

$$y(\tau) \to x(\tau) = y(\tau) + D(\tau) \int_0^t D'(\sigma)/[D^2(\sigma)] y(\sigma) d\sigma$$

and is also continuous), then (1.18) immediately follows from formula (3.7) by substituting into this formula $F[x(\tau)]=1$ and $S=C$,

$$1 = \int_C d_w x = |D| \int_C \exp\left\{-\int_0^t \left[\frac{d}{d\tau} Ax(\tau)\right]^2 d\tau \right.$$

$$\left. -2\int_0^t \left[\frac{d}{d\tau} Ax(\tau)\right] dx(\tau)\right\} d_w x. \quad (3.15)$$

It was just in this manner that formula (1.18) was first derived in reference [9c].

In an analogous way, formula (3.3) can also be utilized for the straightforward evaluation of certain special functional integrals (see for example reference [9a]).

We note that similar formal transformations can also be utilized for proving a theorem on the connection of functional integrals of

$$\exp\left\{-\int_0^t V[x(\tau)] d\tau\right\}$$

with the solution of the partial differential Eq. (1.33), of which we have spoken in the preceding section. (See in this connection the work of Cameron, reference [10b], where a representation for the solution to Cauchy's problem for Eq. (1.33) in the form of a functional integral is obtained in this very way by the use of a certain special nonlinear transformation $A[x; s]$.)

Use of Functional Integrals for Obtaining Asymptotics of Characteristic Functions and Characteristic Values of an Elliptic Differential Equation

To end the present section, we present one simple example which shows that in certain respects the use of functional integrals can be very helpful for the study of differential equations. In particular, following the work of Ray [5], we examine the question of obtaining with the aid of functional integrals a series of results which concern the asymptotic behavior of the characteristic functions and characteristic values of the elliptic differential equation (1.39′) for a nonnegative $V(x)$.

As we have already seen, the characteristic functions $\varphi_n(X)$ of Eq. (1.39′) are connected with the functional integral (1.35) over conditional Wiener measure $d_{w(t,X)}x$ by the relation (1.42); in this, it is supposed that Eq. (1.39′) has a discrete spectrum,[18] but it is assumed that (1.39′) is studied only in a certain region Ω of an N dimensional space (with zero boundary conditions); in the last case, it is only necessary to extend the integration over $d_{w(t,X)}x$ in (1.35) only over con-

[18] We note that the discreteness of the spectrum for the case when $V(X) \to \infty$ for $X \to \infty$ or when the equation is studied only in a finite region can also be established from probability theoretic considerations relevant to the introduction of functional integrals (see reference [5]).

tinuous curves $x(\tau)$ which do not leave the region Ω. We note now that

$$\int_{t,0} d_{w(t,0)}x = 1/[(\pi t)^{N/2}]$$

for arbitrary $t>0$[19]; from this, it is easily obtained that

$$\psi(X,t;X)=\int_{C_{t,0}} \exp\left\{-\int_0^t V[X+x(\tau)]d\tau\right\} d_{w(t,0)}x \sim \frac{1}{(\pi t)^{N/2}} \quad \text{as } t\to 0. \quad (3.16)$$

(In the case of integration over only the set of curves which are distributed in the interior and on the boundary of Ω, the last expression will be exact for any arbitrary internal point X of the region Ω.) In this way, we obtain

$$\sum_n \varphi_n^2(X)e^{-E_n t}=\psi(X,0;X)\sim\frac{1}{(\pi t)^N} \quad \text{as } t\to 0. \quad (3.17)$$

By making use of the Tauberian theorem of Hardy-Littlewood-Karamata (see, for example, reference [30] page 208), we immediately obtain from this the following important asymptotic relation,

$$\sum_{E_n<E} \varphi_n^2(X)\sim\frac{E^{N/2}}{\pi^{N/2}\Gamma[(N/2)+1]} \quad \text{as } E\to\infty. \quad (3.18)$$

This relation can also be obtained by other means; however, the derivation presented here seems to be the simplest and shortest.

For Eq. (3.18) in a finite region Ω, the sum on the right-hand side of (3.18) can be integrated term by term over dX (or we can integrate both sides of (3.17) over dX and apply the corresponding Tauberian theorem to the result); in this way we come to an asymptotic formula for the number $N(E)$ of characteristic values smaller than E,

$$N(E)=\sum_{E_n<E} 1\sim\frac{E^{N/2}\,\mathrm{mes}\Omega}{\pi^{N/2}\Gamma[(N/2)+1]} \quad \text{as } E\to\infty, \quad (3.19)$$

where mesΩ is the Lebesgue measure of the region Ω. Thus, we have very rapidly obtained the most important asymptotic relation of the theory of elliptic differential equations and immediately in a very general form (see the derivation of particular cases of (3.19) in reference [31], chap. 6).

In the case of an unbounded region Ω (in particular for an equation in all space), the considerations presented here can also be made use of for obtaining [under certain conditions concerning the smoothness of the function $V(X)$] the asymptotic form of the function $N(E)$ for large E, but here it is already not as simple (see reference [5]).

[19] This relation was written above only for the case $N=1$; it is not difficult to see, however, that it will also be valid for arbitrary N.

4. THE UTILIZATION OF FUNCTIONAL INTEGRALS IN QUANTUM PHYSICS

In Sec. 2, it was shown that the wave function of a quantum system can be represented in the form of a certain functional integral, whereupon its determination is completely equivalent to the usual determination of a wave function as a solution of Schrödinger's equation. From this, it is clear that all results obtained with the aid of functional integrals will exactly coincide with the results obtained by older methods. It is natural to think, however, as is usual in similar situations, that in investigations of certain questions one of the indicated approaches will appear more convenient, while in investigations of other problems another, so that both of these approaches can be of definite practical interest. At the present time, the utilization of the functional formulation in the solution of problems of quantum physics is only just beginning, and moreover, it is strongly held back by a general lack of development of the corresponding mathematical apparatus; none the less, a series of questions in the study of which this formulation offers notable advantages has already been cleared up now. In what follows, we examine several simple examples of this type; other examples of obtaining physically interesting conclusions with the aid of functional integrals can be found in references [1a]–[1d], [13]–[23a].

Classical Limit and Quasi-Classical Approximation

It is well known that the square of the modulus of a wave function of a particle at any moment of time t gives the probability density for finding the particle at this moment at different points of the space (in the case of an arbitrary system it is necessary here only to replace the ordinary three dimensional physical space by a space of generalized coordinates). Formula (2.3) or, more generally, (2.5) presents this probability density in the form of a sum of contributions over different trajectories of the motion of the particle which end at the investigated point. If in the above mentioned formula the integrals are restricted only to the trajectories which lie in a certain portion Ω of the space, then the square of this integral gives us the probability density for the particle, at the point X_0 at time t_0, to appear at time t at the point X, and, moreover, for all times between t_0 and t not once to go outside the limits of Ω.[20] This circumstance enables us, proceeding from

[20] The neglect in (2.5) of the trajectories which go outside the limits of Ω, in view of the meaning of this functional integral indicated on p. 57, is equivalent to the assumption that $V(X)=\infty$ outside Ω, i.e., to the assumption of the existence along the boundary of Ω of an absolutely inpenetrable barrier (compare with the remarks on the representation of the fundamental solution of Eq. (1.33′) in a finite region in the form of a functional integral).

formula (2.5) to establish in the shortest way the connection between classical and quantum mechanics. Since for small $\hbar$ the exponential function $\exp\{(i/\hbar)\times S[x(\tau)]\}$ oscillates extremely rapidly even for small changes of S, for such $\hbar$ the contributions from neighboring trajectories, generally speaking, will mutually cancel out; the exception is that trajectory in the neighborhood of which the "phase" $S[x(\tau)]$ changes most slowly, i.e., for which $\delta S[x(\tau)]=0$ (compare with the method of stationary phase in mathematical physics). But the last condition determines precisely the trajectory of classical mechanics, and moreover constitutes Hamilton's principle.

In this way, the terms which correspond to the classical trajectory, or which are close to it play the basic role in the integral (2.5) for very small $\hbar$ (practically speaking the entire integral arises from the contributions of those trajectories which are enclosed in a narrow tube which surrounds the classical trajectory); in the limit $\hbar \to 0$, the particle definitely will move along some arbitrary classical trajectory. We note further that it is clear from this that in the "classical" approximation it is necessary to regard $\psi(X,t)\sim e^{S_{cl}}$, where S_{cl} is the classical action of the system.

This result, clearly, can also be obtained from Schrödinger's equation (see for example [32], p. 149), but there it is much less simple and obvious.

The foregoing considerations also enable us to obtain the further corrections to the classical result, i.e., the succeeding term in an expansion of the wave function in a series of powers of $\hbar$. Indeed, expanding the functional $S[x(\tau)]$ about the classical trajectory $x_{cl}(\tau)$ in a functional "Taylor series" we have

$$S[x(\tau)]=S_{cl}+\tfrac{1}{2}\delta^2 S[x_{cl}(\tau)]+\tfrac{1}{6}\delta^3 S[x_{cl}(\tau)]+\cdots, \quad (4.1)$$

where $\delta^2 S[x_{cl}(\tau)]$ is a quadratic functional with regard to the function

$$y(\tau)=x(\tau)-x_{cl}(\tau), \quad (4.2)$$

namely,

$$\delta^2 S[x_{cl}(\tau)]=\int_{t_0}^{t_1}\left[\left.\frac{\delta^2 L}{\delta x^2}\right|_{x=x_{cl}} y^2(\tau)+2\left.\frac{\delta^2 L}{\delta x\delta\dot{x}}\right|_{x=x_{cl}} y(\tau)\dot{y}(\tau)+\left.\frac{\delta^2 L}{\delta\dot{x}^2}\right|_{x=x_{cl}}\dot{y}^2(\tau)\right]d\tau, \quad (4.3)$$

where L is Lagrange's function. Furthermore, taking $e^{S_{cl}}$ outside the functional integral sign and retaining only two terms of the "Taylor's series" (4.1), we obtain the functional integral of an exponential which contains in the exponent a quadratic functional (it is clearly necessary first of all to carry out the "parallel translation" (4.2), p. 61). The last integral can be evaluated with the aid of the formulas for finite dimensional integrals, (1.11) and (1.28); it gives the correction of the next order in $\hbar$ to the wave function equivalent to the so called "quasi-classical approximation" of quantum mechanics.[21]

In courses on quantum mechanics the quasi-classical approximation is usually studied primarily for problems of determining characteristic values; it is just here that it gives the most concrete and practically interesting results (it leads to the quantization conditions of Bohr). Mathematically, these results coincide with the asymptotic formulas for the large characteristic values of the differential equation (1.39′); in this way, the theory of the quasi-classical approximation ties in here with the question studied on p. 63. In particular, if in Eq. (1.39′) we multiply all terms by $2\hbar^2/m$, then the result (3.19) for a three-dimensional space transforms into the following: in the case of Schrödinger's equation in a finite region Ω for small $\hbar$, the number of characteristic values smaller than E will be equal to $(2)^{\frac{1}{2}}m^{\frac{3}{2}}/(3\pi^2\hbar^3)E^{\frac{3}{2}}\,\mathrm{mes}\Omega$ where $\mathrm{mes}\Omega$ is the volume of the region Ω. Since for a region of volume $\mathrm{mes}\Omega$, a volume $8(2)^{\frac{1}{2}}\pi m^{\frac{3}{2}}E^{\frac{3}{2}}/(3)^{-1}\,\mathrm{mes}\Omega$ corresponds to the values of energy less than E in the six dimensional phase space of a particle, the space of the pair $(\boldsymbol{x},\boldsymbol{p})$, where $\boldsymbol{x}$ is the coordinate and $\boldsymbol{p}$ the momentum, so that $p=(2mE)^{\frac{1}{2}}$, we see from here that to each characteristic value there corresponds a cell of volume $\Delta v=1/(2\pi\hbar)^3$ in phase space, while the number of characteristic values related to the element of volume $\Delta p\cdot\Delta x$ in phase space is equal to $\Delta p\cdot\Delta x/(2\pi\hbar)^3$. This result is one of the important conclusions of the quasi-classical theory; it would be very interesting to consider the possibility of extending it simply, with the aid of "functional" methods, to the case of motion in an unbounded space (the first steps in this direction are contained in reference [5]).

The Evaluation of the Statistical Sum of Quantum Statistics [37]

Statistical integrals in the field of statistical physics can have applications similar to those investigated above. As is well known, in a statistical study of the thermodynamic properties of quantum systems a basic role is played by the "statistical sum,"

$$Z=\sum_n e^{-\beta E_n}, \quad \beta=1/kT, \quad (4.4)$$

where k is Boltzmann's constant, T is the absolute temperature, and the sum is carried out over all charac-

[21] A proof of the fact that the approximation obtained with the aid of two terms of the series (4.1) gives us a wave function exact to terms of order $\hbar$ (i.e., equivalent to the "quasi-classical approximation") is contained in reference [33]; some examples are also given there. We note, however, that the strictness of the quoted proof is substantially lowered due to the fact that the question of the precise meaning of the functional integrals studied was not discussed.

Further examples of the "quasi-classical approximation" obtained in this manner for concrete quantum mechanical problems are contained in a recent large work [36].

teristic values E_n of the energy operator H of the system under investigation.[22] As an example, we investigate the simplest problem of the one dimensional motion of a particle of mass m in a force field with a potential $V(X)$; here clearly,

$$H=-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial X^2}+V(X). \tag{4.5}$$

By making a change of variable $X\to[(2)^{\frac{1}{2}}\hbar/(m)^{\frac{1}{2}}]X$, we bring this operator into the form

$$H_1=-\frac{1}{4}\frac{\partial^2}{\partial X^2}+V\left(\frac{(2)^{\frac{1}{2}}\hbar}{(m)^{\frac{1}{2}}}X\right), \tag{4.6}$$

which coincides with the form of the operator in Eq. (1.39). According to Eqs. (1.35) and (1.42), we obtain from this the result that

$$\begin{aligned}\psi(X,\beta;X)&=\int_{C_{\beta;0}}\exp\left\{-\int_0^\beta V\left[\frac{(2)^{\frac{1}{2}}\hbar}{(m)^{\frac{1}{2}}}(X+x(\tau))\right]d\tau\right\}d_{w(\beta,0)}x\\&=\sum_n\varphi_n^2(X)e^{-\beta E_n}\end{aligned} \tag{4.7}$$

and (since the functions $\varphi_n(X)$ are normalized),

$$Z=\sum_n e^{-\beta E_n}=\int_{-\infty}^{\infty}\left\{\int_{C_{\beta;0}}\exp\left\{-\int_0^\beta V\left[\frac{(2)^{\frac{1}{2}}\hbar}{(m)^{\frac{1}{2}}}\times(X+x(\tau))\right]d\tau\right\}d_{w(\beta,0)}x\right\}dX. \tag{4.8}$$

The formula obtained for Z appears at first glance to be comparatively complicated, but nevertheless on solving some questions it turns out to be convenient and significantly simplifies derivations.

In a number of cases, the system can be regarded as classical in a first approximation, while quantum corrections can be regarded as small additions to the classical results; to this approach corresponds the evaluation of Z with the aid of expanding this quantity in a series of powers of $\hbar$ (analogous to the quasi-classical approximation). If we proceed from formula (4.8), then with this aim it is natural first of all to make a change of variable $(2)^{\frac{1}{2}}\hbar/(m)^{\frac{1}{2}}X=X_0$, and then, in the formula so obtained,

$$Z=\frac{(m)^{\frac{1}{2}}}{(2)^{\frac{1}{2}}\hbar}\int_{-\infty}^{\infty}\left\{\int_{C_{\beta;0}}\exp\left[-\int_0^\beta V\left(X_0+\frac{(2)^{\frac{1}{2}}\hbar}{(m)^{\frac{1}{2}}}\times x(\tau)\right)d\tau\right]d_{w(\beta,0)}x\right\}dX_0 \tag{4.9}$$

to make use of the expansion of the function

$$\exp\left\{-\int_0^\beta V\left[X_0+\frac{(2)^{\frac{1}{2}}\hbar}{(m)^{\frac{1}{2}}}x(\tau)\right]d\tau\right\}$$

in a Taylor series in $\hbar$:

$$\begin{aligned}&\exp\left[-\int_0^\beta V\left(X_0+\frac{(2)^{\frac{1}{2}}\hbar}{(m)^{\frac{1}{2}}}x(\tau)\right)d\tau\right]\\&=e^{-\beta V(X_0)}-\frac{(2)^{\frac{1}{2}}\hbar}{(m)^{\frac{1}{2}}}\int_0^\beta x(\tau)d\tau\cdot V'(X_0)e^{-\beta V(X_0)}\\&\quad+\frac{\hbar^2}{2m}\left\{\int_0^\beta\int_0^\beta x(\tau)x(\sigma)d\tau d\sigma\cdot V'^2(X_0)\right.\\&\quad\left.-\int_0^\beta x^2(\tau)d\tau\cdot V''(X_0)\right\}e^{-\beta V(X_0)}+\cdots.\end{aligned} \tag{4.10}$$

The zero-order term of this expansion after substitution into (4.9) gives

$$Z\approx\frac{(m)^{\frac{1}{2}}}{(2\pi\beta)^{\frac{1}{2}}\hbar}\int_{-\infty}^{\infty}e^{-\beta V(X_0)}dX_0, \tag{4.11}$$

so that apart from a nonessential multiplicative factor[23] we arrive at the statistical integral of classical statistics,

$$Z_0=\int_{-\infty}^{\infty}e^{-\beta V(X_0)}dX_0. \tag{4.12}$$

All odd terms in (4.10) after substitution into (4.9) drop out in view of (1.24′). Therefore, the succeeding terms in the expansion of (4.10) give us a representation for Z in the form of the series

$$Z=\frac{(m)^{\frac{1}{2}}}{(2\pi\beta)^{\frac{1}{2}}\hbar}(Z_0+\hbar^2Z_2+\hbar^4Z_4+\cdots); \tag{4.13}$$

moreover, each term of this series contains only moments (1.26) of conditional Wiener measure $d_{w(\beta,0)}x$ and can be evaluated very simply. In particular, the

[22] The quantities $e^{-\beta E_n}$ in quantum statistics determine the probabilities of finding the system in a state with energy E_n, so that Z will enter into the mean value $\bar{A}$ of any physical quantity A:

$$\bar{A}=Z^{-1}\Sigma_n A_{nn}e^{-\beta E_n},$$

where A_{nn} are matrix elements of A. The connection of Z with thermodynamic quantities is determined by the formula, $-kT\times\log Z=F$, where F is the free energy of the system (see [34]).

[23] In the formula for $\bar{A}$, the constant factor in Z will be cancelled in the numerator and denominator; in the formula for F, it gives an unessential additive constant in the expression for the free energy.

second term of (4.13) in view of (1.25) gives

$$Z_2=\frac{1}{m}\Big[\int_{-\infty}^{\infty}V'^2(X_0)e^{-\beta V(X_0)}dX_0\cdot\int_0^\beta\int_0^\beta\Big\{\int_{C_{\beta;0}}x(\tau)x(\sigma)d^*_{w(\beta,0)x}\Big\}d\tau d\sigma-\int_{-\infty}^{\infty}V''(X_0)e^{-\beta V(X_0)}dX_0$$
$$\times\int_0^\beta\Big\{\int_{C_{\beta,0}}x^2(\tau)d^*_{w(\beta,0)x}\Big\}d\tau\Big]$$
$$=\frac{1}{m}\Big\{\int_{-\infty}^{\infty}V'^2(X_0)e^{-\beta V(X_0)}dX_0\cdot 2\int_0^\beta\int_0^\tau\frac{\tau(\beta-\sigma)}{2\beta}d\tau d\sigma-\int_{-\infty}^{\infty}V''(X_0)e^{-\beta V(X_0)}dX_0\cdot\int_0^\beta\frac{\tau(\beta-\tau)}{2\beta}d\tau\Big\}$$
$$=\frac{\beta^3}{24m}\int_{-\infty}^{\infty}V'^2(X_0)e^{-\beta V(X_0)}dX_0-\frac{\beta^2}{12m}\int_{-\infty}^{\infty}V''(X_0)e^{-\beta V(X_0)}dX_0=-\frac{\beta^3}{24m}\int_{-\infty}^{\infty}V'^2(X_0)e^{-\beta V(X_0)}dX_0; \tag{4.14}$$

The further terms of the series (4.13) are evaluated in slighly more difficult fashion (see reference [37]).

The results obtained here coincide with the results already obtained at the beginning of the thirties by Wigner and Kirkwood proceeding from Schrödinger's equation (see [34], Sec. 33); yet another elegant derivation of these same results was recently put forth by I. M. Khalatnikov [23b]. The method indicated by us for obtaining these results with the aid of functional integrals appears, however, to be substantially simpler than all the others.

We note further that formula (4.8) for Z (in application to significantly more complicated systems) was applied by R. Feynman [1c]–[1d] to the problem of the behavior of liquid helium at low temperatures.

Determination of the Smallest Characteristic Value of Schrödinger's Equation

Functional integrals can be made use of also for the evaluation of the lowest characteristic value of Schrödinger's equation (we note that applying the "quasi-classical approximation" we on the other hand obtained estimates for large characteristic values). Let

$$i\partial\psi/\partial t=H\psi \tag{4.15}$$

be such an equation in a system of units where $\hbar=1$ (for the time being we do not specify the form of the operator H); then, an arbitrary solution of this equation can be represented in the form

$$\psi=\sum_n C_n\varphi_n e^{-iE_nt}, \tag{4.16}$$

where the E_n are the characteristic values of the operator H, and the φ_n are the corresponding characteristic functions. In formula (4.16), it is comparatively difficult to separate the lowest characteristic value from all the rest, but if we proceed from (4.15), to the corresponding equation of the "heat conduction type,"

$$\partial\psi/\partial t=-H\psi, \tag{4.17}$$

and notice that the characteristic values and functions for (4.15) and (4.17) are identical and that the solution to (4.17) can be represented in the form

$$\psi=\sum_n C_n\varphi_n e^{-E_nt}, \tag{4.18}$$

then, it is clear that to find the lowest E_n it is only necessary to study the asymptotic behavior of the solution (4.18) for large t. But (4.15) is transformed into (4.17) by a change of t into $-it$. By carrying out the same transformation in formula (2.5), we find that the solution to Eq. (4.17) will be given by a functional integral over Wiener measure

$$\psi=\int_{C_{t,X}}e^{S_0[t_0,t]}d_{w(t,X)}x, \tag{4.19}$$

where $S_0[t_0,t]$ is that part of the classical action of the system remaining after the evaluation of the integral of the "kinetic energy" (the latter after the transformation $t\to -it$ in (2.5) immediately gives Wiener measure). For the investigation of the asymptotic behavior of this integral for large t Feynman in the work of reference [1e] applied the following method. We choose in the role of an approximation to the functional S_0 some simple real functional S_1 whose integral over Wiener measure can readily be evaluated (for example, with the aid of the reduction to the solution of some arbitrary well-studied differential equation). We rewrite (4.19) in the form

$$\psi=\int_{C_{t,X}}e^{(S_0-S_1)}e^{S_1}d_{w(t,X)}x. \tag{4.19'}$$

Then this expression can be considered as some mean of the functional $e^{S_0-S_1}$, evaluated with a positive "weight function" e^{S_1}. But in view of the general inequality between the geometric mean and the arithmetic mean ("Jensen's inequality"), the mean value of $e^{S_0-S_1}$ is never less than the exponential of the mean value of S_0-S_1; therefore, replacing S_0-S_1 in the exponent in (4.19') by the mean value $(S_0-S_1)_{Av}$ of this functional over the measure $e^{S_1[x]}d_wx$ and then removing $e^{(S_0-S_1)_{Av}}$ outside the functional integral sign, we obtain an upper bound on the value of ψ, i.e.,

(passing to the case of large t) the upper bound therefore for E_1. If the functional S_1 contains in addition several numerical parameters, then to obtain the best approximation to E_1 we must certainly solve a variational problem—to choose the parameters in such a way that the value obtained is the smallest.

In reference [1e] this method was applied to the problem of the motion of an electron in a polar crystal and led after several transformations to the investigation of the functional integral

$$\int_{C_{t,x}} \exp\left[\alpha \int\int \frac{e^{-|t-s|}}{|x(t)-x(s)|}dtds\right]d_{w(t,x)}x. \quad (4.20)$$

In the role of "approximating functionals" S_1 in [1e] are studied functionals of the form $\exp\{-\int V[x(t)]dt\}$, where either $k/|x|$ or kx^2 is assumed for $V(x)$. The asymptotic form of integrals over Wiener measure of $\exp[-\int V(x)dt]$ for such V is naturally related to the determination of the characteristic values of Schrödinger's equation for an electron in a Coulomb field and for a harmonic oscillator. The results obtained appear to be very accurate; comparison with the exact value for E_1 obtained by S. I. Pekar [35] for the case of very large α shows that choosing for $V(x)$ the functional kx^2 and determining the value of the parameter k from the solution of a variational problem we obtain E_1 with an error of less than 3%.

Perturbation Theory

We begin with an investigation of the simplest problem of the motion of a particle in a potential field. Let the potential $V[x(\tau)]$ be represented in the form,

$$V[x(\tau)]=V_0[x(\tau)]+\epsilon V_1[x(\tau)], \quad (4.21)$$

where $V_0[x(\tau)]$ is the "unperturbed" potential energy for which the solution of Schrödinger's equation is known [equal, say, to $\psi_0(X,t)$] while $\epsilon V_1(x)$ is a small perturbation (ϵ is a parameter of smallness). Then, according to (2.3),

$$\psi(X,t;X_0,t_0)=\frac{1}{N}\int_{t_0}^{t}\exp\left(\frac{i}{\hbar}\int_{t_0}^{t}\{L_0[x(\tau)]-\epsilon V_1[x(\tau)]\}d\tau\right)\prod_{\tau=t_0}^{t}dx(\tau), \quad (4.22)$$

where $L_0[x(\tau)]$ is the unperturbed Lagrangeian. From here, we easily obtain an expansion of the function ψ in a series of powers of ϵ:

$$\begin{aligned}\psi(X,t;X_0,t_0)&=\frac{1}{N}\int\exp\left\{\frac{i}{\hbar}\int_{t_0}^{t}L_0[x(\tau)]d\tau\right\}\left\{1-\frac{i\epsilon}{\hbar}\int_{t_0}^{t}V_1[x(\tau_1)]d\tau_1\right.\\&\left.+\frac{1}{2!}\left(\frac{i\epsilon}{\hbar}\right)^2\int_{t_0}^{t}\int_{t_0}^{t}V_1[x(\tau_1)]V_1[x(\tau_2)]d\tau_1d\tau_2+\cdots\right\}\prod_{t_0}^{t}dx(\tau)\\&=\psi_0(X,t;X_0,t_0)-\frac{i\epsilon}{\hbar}\int_{-\infty}^{\infty}\int_{t_0}^{t}\psi_0(x_1,\tau_1;X_0,t_0)V_1(x_1)\psi_0(X,t;x_1,\tau_1)dx_1d\tau_1\\&+\left(\frac{i\epsilon}{\hbar}\right)^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{t_0}^{t}\int_{t_0}^{\tau_1}\psi_0(x_2,\tau_2;X_0,t_0)V_1(x_2)\psi_0(x_1,\tau_1;x_2,\tau_2)V_1(x_1)\\&\times\psi_0(X,t;x_1,\tau_1)dx_1dx_2d\tau_1d\tau_2+\cdots \quad (4.23)\end{aligned}$$

and in general

$$\begin{aligned}\psi(X,t;X_0,t_0)&=\psi_0(X,t;X_0,t_0)+\sum_{n=1}^{\infty}\left(-\frac{i\epsilon}{\hbar}\right)^n\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\int_{0}^{t}\int_{0}^{\tau_1}\cdots\int_{0}^{\tau_{n-1}}\psi_0(x_n,\tau_n;X_0,t_0)V_1(x_n)\\&\times\psi_0(x_{n-1},\tau_{n-1};x_n,\tau_n)V_1(x_{n-1})\cdots V_1(x_1)\psi(X,t;x_1,\tau_1)dx_1dx_2\cdots dx_nd\tau_1d\tau_2\cdots d\tau_n. \quad (4.23')\end{aligned}$$

This is the usual result of the nonstationary perturbation theory of quantum mechanics.

The perturbation theory developed here has received a particularly wide application in problems of quantum field theory. We take as an example the same model of Thirring of a quantum field theory which we studied on page 59. Here the wave functional $\psi[u(x)]$ is represented by formula (2.14′). This functional is given as an integral of an exponential function in the exponent of which is a biquadratic form in $u(x)$. The evaluation of such an integral presents a very great difficulty, and at the present time there exist no general approaches to the solution of this problem. However, if we consider the term of the fourth order in the exponential to be small and make use of the expansion of the wave functional in a series of powers of λ (we note that the parameter λ in our theory plays the role of an elementary charge), then all terms of the resulting series will represent functional integrals of the type of moments of a Gaussian distribution which can be evaluated directly.[24] It

[24] We do not concern ourselves here with complications connected with the fact that in concrete evaluations very often infinite expressions and divergent integrals appear which must be treated by a special regularization; these complications have a physical but not a mathematical origin.

is obvious that the last circumstance possesses a general character and is not connected with the choice of a special model of a quantum field theory. By their nature, almost all investigations of a quantum field theory published up to the present time are concerned with the investigation of perturbation series which are obtained in this way.[25] For evaluating the terms of these series a very unusual and interesting mathematical technique has been developed in recent times (the so called "Feynman diagrams"); however, we cannot here go more deeply into these important but not simple questions.

BIBLIOGRAPHY

1. R. P. Feynman, (a) "Space-time approach to nonrelativistic quantum mechanics," Revs. Modern Phys. **20**, No. 2 (1948), pp. 367–387; see also the Russian translation in the collection "Questions of Causality in Quantum Mechanics," IL (1955); (b) "Mathematical formulation of the quantum theory of electromagnetic interaction," Phys. Rev. **80**, 440–457 (1950); (c) "Atomic theory of the λ transition in helium," Phys. Rev. **91**, 1291–1301 (1953); (d) "Atomic theory of liquid helium near absolute zero," Phys. Rev. **91**, 1301–1308 (1953); (e) "Slow electrons in a polar crystal," Phys. Rev. **97**, 660–665 (1955).
2. M. Kac, (a) On distributions of certain Wiener functionals, Trans. Am. Math. Soc. **65**, No. 1, 1–13 (1949); (b) On some connections between probability theory and differential and integral equations, Proc. 2nd Berkeley Symposium Math. State. and Probab., Berkeley, 189–215 (1951).
3. N. Wiener, (a) Differential Space, J. Math. and Phys. **2**, 131–174 (1923); (b) "The average value of a functional," Proc. London Math. Soc. Ser. 2, **22**, No. 6, 454–467 (1924); (c) "Generalized harmonic analysis," Acta Math. **55**, 117–258 (1930).
4. M. Rosenblatt, "On a class of Markov processes," Trans. Am. Math. Soc. **77**, No. 1, 120–135 (1951).
5. D. Ray, On spectra of second-order differential operators, Trans. Am. Math. Soc. **77**, No. 2, 299–321 (1954).
6. A. Blanc-Lapierre and R. Fortet, "Theorie des fonctions aleatoires," Paris, 1953, Chap. VII, Sec. III.
7. E. B. Dynkin, (a) "On certain limit theorems for Markov chains," Ukr. Math. Zhur. **6**, No. 1, 21–27 (1954); (b) "Functionals of the trajectories of Markov random processes," DAN **104**, No. 5, 691–694 (1955).
8. R. Z. Hac'minskii, "Distribution of probabilities for functionals of trajectories of a random process of the diffusion type," DAN **104**, No. 1, 22–25 (1955).
9. R. H. Cameron and W. T. Martin, (a) "Transformations of Wiener integrals under translations," Ann. Math. **45**, No. 2 386–396 (1944); (b) "Transformations of Wiener integrals under a general class of linear transformations," Trans. Am. Math. Soc. **58**, No. 2, 184–219 (1945); (c) "Evaluation of various Wiener integrals by use of certain Sturm-Liouville differential equations," Bull. Am. Math. Soc. **51**, No. 2, 73–90 (1945); (d) "Transformations of Wiener integrals by nonlinear transformations," Trans. Am. Math. Soc. **66**, No. 2, 253–283 (1949).
10. R. H. Cameron, (a) "A Simpson rule for the numerical evaluation of Wiener integrals in function space," Duke Math. J. **18**, No. 1, 111–130 (1951); (b) "The general heat flow equation and a corresponding Poisson formula," Ann. Math. **59**, No. 3, 434–461 (1954); (c) "The translation pathology of Wiener space," Duke Math. J. **21**, No. 4, 623–627 (1954).
11. R. H. Cameron and R. E. Fagen, "Nonlinear transformations of Volterra type in Wiener space," Trans. Am. Math. Soc. **75**, No. 3, 552–575 (1953).
12. E. Hopf, "Statistical hydromechanics and functional calculus," J. ration. mech. and analysis **1**, No. 1, 87–123 (1952).
13. S. F. Edwards and R. E. Peierls, "Field equations in functional form," Proc. Roy. Soc. London, **A224**, No. 1156, 24–33 (1954) [see also the Russian translation in the collection "Problems of contemporary physics," IL, No. 3 (1955)].
14. I. M. Gel'fand and R. A. Minlos, "The solution of the equations of quantized fields," DAN U.S.S.R. **97**, No. 2, 209–212 (1954).
15. N. N. Bogolyubov, "On the representation of Green-Schwinger functions with the aid of functional integrals," DAN U.S.S.R. **99**, No. 1, 225–226 (1954).
16. E. C. Fradkin, "Green's function for the interaction of nucleons with mesons," DAN **98**, No. 1, 47–50 (1954).
17 P. T. Matthews and A. Salam, "The Green's functions of quantized fields," Nuovo cimento **12**, No. 4, 563–565 (1954).
18. N. P. Kelpikov, (a) "On the theory of a vacuum functional, DAN **98**, No. 6, 937–940 (1954); (b) "The solution of systems of equations for a vacuum functional," DAN **100**, No. 6, 1057–1059 (1955).
19. U. A. Gol'fand, "Construction of distribution functions by the method of quasi-fields," J. Exptl. Theoret. Phys. **28**, No. 2, 140–150 (1955).
20. S. F. Edwards, (a) "The nucleon Green function in charged meson theory, Proc. Roy. Soc. London **A228**, No. 1174, 411–424 (1955); (b) "The nucleon Green function in pseudoscalar meson theory," I–II, Proc. Roy. Soc. London **A232**, No. 1190, 371–389 (1955).
21. J. Hamilton, "Functional analysis and strong-coupling theory," Phys. Rev. **97**, No. 5, 1390–1391 (1955).
22. T. H. R. Skyrme, "Quantum field theory," Proc. Roy. Soc. (London) **A231**, 321–335 (1955).
23. I. M. Khalatnikov, (a) "The representation of Green's function in quantum electrodynamics in the form of continual integrals," J. Exptl. Theoret. Phys. **28**, No. 5, 633–636 (1955); (b) "On one method of evaluating the statistical sum," DAN **87**, No. 4, 539–542 (1952).
24. A. N. Kolmogorov, "Basic ideas of the theory of probability," M.-L., ONTI (1936).
25. E. W. Montroll, "Markoff chains, Wiener integrals, and quantum theory," Commun. pure appl. math. **5**, 415–453 (1952).
26. B. V. Gnedenko, "A course in the theory of probability," M.-L., Gostekhizdat (1954).
27. U. V. Prohorov, "The distribution of probabilities in functional spaces," UMN **8**, No. 3 (55) 165–167 (1953).
28. W. Feller, "Introduction to the theory of probability and its applications," M., IL, (1952).
29. E. Titchmarsh, "Introduction to the theory of Fourier integrals," M.-L. Gostekhizdat (1948).
30. G. Doetsch, "Theorie und Anwendung der Laplace-Transformationen," Berlin, (1937).
31. R. Courant and D. Hilbert, "Methods of mathematical Physics," M.-L., GTTI (1934), Vol. 1.
32. W. Pauli, "Basic principles of Wave Mechanics," M.-L., Gostekhizdat (1947).
33. C. Morette, "On the definition and approximation of Feynman's path integrals," Phys. Rev. **81**, 848–852 (1951).
34. L. D. Landau and E. M. Lifshitz, "Statistical physics," M.-L., Gostekhizdat (1951).
35. S. I. Pekar, "Theory of polarons," J. Exptl. Theoret. Phys. **19**, No. 9, 796–806 (1949).
36. P. Choquard, "Traitement semi-classique des forces generales dans la representation de Feynman," Helv. Phys. Acta **28**, No. 2–3, 89–157 (1955).
37. A. M. Yaglom, "Application of functional integrals to the evaluation of the statistical sum of quantum statistics," Theory of probability and its applications **1**, No. 1 (1956).

[25] Exceptions to this are found in certain works in which the result is sought in the form of a series in powers of the constant $1/\lambda$ (the "method of strong coupling"). See also reference [13], in which for one very schematic model problem the evaluation of the functional integrals which arise is carried out in a general form.

Poincaré Cycles, Ergodicity, and Irreversibility in Assemblies of Coupled Harmonic Oscillators*

PETER MAZUR
Institute Lorentz, Leiden, The Netherlands

AND

ELLIOTT MONTROLL
Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Maryland

(Received January 10, 1960)

The transport coefficients (diffusion constant, electrical conductivity, etc.) associated with irreversible processes in an assembly of particles can be expressed as integrals over certain time relaxed correlation functions between small numbers of variables of the assembly. The scattering of slow neutrons is also a measure of time relaxed correlation functions.

Irreversibility is a consequence of the vanishing of the correlation coefficients as the relaxation time becomes infinite. On the other hand these coefficients have Poincaré cycles so that any value which they take on is repeated an infinite number of times. It is shown that, in the case of fluctuations of $0(N^{-\frac{1}{2}})$ from zero (N being the number of degrees of freedom), the period of Poincaré cycles is of the order of the mean period of normal mode vibrations while for fluctuations of a magnitude independent of N the period is of the order of C^N where C is a constant which is greater than 1.

The time relaxed correlation coefficients of a pair of particles separated by r lattice spacings decays as $t^{-m/2}$, m being the number of dimensions of the assembly. The statistics of the decay of the momentum of a particle from a preassigned initial value to its equipartition value are discussed.

1. INTRODUCTION

THIS year marks the 100th anniversary of the publication of the great Maxwell paper[1] entitled "Illustrations of the Dynamical Theory of Gases." The Maxwell velocity distribution was there first exhibited and the elementary Maxwell theory of transport processes of dilute gases there first presented. Maxwell's work was inspired by the now 101 year old paper of Clausius[2] on transport coefficients.

In celebration of this important anniversary the authors wish to make a small contribution to the development of the theory of irreversibility in statistical mechanics. A detailed analysis will be made of the manner in which irreversibility appears in the time development of a small number of variables embedded in an assembly of a large number of coupled harmonic oscillators. Some of the results and ideas presented are similar to those which appear in the works of van Hove[3] and Prigogine[4] and his collaborators. While this manuscript was in preparation, the authors received a copy of a doctoral dissertation by Hemmer[5] which also contains similar material. However, Hemmer's analysis is one dimensional and depends strongly on the analytical characteristics of linear chains of coupled oscillators. We have attempted to use rather general dynamical and statistical arguments which are independent of the dimensionality of the assembly under consideration.

At the turn of the century, certain sharp criticisms of the work of Boltzmann and Gibbs were made by Zermelo who claimed that a state of equilibrium of a mechanical assembly could not exist and therefore that statistical mechanics was nonsense. His argument was based on the existence of Poincaré cycles in closed dynamical assemblies. We shall now give a brief qualitative review of Zermelo's remarks, a more detailed proof being in Chandrasekhar's[6] review of the theory of stochastic processes.

Consider the total phase space available to an isolated closed mechanical assembly of total energy E, and fix attention on a subset of finite measure, or volume, $A(0)$ of the phase space. $A(0)$ represents a set of phase points associated with an ensemble of all assemblies which satisfy some special conditions: for example, all assemblies of N particles in the box (with perfectly reflecting walls) given in Fig. 1(b), which can be constructed so that all particles are in the cube marked A at time $t=0$ and the total energy is E. As time goes on, the phase points which were originally in $A(0)$ move through phase space. The shape of the envelope changes, but Liouville's equation (the equation of continuity in phase space) insures that the volume occupied by the

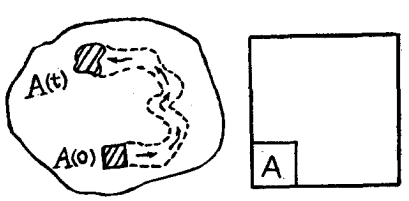

FIG. 1.

* This research was partially supported by the U. S. Air Force ARDC, through its European Office, and the Office of Scientific Research.

[1] J. C. Maxwell, Phil. Mag. (January, 1860).
[2] R. Clausius, Pogg. Ann. 105, 239 (1859).
[3] L. van Hove, Phys. Rev. 95, 249 (1954).
[4] G. Klein and I. Prigogine, Physica 19, 74, 89, 1053 (1953); I. Prigogine and R. Bingen, Physica 21, 299 (1955).
[5] P. C. Hemmer, "Dynamic and stochastic types of motion in the linear chain," Thesis, Trondheim, Norway, (1959).
[6] S. Chandrasekhar, Revs. Modern Phys. 15, 1 (1943).

points originally in $A(0)$ remains invariant. The trajectory of a phase point never crosses itself or any other trajectory in phase space (for otherwise the positions and momenta of an assembly at a given time would not yield a unique solution to the dynamical equations). Hence, eventually the set $A(t)$ must return to be identical with $A(0)$ and repeat its former course in a periodic way. At worst, the entire space would have to be filled with points generated by $A(t)$ before the repetition occurs. Since the entire phase space is finite and since the invariant volume of $A(t)$ is nonvanishing, the space eventually would be filled or a repetition of $A(0)$ would first occur. The length of time for a repetition is the Poincaré recurrence time of the set $A(0)$.

The existence of Poincaré cycles implied to Zermelo that no stationary equilibrium state was possible. On the other hand, through the investigation of certain probabilistic (nonmechanical) models, Ehrenfest and Smoluchowski showed that the existence of an equilibrium state with fluctuations need not be inconsistent with Poincaré cycles and indeed that small fluctuations from equilibrium should have small Poincaré cycles while large fluctuations should have very long Poincaré cycles.

Although the general discussion of Poincaré cycles of an entire assembly is philosophically interesting, we shall now exhibit several formulas which show that the Poincaré cycles which are most relevant in the analysis of typical experimental situations involve directly only a small number of variables.

Time dependent correlations exist between individual atoms even in an assembly at equilibrium. These can be observed through the scattering of slow neutrons or through certain magnetic resonance experiments. Van Hove[3] has shown that the scattering cross section depends only on a time relaxed *pair* distribution function which reflects the correlation of the position of two particles at different times. Two particle spin correlations can also be investigated by scattering experiments.

If time dependent external forces are applied to an assembly (electric or magnetic fields for example) or if flows, temperature gradients, concentration gradients, etc., are set up, the response of the assembly can be described through the solution of hydrodynamic equations. These equations are derived from conservation laws. The connection between the hydrodynamic equations and molecular dynamics is made through recently derived formulas, which relate the transport coefficients (diffusion constant, viscosity coefficients, electrical conductivity, etc.) of the hydrodynamic equations to averages over molecular motions. For example, the self-diffusion constant is[7–9]

$$D=\frac{1}{\beta nZ}\,\mathrm{trace}\int_0^\infty d\tau\int_0^\beta e^{-\beta H}J_z(-i\lambda\hbar)J_z(\tau)d\lambda, \quad (1)\dagger$$

where J_z is the z component of the current operator,

$$J=\textstyle\sum_j p_j e_j/m_j, \quad (2)$$

where p_j is the momentum of the jth particle, e_j its charge, and m_j its mass. Also Z is the partition function, n the number density and $\beta=1/kT$. The operator $J(\tau)$ is related to $J(0)$ by the well-known formula,

$$J(\tau)=e^{iH\tau/\hbar}J(0)e^{-iH\tau/\hbar}, \quad (3)$$

where H is the Hamiltonian of the assembly. The current at an imaginary time $(-i\lambda\hbar)$ is defined by an application of (3)†

$$J(-i\lambda\hbar)=e^{\lambda H}J(0)e^{-\lambda H}.$$

If (2) is substituted into (1), it can be shown (since all particles are equivalent in a one component assembly) that

$$\begin{aligned}D=&\frac{kTv}{Zm^2}\,\mathrm{trace}\int_0^\infty d\tau\int_0^\beta\sum_{j,k}e^{-i\tau H/\hbar}\\&\times p_{j\mu}e^{i\tau H/\hbar}e^{-\lambda H}p_{k\mu}e^{-(\beta-\lambda)H}d\lambda\\=&\frac{kTv}{Zm^2}N\,\mathrm{trace}\int_0^\infty d\tau\int_0^\beta e^{-i\tau H/\hbar}\\&\times p_{1\mu}e^{i\tau H/\hbar}e^{-\lambda H}p_{1\mu}e^{-(\beta-\lambda)H}d\lambda\\&+\frac{N(N-1)kTv}{m^2Z}\,\mathrm{trace}\int_0^\infty d\tau\int_0^\beta e^{-i\tau H/\hbar}\\&\times p_{1\mu}e^{i\tau H/\hbar}e^{-\lambda H}p_{2\mu}e^{-(\beta-\lambda)H}d\lambda.\end{aligned}$$

The first integral is merely an auto-correlation function and the second a pair correlation function. Hence the diffusion constant is a reflection of the direct correlation of at most *two* particles. This is also true for the electrical conductivity.[10] The viscosity coefficient is expressible in terms of *four* particle time relaxed correlations and the thermal conductivity six particle correlations. Although the more exotic transport coefficients such as thermal diffusion coefficient involve higher correlations, all the measureable coefficients involve direct correlations between very few variables. M. Green[11] has also derived formulas for the transport coefficients in terms of correlations.

In summary, if we wish to examine the effect of Poincaré cycles on observable quantities we need only be concerned with the cycles associated with small numbers of variables.

Fortunately, the complete dynamics of a set of coupled harmonic oscillators can be exhibited in a simple mathematical form. This paper is concerned with the calculation of time dependent harmonic oscillator correlation functions and the demonstration of how in the limit of a large number of degrees of freedom Poin-

[7] S. Nakajima, Progr. Theoret. Phys. **21**, 948 (1958).
[8] E. W. Montroll, Il Nuovo cimento (to be published).
[9] H. Mori, Phys. Rev. **112**, 1829 (1958).
† Equations are numbered beginning with (1) in each section. If an equation in another section is referred to, it will contain the section number, for example, (1.1).
[10] R. Kubo, J. Phys. Soc. (Japan) **12**, 570 (1957).
[11] M. Green, J. Chem. Phys. **22**, 398 (1954).

caré cycles behave in the manner observed by Ehrenfest for special nonmechanical models.

It is well known from the theory of stochastic processes that the character of a stochastic process can be deduced from auto or joint correlation functions of variables whose time development is generated by the process. We shall show that the momentum of a single oscillator is generated by a Gaussian random process when the oscillator is coupled properly to an infinite set of other oscillators.

Before proceeding with the required analysis a few relevant theorems on ergodic and stochastic functions will be listed.

2. THEOREM ON ERGODIC FUNCTIONS

Consider a stationary stochastic process $x(t)$. Without restricting the generality of the argument, we shall assume that its average value is equal to zero,

$$E\{x(t)\}=0, \tag{1}$$

where E is the symbol for the mathematical expectation or average value of $x(t)$. Averages are determined from an ensemble of observations. The process shall be normalized in such a way that the dispersion is equal to unity,

$$E\{|x(t)|^2\}=1. \tag{2}$$

The stationarity of the process implies that the correlation function (x^* is complex conjugate of x)

$$\rho(\tau)=E\{x(t+\tau)x^*(t)\} \tag{3}$$

depends only on τ and is an even function of τ. The process $x(t)$ is called continuous if $\rho(\tau)$ is continuous at $\tau=0$; i.e., if

$$\rho(0+)=1. \tag{4}$$

Then $\rho(\tau)$ is an everywhere continuous function of τ as follows from the inequality of Schwarz. Khinchine[12] has shown that the correlation function of a continuous stationary process $x(\tau)$ may be represented as a Fourier-Stieltjes integral,

$$\rho(\tau)=\int_{-\infty}^{\infty} e^{i\tau\omega}dF(\omega), \tag{5}$$

where $F(\omega)$, the spectrum of the process $x(\tau)$, is a never decreasing function of bounded variation. It follows from (5) that

$$\rho(0)=F(\infty)-F(-\infty)=1. \tag{6}$$

On the other hand Cramer[13] has shown that the stationary process $x(t)$ itself has a spectral representation

$$x(t)=\int_{-\infty}^{\infty} e^{i\omega t}dy(\omega), \tag{7}$$

where the process $y(\omega)$ has orthogonal increments, with (see Appendix I)

$$E\{|y(b)-y(a)|^2\}=F(b)-F(a). \tag{8}$$

From (5) and (7) the following results can be established (see Appendix II)

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T x(t)dt=y(0+)-y(0-) \tag{9}$$

$$\lim\frac{1}{T}\int_0^T \rho(\tau)d\tau=F(0+)-F(0-)$$
$$=E\{|y(0+)-y(0-)|^2\}. \tag{10}$$

The result (9) expresses the so-called "law of large numbers" for the stationary process $x(t)$. In deriving (10) we have employed (8).

The following theorem follows immediately from (9) and (10):

Theorem: If $\rho(\tau)\to 0$ as $\tau\to\infty$ the function $x(t)$ is ergodic. Indeed if $\rho(\tau)\to 0$ as $\tau\to\infty$ then from (10)

$$E\{|y(0+)-y(0-)|^2\}=0, \tag{11}$$

and consequently from (9), for almost all initial conditions of the representative ensemble,

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T x(t)dt=0. \tag{12}$$

This proves the ergodic nature of the function $x(t)$ since now the time average (12) of $x(t)$ is equal, for almost all initial conditions, to its ensemble average.

In statistical mechanics one is interested in the ergodic nature of some function $x\{p(t),q(t)\}$ depending on the canonical momenta $p(t)$ and coordinates $q(t)$ of a conservative mechanical system. If in the preceding discussion averages denoted by E represent averages over a surface of constant energy in phase space (over a micro-canonical ensemble), the foregoing theorem states that the dynamical function $x(p,q)$ is ergodic when its phase correlation function $\rho(\tau)$ (on the surface of constant energy) tends to zero as $t\to\infty$. In this last form, the theorem is contained in Khinchine's[14] monograph on the mathematical foundation of statistical mechanics.

3. DYNAMICS OF HARMONIC OSCILLATOR ASSEMBLIES

Most of the analysis in this paper will be concerned with harmonic oscillators coupled in periodic arrays. For simplicity, we associate one degree of freedom with each lattice point of an array. Generally, each atom in a crystal has three degrees of freedom. However, it is well known that the motions in the x, y, and z direction

[12] A. I. Khinchine, Math. Ann. **190**, 604 (1934).
[13] H. Cramer, Ack. Mat. Astr. Fys. **28B**, No. 12 (1942).
[14] A. I. Khinchine, *Statistical Mechanics* (Dover Publications, New York, 1949).

of the atoms in simple cubic lattices are independent of each other.[15,16] Hence the one degree of freedom model would be applicable to a discussion of relaxation processes in such lattices. Motions in each direction would be treated separately. Actually, most of the results given in later sections can be generalized to other Bravais lattices in which motions in all directions are coupled.

We shall always discuss linear chains, square lattices, and simple cubic lattices by applying the Born-Karman periodic boundary conditions. The normal mode frequencies of an $N=M^n$, n-dimensional lattice are [15,16]

$$m\omega^2=2\sum_{j=1}^{n}\gamma_j\left(1-\cos\frac{2\pi k_j}{M}\right) \quad k_j=1, 2, \cdots, M$$
$$=4\sum_j \gamma_j \sin^2\pi k_j/M, \tag{1}$$

where γ is the force constant in various directions and m is the mass of each particle. In an isotropic $3D$ lattice, γ_1 can be chosen to be the central force constant and $\gamma_2=\gamma_3$ the noncentral force constants (with $\gamma_1\gg\gamma_2=\gamma_3$).

The frequency spectra[16] or normal mode distribution functions (as $N\rightarrow\infty$) associated with (1) are sketched in Fig. 2(a). Here $g(\omega)d\omega$ is the fraction of frequencies between ω and $(\omega+d\omega)$, while ω_L is the largest frequency. When $n=1$ (linear lattice)

$$g(\omega)=2/\pi(\omega_L^2-\omega^2)^{\frac{1}{2}} \quad m\omega_L^2=4\gamma_1. \tag{2}$$

When $n=2$ (square lattice) with $\gamma_1>\gamma_2$,

$$g(\omega)=\begin{cases}\dfrac{4\omega}{\pi^2[\omega^2(\omega_1^2+\omega_2^2-\omega^2)]^{\frac{1}{2}}} \\ \quad\times K\left(\dfrac{\omega_1\omega_2}{[\omega^2(\omega_1^2+\omega_2^2-\omega^2)]^{\frac{1}{2}}}\right) & \omega_2^2<\omega^2<\omega_1^2 \\ \dfrac{4\omega}{\pi^2\omega_1\omega_2}K\left(\dfrac{[\omega^2(\omega_1^2+\omega_2^2-\omega^2)]^{\frac{1}{2}}}{\omega_1\omega_2}\right) \\ \qquad\qquad \omega^2<\omega_2^2;\quad \omega_1^2<\omega^2<\omega_L^2,\end{cases} \tag{3}$$

where

$$m\omega_1^2=4\gamma_1, \quad m\omega_2^2=4\gamma_2, \quad \text{and} \quad m\omega^2=4(\gamma_1+\gamma_2) \tag{4}$$

and $K(k)$ is the complete elliptic integral of the second kind.

The $3D$ spectrum is sketched in Fig. 2(c). No simple formula exists in terms of standard functions. However,

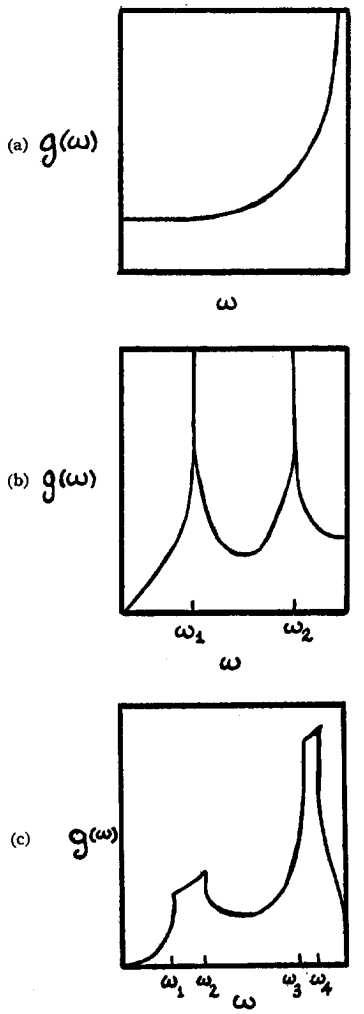

FIG. 2. Schematic frequency spectra of simple cubic lattice.

singularities exist at points $\omega_1, \cdots, \omega_4$, and ω_L, where

$$\begin{aligned} &m\omega_1^2=4\gamma_2 && m\omega_3^2=4\gamma_1 \\ &m\omega_2^2=4(\gamma_3+\gamma_2) && m\omega_4^2=4(\gamma_1+\gamma_3) \\ &m\omega_L^2=4(\gamma_1+\gamma_2+\gamma_3) && \gamma_1>\gamma_2>\gamma_3. \end{aligned} \tag{5}$$

The asymptotic form near $\omega=0$, ω_L, and ω_j $(j=1, \cdots, 4)$ is summarized in

$$g(\omega)=\frac{m^{\frac{3}{2}}}{\pi^2(\gamma_1\gamma_2\gamma_3)^{\frac{1}{2}}}\{u(\omega)-(\omega_1^2-\omega^2)^{\frac{1}{2}}\}\omega^2/\omega_1 \qquad 0<\omega<\omega_1 \tag{6a}$$

$$g(\omega)=\frac{\omega m^{\frac{3}{2}}u(\omega)}{\pi^2(\gamma_1\gamma_2\gamma_3)^{\frac{1}{2}}} \quad \omega_1<\omega<\omega_2 \quad \text{and} \quad \omega_3<\omega<\omega_4 \tag{6b}$$

$$g(\omega)=\frac{\omega m^{\frac{3}{2}}}{\pi^2(\gamma_1\gamma_2\gamma_3)^{\frac{1}{2}}}\left\{u(\omega)-\frac{1}{2}\frac{(\omega^2-\omega_2^2)^{\frac{1}{2}}(\omega_3^2-\omega^2)^{\frac{1}{2}}}{(\omega_3^2-\omega_2^2)^{\frac{1}{2}}}\right\} \qquad \omega_2<\omega<\omega_3 \tag{6c}$$

$$g(\omega)=\frac{\omega m^{\frac{3}{2}}}{\pi^2(\gamma_1\gamma_2\gamma_3)^{\frac{1}{2}}}\{u(\omega)-(\omega^2-\omega_4^2)^{\frac{1}{2}}\}\frac{(\omega_L^2-\omega^2)^{\frac{1}{2}}}{\omega_1^2}, \qquad \omega_4<\omega<\omega_L. \tag{6d}$$

[15] H. B. Rosenstock and G. F. Newell, J. Chem. Phys. **21**, 1607 (1953).

[16] E. W. Montroll, Proc. Third Berkeley Symposium on Math. Stat. and Prob. **3**, 209 (1957).

Here $u(\omega)$ is a continuous function of ω with continuous derivatives.

The Hamiltonian of our "one degree of freedom per lattice point" model is

$$H=\frac{1}{2m}\sum_{j=1}^{N}|p_j|^2+\sum \tfrac{1}{2}q_jA_{jk}q_k^*, \tag{7}$$

where p_j is the momentum of the jth particle and q_j its displacement from its equilibrium position. The A_{jk} matrix depends on force constants and dimensionality of the lattice. Although the p's and q's are real, it is convenient in the case of periodic boundary conditions to use complex normal coordinates.

By transforming to normal coordinates $\{P_k\}$ and $\{Q_k\}$ through

$$p_j=\sum_k C_{jk}P_k \tag{8a}$$

$$q_j=\sum_k C_{jk}Q_k, \tag{8b}$$

one obtains a new Hamiltonian,

$$H=\tfrac{1}{2}\sum_{j=1}^{N}\left(\frac{1}{m}|P_j|^2+m\omega_j^2|Q_j|^2\right), \tag{9}$$

where ω_j is the jth normal mode frequency and the matrix $C=(C_{jk})$ satisfying

$$\sum_j C_{jk}C_{jl}^*=\delta_{jl}. \tag{10}$$

The variation of the normal mode Q_k's and P_k's with time is obtained by solving the equations of motion,

$$m\ddot{Q}_k+\omega_k^2Q=0. \tag{11}$$

One finds

$$Q_k(t)=(P_k(0)/m\omega_k)\sin t\omega_k+Q_k(0)\cos t\omega_k \tag{12}$$

$$P_k(t)=P_k(0)\cos t\omega_k-m\omega_kQ_k(0)\sin t\omega_k. \tag{13}$$

The momentum of the jth particle at time t is

$$p_j(t)=\sum_k a_{jk}p_k(0)+\sum_k b_{jk}Q_k(0), \tag{14}$$

where

$$a_{jk}=\sum_s C_{js}C_{ks}^*\cos t\omega_s, \tag{15}$$

$$b_{jk}=-C_{jk}m\omega_k\sin t\omega_k. \tag{16}$$

The orthogonality condition (10) implies that

$$\sum_j a_{jj}=\sum_k \cos t\omega_k. \tag{17}$$

It also implies that

$$\sum_k\{|a_{jk}|^2+(m\omega_k)^{-2}|b_{jk}|^2\}=1. \tag{18}$$

In an n-dimensional simple cubic lattice the C_{jk}'s have the form

$$C_{jk}=\frac{1}{M^{n/2}}\exp 2\pi i\Big(\sum_{\alpha=1}^{n}j_\alpha k_\alpha\Big)/M, \tag{19}$$

where j and k are to be interpreted as vectors with components $\{j_\alpha\}$ and $\{k_\alpha\}$. Each k and j runs through the integers

$$j_\alpha=1, 2, 3, \cdots, M,$$
$$k_\alpha=1, 2, 3, \cdots, M.$$

It is sometimes convenient to transform to normal coordinates through an orthogonal transformation

$$(p_j,q_j)=\sum_k U_{jk}(P_k,Q_k) \tag{20}$$

such that

$$\sum_j U_{jk}U_{jl}=\delta_{kl}. \tag{21}$$

The diagonalized Hamiltonian is then

$$H=\tfrac{1}{2}\sum_{j=1}^{N}\left(\frac{1}{m}P_j^2+m\omega_j^2Q_j^2\right). \tag{22}$$

Equation (14) is still valid, but the coefficients a_{jk} and b_{jk} are now defined by

$$a_{jk}=\sum_s U_{js}U_{ks}\cos t\omega_s \tag{23}$$

$$b_{jk}=-U_{jk}m\omega_k\sin t\omega_k. \tag{24}$$

The analog of (18) is

$$\sum_k\{a_{jk}^2+(m\omega_k)^{-2}b_{jk}^2\}=1. \tag{25}$$

4. CLASSICAL STATISTICS OF HARMONIC OSCILLATOR ASSEMBLIES

The time relaxed correlation function between two particles separated by a lattice vector r in an assembly of $N=M^n$ oscillators is

$$\rho_N(t,r)=F_N(t,r)/F_N(0,0), \tag{1}$$

where [using periodic boundary conditions and Eqs. (3.10) and (3.13)]

$$F_N(t,r)=\tfrac{1}{2}E\{p_{s+r}(t)p_s^*(0)+p_s(t)p_{s+r}^*(0)\} \tag{2}$$

$$=\tfrac{1}{2}E\{\sum_{k,l}(C_{s+r,k}C_{s,l}^*+C_{s,k}C_{s+r,l}^*)\times[P_k(0)\cos t\omega_k-m\omega_kQ_k(0)\sin t\omega_k]P_l^*(0)\}. \tag{3}$$

If, as is the case in an initially canonical or microcanonical ensemble, a particle has the same probability of possessing a momentum p as $-p$, and if initially equipartition exists,

$$E\{P_k(0)P_l^*(0)\}=mkT\delta_{kl}, \tag{4}$$

$$E\{Q_k(0)P_l^*(0)\}=0 \tag{5}$$

so that, by use of (3.19),

$$F_N(t,r)=\frac{mkT}{M^n}\sum_k\cos\frac{2\pi r\cdot k}{M}\cos t\omega_k \tag{6}$$

and

$$\rho_N(t,r)=\frac{1}{M^n}\sum_k\cos\frac{2\pi r\cdot k}{M}\cos t\omega_k. \tag{7}$$

In particular, the autocorrelation function of any oscillator is found by setting $r=0$,

$$\rho_N(t)=M^{-n}\sum_k\cos t\omega_k. \tag{8}$$

The remainder of this section is devoted to general remarks about the autocorrelation function. The properties of the joint correlation function in the limit as $N \to \infty$ are discussed in Sec. 5. The summation extends over all normal mode frequencies $\{\omega_k\}$. The only statistical hypothesis made in the derivation of (8) concerns the nature of the ensemble at time $t=0$. Dynamics relates the properties at time t to those at $t=0$. Note that (8) is valid in both the initial microcanonical and canonical ensembles.

Various remarks can be made about $\rho_N(t)$ without employing detailed information about the ω's but only by using statistical properties of the set $\{\omega_j\}$. First, if the ω_k's are distinct,

$$f_N(t)=\sum_{k=1}^{N}\cos t\omega_k \tag{9}$$

is almost periodic so that any value which is achieved once will be achieved an infinite number of times. The average frequency with which any value of such a sum of cosines will be achieved has been calculated by Kac[17] (a shortened proof for physicists is given in Appendix III).

Let $N_{\Delta T}(q)$ be the number of zeros of $(f_N(t)-q)$ in the interval ΔT. Then the mean frequency for the achievement of q by $f_N(t)$ is

$$L(q)=\lim_{\Delta T\to\infty}\frac{1}{\Delta T}N_{\Delta T}(q). \tag{10}$$

For large N Kac has shown that

$$L(bN^{\frac{1}{2}})\sim\frac{2}{\pi}\omega_0\exp(-\tfrac{1}{2}b^2), \tag{11}$$

where

$$\omega_0^2=\frac{1}{N}\sum_{k=1}^{N}\omega_k^2. \tag{12}$$

The formula being valid if as $N \to \infty$

$$\lim_{N\to\infty}\frac{1}{N^2}\sum_{k=1}^{N}\omega_k^4=0. \tag{13}$$

On the other hand, N. B. Slater[18] has shown that when $(N-q)$ is small and N is large

$$L(q)\sim\frac{N^{\frac{1}{2}}\omega_0}{2\pi\Gamma(\frac{1}{2}N+\frac{1}{2})}\left(\frac{N-q}{2\pi}\right)^{\frac{1}{2}(N-1)}. \tag{14a}$$

Application of Stirling's theorem and the introduction of the parameter α (with $0<\alpha<1$)

$$\alpha=(q-1)/(N-1)\sim(q/N) \tag{14b}$$

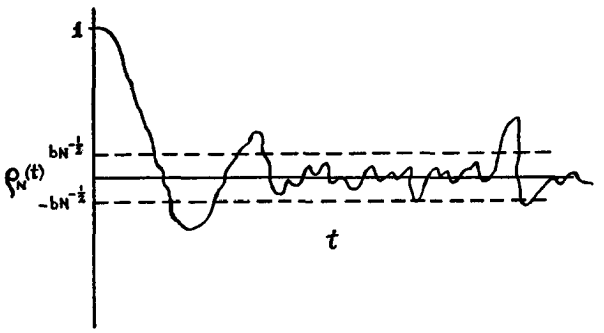

FIG. 3. Schematic autocorrelation function of momentum of particle in a simple cubic lattice.

yields

$$L(q)\sim\frac{\omega_0}{2\pi^{\frac{1}{2}}}\left(\frac{1-\alpha}{\pi e}\right)^{\frac{1}{2}(N-1)}\sim L(\alpha N). \tag{14c}$$

If equilibrium is to be established as $t \to \infty$, $p(t)$ and $p(0)$ must become independent so that $\rho_N(t) \to 0$. On the other hand if N is finite but large, we know that Poincaré cycles exist, and therefore whatever value of $\rho_N(t)$ is achieved at some time is achieved over and over again. The passage of $\rho_N(t) \to 0$ and at the same time the existence of Poincaré cycles can be reconciled in the following manner:

First let N be very large and consider a value of $\rho_N(t)=(1/N)f_N(t)$ in the range $(-bN^{-\frac{1}{2}}, bN^{-\frac{1}{2}})$ where b is of 0(1) and independent of N. From Eq. (11) the frequency of a Poincaré cycle is

$$(2\omega_0/\pi)\exp(-\tfrac{1}{2}b^2). \tag{15}$$

On the other hand, consider a value of $0<\alpha<1$ which is independent of N. Then the mean time between repetitions of a value (Fig. 3) of $f_N(t)=\alpha N$ or of $\rho_N(t)=\alpha$ is (from 14c)

$$\frac{2\pi^{\frac{1}{2}}}{\omega_0}\left(\frac{e\pi}{1-\alpha}\right)^{\frac{1}{2}(N-1)}. \tag{16}$$

Hence as $N \to \infty$ the period of Poincaré cycles becomes enormously large for values of $\rho_N(t)$ outside the "fluctuation range" $\pm bN^{\frac{1}{2}}$ but are of the order of the reciprocal of the root mean square natural frequency ω_0 of an assembly for values of $\rho_N(t)$ inside the fluctuation range. As $N \to \infty$ the fluctuation range approaches zero so that a true equilibrium situation develops with $\rho_N(t)$ approaching zero as $t \to \infty$.

The function $\rho_N(t)$ is that associated with either a canonical or microcanonical ensemble and the mean recurrence times are also associated with the ensembles. As in the case of equilibrium statistical mechanics, one can show that fluctuations from the average ensemble behavior become small as the number of degrees of freedom N becomes large. This will be discussed elsewhere.

[17] M. Kac, Am. J. Math. 65, 609 (1943).
[18] N. B. Slater, Proc. Cambridge Phil. Soc. 35, 56 (1939).

Let us now consider the statistics of the momentum of a particle which is known to have the momentum $p_j(0)$ at time $t=0$ while all other $(2N-1)$ variables required to describe the assembly are distributed initially according to a microcanonical distribution such that the total energy of the assembly is $\mathcal{E}$. Then (3.22) yields

$$\sum_{k\neq j}\frac{1}{2m}p_k^2(0)+\sum_{k=1}^{N}\tfrac{1}{2}m\omega_k^2Q_k^2(0)=R^2=\mathcal{E}-\frac{1}{2m}p_j^2(0). \quad (17)$$

Now

$$p_j(t)-a_{jj}p_j(0)=\sum_{k\neq j}a_{jk}p_k(0)+\sum_k b_{jk}Q_k(0), \quad (18)$$

where in an assembly with periodic boundary conditions we apply (3.17) (3.19), and (3.18) to find

$$a_{jj}=\sum_s C_{js}C_{js}^*\cos l\omega_s=\frac{1}{N}\sum_s\cos l\omega_s$$

$$=\rho_N(l). \quad (19)$$

The mean value (18) when averaged over (17) vanishes because $p_k(0)$ and $Q_k(0)$ have the same probability of achieving positive and negative values of the same absolute value.

The distribution function of

$$Y(l)=p_j(l)-p_j(0)\rho_N(l) \quad (20)$$

is the same as that of the right hand side of (18). We now find this distribution when every value of the set of p_k's and $Q_k(0)$'s consistent with (17) is given equal weight.

To this end we find the distribution function of

$$Y=\sum\alpha_jx_j \quad (21)$$

when the x_j's are distributed uniformly on the ellipsoid

$$\sum_{k=1}^{n}\beta_kx_k^2=R^2. \quad (22)$$

The various α's and β's are to be identified with the parameters of (17) and (18). Our required distribution function is the fourier transform of the characteristic function $f(\alpha)/f(0)$, where

$$f(\alpha)=\int\cdots\int_{\sum\beta_kx_k^2=R^2}\exp(i\alpha\sum\alpha_jx_j)dx_1\cdots dx_n. \quad (23)$$

Let $x_j^2=y_j^2/\beta_j$ and $\gamma_j=\alpha_j/\beta_j^{\frac{1}{2}}$ and convert the resulting spherically restricted integral (over $\sum y_k^2=R^2$) $f(\alpha)$ to an unrestricted integral through the introduction of the Fourier integral representation of the δ function,

$$\delta(R^2-\sum y_k^2)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp i\beta(R^2-\sum y_k^2)d\beta$$

$$=\frac{\exp(bR^2)}{2\pi}\int_{-\infty}^{\infty}\exp(i\beta R^2)$$

$$\times\exp\{-(b+i\beta)\sum y_k^2\}d\beta, \quad (24)$$

where b is any positive number.

Then, if we let $\sigma^2=\sum\gamma_k^2$,

$$f(\alpha)=\frac{e^{bR^2}}{2\pi(\beta_1\cdots\beta_n)^{\frac{1}{2}}}\int_{-\infty}^{\infty}e^{i\beta R^2}d\beta\prod_{j=1}^{n}\int_{-\infty}^{\infty}e^{i\alpha\gamma_jy_j}$$

$$\times\exp\{-y_j^2(b+i\beta)\}dy_j$$

$$=\frac{\pi^{n/2}\exp bR^2}{2\pi(\beta_1\cdots\beta_n)^{\frac{1}{2}}}\int_{-\infty}^{\infty}d\beta\exp\{i\beta R^2-\alpha^2\sigma^2/4(b+i\beta)\}$$

$$\times(b+i\beta)^{-n/2}$$

$$=\frac{\pi^{n/2}}{(\beta_1\cdots\beta_n)^{\frac{1}{2}}}\left(\frac{2R}{\alpha\sigma}\right)^{(n/2)-1}J_{(n/2)-1}(\alpha\sigma R). \quad (25)$$

One finds after employing the power series expansion for the Bessel function $J_n(x)$

$$f(0)=\frac{\pi^{n/2}}{(\beta_1\cdots\beta_n)^{\frac{1}{2}}}\frac{R^{2(\frac{1}{2}n-1)}}{(\frac{1}{2}n-1)!}, \quad (26)$$

so that

$$f(\alpha)/f(0)=(\tfrac{1}{2}n-1)!(2/R\alpha\sigma)^{\frac{1}{2}n-1}J_{(n/2)-1}(\alpha\sigma R)$$

$$=1-(\alpha^2\sigma^2R^2/2n)+\cdots,$$

where the kth term in the expansion is, as $n\to\infty$,

$$\frac{(-1)^k(\alpha R)^{2k}}{2^{2k}k!(n/2)[(n/2)-1]\cdots[(n/2)+(k-1)]}$$

$$\sim\frac{(-1)^k\alpha^{2k}R^{2k}}{2^kn^kk!}.$$

Hence,

$$\lim_{n\to\infty}f(\alpha)/f(0)=\exp(-\tfrac{1}{2}\alpha^2\sigma^2R^2n^{-1}). \quad (27)$$

The distribution function of Y is, as $n\to\infty$,

$$F(Y)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-i\alpha Y}\exp(-\tfrac{1}{2}\alpha^2\sigma^2R^2n^{-1})d\alpha$$

$$=\frac{1}{\sigma R(2\pi/n)^{\frac{1}{2}}}\exp-\frac{ny^2}{2\sigma^2R^2}. \quad (28)$$

The values of α_k of (21) which correspond to the coefficients of $p_k(0)$ of (18) are a_{jk}, while those which are associated with the $Q_k(0)$'s are b_{jk}. The β_k's of (22) are similarly $(2m)^{-1}$ and $\frac{1}{2}m\omega_k^2$, the coefficients respectively

of $p_k^2(0)$ and $Q_k^2(0)$ of 17. Then the $\gamma_k=\alpha_k\beta_k^{-\frac{1}{2}}$'s are

$$(2m)^{\frac{1}{2}}a_{jk} \quad \text{and} \quad b_{jk}/(\tfrac{1}{2}m\omega_k^2)^{\frac{1}{2}},$$

so that

$$\sigma^2=\sum \gamma_k^2 = 2m\sum_k a_{jk}^2+\sum_k b_{jk}^2/(\tfrac{1}{2}m\omega_k^2)-2ma_{jj}^2. \quad (29)$$

The last term is subtracted because the term with $k=j$ is not included in (17) and (18). In view of (3.25) and (19)

$$\sigma^2=2m(1-a_{jj}^2)=2m\{1-[\rho_N(t)]^2\}. \quad (30)$$

Also, from (17),

$$R^2=\mathcal{E}-\frac{1}{2m}p_j^2(0). \quad (31)$$

Since the total energy $\mathcal{E}$ is $\frac{1}{2}(2N-1)kT$, as $N\to\infty$,

$$R^2=NkT, \quad (32)$$

and the n of (28) is $(2N-1)$.

If the Y of (28) is associated with (18), and therefore with (20), the distribution function of $Y(t)=p_j(t)-p_j(0)\rho_N(t)$ is, when N is large,

$$P_N[p_j(t)\,|\,p_j(0)]=\{2\pi mkT[1-\rho_N(t)]\}^{-\frac{1}{2}} \times\exp-\left\{\frac{[p_j(t)-p_j(0)\rho_N(t)]^2}{2mkT\{1-[\rho_N(t)]^2\}}\right\}. \quad (33)$$

This function is just the probability for a transition $p_j(0)\to p_j(t)$ in the time t. It is not surprising that the distribution is Gaussian because as $N\to\infty$ the variables x_j in (21) become independent so that the central limit theorem is applicable.

We shall show below that as $N\to\infty$ and $t\to\infty$, $\rho_N(t)\to 0$. However, before proceeding with the proof of this fact let us investigate some of its consequences.

As $t\to\infty$ $P_n[p_j(t)\,|\,p_j(0)]$ approaches an equilibrium Maxwellian distribution independently of $p_j(0)$. The manner in which the kinetic energy achieves its equipartition value is observed by considering

$$E\left\{\frac{1}{2m}p_j^2(t)\right\}=E\left\{\frac{1}{2m}[p_j(t)-p_j(0)\rho_N(t)]^2\right\} +2p_j(0)\rho_N(t)E\left\{\frac{1}{2m}[p_j(t)-p_j(0)\rho_N(t)]\right\} +\frac{1}{2m}p_j^2(0)\rho_N(t)E\{1\}, \quad (34)$$

where the average is to be taken over the distribution (33). The first expectation value on the right (34) is just the dispersion of the distribution function (33)

$$\tfrac{1}{2}kT\{1-[\rho_N(t)]^2\},$$

the second vanishes, and $E\{1\}=1$. Hence the expectation value of the kinetic energy is

$$E\left\{\frac{1}{2m}p_j^2(t)\right\}=\tfrac{1}{2}kT\{1-[\rho_N(t)]^2\}+\frac{1}{2m}p_j^2(0)\rho_N(t)$$

as $t\to\infty$ this approaches the equipartition value kT, while as $t\to 0$ it approaches $p_j^2(0)/2m$, since $\rho_N(t)\to 1$ as $t\to 0$.

As a further consequence of the fact that $\rho(\tau)\to 0$ as $\tau\to\infty$, we may apply the theorem quoted in Sec. 2 to conclude that $p_j(t)$ is ergodic; i.e., that for almost all initial conditions

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T p_j(t)dt=E\{p_j(0)\}=0,$$

the expectation value being taken over an initial canonical or microcanonical ensemble.

It also can be shown after some calculation that

$$R_n(t)=\frac{E\{(p_j^2(t)-\mathcal{E}_0)(p_j^2(0)-\mathcal{E}_0)\}}{E\{[p_j^2(0)-\mathcal{E}_0]\}^2}=[\rho_N(t)]^2,$$

where

$$\mathcal{E}_0=E\{p_j^2(0)\}.$$

Hence, as $N\to\infty$, $\mathcal{E}_0=mkT$ and

$$\lim_{t\to\infty}R_\infty(t)=0.$$

Thus $p_j^2(t)$ is also ergodic so that

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T p_j^2(t)dt=E\{p_j^2(0)\}=mkT.$$

One can finally show that the Gaussian character of $P_\infty[p_j(t)\,|\,p_j(0)]$ implies that any function which depends only on p_j and whose phase average exists is ergodic. Ergodicity does not appear as a general dynamical property of an assembly but as a property of a special class of functions of a small number of variables in an assembly composed of a large number of coupled degrees of freedom.

The average time that our oscillator spends in a specified momentum range can be obtained from a discussion of the function

$$D(p(t);a,b)=\begin{cases}1 \text{ if } a<p(t)<b\\ 0 \text{ otherwise.}\end{cases}$$

The fraction of time f_{ab} spent by the jth oscillator in the momentum range

$$a<p_j(t)<b$$

is

$$f_{ab}=\lim_{T\to\infty}\frac{1}{T}\int^T D(p_j(t);a,b)dt.$$

Since the function D is ergodic in the limit as $N\to\infty$, f_{ab} can, for almost all initial conditions, be calculated

from an initial ensemble average as well as from the time average. An initial canonical ensemble yields

$$f_{ab}=\frac{1}{(2\pi mkT)^{\frac{1}{2}}}\int_{-\infty}^{\infty}dpD(p;a,b)\exp(-p^2\beta/2m)$$
$$=\frac{1}{(2\pi mkT)^{\frac{1}{2}}}\int_a^b e^{-p^2\beta/2m}dp.$$

We thus see that for almost all initial conditions a single system will in a long time interval, spend most of the time within the "fluctuation region" of the momentum p_j, and spend only a small fraction of time in "improbable states" for which $p_j^2(t)\gg mkT$. This also implies that if a system is in such an improbable state it will on the average rapidly decay towards the "fluctuation region" $[p_j^2(t)=0(mkT)]$ and remain there most of the time. Alternatively, we may say that the mean recurrence time for states within the fluctuation region will be small, whereas they will be very large for improbable states. This illustrates the irreversible behavior of the momentum $p_j(t)$ in a single system.

We note that all these conclusions hold for a single system in the limit $N\to\infty$.

For finite but very large N, the behavior of $p_j(t)$ will be practically the same.

The irreversible behavior of $p_j(t)$ may be further illustrated as follows: According to the Gaussian form of $P_\infty[p_j(t)\,|\,p_j(0)]$, the probability that $p_j(t)$ at time t will differ from its conditional average $\rho(t)p_j(0)$ by an amount much larger than $(mkT)^{\frac{1}{2}}$ is very small. Thus there is a high probability for a single system to decay close to the average path, in agreement with the macroscopic concept of irreversible behavior.

5. DETAILED CALCULATION OF MOMENTUM CORRELATION FUNCTIONS

The one dimensional correlation function is found by combining (4.7) with (3.1) when $n=1$. Then

$$\rho_N(t,r)=\frac{1}{N}\sum_{k=1}^{N}\cos\frac{2\pi rk}{N}\cos\left(\tau\sin\frac{\pi k}{N}\right) \tag{1}$$

with

$$\tau=t\omega_L \quad\text{and}\quad \omega_L=2(\gamma/m). \tag{2}$$

In the limit as $N\to\infty$ let $\vartheta=\pi k/N$ and $d\vartheta=\pi/N$. Then

$$\rho_N(t,r)\sim\frac{1}{\pi}\int_0^\pi\cos2r\vartheta\cos(\tau\sin\vartheta)d\vartheta$$
$$=\frac{2}{\pi}\int_0^{\pi/2}\cos2r\vartheta\cos(\tau\sin\vartheta)d\vartheta=J_{2r}(\tau), \tag{3}$$

the Bessel function of order $2r$. The exact calculation of $\rho_N(t,r)$ is given in Appendix IV for arbitrary N.

It is merely a fortunate accident that $\rho_N(t,r)$ has such a simple form in the one dimensional lattice. In preparation for the $2D$ and $3D$ calculation of $\rho_N(t,0)\equiv\rho_N(t)$ we rewrite

$$\rho_N(t)=\frac{1}{N}\sum_j\cos t\omega_j\sim\int_0^{\omega_L}\cos t\omega\, g(\omega)d\omega, \tag{4}$$

where $g(\omega)d\omega$ is the fraction of circular frequencies between ω and $\omega+d\omega$. In the one-dimensional lattice,

$$g(\omega)=(2/\pi)(\omega_L^2-\omega^2)^{-\frac{1}{2}}. \tag{5a}$$

Then as $N\to\infty$

$$\rho_N(t)\sim\frac{2}{\pi}\int_0^{\omega_L}\frac{\cos t\omega d\omega}{(\omega_L^2-\omega^2)^{\frac{1}{2}}}=J_0(\tau). \tag{5b}$$

As $\tau\to\infty$ [we henceforth write $\rho(\tau)\equiv\rho_\infty(\tau)$],

$$\rho(\tau)\sim(2/\pi\tau)^{\frac{1}{2}}\sin\tau. \tag{6}$$

The value of the ω_0 which appears in the Poincaré cycle formulas (4.15) and (4.16) is, in the n-dimensional case,

$$\omega_0^2=\frac{4}{mM^n}\sum_{j,k_j}\gamma_j\sin^2(\pi k_j/M).$$
$$=\frac{4}{m\pi}\sum_j\gamma_j\int_0^\pi\sin^2\vartheta d\vartheta=\frac{2}{m}\sum\gamma_j$$
$$=\tfrac{1}{2}\omega_L^2. \tag{7}$$

The short time behavior of $\rho(t)$ is found from the expansion

$$\rho_N(t)=\frac{1}{N}\sum_{j=1}^{N}\cos t\omega_j\sim1-t^2\mu_2/2!+t^4\mu_4/4!-\cdots, \tag{8a}$$

where

$$\mu_{2k}=\frac{1}{N}\sum_{j=1}^{N}\omega_j^{2k}. \tag{8b}$$

For a $2D$ lattice of N^2 lattice points,

$$\mu_{2k}=\frac{2^{2k}}{m^kN^2}\sum_{j_1,j_2=1}^{M}\left(\gamma_1\sin^2\frac{\pi j_1}{M}+\gamma_2\sin^2\frac{\pi j_2}{M}\right)^k$$
$$=\frac{2^{2k}}{\pi^2m^k}\int_0^\pi\int(\gamma_1\sin^2\varphi_1+\gamma_2\sin^2\varphi_2)^kd\varphi_1d\varphi_2$$
$$=\frac{2^{2k}}{m^k}\sum_{l=0}^{k}\frac{k!\gamma_1^l\gamma_2^{k-l}}{l!(k-l)!}\left\{\frac{1}{\pi}\int_0^\pi\sin^{2l}\varphi_1d\varphi_1\right\}$$
$$\times\left\{\frac{1}{\pi}\int_0^\pi\sin^{2(k-l)}\varphi_2d\varphi_2\right\}$$
$$=\frac{1}{m^k}\sum_{l=0}^{k}\frac{\gamma_1^l\gamma_2^{k-l}k!(2l)!(2[k-l])!}{l!(k-l)!l!l!(k-l)!(k-l)!}. \tag{9}$$

The first few μ's are

$$\mu_2/2!=(\gamma_1+\gamma_2)/m;\ \mu_4/4!=[\tfrac{1}{4}(\gamma_1^2+\gamma_2^2)+\tfrac{1}{3}\gamma_1\gamma_2]/m^2$$

$$\mu_6/6!=\left[\frac{1}{36}(\gamma_1^3+\gamma_2^3)+\frac{1}{20}\gamma_1\gamma_2(\gamma_1+\gamma_2)\right]\Big/m^3. \quad (10)$$

As is well known in the theory of the asymptotic behavior of Fourier integrals, the asymptotic form of (4) for large t depends on the singularities of $g(\omega)$. As is shown in Appendix IV, the discontinuity of $g(\omega)$ at the end of a two-dimensional frequency spectrum yields, as $t\to\infty$,

$$\rho(t)\sim\frac{2\omega_L\sin t\omega_L}{\pi t\omega_1\omega_2}, \quad (11a)$$

where

$$m\omega_1^2=4\gamma_1 \quad \text{and} \quad m\omega_2^2=4\gamma_2. \quad (11b)$$

In a $3D$ simple cubic lattice of N^3 lattice points,

$$\mu_{2k}=\frac{2^k}{m^kN^3}\sum_{j_1j_2j_3=1}^{M}\left(\gamma_1\sin^2\frac{\pi j_1}{M}+\gamma_2\sin^2\frac{\pi j_2}{M}+\gamma_3\sin^2\frac{\pi j_3}{M}\right)^k. \quad (12)$$

The first few μ's are

$$\mu_2/2!=(\gamma_1+\gamma_2+\gamma_3)/m;$$

$$\mu_4/4!=[\tfrac{1}{4}(\gamma_1^2+\gamma_2^2+\gamma_3^2)+\tfrac{1}{3}(\gamma_1\gamma_2+\gamma_2\gamma_3+\gamma_3\gamma_1)]/m^2$$

$$\mu_6/6!=\left[\frac{1}{36}(\gamma_1^3+\gamma_2^3+\gamma_3^3)+\frac{1}{20}(\gamma_1\gamma_2^2+\gamma_2\gamma_3^2+\gamma_2\gamma_1^2+\gamma_3\gamma_2^2)+\frac{1}{15}\gamma_1\gamma_2\gamma_3\right]\Big/m^3.$$

The long time behavior of $\rho(t)$ is a consequence of the four internal and one end singularity of the $g(\omega)$ function which is plotted in Fig. 2(c).

The three-dimensional correlation function (4.7) is in the limit $N\to\infty$

$$\rho(t,r)=\frac{1}{(2\pi)^3}\int_0^{2\pi}\!\!\int\!\!\int\cos(r\cdot\varphi)\cos[t\omega(\varphi)]d^3\varphi$$

$$=\tfrac{1}{4}\{[I(t,r)+I(t,-r)]+\text{cc}\}, \quad (13)$$

where

$$I(t,r)=\frac{1}{(2\pi)^3}\int_0^{2\pi}\!\!\int\!\!\int\exp i[\varphi\cdot r+t\omega(\varphi)]d^3\varphi. \quad (14)$$

When t is very large this integral can be calculated by the method of stationary phase. The main contribution comes from the neighborhood of the stationary or critical points of $\omega(\varphi)$; those points for which $\text{grad}\omega(\varphi)=0$. These are of course the maxima, minima, and saddle points of the surfaces $\omega(\varphi)=\text{const}$. The critical points associated with the surfaces of constant frequency were first discussed by van Hove.[19]

We suppose that in the neighborhood of a critical point $\varphi^{(0)}$,

$$\omega=\omega_0+(2m\omega_0)^{-1}\sum_{j=1}^{3}\epsilon_j\gamma_j[\varphi_j-\varphi_j^{(0)}]^2+0(\varphi_j-\varphi_j^{(0)})^3, \quad (15)$$

where γ_j' is a force constant and ϵ_j' is ±1 (all are -1 in the neighborhood of a maximum and two are of one sign and one of the other in the neighborhood of a saddle point). Then, after substituting (15) into (14), we obtain the following contribution to I from the region of φ space near the critical point of interest (letting $\xi_j=\varphi_j-\varphi_j^{(0)}$),

$$\delta I(t,r)=\frac{\exp i[r\cdot\varphi^{(0)}+t\omega_0]}{(2\pi)^3}\times\int\cdot\int\exp i[r\cdot\xi+(t/2m\omega_0)\sum\epsilon_j\gamma_j\xi_j^2]d^3\xi.$$

Now let

$$x_j=\xi_j(t\gamma_j/2m\omega_0)^{\frac{1}{2}} \quad \text{and} \quad s_j=r_j(2m\omega_0/t\gamma_j)^{\frac{1}{2}}.$$

Then, as $t\to\infty$,

$$\delta I(t,r)\sim\left(\frac{2m\omega_0}{t}\right)^{\frac{3}{2}}\frac{\exp i[r\cdot\varphi^{(0)}+t\omega_0]}{(\gamma_1\gamma_2\gamma_3)^{\frac{1}{2}}}\times\prod_{j=1}^{3}\int_{-\infty}^{\infty}\exp(i\epsilon_jx_j^2)dx_j.$$

If $\epsilon=\pm1$,

$$\int_{-\infty}^{\infty}\exp(i\epsilon x^2)dx=\pi^{\frac{1}{2}}\exp(i\epsilon\pi/4).$$

Hence,

$$\delta I(t,r)\sim\left(\frac{2m\omega_0}{t}\right)^{\frac{3}{2}}\frac{\exp i[r\cdot\varphi^{(0)}+t\omega_0+\frac{1}{4}\pi\sum\epsilon_j]}{(\gamma_1\gamma_2\gamma_3)^{\frac{1}{2}}}. \quad (16)$$

The contribution to $\rho(t,r)$ from the lth stationary point is, as $t\to\infty$,

$$\rho^{(l)}(t,r)\sim\left(\frac{2m\omega_l}{t}\right)^{\frac{3}{2}}(\gamma_1\gamma_2\gamma_3)^{\frac{1}{2}}\times\cos(r\cdot\varphi^{(l)})\cos(t\omega_0+\tfrac{1}{4}\pi\sum\epsilon_j). \quad (17)$$

Hence the envelope of the correlation functions vanishes as $t^{-\frac{3}{2}}$ as $t\to\infty$.

The values of $\varphi^{(l)}$, $m\omega_l^2$ and $\sum\epsilon_j$ for the various critical points, with the exception of the minimum at (0,0,0) which contributes a term of $0(t^{-2})$ to $\rho(t,r)$, are listed in Table I. The point (π,π,π) is a maximum. It

[19] L. van Hove, Phys. Rev. 89, 1189 (1953).

TABLE I

$\varphi^{(l)}$	$m\omega_l^2$	$\Sigma\,\epsilon_j$
(π,π,π)	$4(\gamma_1+\gamma_2+\gamma_3)$	-3
$(0,0,\pi)$	$4\gamma_2$	1
$(0,\pi,0)$	$4\gamma_3$	1
$(\pi,0,0)$	$4\gamma_1$	1
$(\pi,0,\pi)$	$4(\gamma_1+\gamma_2)$	-1
$(\pi,\pi,0)$	$4(\gamma_1+\gamma_3)$	-1
$(0,\pi,\pi)$	$4(\gamma_2+\gamma_3)$	-1

yields the largest frequency $m\omega_L^2=4(\gamma_1+\gamma_2+\gamma_3)$. We generally assume that $\gamma_1>\gamma_2=\gamma_3$. These critical points were obtained from (3.1) when $n=3$. The total asymptotic expression for $\rho(t,r)$ is

$$\rho(t,r)=\sum_{l=1}^{7}\rho^{(l)}(t,r). \tag{18}$$

Van Hove has made similar calculations in developing the theory of the time relaxed pair distribution function of crystals (see also Prigogine and Bingen[4]).

It is interesting to calculate $\rho(t)$ for the Debye spectrum,

$$g(\omega)=\begin{cases}3\omega^2/\omega_L^3 & \omega<\omega_L\\ 0 & \omega>\omega_L.\end{cases} \tag{19}$$

Here

$$\rho(t)=\frac{3}{\omega_L^3}\int_0^{\omega_L}\omega^2\cos\omega t\,d\omega$$

$$=\frac{3}{\tau^3}\{2\tau\cos\tau+(\tau^2-2)\sin\tau\}$$

with $\tau=t\omega_L$; As $t\to\infty$,

$$\rho(t)\sim(3/\omega_L t)\sin\omega_L t,$$

which is somewhat different from (18), being proportional to t^{-1} rather than $t^{-\frac{1}{2}}$. This has also been observed by Prigogine and Bingen.[4]

We close this section with an analysis of $\rho_N(t)$ for large but finite N in a one dimensional assemble. We use Eulers summation formula[20]

$$\sum_{n>b}^{n\le b}f(n)=\int_a^b f(x)dx+u(b)f(b)-f(a)u(a)-\int_a^b u(x)f'(x)dx,$$

which is valid when $f(x)$ is continuous and has a continuous first derivative in the interval (a,b). Here

$$u(x)=[x]-x+\tfrac{1}{2},$$

where $[x]$ is the integral part of the number x (the largest integer contained in x). In our problem $b=N$ and $a=0$.

[20] c.f., T. V. Uspensky, *Introduction to Mathematical Probability* (McGraw-Hill Book Company, Inc., New York, 1937), p. 347.

We have

$$\frac{1}{N}\sum_{k=1}^{N}\cos\tau\left(\sin\frac{\pi k}{N}\right)=\frac{1}{N}\int_0^N\cos\tau\left(\sin\frac{\pi x}{N}\right)dx+\frac{\tau\pi}{N^2}\int_0^N u(x)\cos\frac{\pi x}{N}\sin\left(\tau\sin\frac{\pi x}{N}\right)dx.$$

If we let $\vartheta=\pi x/N$ and apply Eqs. (1) and (3), we find

$$\rho_N(t)=J_0(\tau)+\frac{\tau}{N}\int_0^\pi u(N\vartheta/\pi)\cos\vartheta\sin(\tau\sin\vartheta)d\vartheta. \tag{20}$$

Now $|u(x)|\le\frac{1}{2}$, and the absolute values of $\sin\varphi$ and $\cos\varphi$ are ≤ 1. Hence, the integral component of $\rho_N(t)$ is $\le\tau\pi/2N$. As long as $\tau\ll N$, $\rho_N(t)$ is well represented by $J_0(\tau)$. Actually it is still a good representation for much larger τ because when $\tau=0(N)$, $\sin(\tau\sin\vartheta)$ and $u(N\vartheta/\pi)$ oscillate and change sign very rapidly so that the integral itself approaches zero as τ or as $N\to\infty$ unless as occasionally happens sign changes of $u(N\vartheta/\pi)$ and $\sin(\tau\sin\vartheta)$ cancel each others. It is just this occasional proper phasing of these two functions which leads to Poincaré cycles in $\rho_N(t)$ for values of $\rho_N(t)$ of 0(1). A complete expansion of (20) is given in Appendix V.

6. CALCULATION OF QUANTUM MECHANICAL AUTOCORRELATION FUNCTION

The ensemble average of a dynamical operator B is

$$E\{B\}=\text{trace}B\rho, \tag{1}$$

where ρ is the density matrix associated with the ensemble. If we wish to find the quantum mechanical analog of the autocorrelation function (4.1) we must then find

$$E\{p_j(t)p_j(0)\}=\text{trace}\{p_j(t)p_j(0)\rho\}, \tag{2}$$

where ρ is the density matrix associated with the appropriate ensemble. Since, H is the Hamiltonian of our assembly,

$$p_j(t)=e^{iHt/\hbar}p_j(0)e^{-iHt/\hbar}, \tag{3}$$

the only operators which require an ensemble averaging are $p_j(0)$. If the assembly of interest belongs to a canonical ensemble at temperature T (with $\beta=1/kT$), the density matrix is $\rho=Z^{-1}\exp(-\beta H)$, Z being the partition function. Then, if we abbreviate $p_j\equiv p_j(0)$ and employ periodic boundary conditions so that all particles are equivalent, we have

$$F_N(t)=\frac{1}{N}\sum_i E\{p_j(t)p_j(0)\}$$

$$=\frac{1}{N}\sum_i E\{e^{itH/\hbar}p_j(0)e^{-itH/\hbar}p_j(0)\}$$

$$=\frac{1}{NZ}\sum_i\text{trace}\{e^{iHt/\hbar}p_je^{-iHt/\hbar}p_je^{-\beta H}\}. \tag{4}$$

As in the classical case, we transform to normal coordinates to find

$$F_N(t)$$

$$=\frac{1}{NZ}\sum_j \text{trace}\{e^{iHt/\hbar}\sum_k U_{jk}P_k e^{-iHt/\hbar}\sum_m U_{jm}P_m e^{-\beta H}\}$$

$$=\frac{1}{NZ}\sum_{k,m}\text{trace}\{e^{iHt/\hbar}P_k e^{-iHt/\hbar}P_m e^{-\beta H}\sum_j U_{jk}U_{jm}\}$$

$$=\frac{1}{NZ}\sum_k \text{trace}\{e^{iHt/\hbar}P_k e^{-iHt/\hbar}P_k e^{-\beta H}\} \quad (5)$$

(using 3.10). The following three identities:

$$\exp(-sH)\psi_j=\exp(-sE_j)\psi_j,$$

$$P_k\psi_j(P)=\sum_l \psi_l(P)\int\psi_l^*(P')P_k'\psi_j(P')d^NP',$$

and

$$\sum_m \psi_m^*(P')\psi_m(P'')=\delta(P'-P''),$$

yield

$$F_N(t)=\frac{1}{NZ}\int P_k''P_k'\{\sum_l e^{-iEit/\hbar}\psi_l(P'')\psi_l^*(P')$$

$$\times\{\sum_j e^{-(\beta-it/\hbar)E_j}\psi_j^*(P'')\psi_j(P')\}d^NP''d^NP'. \quad (6)$$

In view of the fact that the complete partition function Z factors into N single normal mode partition functions and that the wave functions factor into N independent wave function, we find, remembering that the energy levels of the kth oscillator are $\hbar\omega_k(j+\frac{1}{2})$ with $j=0, 1, 2\cdots$,

$$F_N(t)=\frac{1}{N}\sum_k Z_k^{-1}\sum_{l,j}\exp\{-i\omega_k t(l-j)-\beta\hbar\omega_k(j+\tfrac{1}{2})\}$$

$$\times(j|P_k|l)(l|P_k|j), \quad (7)$$

where (in a position representation)

$$(j|P|l)=\frac{\hbar}{i}\int_{-\infty}^{\infty}\varphi_j^*(x)\frac{\partial}{\partial x}\varphi_l(x)dx \quad (8)$$

is the matrix element of the momentum operator $P\equiv\hbar i^{-1}\partial/\partial x$. Here $\varphi_l(x)$ is the lth harmonic oscillator wave function. It is well known that

$$(\hbar/m\omega)^{\frac{1}{2}}d\varphi_n/dx=-[\tfrac{1}{2}(n+1)]^{\frac{1}{2}}\varphi_{n+1}(x)$$

$$+(\tfrac{1}{2}n)^{\frac{1}{2}}\varphi_{n-1}(x). \quad (9)$$

Substitution of this expression into (8) and application of the orthonormality condition $\int\varphi_j^*(x)\varphi_l(x)dx=\delta_{jl}$ yields

$$(j|p|l)=(\hbar/i)(m\omega/\hbar)^{\frac{1}{2}}$$

$$\times\{-[\tfrac{1}{2}(l+1)]^{\frac{1}{2}}\delta_{j,l+1}+[\tfrac{1}{2}l]^{\frac{1}{2}}\delta_{j,l-1}\}. \quad (10)$$

The matrix element $(l|p|j)$ is obtained by interchanging j and l in this formula. Substitution of these matrix elements into (7) yields

$$F_N(t)=\frac{m\hbar}{2N}\sum_k(\omega_k/Z_k)e^{-\frac{1}{2}\beta\hbar\omega_k}$$

$$\times[e^{it\omega_k}+e^{\beta\hbar\omega_k}e^{-it\omega_k}]\sum_{j=0}^{\infty}je^{-j\beta\hbar\omega_k}. \quad (11)$$

Since

$$\sum je^{-j\theta}=-(\partial/\partial\theta)\sum_j e^{-j\theta}$$

$$=-(\partial/\partial\theta)(1-e^{-\theta})^{-1}=e^{-\theta}/(1-e^{-\theta})^2$$

and since the harmonic oscillator partition function is $Z=\exp(-\frac{1}{2}\hbar\omega\beta)/[1-\exp(-\hbar\beta\omega)]$, we finally obtain

$$F_N(t)=\frac{m\hbar}{2N}\{\sum_k\omega_k\coth(\tfrac{1}{2}\beta\hbar\omega_k)\cos t\omega_k$$

$$-i\sum\omega_k\sin t\omega_k\}, \quad (12)$$

which is complex, as is generally the case for quantum mechanical correlation functions.

Now the required autocorrelation function is

$$\rho_N(t)=F_N(t)/F_N(0)$$

$$=\frac{\{\sum_k\omega_k\coth(\tfrac{1}{2}\beta\hbar\omega_k)\cos t\omega_k-i\sum\omega_k\sin t\omega_k\}}{\sum_k\omega_k\coth(\tfrac{1}{2}\beta\hbar\omega_k)}. \quad (13)$$

It reduces to the classical form (4.8) as $\hbar\to 0$.

We hope to discuss quantum mechanical Poincaré cycles at a later date.

APPENDIX I

Let $y(\omega)$ and $x(t)$ be related in the manner described in Eq. (2.7). Then

$$y(b)-y(a)=\frac{1}{2\pi}\int_{-\infty}^{\infty}x(t)dt\int_a^b e^{-i\omega t}d\omega.$$

Hence,

$$|y(b)-y(a)|^2=\frac{1}{4\pi^2}\int\int x(t)x^*(t')dtdt'$$

$$\times\int_a^b\int e^{-i(\omega t-\omega' t')}d\omega d\omega',$$

so that [in view of (2.1) and (2.5)]

$$E\{|y(b)-y(a)|^2\}$$

$$=\frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int\rho(t-t')dtdt'\int_a^b\int e^{-i(\omega t-\omega' t')}d\omega d\omega'$$

$$=\frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_a^b\int dF(\omega'')d\omega d\omega'$$

$$\times\int\int e^{-i(t-t')(\omega-\omega'')}e^{it'(\omega'-\omega)}dtdt'.$$

Let $\tau=t-t'$ and integrate first with respect to τ. Then,

$$E\{|y(b)-y(a)|^2\} = \int_{-\infty}^{\infty} dF(\omega'') \int_a^b \int_a^b d\omega d\omega' \delta(\omega-\omega'')\delta(\omega-\omega').$$

Since both ω and ω' are restricted to the same interval (a,b),

$$E\{|y(b)-y(a)|^2\} = \int_{-\infty}^{\infty} dF(\omega'') \int_a^b \delta(\omega-\omega'')d\omega.$$

Now the integral over ω is 1 if $a<\omega''<b$, and 0 otherwise. Hence

$$E\{|y(b)-y(a)|^2\} = F(b)-F(a).$$

In particular, as $\mathcal{E}\to 0$,

$$E\{|y(\omega+\mathcal{E})-y(\omega-\mathcal{E})|^2\} = F(\omega+\mathcal{E})-F(\omega-\mathcal{E}) = dF(\omega).$$

APPENDIX II

Let

$$H(t)=\begin{cases}1 & \text{if } |t|<T\\ 0 & \text{if } |t|>T.\end{cases}$$

Then $H(t)$ has the fourier integral representation

$$H(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin\omega T}{\omega} e^{-i\omega t} d\omega.$$

Hence,

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x(t)dt = \lim_{T\to\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty} x(t)dt \int_{-\infty}^{\infty} d\omega \frac{\sin\omega T}{\omega T} e^{-i\omega T}.$$

Now we employ the fourier integral representation (2.7) of $x(t)$ and interchange orders of integration. Then

$$\lim_{T\to\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega \frac{\sin\omega T}{\omega T}\int_{-\infty}^{\infty} dy(u)\int_{-\infty}^{\infty} e^{-it(u-\omega)}dt$$

$$= \lim_{T\to\infty}\int_{-\infty}^{\infty} d\omega \frac{\sin\omega T}{\omega T}\int_{-\infty}^{\infty}\delta(u-\omega)dy(u)$$

$$= \lim_{T\to\infty}\int_{-\infty}^{\infty}\frac{\sin uT}{uT}dy(u) = \lim_{T\to\infty} I_1+I_2+I_3,$$

where, if we let $h=T^{-(1+\mathcal{E})}$,

$$(I_1,I_2,I_3)=\left(\int_{-h}^{h},\int_{h}^{\infty},\int_{-\infty}^{-h}\right)\frac{\sin uT}{uT}dy(u).$$

Since $|uT|<|T^{-\mathcal{E}}|$ in the interval $(-T^{-(1+\mathcal{E})}, T^{-(1+\mathcal{E})})$, one can by choosing T sufficiently large make $(\sin ut)/uT$ as close to unity as is desired throughout the entire interval. Hence,

$$\lim_{T\to\infty} I_1 = \lim_{h\to 0}\int_{-h}^{h} dy(u) = y(0+)-y(0-).$$

If $y(u)$ is of bounded variation it is easy to show by standard methods that $\lim(I_2+I_3)=0$. Hence

$$\lim\frac{1}{2T}\int_{-T}^{T} x(t)dt = y(0+)-y(0-).$$

APPENDIX III. STATISTICAL PROPERTIES OF FUNCTIONS WHICH CAN BE EXPRESSED AS FINITE FOURIER SERIES

We now derive the theorem of Kac concerning the average number of times a function

$$x(t)=\sum_{k=1}^{n} a_k\cos(t\omega_k+\alpha_k)$$

achieves a value q in a time interval of a given length. The a_k's and ω_k's are postulated to be real and the ω_k's to be linearly independent.

The required mean frequency can be written as

$$L(q)=\lim_{T\to\infty}\frac{1}{2T}N_T(q), \tag{2}$$

where $N_T(q)$ is the number of zeros of

$$F(t)=x(t)-q \tag{3}$$

in the interval $-T<t<T$.

In the calculation of $L(q)$ we use the Kronecker-Weyl theorem[21] which allows one to replace time averages by phase averages. Let a multiply periodic function $f(\vartheta_1,\cdots,\vartheta_n)\equiv f(\vartheta)$ be defined on an n-dimensional torus on which each ϑ_j ranges from 0 to 2π. $f(\vartheta')\equiv f(\vartheta'')$ if $\vartheta_j'\equiv\vartheta_j''(\mathrm{mod}2\pi)$ for all j. Finally, if

$$\vartheta_j = t\omega_j+\alpha_j \tag{4}$$

and the ω_j's are linearly independent, i.e., if there exists no set of integral m_k's (positive or negative) such that

$$\sum \omega_k m_k = 0, \tag{5}$$

then one can prove the result of Weyl that

$$\frac{1}{2T}\int_{-T}^{T} f[\vartheta(t)]dt \to \frac{1}{(2\pi)^n}\int_0^{2\pi}\cdots\int f(\vartheta_1\cdots\vartheta_n)d\vartheta_1\cdots d\vartheta_n \tag{6}$$

as $T\to\infty$.

We now show that in the case of linearly independent ω_k's the Kronecker-Weyl theorem and some delta function arguments lead immediately to a formula

[21] H. Weyl, Am. J. Math. **60**, 889 (1938).

derived more rigorously by Kac,

$$L(q)=\frac{1}{2\pi^2}\int_{-\infty}^{\infty}\int \eta^{-2}\cos q\alpha\{\prod_{k=1}^{n}J_0(|a_k|\alpha) - \prod_{k=1}^{n}J_0(|a_k|[\alpha^2+\eta^2\omega_k^2]^{\frac{1}{2}})\}d\alpha d\eta. \quad (7)$$

Let $F(t)$ be a function of t which is real for real values of t. Let its zeros be at $t_1, t_2\cdots$. Then the number of zeros in the interval $(-T, T)$ is

$$\int_{-T}^{T}\sum_i \delta(t-t_j)dt. \quad (8)$$

Since complex roots occur in pairs, say at points $a+ib$ and $a-ib$ the delta function sum which go with the pair is

$$\delta([t-a]-ib)+\delta([t-a]+ib)=\text{const}\delta([t-a]^2+b^2)$$

and

$$\int_{-\infty}^{\infty}\delta[(t-a)^2+b^2]dt=0,$$

(since the $[(t-a)^2+b^2]$ cannot vanish for any real value of t). Hence, as we would expect, complex roots never contribute to (8).

The well-known delta function formula

$$F(t)=\sum_i \frac{\delta(t-t_j)}{|F'(t_j)|}$$

[the t_j's being the zeros of $F(t)$] implies that (8) is equivalent to

$$\int_{-T}^{T}|F'(t)|\sum_i\frac{\delta(t-t_j)}{|F'(t_j)|}dt=\int_{-T}^{T}|F'(t)|\delta[F(t)]dt. \quad (9)$$

Then, if we employ the Fourier integral representation of $\delta|F(t)|$, and interchange orders of integration, (2) becomes

$$L(q)=\lim_{T\to\infty}\frac{1}{4\pi T}\int_{-\infty}^{\infty}\int_{-T}^{T}|F'(t)|\exp i\alpha F(t)dtd\alpha. \quad (10)$$

Finally, the formula

$$|u|=-\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1-\cos\eta u}{\eta^2}d\eta$$

yields

$$L=\lim_{T\to\infty}\frac{1}{4\pi^2 T}\int_{-\infty}^{\infty}\int\int_{-T}^{T}\eta^{-2}d\eta d\alpha\{1-\cos[\eta F'(t)]\}\times\exp[i\alpha F(t)]dtd\alpha$$

$$=\frac{1}{2\pi^2}\int_{-\infty}^{\infty}\int\eta^2\{U(\alpha,0)-\tfrac{1}{2}U(\alpha,\eta)-\tfrac{1}{2}U(\alpha,-\eta)\}d\eta d\alpha,$$

where

$$U(\alpha,\beta)=E\{\exp i[\alpha F(t)+\beta F'(t)]\}. \quad (12)$$

This characteristic function is readily obtained for our special case $F(t)=x(t)-q$ since the Kronecker-Weyl theorem allows us to replace time averages by averages over our ϑ torus. Then

$$U(\alpha,\beta)=(\exp i\alpha q)\prod_{j=1}^{n}\frac{1}{2\pi}\int_0^{2\pi}\times\exp i[\alpha a_j\cos\vartheta_j-\beta a_j\omega_j\sin\vartheta_j]d\vartheta_j$$

$$=(\exp i\alpha q)\prod_{j=1}^{n}J_0[|a_j|(\alpha^2+\beta^2\omega_j^2)^{\frac{1}{2}}]. \quad (13)$$

Substitution of (13) and (12) into (11) gives us the Kac formula (7).

Consider the special case $a_1=a_2=\cdots=a=1$. Since

$$J_0(z)=\{1-(\tfrac{1}{2}z)^2+\tfrac{1}{2}\cdot\tfrac{1}{2}(\tfrac{1}{2}z)4-\cdots\}$$
$$=\{1-\tfrac{1}{2}\cdot\tfrac{1}{2}(\tfrac{1}{2}z)^4+0(z^6)\}\exp-(\tfrac{1}{2}z)^2 \quad (14)$$

is peaked at $z=0$, the product (13) of Bessel functions (14) becomes even more peaked in the limit as $n\to\infty$. Then (13) becomes

$$U(\alpha,\beta)=\exp[i\alpha q-(n/4)(\alpha^2+\beta^2\omega_0^2)]\times\left\{1-\frac{1}{64}\sum_j(\alpha^2+\beta^2\omega_j^2)^2+\cdots\right\},$$

where

$$\omega_0^2=\frac{1}{n}\sum\omega_j^2.$$

Now let $q=bn^{\frac{1}{2}}$, $\alpha n^{\frac{1}{2}}=x$, $\eta n^{\frac{1}{2}}=y$, and let $n\to\infty$. Then, if

$$\lim\frac{1}{n^2}\sum\omega_j^4=0.$$

We have, as $n\to\infty$,

$$L(bn^{\frac{1}{2}})\sim\frac{1}{2\pi^2}\int_{-\infty}^{\infty}\int\exp\{ibx-\tfrac{1}{4}x^2\}\times(1-\exp[-\tfrac{1}{4}y^2\omega_0^2])y^{-2}dxdy=(\omega_0/\pi)\exp-b^2,$$

which has also been derived by Kac.

APPENDIX IV. AUTOCORRELATION FUNCTION OF SQUARE LATTICE

The behavior of the autocorrelation function $\rho(t)$ as $t\to\infty$ can be obtained from (5.4) by employing the form (3.3) for the frequency spectrum (with $\gamma_1>\gamma_2$). Logrithmic singularities exist at $\omega=\omega_1$, and $\omega=\omega_2$. If we substitute (3.3) into (5.4), break up the integral into three integrals of ranges $(0,\omega_1)$, (ω_1,ω_2), and (ω_2,ω_L),

and integrate by parts we find

$$\rho(t)=\lim_{\mathcal{E}\to 0}\Big\{\frac{1}{t}\sin t\omega_2[g([\omega_2^2-\mathcal{E}]^{\frac{1}{2}})-g([\omega_2^2+\mathcal{E}]^{\frac{1}{2}})]$$
$$+\frac{1}{t}\sin t\omega_1[g([\omega_1^2-\mathcal{E}]^{\frac{1}{2}})-g([\omega_1^2+\mathcal{E}]^{\frac{1}{2}})]$$
$$+\frac{1}{t}\sin t\omega_L g([\omega_L^2-\mathcal{E}]^{\frac{1}{2}})-\int_0^{\omega_L} g'(\omega)\sin t\omega dt\Big\}.$$

However, it is easy to show that, as $\mathcal{E}\to 0$,

$$g([\omega_2^2-\mathcal{E}]^{\frac{1}{2}})\sim\frac{4}{\pi^2\omega_1}K\{[(1-\mathcal{E}\omega_2^{-2})(1+\mathcal{E}\omega_1^{-2})]^{\frac{1}{2}}\}$$
$$=\frac{4}{\pi^2\omega_1}K(1+\tfrac{1}{2}\mathcal{E}[\omega_1^{-2}-\omega_2^{-2}])$$

$$g([\omega_2^2+\mathcal{E}]^{\frac{1}{2}})\sim\frac{4}{\pi^2\omega_1}K([1+\mathcal{E}\omega_2^{-2})(1-\mathcal{E}\omega_2^{-2})]^{-\frac{1}{2}})$$
$$=\frac{4}{\pi^2\omega_1}K(1+\tfrac{1}{2}\mathcal{E}[\omega_1^{-2}-\omega_2^{-2}])$$

so that as $\mathcal{E}\to 0$

$$g([\omega_2^2-\mathcal{E}]^{\frac{1}{2}})-g([\omega_2^2-\mathcal{E}]^{\frac{1}{2}})\to 0.$$

Also

$$g([\omega_1^2-\mathcal{E}]^{\frac{1}{2}})\sim\frac{4}{\pi^2\omega_2}K([(1-\mathcal{E}\omega_1^{-2})(1+\mathcal{E}\omega_2^{-2})]^{-\frac{1}{2}})$$
$$=\frac{4}{\pi^2\omega_2}K(1+\tfrac{1}{2}\mathcal{E}[\omega_1^{-2}-\omega_2^{-2}])$$

$$g([\omega_1^2+\mathcal{E}]^{\frac{1}{2}})\sim\frac{4}{\pi^2\omega_2}K([(1+\mathcal{E}\omega_1^{-2})(1-\mathcal{E}\omega_2^{-2})]^{\frac{1}{2}})$$
$$=\frac{4}{\pi^2\omega_2}K[1+\tfrac{1}{2}\mathcal{E}(\omega_1^{-2}-\omega_2^{-2})],$$

so that, as $\mathcal{E}\to 0$,

$$g([\omega_1^2-\mathcal{E}]^{\frac{1}{2}})-g([\omega_1^2-\mathcal{E}]^{\frac{1}{2}})\to 0.$$

Then, since, as $\mathcal{E}\to 0$,

$$g([\omega_L^2-\mathcal{E}]^{\frac{1}{2}})\sim\frac{4\omega_L}{\pi^2\omega_1\omega_2}K(0)=\frac{2\omega_L}{\pi\omega_1\omega_2},$$

we have

$$\rho(t)=\frac{2\omega_L\sin t\omega_L}{\pi t\omega_1\omega_2}-\frac{1}{t}\int_0^{\omega_L} g'(\omega)\sin\omega t dt$$
$$\sim(2\omega_L\sin t\omega_L)/\pi t\omega_1\omega_2.$$

APPENDIX V

The correlation function,

$$\rho_N(t,r)=\frac{1}{M^n}\sum_k\cos\frac{2\pi r\cdot k}{M}\cos t\omega_k \tag{1}$$

has the property

$$\rho_N(t,0)=\delta_{r,0}. \tag{2}$$

It easily can be verified to satisfy the equation

$$m\frac{d^2\rho_N}{dt^2}=\sum_i\gamma[_j\rho_N(t,\cdots,r_{j+1},\cdots)$$
$$-2\rho_N(t,\cdots r_j\cdots)+\rho_N(t,\cdots r_{j-1}\cdots)] \tag{3}$$

with boundary conditions,

$$\rho_N(t,r+sM)=\rho_N(t,r), \tag{4}$$

where s is any vector with integral components.

One can easily verify that

$$\rho_N(t,r)=\sum_{k=-\infty}^{\infty}J_{2(r+kN)}(t\omega_L) \tag{5}$$

is the solution of (2), (3), and (4) when $n=1$ (linear chain). Here $J_\alpha(t)$ is the Bessel function of order α.

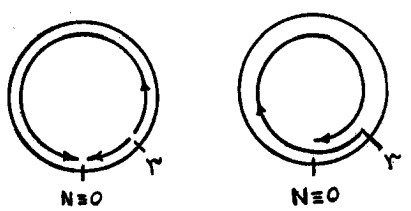

FIG. 4.

This result has a simple physical interpretation. The point 0 correlates with r on an infinite lattice through $J_{2r}(t\omega_L)$ (see Eq. 5.3). Hence, with our periodic boundary conditions the correlation of point 0 (which we identify with N) with r in a clockwise direction is also $J_{2r}(t\omega_L)$. In the counterclockwise direction, it is

$$J_{2(N-r)}(t\omega_L)\equiv J_{2(r-N)}(t\omega_L).$$

It is also correlated through multicirculation paths (see Fig. 4). The total correlation length in Fig. 4(b) is $(N+r)$ so that the term associated with this diagram is

$$J_{2(N+r)}(t\omega_L).$$

Finally, if one adds the contribution of all paths (clockwise and counterclockwise) which encircle our one dimensional lattice any number of times one merely obtains (5).